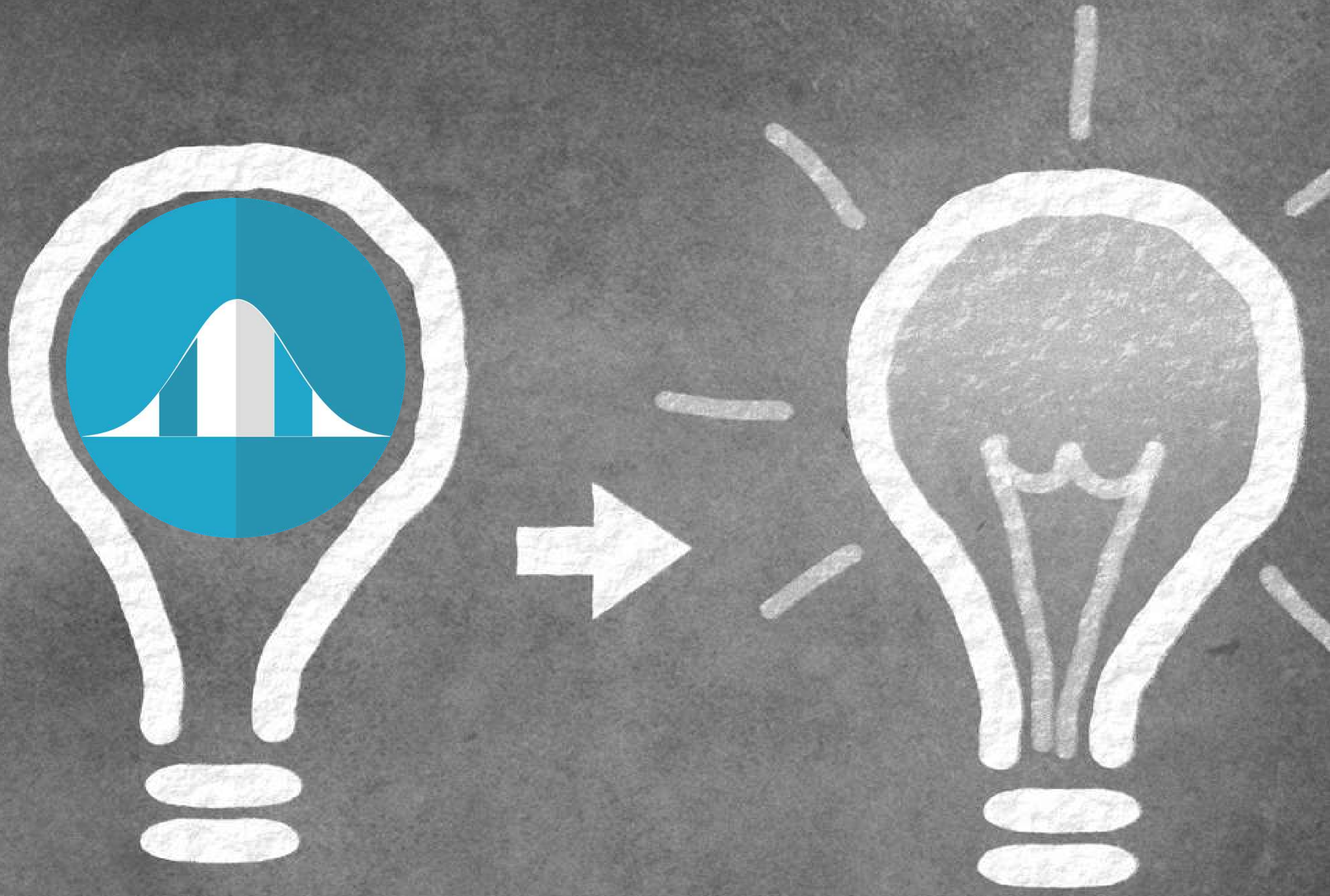
# DISEÑO DE CUESTIONARIOS Y CREACIÓN DE ESCALAS



*Uso del EQS. en las ciencias económico-administrativas*

*Juan Mejía Trejo*

# DISEÑO DE CUESTIONARIOS y CREACIÓN DE ESCALAS

## Uso de EQS en las ciencias económico-administrativas

# DISEÑO DE CUESTIONARIOS y CREACIÓN DE ESCALAS
# Uso de EQS en las ciencias económico-administrativas

.Investigación y fuentes de información

.Importancia de realizar mediciones

.Dimensionalidad

.Confiabilidad

.Validez

.Constructo y contenido

.Análisis factorial exploratorio

.Análisis factorial confirmatorio

**Juan Mejía Trejo**

**Este libro fue sometido a un proceso de evaluación por pares , en condiciones de anonimato, de acuerdo con la normatividad que regula la producción editorial de Centro Universitario de Ciencias Económico Administrativas de la Universidad de Guadalajara.**

Primera edición 2019
D.R. @ Universidad de Guadalajara
Centro Universitario de Ciencias Económico
Administrativas Periférico Nte. 799, núcleo
universitario Los Belenes
45100, Zapopan, Jalisco
ISBN: **9786075384672**
Impreso y hecho en México
Printed and made in Mexico

# Índice

**INTRODUCCIÓN** ..... **1**

**INVESTIGACIÓN y FUENTES DE INFORMACIÓN. ANTECEDENTES** ..... **3**

**¿Qué es investigar?** ..... **3**
**Enfoques cuantitativo y cualitativo. Sus diferencias** ..... **8**
**¿Existe algún enfoque que sea mejor?** ..... **14**
**El enfoque mixto** ..... **14**
**Alcances de una investigación** ..... **20**
**Los alcances ¿todos pueden estar en una misma investigación?** ..... **23**
**La idea, el conocimiento del tema de investigación y su perspectiva** ..... **24**
**¿Cuál de los cuatro alcances para un estudio es el mejor?** ..... **26**
**El problema y su planteamiento** ..... **26**
**El marco teórico** ..... **27**
¿Cómo organizar el marco teórico? ..... 30
Confirmando una revisión adecuada de la literatura ..... 30
**Fuentes de información** ..... **32**
Por su disponibilidad ..... 32
Por su tendencia a tratamiento cuantitativo ..... 33
Por su grado de dinamismo: estáticas y dinámicas ..... 33
Por su naturaleza ..... 33
Por su grado de contacto ..... 50
Por su grado de colaboración ..... 50
Por su proceso de investigación ..... 50
**Características de las fuentes de información en las ciencias económico-administrativas** ..... **51**
**Fuentes descriptivas de información** ..... **52**
**El Cuestionario** ..... **54**
**Etapas y proceso de diseño** ..... **55**
**Medición. Escalas y sus propiedades** ..... **72**
**Clases de Escalas** ..... **72**
Escalas Primarias ..... 73
Escalas comparativas ..... 76
Escalas no comparativas ..... 78
Escalas Itemizadas ..... 79
**Confiabilidad de las escalas** ..... **86**
**Medición de la confiabilidad de las escalas** ..... **87**
**Influencia del diseño de una escala sobre la confiabilidad** ..... **90**
**Validez de las escalas. Primera aproximación** ..... **91**
**Criterios de valoración de las escalas** ..... **92**
**Fuentes aleatorias de información** ..... **93**
**Formulación de hipótesis** ..... **119**

¿Qué y cuántas hipótesis deben formularse en una investigación ? ..... 127
Probando las hipótesis ..... 128
**Diseñando la investigación ..... 129**
**Muestreo cuantitativo ..... 134**
**Muestreo cualitativo ..... 142**

**CAPÍTULO 1. LA IMPORTANCIA DE REALIZAR MEDICIONES EN RELACIÓN A LAS CIENCIAS ECONÓMICO-ADMINISTRATIVAS ..... 147**
**¿Qué es medir? ..... 147**
**¿Qué es una escala de medición? ..... 149**
**La utilidad de realizar mediciones en las ciencias económico-administrativas ..... 150**
**Realizando la escala de constructos subyacentes con múltiples elementos ..... 151**
**Constructo latente. ..... 152**
**Diagramas de trayectorias ..... 155**
**La medición y su error ..... 158**
**Tendencias en la dimensionamiento de los constructos para las ciencias económico-administrativas ..... 159**
**Marco teórico y validez de los constructos subyacentes ..... 160**
**Importancia de la revisión de la literatura del estado del arte en el marco teórico ..... 161**
**Propiedades de una medición: dimensionalidad, confiabilidad y validez ..... 162**
Dimensionalidad ..... 163
Confiabilidad ..... 163
Validez ..... 164
**El rol del investigador en las ciencias económico-administrativas ..... 167**
El Cientificismo ..... 168
Determinación del problema a investigar ..... 168
**La medición en las ciencias económico-administrativas ..... 169**
**Concepto y variables: definición y tipos ..... 172**
**Etapas para el desarrollo de una escala ..... 174**
**Conclusión ..... 176**

**CAPÍTULO 2. DIMENSIONALIDAD ..... 177**
**Unidimensionalidad de constructo, de indicadores y de conjunto de indicadores ..... 178**
**Multidimensionalidad de constructo, de indicadores y de conjunto de indicadores ..... 180**
**Implicaciones de la unidimensionalidad ..... 182**
**Importancia de la unidimensionalidad ..... 185**
**Evaluando la dimensionalidad de los constructos ..... 186**
**Análisis Factorial Exploratorio (EFA. *Exploratory Factor Analysis*) ..... 187**
**Análisis Factorial Confirmatorio (CFA. *Confirmatory Factor Analysis*) ..... 194**
**Conclusión ..... 196**

**CAPÍTULO 3. CONFIABILIDAD ..... 197**
**La importancia de medir la confiabilidad ..... 198**

**Confiabilidad *test-retest*** ..... **200**
**Confiabilidad alternativa** ..... **201**
**Confiabilidad por consistencia interna** ..... **202**
Confiabilidad por consistencia interna: división en mitades ..... 202
Confiabilidad por consistencia interna: coeficiente de alfa (α) ..... 204
Coeficiente alfa (α). Explicación 2. ..... 210
**Coeficiente de alfa (α) y dimensionalidad** ..... **219**
**Coeficiente de alfa (α), longitud de escala, correlación y redundancia entre indicadores** ..... **220**
**Índice de fiabilidad compuesta (IFC)** ..... **223**
**Índice de la varianza compuestas extraída (IVE/AVE)** ..... **227**
**Conceptos adicionales de la confiabilidad** ..... **229**
**Conclusión** ..... **229**

**CAPÍTULO 4. VALIDEZ** ..... **230**
**Validez de versión (*translation validity*)** ..... **231**
Validez aparente (face validity) ..... 231
Validez de contenido (content validity) ..... 232
**Validez relacionada con el criterio (*criterion-related validity*)** ..... **236**
Validez predictiva y post-dictiva (predictive and post-dictive validity) ..... 236
Validez concurrente (concurrent validity) ..... 236
Validez convergente (convergent validity). ..... 237
Validez discriminante (discriminant validity) ..... 241
Validez del grupo conocido (known-group validity) ..... 246
**Validez nomológica (*nomological validity*)** ..... **248**
Estudio de caso: administración del conocimiento OECD(2003) ..... 249
Interpretación ..... 251
**Ejemplo: respuesta socialmente deseable (*socially desirable response bias*)** ..... **251**
**Conclusión** ..... **253**

**CAPÍTULO 5. DEFINIENDO EL CONSTRUCTO Y SU CONTENIDO (Etapas: 1 y 2)** ..... **256**
**Etapa 1: El Constructo** ..... **257**
**La importancia de la definición de un constructo claro y preciso** ..... **257**
**La importancia de una teoría sólida, revisión de la literatura y su análisis** ..... **258**
**El enfoque de indicadores reflectivos (*de efecto*) vs. indicadores formativos (*de causa*)** ..... **260**
**El papel de la teoría al especificar la dimensionalidad** ..... **261**
**Etapa 2: Generando y Analizando los Indicadores de Contenido** ..... **263**
**La muestra y su contenido** ..... **264**
**Generando el *pool* de indicadores** ..... **264**
**Fuentes de indicadores** ..... **265**
**Escritura de los indicadores** ..... **266**
Claridad de la redacción ..... 266
La redacción redundante ..... 267
La redacción en sentido positivo o negativo ..... 268

La elección de formato de respuesta ..... 268
Número de indicadores del pool inicial ..... 270
Validez de contenido (content validity) y validez aparente (face validity) ..... 271
Guía para validez de ambas ..... 271
**Guía para el desarrollo de escalas ..... 272**
Determinar qué es lo que se pretende medir ..... 272
Generar un listado de ítems ..... 273
Determinar el formato de medida ..... 274
El listado de ítems debe ser revisado por expertos ..... 277
Considerar la inclusión de ítems de validación ..... 278
Administrar la escala a una muestra de prueba ..... 278
Evaluando los ítems ..... 278
Optimizar la longitud de la escala ..... 279
**Conclusión ..... 279**

**CAPÍTULO 6. DISEÑANDO LA ESCALA. ANÁLISIS FACTORIAL EXPLORATORIO (Etapa 3)..... 280**
**¿Qué es el análisis factorial? ..... 282**
**Prueba piloto ..... 284**
**Análisis factorial y el proceso de decisión ..... 286**
**Objetivos ..... 289**
La identificación de estructura mediante el resumen de datos ..... 289
Reducción de datos ..... 289
El uso del análisis factorial con otras técnicas multivariantes ..... 290
Selección de variables ..... 290
**Diseño ..... 291**
Las correlaciones entre las variables o los encuestados ..... 292
La selección de variables y cuestiones de medición ..... 293
Tamaño muestra ..... 293
**Supuestos ..... 294**
**Estimación y Ajuste ..... 296**
Análisis factorial común vs. Análisis de componentes principales ..... 296
Problemas ..... 298
Criterios para el cálculo del número de factores a ser extraídos ..... 298
**Interpretación de factores ..... 303**
Métodos de rotación ..... 304
Criterios de rotación de factores ..... 305
Métodos de rotación ortogonal ..... 307
Métodos de rotación oblicua ..... 308
Selección del método de rotación ..... 308
Retención de indicadores ..... 309
Criterios para la significación de la carga factorial ..... 309
Interpretación de la matriz de factores ..... 311

**Validación** ..... **313**
Prueba inicial de indicadores y de confiabilidad ..... 314
**Usos adicionales de los resultados** ..... **315**
Selección de variables suplentes para el análisis subsiguiente ..... 316
Problemas ..... 316
Creación de escalas aditivas ..... 317
Cálculo de la puntuación factorial ..... 320
Selección entre los 3 métodos ..... 320
**Análisis factorial: Resumen para aplicar** ..... **321**
**Caso de estudio** ..... **323**
Problema 1 ..... 323
Problema 2 ..... 326
Problema 3 ..... 328
Problema 4 ..... 331
Problema 5 ..... 333
Problema 6 ..... 334
Problema 7 ..... 348
**Conclusión** ..... **357**

**CAPÍTULO 7. FINALIZANDO LA ESCALA. ANÁLISIS FACTORIAL CONFIRMATORIO Y ECUACION ESTRUCTURAL (Etapa 4)** ..... **358**
**Variables latentes vs. observadas** ..... **363**
**Variables latentes exógenas vs endógenas** ..... **363**
**El modelo de análisis factorial** ..... **364**
**Modelo completo de variable latente** ..... **366**
**Modelaje estadístico de propósito general y proceso** ..... **366**
**Notación simbólica** ..... **368**
**Diagrama de trayectoria** ..... **369**
**La ecuación estructural (SEM) y su historia** ..... **370**
**Componentes no visibles de un modelo** ..... **380**
**Modelo SEM. Composición básica** ..... **380**
**Notación** ..... **409**
**Diagrama de trayectoria** ..... **409**
**Representación del sistema Bentler-Weeks** ..... **410**
**EQS. Uso del programa** ..... **412**
**EQS. Componentes del archivo de entrada** ..... **412**
**EQS. Reglas básicas en la creación de archivos de entrada** ..... **412**
Keywords. (Palabras clave) ..... 412
Descriptive Statements ..... 413
File Editors ..... 413
Basic Components of the EQS Input File ..... 414
/TITLE (Optional) ..... 414

/SPECIFICATIONS (SPEC; Obligatorio) ..... 414
Data (DA)..... 415
Variables (VAR) ..... 415
Cases (CAS)..... 415
**Method (ME)** ..... 415
Analysis. (ANAL). ..... 416
Matrix (MA)..... 417
/LABELS (Opcional)..... 418
/EQUATIONS (EQU; Obligatorio)..... 418
**EQS. Cómo escribir ecuaciones..... 419**
**EQS. /VARIANCES (VAR; Obligatorio) ..... 420**
**EQS. /COVARIANCES (COV; Opcional)..... 421**
**EQS.Casos de ejemplos de archivos de entrada ..... 422**
Modelo de primer orden ..... 422
Modelo de segundo orden..... 427
SEM. Modelo completo ..... 429
**Modelo de primer orden. Análisis de caso de estudio 1..... 431**
Problema..... 431
Formulación de matriz ..... 432
Condiciones de aplicabilidad..... 435
Identificación ..... 436
Estimación..... 438
Entrada de datos al programa ..... 439
EQS. Cómo ejecutar el programa..... 442
EQS. Análisis de resultados ..... 444
Discusión..... 455
Reespecificación del modelo ..... 457
Confiabilidad ..... 458
Cálculo IFC..... 461
Redefiniendo el modelo..... 463
Recálculo de IFC y cálculo IVE ..... 467
Determinando la validez ..... 468
Calculando la ecuación estructural (SEM)..... 477
**Modelo de segundo orden. Análisis de caso de estudio 2. ..... 488**
EQS. Entrada de datos al programa ..... 490
Análisis de resultados ..... 493
Eliminando V1 ..... 506
Discusión rápida de resultados, sin V1..... 510
Eliminando V3 ..... 510
Discusión rápida de resultados, sin V1 y V3..... 515
Cálculo del alfa de Cronbach, sin V1 y V3 ..... 515
Cálculo de IFC, sin V1 y V3 ..... 518

Cálculo de IVE, sin V1 y V3 ..... 519
Análisis de la validez de la escala, sin V1 y V3 ..... 520
Reportando resultados: consistencia interna y validez convergente ..... 525
Conclusión ..... 526
**Modelo de segundo orden. Análisis de caso de estudio 3 ..... 529**
EQS. Entrada de datos al programa ..... 531
Discusión rápida de resultados. ..... 537
Discusión rápida de resultados sin V14, V15 y V16 ..... 538
Cálculo del alfa de Cronbach, sin V14,V15 y V16 ..... 540
Cálculo de IFC, sin V14,V15 y V16 ..... 543
Cálculo de IVE, sin V14,V15 y V16 ..... 545
Análisis de la validez de la escala, sin V14,V15 y V16 ..... 546
**Conclusión final SEM y CFA ..... 558**

**Apéndice 1 ..... 569**
**Análisis de validez y confiabilidad: ¿Qué es? ..... 569**
El Alfa de Cronbach ..... 570
Ejemplo ..... 571

**Apéndice 2 ..... 578**
Antecedentes ..... 578
Consideraciones previas ..... 579
Cálculo de la muestra ..... 579
Fórmulas para calcular la muestra en estudios descriptivos ..... 581
Para estudios cuya variable principal es de tipo cuantitativo ..... 581
Para estudios cuya variable principal es de tipo cualitativo: ..... 581
Fórmulas para calcular la muestra en estudios explicativos ..... 582
Cálculo de muestra para prueba de hipótesis en una proporción de población ..... 583
Cálculo de muestra para prueba de hipótesis en estudios de casos y controles ..... 584
Cálculo de muestra para probar hipotesis en estudios de cohortes ..... 585
Cálculo de muestra para asociar dos variables cuantitativas utilizando el coeficiente de correlación de Pearson ..... 585

**BIBLIOGRAFÍA ..... 587**

# INTRODUCCIÓN

La investigación es una actividad permanente del ciclo: entre análisis, prueba de campo y síntesis que requiere de una gran agudeza por parte de quien lo practica, para discernir de las fuentes de información y recursos que empleará en la misma. Aún más, obtener reconocimiento y credibilidad en los resultados de investigación científica depende de qué estrategia esclarece y justifica mejor, la técnica de medición empleada.

Así, es posible plantear diversas estrategias que permiten desarrollar y afinar las mediciones, sin embargo, su verdadero impacto dependerá del tipo de fenómeno científico que esté siendo medido, tanto con variables observables directamente como de sus relaciones subyacentes. De ésta manera y basados en el enfoque y alcances que su investigación demanda es que las ciencias económico-administrativas deberán centrar su interés en el planteamiento, diseño, desarrollo y validación de constructos tanto de relaciones directas como subyacentes, Particularmente, éste último tipo de relaciones implica constructos que al representar abstracciones sólo pueden evaluarse de manera indirecta.

La evaluación indirecta implica el diseño y uso de múltiples elementos (en nuestro caso, indicadores) que miden al constructo, es decir, plantean la escala para medir. Por lo tanto, partiendo de un problema de investigación determinado y caracterizados las factores, variables e indicadores que mejor la describan en forma de constructos, el objetivo principal de este documento es el de plantear y diseñar las relaciones de dichos factores, variables e indicadores para discutir cómo medirlos y validarlos a través de la creación de escalas.
Para lograr lo anterior, se proporciona al lector:

INVESTIGACIÓN y FUENTES DE INFORMACIÓN. ANTECEDENTES con breves antecedentes de los que consta una idea de investigación, sus enfoques y alcances, el planteamiento del problema, los objetivos a alcanzar, y cómo conformar el marco teórico. Se ofrecen recomendaciones para hacer una mejor captura de la información. Se hace especial énfasis en la descripción de las fuentes de información, cómo tratarlas según el tipo de investigación a realizar y se exponen diversas características y condiciones de instrumentos tales como: el diseño de páneles, las entrevistas, los cuestionarios, etc.

CAPÍTULO 1. LA IMPORTANCIA DE REALIZAR MEDICIONES EN RELACIÓN A LAS CIENCIAS ECONÓMICO-ADMINISTRATIVAS el cual, justifica y enfatiza la importancia de la teoría del desarrollo de escalas a partir de explicar conceptos clave y básicos, como lo son la: dimensionalidad, confiabilidad y validez.
Se destaca, la importancia del enfoque de cuatro etapas (el constructo; el contenido; el análisis factorial exploratorio y el análisis factorial confirmatorio) en el desarrollo de una escala que involucra a estos tres conceptos.

CAPÍTULO 2. DIMENSIONALIDAD. Donde se esclarece cómo ésta etapa permite definir el alcance del constructo y el modelo final, con el fin de entender dos técnicas base en toda investigación de las ciencias económico-administrativas: el análisis factorial exploratorio y el análisis factorial confirmatorio.

CAPÍTULO 3. CONFIABILIDAD. En este apartado, se explica la importancia del concepto y los diferentes tipos de confiabilidad existentes, la implicación de lo que es la consistencia interna, la fiabilidad y varianza compuestas en el diseño de la escala.
CAPÍTULO 4. VALIDEZ. Es aquí donde se expone el impacto que tiene la validez del instrumento a diseñar, los diferentes tipos enfatizando la impirtancia de la validez discriminante y nomológica para el análisis factorial confirmatorio.

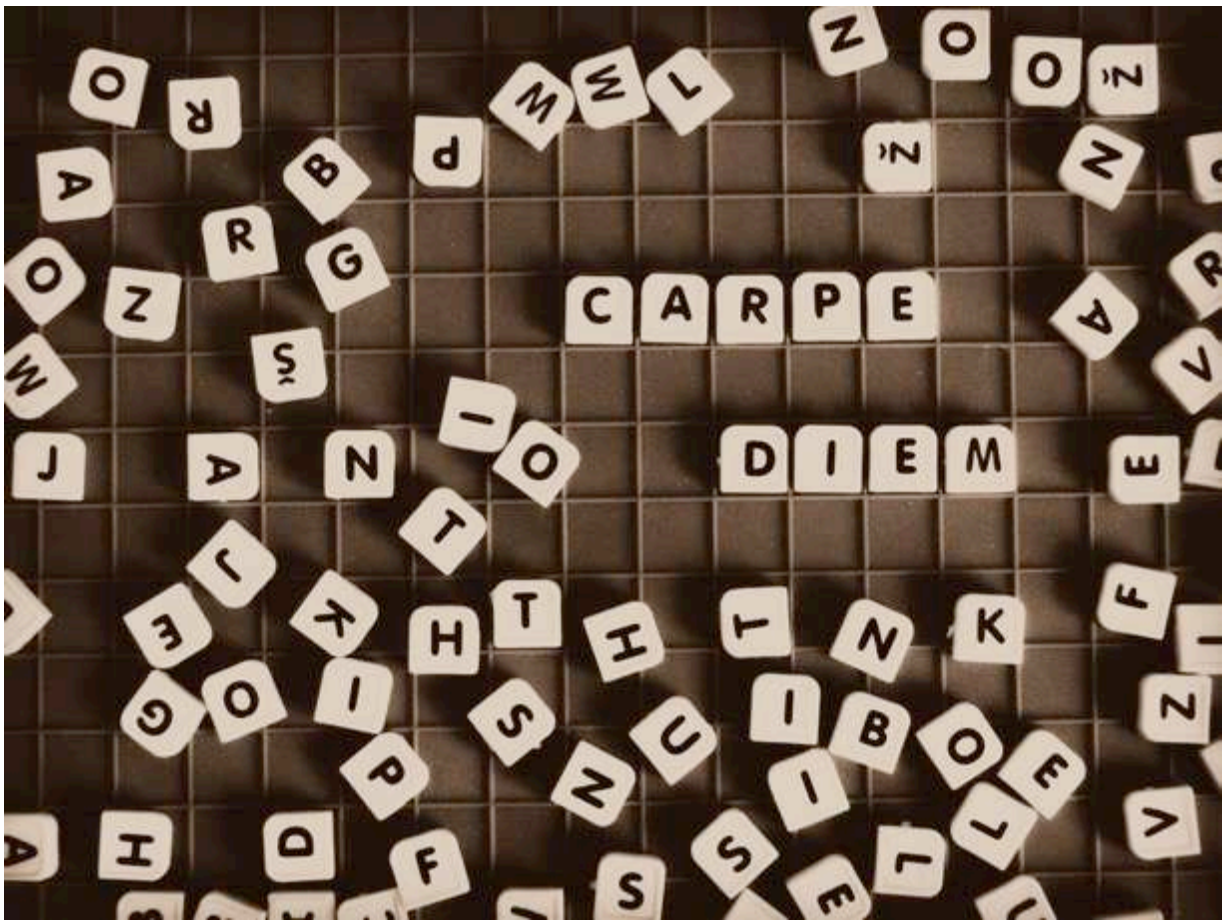
CAPÍTULO 5. DEFINIENDO EL CONSTRUCTO Y SU CONTENIDO (Etapas: 1 y 2). Aquí, se inicia con el enfoque de cuatro etapas, haciendo recomendaciones profusas sobre la importancia de diseñar un marco teórico amplio y suficiente que soporte el diseño de la escala a través de la creación de indicadores.

CAPÍTULO 6. DISEÑANDO LA ESCALA. ANÁLISIS FACTORIAL EXPLORATORIO (Etapa 3). Ésta técnica permite hacer las primeras reducciones de las variables y sus indicadores en el constructo y modelo final. Para su mayor comprensión, se explica a través de problemas basados en el uso del software estadístico SPSS.

CAPÍTULO 7. FINALIZANDO LA ESCALA. ANÁLISIS FACTORIAL CONFIRMATORIO Y ECUACION ESTRUCTURAL (Etapa 4). Es la última etapa de diseño y da cuenta del uso de software de análisis estructural EQS 6. El objetivo es hacer una demostración clara y concisa de cómo ingresar e interpretar los resultados que ésta herramienta aporta con el fin de confirmar el modelo e identificar relaciones subyacentes que el marco teórico no contempló.

Por último, cabe destacar, que el diseño de una escala basado en constructos que definen finalmente un modelo, son el cimiento de la investigación científica, los cuales, son de carácter más que nada iterativo, en vez de lineal. Esto quiere decir, que en lugar de seguir pasos y actividades lineales o consecutivas, como todo proceso lógico y secuencial, el proceso sugerido de desarrollo de la escala tiende fuertemente a ser un procedimiento iterativo y continuo en el que se reinician todos los procedimientos de creación de escalas. Esto es así debido a que los investigadores conscientes, aprenden de sus esfuerzos y errores en el campo económico-administrativo, por lo que se hacen necesarias las revisiones, incluyendo las que sean de etapas tempranas como la definición de factores, variables e indicadores a nivel conceptual, así como la definición de la dimensionalidad, confiabilidad y validez del constructo, nuevamente planteado.

# INVESTIGACIÓN y FUENTES DE INFORMACIÓN. ANTECEDENTES




## ¿Qué es investigar?

La palabra investigar proviene del latín *investigare*, la cual deriva de *vestigium* que significa en *pos de la huella de*, es decir *ir en busca de una pista*. La palabra investigar es muy similar a indagar, en el sentido de seguir la pista de algo.

Esto es, implica el *realizar actividades intelectuales y experimentales de modo sistemático con el propósito de aumentar los conocimientos sobre una determinada materia* (DRAE, 2017). Así, podemos concluir que una investigación *es un conjunto de procesos sistemáticos, críticos y empíricos que se aplican al estudio de un fenómeno* (Hernández-Sampieri et. Al, 2010)

Presenta dos enfoques: *cuantitativo* y *cualitativo* y utilizan, en términos generales, cinco fases similares y relacionadas entre sí (Grinnell, 1997):

- Llevan a cabo la observación y evaluación de fenómenos.

- Establecen suposiciones o ideas como consecuencia de la observación y evaluación realizadas.
- Demuestran el grado en que las suposiciones o ideas tienen fundamento.
- Revisan tales suposiciones o ideas sobre la base de las pruebas o del análisis.
- Proponen nuevas observaciones y evaluaciones para esclarecer, modificar y fundamentar las suposiciones e ideas; o incluso para generar otras.

Se destaca que, aunque ambos enfoques comparten estas fases generales, cada una tiene sus propias características. Ver **Tabla A.1.**

**Tabla A.1. Características investigación cuantitativa**

| Características investigación cuantitativa |
|---|
| • El investigador *plantea* un problema de estudio delimitado y concreto. Sus preguntas de investigación versan sobre cuestiones específicas.<br>• Una vez planteado el problema de estudio, el investigador o investigadora considera lo que se ha investigado anteriormente (la *revisión de la literatura*) y construye un *marco teórico* (la teoría que habrá de guiar su estudio), del cual deriva una o varias *hipótesis* (cuestiones que va a examinar si son ciertas o no) y las somete a prueba mediante el empleo de los diseños de investigación apropiados. Si los resultados corroboran las hipótesis o son congruentes con éstas, se aporta evidencia en su favor. Si se refutan, se descartan en busca de mejores explicaciones y nuevas hipótesis. Al apoyar las hipótesis se genera confianza en la teoría que las sustenta. Si *no* es así, se descartan las hipótesis y, eventualmente, la teoría.<br>• Así, las hipótesis (por ahora denominémoslas creencias) se generan antes de recolectar y analizar los datos.<br>• La *recolección de los datos* se fundamenta en la medición (se miden las variables o conceptos con- tenidos en las hipótesis). Esta recolección se lleva a cabo al utilizar procedimientos estandarizados y aceptados por una comunidad científica. Para que una investigación sea creíble y aceptada por otros investigadores, debe demostrarse que se siguieron tales procedimientos. Como en este enfoque se pretende *medir*, los fenómenos estudiados deben poder observarse o *referirse* en el *mundo real.*<br>• Debido a que los datos son producto de mediciones se representan mediante números (cantidades) y se deben *analizar* a través de *métodos estadísticos*.<br>• En el proceso se busca el máximo control para lograr que otras explicaciones posibles distintas o *rivales* a la propuesta del estudio (hipótesis), sean desechadas y se excluya la incertidumbre y minimice el error. Es por esto que se confía en la experimentación y/o las pruebas de causa-efecto.<br>• Los análisis cuantitativos se interpretan a la luz de las predicciones iniciales (hipótesis) y de estudios previos (teoría). La interpretación constituye una explicación de cómo los resultados encajan en el conocimiento existente.<br>• La investigación cuantitativa debe ser lo más *objetiva* posible. Los fenómenos que se observan y/o miden no deben ser afectados por el investigador. Éste debe evitar en lo posible que sus temores, creencias, deseos y tendencias influyan en los resultados del estudio o interfieran en los procesos y que tampoco sean alterados por las tendencias de otros . |

- Los estudios cuantitativos siguen un patrón predecible y estructurado (el proceso) y se debe tener presente que las decisiones críticas se efectúan antes de recolectar los datos.
- En una investigación cuantitativa se pretende generalizar los resultados encontrados en un grupo o segmento (muestra) a una colectividad mayor (universo o población). También se busca que los estudios efectuados puedan replicarse.
- Al final, con los estudios cuantitativos se intenta explicar y predecir los fenómenos investigados, buscando regularidades y relaciones causales entre elementos. Esto significa que la meta principal es la construcción y demostración de teorías (que explican y predicen).
- Para este enfoque, si se sigue rigurosamente el proceso y, de acuerdo con ciertas reglas lógicas, los datos generados poseen los estándares de validez y confiabilidad, y las conclusiones derivadas contribuirán a la generación de conocimiento.
- Esta aproximación utiliza la lógica o razonamiento deductivo, que comienza con la teoría y de ésta se derivan expresiones lógicas denominadas hipótesis que el investigador busca someter a prueba.
- La investigación cuantitativa pretende identificar leyes universales y causales.
- La búsqueda cuantitativa ocurre en la realidad externa al individuo. Esto nos conduce a una explicación sobre cómo se concibe la realidad con esta aproximación a la investigación.

Para este último fin utilizaremos la explicación de Grinnell (1997) y Creswell (2007)

- Hay dos realidades: la primera es *interna* y consiste en las creencias, presuposiciones y experiencias *subjetivas* de las personas. Éstas llegan a variar: desde ser muy vagas o generales (intuiciones) hasta ser creencias bien organizadas y desarrolladas lógicamente a través de teorías formales.
- La segunda realidad es *objetiva*, *externa* e *independiente* de las creencias que tengamos sobre ella (la autoestima, una ley, los mensajes televisivos, una edificación, el sida, etc., ocurren, es decir, cada una constituye una realidad a pesar de lo que pensemos de ella).
- Esta realidad objetiva es susceptible de conocerse. Bajo esta premisa, resulta posible investigar una realidad externa y autónoma del investigador.
- Se necesita comprender o tener la mayor cantidad de información sobre la realidad objetiva. Conocemos la realidad del fenómeno y los eventos que la rodean a través de sus manifestaciones, y para entender cada realidad (el porqué de las cosas) es necesario registrar y analizar dichos eventos. Desde luego, en el *enfoque cuantitativo* lo subjetivo existe y posee un valor para los investiga- dores; pero de alguna manera este enfoque se aboca a demostrar qué tan bien se adecua el conocimiento a la realidad objetiva. Documentar esta coincidencia constituye un propósito central de muchos estudios cuantitativos (que los efectos que consideramos que provoca una enfermedad sean verdaderos, que captemos la relación *real* entre las motivaciones de un sujeto y su conducta, que un material que se supone posea una determinada resistencia auténticamente la tenga, entre otros). • Cuando las investigaciones creíbles establezcan que la *realidad objetiva* es diferente de nuestras creencias, éstas deben modificarse o adaptarse a tal realidad. En el caso de las ciencias sociales, el enfoque

cuantitativo parte de que el mundo (*social*) es intrínsecamente cognoscible y todos podemos estar de acuerdo con la naturaleza de la realidad social.

Así, se encuentran 10 fases de investigación desde este enfoque:
Fase 1. Idea
Fase 2. Planteamiento del problema
Fase 3. Revisión de la literatura y desarrollo del marco teórico
Fase 4. Visualización del alcance del estudio
Fase 5. Elaboración de hipótesis y definición de variables
Fase 6. Desarrollo del diseño de investigación
Fase 7. Definición y selección de la muestra
Fase 8. Recolección de los datos
Fase 9. Análisis de datos
Fase 10. Elaboración del reporte de resultados

**Características investigación cualitativa**

- Aunque ciertamente hay una revisión inicial de la literatura, ésta puede complementarse en cualquier etapa del estudio y apoyar desde el planteamiento del problema hasta la elaboración del reporte de resultados
- En la investigación cualitativa con frecuencia es necesario regresar a etapas previas. Por ello, las flechas de las fases que van de la inmersión inicial en el campo hasta el reporte de resultados se visualizan en dos sentidos.
- El investigador plantea un problema, pero no sigue un proceso claramente definido. Sus planteamientos *no* son tan específicos como en el enfoque cuantitativo y las preguntas de investigación *no* siempre se han conceptualizado ni definido por completo.
- En lugar de iniciar con una teoría particular y luego *voltear* al mundo empírico para confirmar si ésta es apoyada por los hechos, el investigador comienza examinando el mundo social y en este proceso desarrolla una teoría coherente con los datos, de acuerdo con lo que observa, frecuentemente denominada *teoría fundamentada*, con la cual observa qué ocurre. Dicho de otra forma, las *investigaciones cualitativas* se basan más en una lógica y proceso inductivo (explorar y describir, y luego generar perspectivas teóricas). Van de lo particular a lo general. Por ejemplo, en un típico estudio cualitativo, el investigador entre- vista a una persona, analiza los datos que obtuvo y saca algunas conclusiones; posteriormente, entrevista a otra persona, analiza esta nueva información y revisa sus resultados y conclusiones; del mismo modo, efectúa y analiza más entrevistas para comprender lo que busca. Es decir, procede caso por caso, dato por dato, hasta llegar a una perspectiva más general.
- En la mayoría de los estudios cualitativos no se prueban hipótesis, éstas se generan durante el proceso y van refinándose conforme se recaban más datos o son un resultado del estudio.
- El enfoque se basa en métodos de recolección de datos *no* estandarizados ni completamente predeterminados. No se efectúa una medición numérica, por lo cual el análisis no es estadístico. La recolección de los datos consiste en obtener las perspectivas y puntos de vista de los participantes (sus emociones, prioridades, experiencias, significados y otros aspectos subjetivos). También resultan de interés las

interacciones entre individuos, grupos y colectividades. El investigador pregunta cuestiones abiertas, recaba datos expresados a través del lenguaje escrito, verbal y no verbal, así como visual, los cuales describe y analiza y los convierte en temas que vincula, y reconoce sus tendencias personales. Debido a ello, la preocupación directa del investigador se concentra en las vivencias de los participantes tal como fueron (o son) sentidas y experimentadas. Se define los datos cualitativos como descripciones detalladas de situaciones, eventos, personas, interacciones, conductas observadas y sus manifestaciones.

- Por lo expresado en los párrafos anteriores, el investigador cualitativo utiliza técnicas para recolectar datos, como la observación no estructurada, entrevistas abiertas, revisión de documentos, discusión en grupo, evaluación de experiencias personales, registro de historias de vida, e interacción e introspección con grupos o comunidades.

- El proceso de indagación es más flexible y se mueve entre las respuestas y el desarrollo de la teoría. Su propósito consiste en *reconstruir* la realidad, tal como la observan los actores de un sistema social previamente definido. A menudo se llama *holístico*, porque se precia de considerar el *todo* sin reducirlo al estudio de sus partes.

- El enfoque cualitativo evalúa el desarrollo natural de los sucesos, es decir, no hay manipulación ni estimulación con respecto a la realidad.
- La investigación cualitativa se fundamenta en una perspectiva interpretativa centrada en el entendimiento del significado de las acciones de seres vivos, sobre todo de los humanos y sus instituciones (busca interpretar lo que va captando activamente).
- Postula que la *realidad* se define a través de las interpretaciones de los participantes en la investigación respecto de sus propias realidades. De este modo convergen varias realidades, por lo menos la de los participantes, la del investigador y la que se produce mediante la interacción de todos los actores. Además son realidades que van modificándose conforme transcurre el estudio y son las fuentes de datos.
- Por lo anterior, el investigador se introduce en las experiencias de los participantes y construye el conocimiento, siempre consciente de que es parte del fenómeno estudiado. Así, en el centro de la investigación está situada la diversidad de ideologías y cualidades únicas de los individuos.
- Las indagaciones cualitativas no pretenden generalizar de manera probabilística los resultados a poblaciones más amplias ni necesariamente obtener muestras representativas; incluso, regular- mente no buscan que sus estudios lleguen a replicarse.
- El enfoque cualitativo puede concebirse como un conjunto de prácticas interpretativas que hacen al mundo *visible*, lo transforman y convierten en una serie de representaciones en forma de observaciones, anotaciones, grabaciones y documentos. Es *naturalista* (porque estudia a los objetos y seres vivos en sus contextos o ambientes naturales y cotidianidad) e *interpretativo* (pues intenta encontrar sentido a los fenómenos en función de los significados que las personas les otorguen).

Así, se encuentran 9 fases de investigación desde este enfoque:

Fase común a cada una . Literatura existente

Fase 1. Idea
Fase 2. Planteamiento del problema
Fase 3. Inmersión inicial en el campo
Fase 4. Concepción del diseño de estudio
Fase 5. Definición de la muestra inicial del estudio y acceso a ésta
Fase 6. Recolección de los datos
Fase 7. Análisis de los datos
Fase 8. Interpretación de los resultados
Fase 9. Elaboración del reporte de resultados

Fuente: Henández-Sampieri et al. (2010) con adaptación propia

# Enfoques cuantitativo y cualitativo. Sus diferencias

La pertinencia y oportunidad que tiene una investigación depende de precisar las diferencias esenciales de ambos enfoques. (Ver **Tabla A.2**) así como las etapas de investigación (Ver **Tabla A.3**)

**Tabla A.2. Diferencias entre los enfoques cuantitativo y cualitativo**

| Dimensiones | Cuantitativo | Cualitativo |
|---|---|---|
| • Marcos generales de referencia básicos | • Positivismo, neopositivismo y pospositivismo. | • Fenomenología, constructivismo, naturalismo, interpretativismo. |
| • Punto de partida | • Hay una realidad que conocer. Esto puede hacerse a través de la mente. | • Hay una realidad que descubrir, construir e interpretar. La realidad es la mente. |
| • Realidad a estudiar | • Existe una realidad objetiva única. El mundo es concebido como externo al investigador. | • Existen varias realidades subjetivas construidas en la investigación, las cuales varían en su forma y contenido entre individuos, grupos y culturas.<br>• El investigador cualitativo parte de la premisa de que el mundo social es *relativo* y sólo puede ser entendido desde el punto de vista de los actores estudiados.<br>• Dicho de otra forma, el mundo es construido por el investigador. |
| • Naturaleza de la realidad | • La realidad no cambia por las observaciones y mediciones realizadas | • La realidad sí cambia por las observaciones y la recolección de datos. |
| • Objetividad | • Busca ser objetivo | • Admite subjetividad. |
| • Metas dela | • Describir, explicar y | • Describir, comprender e interpretar |

| | | |
|---|---|---|
| investigación | predecir los fenómenos (causalidad). | los fenómenos, a través de las percepciones y significados producidos por las experiencias de los participantes. |
| • Lógica | • Se aplica la lógica deductiva. De lo general a lo particular (de las leyes y teoría a los datos). | • Se aplica la *lógica inductiva*. De lo particular a lo general (de los datos a las generalizaciones —no estadísticas— y la teoría). |
| • Relación entre ciencias físicas/naturales y sociales | • Las ciencias físicas/naturales y las sociales son una unidad. A las ciencias sociales pueden aplicárseles los principios de las ciencias naturales. | • Las ciencias físicas/naturales y las sociales son diferentes. No se aplican los mismos principios. |
| • Posición personal del investigador | • Neutral. El investigador *hace a un lado* sus propios valores y creencias.<br>• La posición del investigador es *imparcial*, intenta asegurar procedimientos rigurosos y *objetivos* de recolección y análisis de los datos, así como evitar que sus sesgos y tendencias influyan en los resultados. | • Explícita. El investigador reconoce sus propios valores y creencias, incluso son parte del estudio. |
| • Interacción física entre el investigador y el fenómeno | • Distanciada, separada. | • Próxima, suele haber contacto. |
| • Interacción psicológica entre el investigador y el fenómeno | • Distanciada, lejana, neutral, sin involucramiento. | • Cercana, próxima, empática, con involucramiento. |
| • Papel de los fenómenos estudiados (objetos, seres vivos, etcétera) | • Los papeles son más bien pasivos. | • Los papeles son más bien activos. |
| • Relación entre el investigador y el fenómeno estudiado | • De independencia y neutralidad, no se afectan. Se separan. | • De interdependencia, se influyen. No se separan. |
| • Planteamiento del problema | • Delimitado, acotado, específico. Poco flexible. | • Abierto, libre, no es delimitado o acotado. Muy flexible. |
| • Uso de la teoría | • La teoría se utiliza para ajustar sus postulados al | • La teoría es un marco de referencia. |

| | | |
|---|---|---|
| | mundo empírico. | |
| • Generación de la teoría | • La teoría es generada a partir de comparar la investigación previa con los resultados del estudio. De hecho, éstos son una extensión de los estudios antecedentes. | • La teoría no se fundamenta en estudios anteriores, sino que se genera o construye a partir de los datos empíricos obtenidos y analizados. |
| • Papel de la revisión de la literatura | • La literatura representa un papel crucial, guía a la investigación. Es fundamental para la definición de la teoría, las hipótesis, el diseño y demás etapas del proceso. | • La literatura desempeña un papel menos importante al inicio, aunque sí es relevante en el desarrollo del proceso.<br>• En ocasiones, provee de dirección, pero lo que principalmente señala el rumbo es la evolución de eventos durante el estudio y el aprendizaje que se obtiene de los participantes.<br>• El marco teórico es un elemento que ayuda a justificar la necesidad de investigar un problema planteado.<br>• Algunos autores del enfoque cualitativo consideran que su rol es únicamente auxiliar. |
| • La revisión de la literatura y las variables o conceptos de estudio | • El investigador hace una revisión de la literatura principalmente para buscar variables significativas que puedan ser medidas. | • El investigador, más que fundamentarse en la revisión de la literatura para seleccionar y definir las variables o conceptos clave del estudio, confía en el proceso mismo de investigación para identificar- los y descubrir cómo se relacionan. |
| • Hipótesis | • Se prueban hipótesis. Éstas se establecen para aceptarlas o rechazarlas dependiendo del grado de certeza (probabilidad). | • Se generan hipótesis durante el estudio o al final de éste. |
| • Diseño de la investigación | • Estructurado | • Abierto, flexible, construido durante el trabajo de campo o realización del estudio. |
| • Población-muestra | • El objetivo es generalizar los datos de una muestra a una población (de un grupo pequeño a uno mayor). | • Regularmente no se pretende generalizar los resultados obtenidos en la muestra a una población. |
| • Muestra | • Se involucra a muchos sujetos en la investigación | • Se involucra a unos cuantos sujetos porque no se pretende |

| | | |
|---|---|---|
| | porque se pretende generalizar los resultados del estudio. | necesariamente generalizar los resultados del estudio. |
| • Composición de la muestra | • Casos que en conjunto son estadística- mente representativos. | • Casos individuales, representativos no desde el punto de vista estadístico. |
| • Naturaleza | • La naturaleza de los datos es cuantitativa (datos numéricos). | • La naturaleza de los datos es cualitativa (textos, narraciones, significados, etcétera). |
| • Tipo de datos | • La recolección se basa en instrumentos estandarizados. • Es uniforme para todos los casos. Los datos se obtienen por observación, medición y documentación de mediciones. • Se utilizan instrumentos que han demostrado ser válidos y confiables en estudios previos o se generan nuevos basados en la revisión de la literatura y se prueban y ajustan. • Las preguntas o ítems utilizados son específicos con posibilidades de respuesta predeterminadas. | • La recolección de los datos está orientada a proveer de un mayor entendimiento de los significados y experiencias de las personas. • El investigador es el instrumento de recolección de los datos, se auxilia de diversas técnicas que se desarrollan durante el estudio. • No se inicia la recolección de los datos con instrumentos preestablecidos, sino que el investigador comienza a aprender por observación y descripciones de los participantes y concibe formas para registrar los datos que se van refinando conforme avanza la investigación. |
| • Concepción de los participantes en la recolección de datos | • Los participantes son fuentes externas de datos. | • Los participantes son fuentes internas de datos. El investigador también es un participante. |
| • Finalidad del análisis de los datos | • Describir las variables y explicar sus cambios y movimientos. | • Comprender a las personas y sus contextos. |
| • Características del análisis de los datos | • Sistemático. Utilización intensiva de la estadística (descriptiva e inferencial). • Basado en variables. Impersonal. • Posterior a la recolección de los datos. | • El análisis varía dependiendo del modo en que hayan sido recolectados los datos. • Fundamentado en la inducción analítica. • Uso moderado de la estadística (conteo, algunas operaciones aritméticas). • Basado en casos o personas y sus manifestaciones. |

| | | |
|---|---|---|
| | | •Simultáneo a la recolección de los datos. •El análisis consiste en describir información y desarrollar temas. |
| •Forma de los datos para analizar | •Los datos son representados en forma de números que son analizados estadísticamente. | •Datos en forma de textos, imágenes, piezas audio- visuales, documentos y objetos personales. |
| •Proceso del análisis de los datos | •El análisis se inicia con ideas pre- concebidas, basadas en las hipótesis formuladas. Una vez recolectados los datos numéricos, éstos se transfieren a una matriz, la cual se analiza mediante procedimientos estadísticos. | •Por lo general, el análisis no se inicia con ideas preconcebidas sobre cómo se relacionan los conceptos o variables.<br>•Una vez reunidos los datos verbales, escritos y/o audiovisuales, se integran en una base de datos compuesta por texto y/o elementos visuales, la cual se analiza para determinar significados y describir el fenómeno estudiado desde el punto de vista de sus actores.<br>•Se integran descripciones de personas con las del investigador.¡ |
| •Perspectiva del investigador en el análisis de los datos | •Externa (al margen de los datos).<br>•El investigador no involucra sus antecedentes y experiencias en el análisis. Mantiene distancia de éste. | •Interna (desde los datos).<br>•El investigador involucra en el análisis sus propios antecedentes y experiencias, así como la relación que tuvo con los participantes del estudio. |
| •Principales criterios de evaluación en la recolección y análisis de los datos | •Objetividad, rigor, confiabilidad y validez. | •Credibilidad, confirmación, valoración y transferencia. |
| •Presentación de resultados | •Tablas, diagramas y modelos estadísticos. El formato de presentación es estándar. | •El investigador emplea una variedad de formatos para reportar sus resultados: narraciones, fragmentos de textos, videos, audios, fotografías y mapas; diagramas, matrices y modelos conceptuales.<br>•Prácticamente, el formato varía en cada estudio. |

| | | |
|---|---|---|
| • Reporte de resultados | • Los reportes utilizan un tono objetivo, impersonal, no emotivo. | • Los reportes utilizan un tono personal y emotivo. |

Fuente: Henández-Sampieri et al. (2010) con adaptación propia

**Tabla A.3. Comparación de las etapas de investigación de los procesos cuantitativo y cualitativo**

| Características Cuantitativas | Procesos fundamentales del proceso general de investigación | Características Cualitativas |
|---|---|---|
| • Orientación hacia la descripción, predicción y explicación<br>• Específico y acotado<br>• Dirigido hacia datos medibles u observables | • Planteamiento del problema | • Orientación hacia la exploración, la descripción y el entendimiento<br>• General y amplio<br>• Dirigido a las experiencias de los participantes |
| • Rol fundamental<br>• Justificación para el planteamiento y la necesidad del estudio | • Revisión de la literatura | • Rol secundario<br>• Justificación para el planteamiento y la necesidad del estudio |
| • Instrumentos predeterminados<br>• Datos numéricos<br>• Número considerable de casos | • Recolección de los datos | • Los datos emergen poco a poco<br>• Datos en texto o imagen<br>• Número relativamente pequeño de casos |
| • Análisis estadístico<br>• Descripción de tendencias, comparación de grupos o relación entre variables<br>• Comparación de resultados con predicciones y estudios previos | • Análisis de datos | • Análisis de textos y material audiovisual<br>• Descripción, análisis y desarrollo de temas<br>• Significado profundo de los resultados |
| • Estándar y fijo<br>• Objetivo y sin tendencias | • Reportes de resultados | • Emergente y flexible<br>• Reflexivo y con aceptación de tendencias |

Fuente: Henández-Sampieri et al. (2010) con adaptación propia

# ¿Existe algún enfoque que sea mejor?

Ninguno es explícitamente mejor que otro. Simplemente son diferentes acercamientos al estudio de un fenómeno.

La *investigación cuantitativa* representa la posibilidad de *generalizar* los resultados de manera amplia, con control sobre los fenómenos, abonando a definir las magnitudes de éstos. Brinda una gran posibilidad de réplica y un enfoque sobre puntos específicos de tales fenómenos, además de que facilita la comparación entre estudios similares. Ha sido el más enfoque más usado por ciencias como la física, química y biología y es más propio para las ciencias llamadas *exactas o naturales*.

La *investigación cualitativa* proporciona profundidad a los datos, dispersión, riqueza interpretativa, contextualización del ambiente o entorno, detalles y experiencias únicas. También aporta un punto de vista *holístico* y flexible de los fenómenos, Se ha empleado más bien en disciplinas humanísticas como la antropología, la etnografía y la psicología social.

No obstante, ambos tipos de estudio son de utilidad para todos los campos como las ciencias económico-administrativas. Por ejemplo, un administrador de tecnología puede llevar un estudio sobre cómo construir cierto producto. Emplearía estudios cuantitativos y cálculos matemáticos para realizar desde el diseño del mismo hasta la calidad y desempeño del mismo. Asimismo, realizar un estudio cualitativo para detectar el sentir del consumidor sobre la introducción de dicho producto, el grado de aceptación, el nivel de asimilación y adaptación al mismo, etc.

En el pasado se consideró que los enfoques *cuantitativo y cuali*tativo eran perspectivas opuestas, irreconciliables y que no debían mezclarse. Los críticos del *enfoque cuantitativo* lo acusaron de ser *impersonal, frío, reduccionista, limitativo, cerrado y rígido*. Además, consideraron que se estudiaba a las personas como *objetos* y que las diferencias individuales y culturales entre grupos *no podían promediarse ni agruparse estadísticamente.* Por su parte, los detractores del *enfoque cualitativo* lo consideraron *vago, subjetivo, inválido, meramente especulativo, sin posibilidad de réplica y sin datos sólidos que apoyaran las conclusiones*. Argumentaban que no se tiene control sobre las variables estudiadas y que se carece del poder de entendimiento que generan las mediciones (Hernández-Sampieri et al. 2010).

Se llegó a considerar la idea que ambos enfoque se podrían entre sí. La posición asumida en la mayoría de los estudios actuales, es la de ser *enfoques complementarios*. El investigador debe ser metodológicamente y pragmático guiándose por el contexto, la situación, los recursos de que dispone, sus objetivos y el problema de estudio.

# El enfoque mixto

Estos representan un conjunto de procesos sistemáticos, empíricos y críticos de investigación e implican la recolección y el análisis de datos cuantitativos y cualitativos, así como su integración y discusión conjunta, para realizar inferencias producto de toda la información recabada (*metainferencias*) y lograr un mayor entendimiento del fenómeno bajo estudio. Antes de definir propiamente los componentes del enfoque mixto, hemos de comentar que éstos agregan cada año

más adeptos y su desarrollo durante la primera década del siglo XXI ha sido vertiginoso. Han recibido varias denominaciones tales como *investigación integrativa*), *investigación multimétodos, métodos múltiples*, *estudios de triangulación* e *investigación mixta* (Hernández Sampieri y Mendoza, 2008).
Hoy en día, se tienen aún posiciones desde el completo rechazo hasta la total aceptación, por lo que corresponde al investigador convencer tanto en su enfoque como perspectiva a trabajar. Ver **Tabla A.4**

**Tabla A.4. Características del enfoque mixto**

| Características del enfoque mixto |
|---|
| •Lograr una perspectiva más amplia y profunda del fenómeno. La percepción de éste resulta más integral, completa y holística. Además, si son empleados dos métodos (con fortalezas y debilidades propias) que llegan a los mismos resultados, se incrementa la confianza en que éstos son una representación fiel, genuina y fidedigna de lo que ocurre con el fenómeno estudiado.<br>•La investigación se sustenta en las fortalezas de cada método y no en sus debilidades potenciales posible explorar distintos niveles del problema de estudio. Incluso, evaluar más extensamente las dificultades y problemas de las indagaciones, ubicados en todo el proceso de investigación y en cada una de sus etapas.<br>•Los diseños mixtos logran obtener una mayor variedad de perspectivas del problema: frecuencia, amplitud y magnitud (*cuantitativa*), así como profundidad y complejidad (*cualitativa*); generalización (*cuantitativa*) y comprensión (*cualitativa*). •Hernández Sampieri y Mendoza (2008) la denominan: *riqueza interpretativa*. Cada método (*cuantitativo y cualitativo*) proporcionan una visión de la realidad .<br>•Formular el planteamiento del problema con mayor claridad, así como las maneras más apropiadas para estudiar y teorizar los problemas de investigación<br>•Con un solo enfoque, el investigador regularmente se esfuerza menos en considerar estos aspectos con una profundidad suficiente<br>•A través de una perspectiva mixta, el investigador debe confrontar las *tensiones* entre distintas concepciones teóricas y al mismo tiempo, considerar la vinculación entre los conjuntos de datos emanados de diferentes métodos.<br>•Producir datos más *ricos* y variados mediante la multiplicidad de observaciones, ya que se consideran diversas fuentes y tipos de datos, contextos o ambientes y análisis. Se rompe con la investigación *uniforme*.<br>•Potenciar la creatividad teórica por medio de suficientes procedimientos críticos de valoración .<br>•Efectuar indagaciones más dinámicas.<br>•Apoyar con mayor solidez las inferencias científicas, que si se emplean aisladamente.<br>•Posibilidad de tener mayor éxito al presentar resultados a una audiencia hostil.<br>•Oportunidad para desarrollar nuevas destrezas o competencias en materia de investigación, o bien, reforzarla<br><br>•Existen combinaciones:<br>CUALITATIVA-Cuantitativa. Preponderancia cualitativa |

CUALITATIVA-CUANTITATIVA.Mixto Puro. Mismo estatus
Cualitativa-CUANTITATIVA. Preponderancia cuantitativa
Con nueve pretensiones básicas
1.*Triangulación* (corroboración): lograr convergencia, confirmación y/o correspondencia o no, de métodos cuantitativos y cualitativos. El énfasis es en el contraste de ambos tipos de datos e información.
2.*Complementación:* mayor entendimiento, ilustración o clarificación de los resultados de un método sobre la base de los resultados del otro método.
3.*Visión holística:* obtener un abordaje más completo e integral del fenómeno estudiado usando información cualitativa y cuantitativa (la visión completa es más significativa que la de cada uno de sus componentes).
4.*Desarrollo:* usar los resultados de un método para ayudar a desplegar o informar al otro método en diversas cuestiones, como el muestreo, los procedimientos, la recolección y el análisis de los datos. Incluso, un enfoque puede proveerle al otro de hipótesis y soporte empírico.
5.*Iniciación:* descubrir contradicciones y paradojas, así como obtener nuevas perspectivas y marcos de referencia, y también a la posibilidad de modificar el planteamiento original y resultados de un método con interrogantes y resultados del otro método.
6.*Expansión:* extender la amplitud y el rango de la indagación usando diferentes métodos para distintas etapas del proceso investigativo. Un método puede expandir o ampliar el conocimiento obtenido en el otro. 7.
7:*Compensación:* un método puede visualizar elementos que el otro no, las debilidades de cada uno puede ser subsanadas por su *contraparte*.
8.*Diversidad:* obtener puntos de vista variados, incluso divergentes, del fenómeno o planteamiento bajo estudio. Distintas ópticas (*lentes*) para estudiar el problema.
9. *Pragmatismo.* sugiere usar el método más apropiado para un estudio específico. Es una orientación filosófica y metodológica, como el positivismo, pospositivismo o constructivismo. Incluye el uso de la inducción (o descubrimiento de patrones), deducción (prueba de teorías e hipótesis) y de la abducción (apoyarse y confiar en el mejor conjunto de explicaciones para entender los resultados).

Las *hipótesis* se incluyen :
- *En y para* la parte o *enfoque cuantitativo*, si la investigación pretende algún fin *confirmatorio o probatorio*;
- *Como un producto de la fase cualitativa* (típicamente de *carácter exploratorio* en el enfoque mixto). Se tiene como consecuencia, que emergen nuevas hipótesis a lo largo de la indagación.

Para su uso, se sugiere al investigador responder:
- ¿Qué clase de datos tienen prioridad: los cuantitativos, los cualitativos o ambos por igual?
- ¿Qué resulta más apropiado para el estudio en particular: recolectar los datos cuantitativos y cualitativos de manera simultánea (al mismo tiempo) o secuencial (un tipo de datos primero y luego el otro)?
- ¿Cuál es el propósito central de la integración de los datos cuantitativos y cualitativos? Por ejemplo: triangulación, complementación, exploración o explicación.

• ¿En qué parte del proceso, fase o nivel es más conveniente que se inicie y desarrolle la estrategia mixta?

*CUAN-CUAL Ejecución concurrente*

• Se aplican ambos métodos de manera simultánea (los datos cuantitativos y cualitativos se recolectan y analizan más o menos en el mismo tiempo).

• Regularmente los datos cualitativos requieren de mayor tiempo para su obtención y análisis.

• Los diseños concurrentes implican cuatro condiciones

*i*) Se recaban en paralelo y de forma separada datos cuantitativos y cualitativos.

*ii*) Ni el análisis de los datos cuantitativos ni el análisis de los datos cualitativos se construye sobre la base del otro análisis.

*iii*) Los resultados de ambos tipos de análisis no son consolidados en la fase de interpretación de los datos de cada método, sino hasta que ambos conjuntos de datos han sido recolectados y analiza- dos de manera separada se lleva a cabo la consolidación.

*iv*) Después de la recolección e interpretación de los datos de los componentes CUAN y CUAL, se efectúa una o varias *metainferencias* que integran las inferencias y conclusiones de los datos y resultados cuantitativos y cualitativos realizadas de manera independiente.

*CUAN-CUAL Ejecución secuencial*

• En una primera etapa se recolectan y analizan datos cuantitativos o cualitativos, y en una segunda fase se recaban y analizan datos del otro método. Típicamente, cuando se recolectan primero los datos cualitativos, la intención es explorar el planteamiento con un grupo de participantes en su contexto, para posteriormente expandir el entendimiento del problema en una muestra mayor y poder efectuar generalizaciones a la población

• En los diseños secuenciales, los datos recolectados y analizados en una fase del estudio (CUAN o CUAL) se utilizan para informar a la otra fase del estudio (CUAL o CUAN).

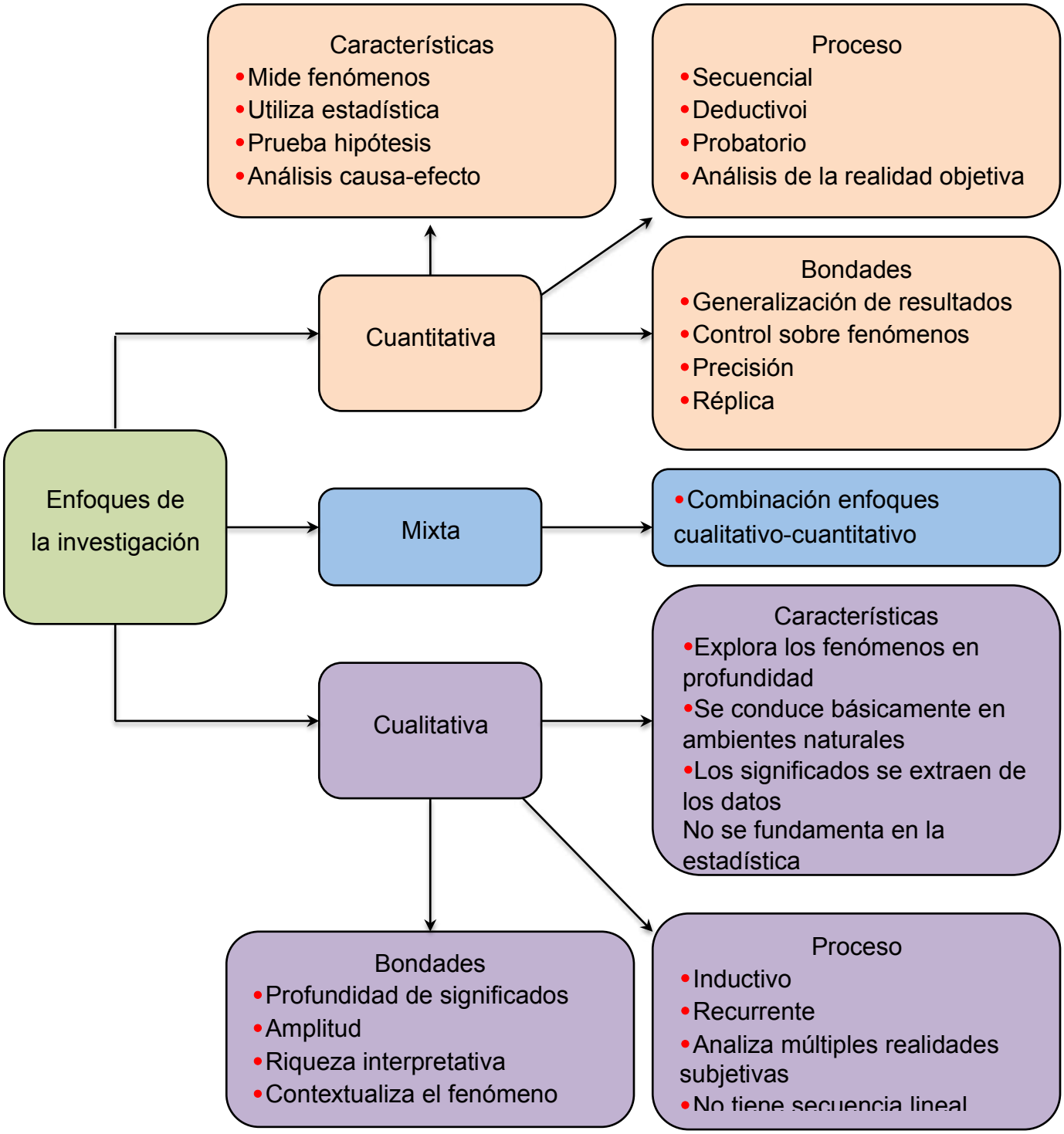
*Elementos para decidir el diseño mixto apropiado*

| Tiempos | Prioridad | Mezcla | Teorización |
|---|---|---|---|
| Concurrente (no hay secuencia) | Igual | Integrar ambos métodos | --- |
| Secuencial: primero el método cualitativo | Cualitativo (CUAL) | Conectar un método con el otro | Explícita |
| Secuencial: primero el método cuantitativo | Cuantitativo (CUAN) | Anidar o incrustar un método dentro de otro | Implícita |

Fuente: Hernández-Sampieri et al. (2010) con sdaptación propia

A continuación, se muestran figuras que desglosan los enfoques de investigación. (Ver **Figura A.1**.) y el desglose del enfoque de investigación mixto (Ver **Figura A.2**.)

**Figura A.1. Enfoques de la investigación**



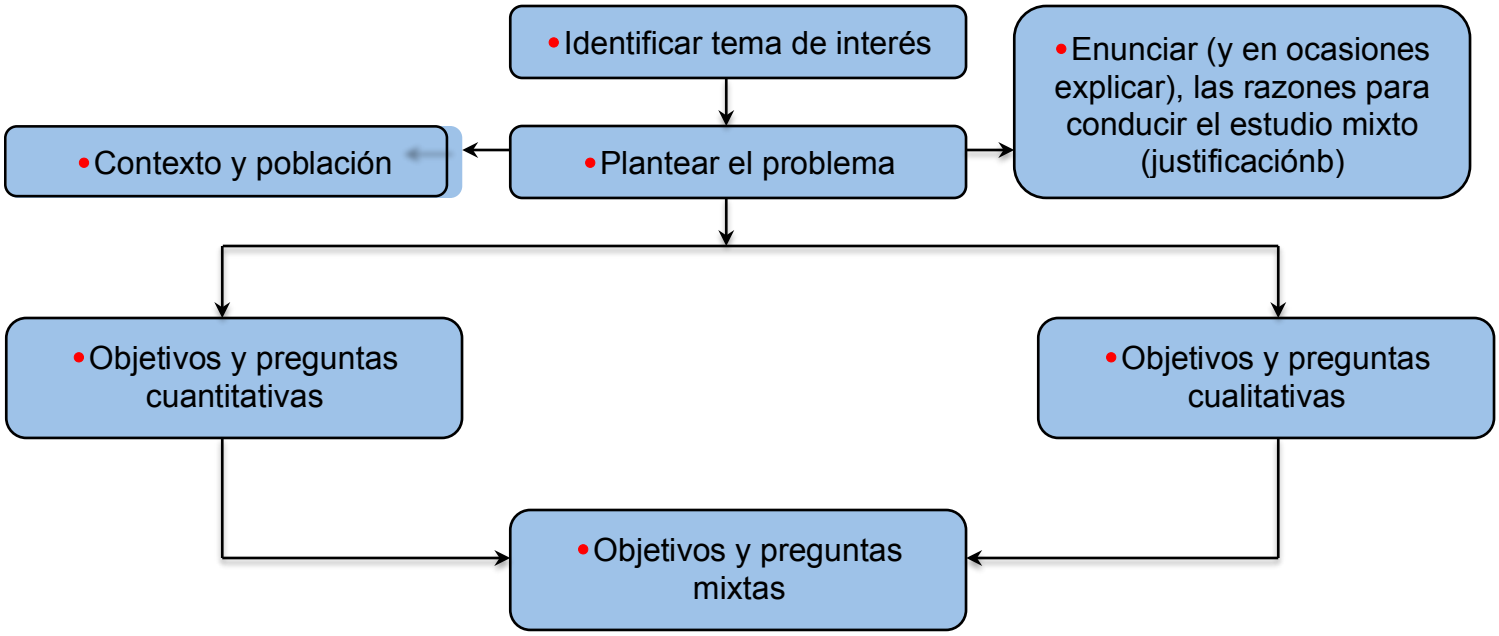

Fuente: Hernández-Sampieri et al. (2010) con sdaptación propia

**Figura A.2. Enfoques de la investigación mixto**

Fuente: Hernández-Sampieri et al. (2010) con sdaptación propia

El enfoque mixto, tine un proceso de planteamiento de problemas, similar a la **Figura A.3.**

**Figura A.3. Proceso de planteamiento de problema**

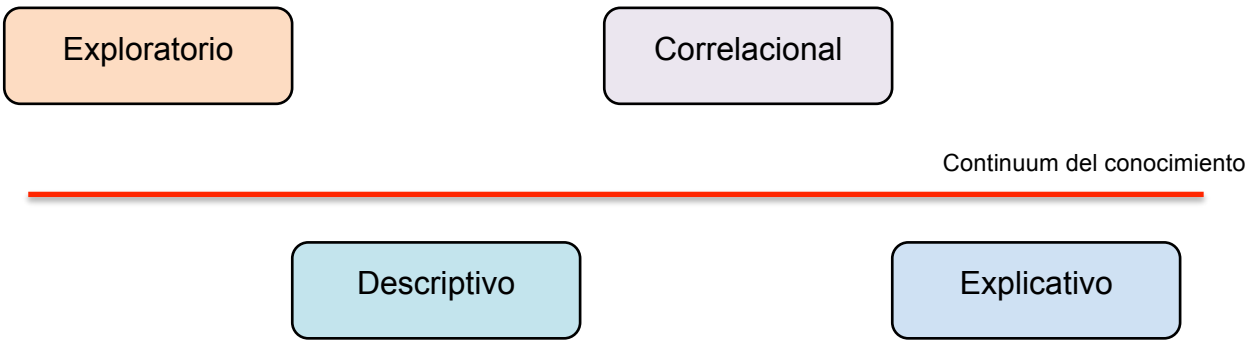



Fuente: Hernández-Sampieri et al. (2010) con sdaptación propia

# Alcances de una investigación

Una vez hecha la revisión de la literatura, que demuestre que la propuesta de investigación es pertinente y se deba realizarse, deberá por tanto, determinar el *alcance* que tendrá. El alcance no se define como un como *tipos* de investigación, sino más bien, constituye el ubicar en qué parte del un *continuum del conocimiento* se encuentra de la determinación de su *causalidad*. Ver **Figura A.1.**

**Figura A.4. Alcances de los estudios cuantitativos**




Fuente: Henández-Sampieri et al. (2010) con adaptación propia

Normalmente, una investigación se inicie como *exploratoria*, después puede ser *descriptiva y correlacional, y terminar como explicativa* . En la práctica, cualquier investigación puede incluir elementos de más de uno de estos cuatro alcances.
Así, a nivel de enfoque cuantitativo, una investigación iniciará en el tipo de alcance del continuum dependiendo de:

- *El estado del conocimiento* sobre el problema de investigación, mostrado por la revisión de la literatura,
- *La perspectiva* que se pretenda dar al estudio. Ver **Tabla A.5.**

**Tabla A.5. Alcances de una investigación**

| Exploratorio |
|---|
| • Se aplican cuando el objetivo es examinar un problema de investigación *poco estudiado*, *del que aún se tienen dudas* o *no se ha abordado antes*. En otras palabras, cuando la revisión de la literatura revela que tan sólo hay guías no investigadas e ideas vagamente relacionadas con el problema de estudio, o bien, si deseamos indagar sobre temas y áreas desde nuevas perspectivas.<br>• En pocas ocasiones son un fin en sí mismos, ya que generalmente se orientan a determinar tendencias, identificando áreas, ambientes, contextos y situaciones de estudio, relaciones potenciales entre variables; o establecen el *tono* de investigaciones posteriores más elaboradas y rigurosas.<br>• Estas indagaciones se caracterizan por ser más flexibles en su método en comparación con las descriptivas, correlacionales o explicativas, y son más amplias y dispersas. Asimismo, implican un mayor *riesgo* y requieren gran paciencia, serenidad y receptividad por parte del investigador. |
| **Descriptivo** |
| • Con frecuencia, la meta del investigador consiste en describir fenómenos, situaciones, contextos y eventos; esto es, detallar cómo son y se manifiestan. Los estudios descriptivos buscan especificar las propiedades, las características y los perfiles de personas, grupos, comunidades, procesos, objetos o cualquier otro fenómeno que se someta a un análisis.<br>• Únicamente pretenden medir o recoger información de manera independiente o conjunta sobre los conceptos o las variables a las que se refieren, esto es, su. objetivo no es indicar cómo se relacionan éstas.<br>• Así como el alcance *exploratorio* sirve fundamentalmente para descubrir y prefigurar, el alcance *descriptivo* es útiles para mostrar con precisión los ángulos o dimensiones de un fenómeno, suceso, comunidad, contexto o situación.<br>• En esta clase de estudios debe ser capaz de definir, o prever, qué se medirá (qué conceptos, variables, componentes, etc.) y sobre qué o quiénes se recolectarán los datos (personas, grupos, comunidades, objetos, animales, hechos, etc.).<br>• La descripción puede ser más o menos profunda, aunque en cualquier caso se basa en la medición de uno o más atributos del fenómeno de interés. |
| **Correlacional** |
| • Este tipo de alcance, tiene como finalidad *determinar la relación, nivel o grado de* |

*asociación que exista entre dos o más conceptos, categorías o variables en una perspectiva en particular.*

- En ocasiones sólo se analiza la relación entre dos variables, pero con frecuencia se ubican en el estudio relaciones entre tres, cuatro o más variables.
- Este alcance se orienta a evaluar el grado de asociación entre dos o más variables, miden cada una de ellas (presuntamente relacionadas) y, después, cuantifican y analizan la vinculación.
- Tales correlaciones se sustentan en hipótesis sometidas a prueba. Se destaca, que en la mayoría de los casos, las mediciones de las variables a correlacionar provienen de los mismos participantes, pues *no es lo común que se correlacionen mediciones de una variable hechas en ciertas sujetos, con mediciones de otra variable realizadas en sujetos distintas.*
- La utilidad de este alcance es *saber cómo se puede comportar un concepto o una variable al conocer el comportamiento de otras variables vinculadas.*
- Intentar *predecir* el valor aproximado que tendrá un grupo de individuos o casos en una variable, a partir del valor que poseen en la o las variables relacionadas.

La correlación puede ser *positiva o negativa*.

- *Si es positiva,* significa que alumnos con valores altos en una variable tenderán también a mostrar valores elevados en la otra variable.
- *Si es negativa*, significa que sujetos con valores elevados en una variable tenderán a mostrar valores bajos en la otra variable.
- *Si no hay correlación* entre las variables, ello nos indica que éstas fluctúan sin seguir un patrón sistemático entre sí.
- *Si dos variables están correlacionadas* y se conoce la magnitud de la asociación, se tienen bases para predecir, con mayor o menor exactitud, el valor aproximado que tendrá un grupo de personas en una variable, al saber qué valor tienen en la otra.
- El alcance *correlacional* se distingue de los *descriptivos* principalmente en que, mientras estos últimos se centran en medir con precisión las variables individuales (algunas de las cuales se pueden medir con independencia en una sola investigación), los primeros evalúan, con la mayor exactitud que sea posible, el *grado de vinculación entre dos o más variables*, pudiéndose incluir varios pares de evaluaciones de esta naturaleza en una sola investigación (comúnmente se incluye más de una correlación.
- El alcance *correlacional* tiene, en alguna medida, un *valor explicativo, aunque parcial,* ya que el hecho de saber que dos conceptos o variables se relacionan aporta cierta información explicativa.
- Desde luego, la *explicación parcial*, es así porque hay otros factores vinculados con la adquisición de conceptos de acuerdo a la perspectiva. Así, *cuanto mayor sea el número de variables que se asocien en el estudio y mayor sea la fuerza de las relaciones, más completa será la explicación.*
- Llega a darse el caso de que dos variables estén aparentemente relacionadas, pero que en realidad no sea así. Esto se conoce en el ámbito de la investigación

| |
|---|
| como *correlación espuria, con tendencia alta ser errónea.* |
| Explicativo |
| •Van más allá de la descripción de conceptos o fenómenos o del establecimiento de relaciones entre conceptos; es decir, están dirigidos a responder por las causas de los eventos y fenómenos físicos o sociales.<br>•Como su nombre lo indica, su interés se centra en explicar por qué ocurre un fenómeno y en qué condiciones se manifiesta, o por qué se relacionan dos o más variables |

Fuente: Henández-Sampieri et al. (2010) con adaptación propia

# Los alcances ¿todos pueden estar en una misma investigación?

Algunas veces una investigación puede caracterizarse como básicamente *exploratoria, descriptiva, correlacional o explicativa*, *pero no situarse únicamente como tal.* Esto es, aunque un estudio sea en esencia exploratorio contendrá elementos descriptivos; o bien, un estudio correlacional incluirá componentes descriptivos, y lo mismo ocurre con los demás alcances. Asimismo, recuerde que es posible que una investigación se inicie como exploratoria o descriptiva y después llegue a ser correlacional y aun explicativa. El estudio se inicia como exploratorio, para después ser descriptivo, correlacional y explicativo (*no se recomienda que se sitúe, únicamente en alguno de los tipos citados*). Ver **Tabla A.6.**

**Tabla A.6. Propósito y valor de los alcances de una investigación**

| Alcance | Propósito | Valor |
|---|---|---|
| Exploratorio | •Se realiza cuando el objetivo es examinar un tema o problema de investigación poco estudiado, del cual se tienen muchas dudas o no se ha abordado antes. | •Ayuda a familiarizarse con fenómenos desconocidos, obtener información para realizar una investigación más completa de un contexto particular, investigar nuevos problemas, identificar conceptos o variables promisorias, establecer prioridades para investigaciones futuras, o sugerir afirmaciones y postulados. |
| Descriptivo | •Busca especificar las propiedades, las características y los perfiles de personas, grupos, comunidades, procesos, objetos o cualquier otro fenómeno que se someta a un análisis. | •Es útil para mostrar con precisión los ángulos o dimensiones de un fenómeno, suceso, comunidad, contexto o situación. |

| | | |
|---|---|---|
| Correlacional | • Su finalidad es conocer la relación o grado de asociación que exista entre dos o más conceptos, categorías o variables en un contexto en particular. | • En cierta medida tiene un valor explicativo, aunque parcial, ya que el hecho de saber que dos conceptos o variables se relacionan aporta cierta información explicativa. |
| Explicativo | • Está dirigido a responder por las causas de los eventos y fenómenos físicos o sociales.<br>• Se enfoca en explicar por qué ocurre un fenómeno y en qué condiciones se manifiesta, o por qué se relacionan dos o más variables. | • Se encuentra más estructurado que las demás investigaciones (de hecho implica los propósitos de éstas)<br>• Proporciona un sentido de entendimiento del fenómeno a que hacen referencia. |

Fuente: Henández-Sampieri et al. (2010) con adaptación propia

# La idea, el conocimiento del tema de investigación y su perspectiva

Las investigaciones nacen del planteamiento de *ideas*, sin importar qué tipo de paradigma fundamente nuestro estudio ni el enfoque que habremos de seguir. Las *ideas* constituyen el primer acercamiento a la *realidad objetiva* (*enfoque cuantitativo*), a la *realidad subjetiva* (*enfoque cualitativo*) o a la *realidad intersubjetiva* (desde *ambos enfoques, o mixto*) que habrá de investigarse.

Es importante notar que una idea de investigación debe partir de un enfoque de análisis (*cuantitativo, cualitativo, mixto*) con *perspectiva* (o la línea de *investigación con la que se abordará*), siendo novedosa e intrigante, Por ejemplo, se puede abordar la mercadotecnia digital desde un *enfoque* cualitativo a partir de la *perspectiva* de la innovación tecnológica del uso de smartphones en niños de 10-15 años (novedosa e intrigante). Ver **Tabla A.7.**

**Tabla A.7. Conocimiento y perspectiva en el problema de investigación**

| Conocimiento actual del tema de investigación |
|---|
| Existen tres posibilidades de influencia.:<br>1.En primer término, la literatura puede revelar que no hay antecedentes sobre el tema en cuestión o que no son aplicables al contexto en el cual habrá de desarrollarse el estudio, entonces la investigación deberá iniciarse como exploratoria. Si la literatura nos revela guías aún no estudiadas e ideas vagamente vinculadas con el problema de investigación, la situación resulta similar, es decir, el estudio se iniciaría como *exploratorio*. Por ejemplo, si pretendemos realizar una investigación sobre el consumo de software de negocios electrónicos en determinados sectores e industrias y quisiéramos saber: ¿en qué medida ocurre?, ¿qué tipos de software se consumen?, ¿cuáles más?, ¿a qué se debe ese consumo?, ¿de qué fuentes proviene?, ¿cómo es que se introducen en el sector?, ¿quiénes intervienen en su distribución?, etc., pero encontramos que no existen |

antecedentes ni tenemos una idea clara y precisa sobre el fenómeno, el estudio se iniciaría como *exploratorio*.

2. La literatura nos puede revelar que hay *piezas* de teoría con apoyo empírico moderado; esto es, estudios descriptivos que han detectado y definido ciertas variables y generalizaciones. En estos casos nuestra investigación puede iniciarse como *descriptiva* o *correlacional*, pues se descubrieron ciertas variables sobre las cuales fundamentar el estudio. Asimismo, es posible adicionar variables a medir. Si pensamos describir el uso que un grupo específico de niños hace de los smartphone, encontraremos investigaciones que nos sugieren variables a considerar: tiempo que dedican diariamente a verlos, contenidos que ven con mayor frecuencia, actividades que realizan mientras los ven, etc. A ellas podemos agregar otras, como el control paterno sobre el uso que los niños hacen del smartphone. El estudio será *correlacional* cuando los antecedentes nos proporcionan generalizaciones que vinculan variables (hipótesis) sobre las cuales trabajar, por ejemplo: a mayor nivel socioeconómico, menor tiempo dedicado a la actividad de usar un smartphone.

3.La literatura nos puede revelar que existe una o varias teorías que se aplican a nuestro problema de investigación; en estos casos, el estudio puede iniciarse como explicativo. Si pretendemos evaluar por qué ciertos ejecutivos están más motivados intrínsecamente hacia su trabajo que otros, al revisar la literatura nos encontraremos con la teoría de la relación entre las características del trabajo y la motivación intrínseca, la cual posee evidencia empírica de diversos contextos. Entonces pensaríamos en llevar a cabo un estudio para explicar el fenómeno en nuestro contexto.

**Perspectiva que se le otorgue al estudio**

- El sentido o perspectiva que el investigador le dé a su estudio determinará cómo iniciar éste. Si piensa en realizar una investigación sobre un tema previamente estudiado, pero quiere darle un sentido diferente, el estudio puede iniciarse como exploratorio.
- De este modo, el liderazgo se ha investigado en muy diversos contextos y situaciones (en organizaciones de distintos tamaños y características, con trabajadores de línea, gerentes, supervisores, etc.; en el proceso de enseñanza-aprendizaje; en diversos movimientos sociales masivos, y muchos ambientes más). Asimismo, las prisiones como forma de organización también se han estudiado.
- Sin embargo, quizás alguien pretenda llevar a cabo una investigación para analizar las características de las mujeres CEO en diversos estados del país, así como qué factores hacen que ejerzan su liderazgo. El estudio se iniciaría como exploratorio, en el supuesto de que no existan antecedentes desarrollados sobre los motivos que provocan este fenómeno (el liderazgo).

Fuente: Henández-Sampieri et al. (2010) con adaptación propia

# ¿Cuál de los cuatro alcances para un estudio es el mejor?

Los cuatro alcances son igualmente válidos e importantes y han contribuido al avance de las diferentes ciencias. Cada uno tiene sus objetivos y razón de ser. En este sentido, no debe preocuparse si su estudio va a ser o iniciarse como exploratorio, descriptivo, correlacional o explicativo; más bien, debe interesarse por hacerlo bien y contribuir al conocimiento de un fenómeno. Que la investigación sea de un tipo u otro, o incluya elementos de uno o más de éstos, depende de cómo se plantee el problema de investigación y los antecedentes previos. La investigación debe hacerse *a la medida* del problema que se formule; ya que no se dice de manera *a priori*: *voy a llevar a cabo un estudio exploratorio o descriptivo*, sino que primero, debe plantear el problema y revisar la literatura y, después, analizar si la investigación va a tener uno u otro alcance.

Después de la revisión de la literatura*, el planteamiento del problema puede permanecer sin cambios, modificarse radicalmente o experimentar algunos ajustes*. Lo mismo ocurre una vez que haya definido el alcance o los alcances de su investigación.

# El problema y su planteamiento

Esta etapa corresponde a *afinar y estructurar más formalmente la idea de investigación*. El paso de la idea al planteamiento del problema en ocasiones puede ser inmediato, casi automático, o bien llevar una considerable cantidad de tiempo; ello depende de cuán familiarizado esté el investigador con el tema a tratar, la complejidad misma de la idea, la existencia de estudios antecedentes, el empeño del investigador y sus habilidades personales. Seleccionar un tema o una idea no lo coloca inmediatamente en la posición de considerar qué información habrá de recolectar con cuáles métodos y cómo analizará los datos que obtenga. Antes necesita formular el *problema específico* en términos concretos y explícitos, de manera que sea susceptible de investigarse con procedimientos científicos. *Delimitar* es la esencia de los planteamientos cuantitativos (Hernández-Sampieri et al. 2010)

U*n problema bien planteado está parcialmente resuelto*; a mayor exactitud corresponden más posibilidades de obtener una solución satisfactoria. El investigador debe ser capaz no sólo de conceptuar el problema, sino también de escribirlo en forma clara, precisa y accesible. En algunas ocasiones sabe lo que desea hacer, pero no cómo comunicarlo a los demás y es necesario que realice un mayor esfuerzo por traducir su pensamiento a términos comprensibles, pues en la actualidad la mayoría de las investigaciones requieren la colaboración de varias personas. **Ver Tabla A.8.**

**Tabla A.8. Planteando el problema de investigación**

| Criterios para planterar el problema |
|---|
| • El problema debe expresar una relación entre dos o más conceptos o variables.<br>• El problema debe estar formulado como pregunta, claramente y sin ambigüedad; por ejemplo, ¿qué efecto?, ¿en qué condiciones...?, ¿cuál es la probabilidad de...?, |

¿cómo se relaciona con...?

- El planteamiento debe implicar la posibilidad de realizar una prueba empírica, es decir, la factibilidad de observarse en la *realidad única y objetiva*.

**Elementos**

Son seis y están relacionados entre sí:

- Los *objetivos que persigue la investigación,.* Señalan a lo que se aspira en la investigación y deben expresarse con claridad, pues son las guías del estudio.
- Las *preguntas de investigación.* Orientan hacia las respuestas que se buscan con la investigación. Las preguntas no deben utilizar términos ambiguos ni abstractos.
- La *justificación .* Indica el porqué de la investigación exponiendo sus razones. Por medio de la justificación debemos demostrar que el estudio es necesario e importante. Considera: *la conveniencia, la relevancia social, implicaciones prácticas, valor teórico y utilidad metodológica*
- La *viabilidad del estudio, en la que* debe tomar en cuenta la disponibilidad de recursos financieros, humanos, materiales en incluso temporales que determinarán, en última instancia los resultados de la investigación.
- La *evaluación de las deficiencias en el conocimiento del problema.* Este elemento del planteamiento sólo se puede incluir si el investigador ha trabajado anteriormente o se encuentra vinculado con el tema de estudio, y este conocimiento le permite contar con una clara perspectiva del problema a indagar. Se plantea respecto al problema de investigación los siguientes cuestionamientos: ¿qué más necesitamos saber del problema?, ¿qué falta de estudiar o abordar?, ¿qué no se ha considerado?, ¿qué se ha olvidado? Las respuestas a estas interrogantes nos ayudarán a saber dónde se encuentra ubicada nuestra investigación en la evolución del estudio del problema y qué nuevas perspectivas podríamos aportar.
- *Consecuencias de la investigación.* Aunque no sea con fines científicos, pero sí éticos, es necesario que el investigador se cuestione acerca de este punto.

De lo que aquí se habla es de suspender una investigación por cuestiones de ética personal, y no de llevar a cabo un estudio de cuestiones éticas o morales. La decisión de realizar o no una investigación por las consecuencias que ésta pueda acarrear es una decisión personal de quien la concibe.

Fuente: Henández-Sampieri et al. (2010) con adaptación propia

# El marco teórico

Desde el punto de vista de la perspectiva, el conocimiento teórico inicia consiste en *detectar, consultar y obtener la bibliografía* y otros materiales útiles para los propósitos del estudio, de los cuales se extrae y recopila información relevante y necesaria para el problema de investigación. La revisión de la literatura puede iniciarse directamente con el acopio de las referencias o fuentes primarias, situación que ocurre cuando el investigador conoce su localización, se encuentra muy familiarizado con el campo de estudio y tiene acceso a ellas (puede utilizar material de bibliotecas, filmotecas, hemerotecas y bancos de información). Sin embargo, es poco común que suceda así, especialmente en lugares donde se cuenta con un número reducido de centros bibliográficos, pocas revistas académicas y libros. Por

ello, *es recomendable iniciar la revisión de la literatura consultando a uno o varios expertos en el tema* (ver ejemplos: Mejia-Trejo, 2017a en http://ojs.dpi.ulsa.mx/index.php/rci/article/view/1146 y Mejía Trejo, 2017b en: http://novascientia.delasalle.edu.mx/ojs/index.php/Nova/article/view/710/334),buscando (vía internet) fuentes *primarias* en centros o sistemas de información y bases de referencias y datos. Para ello, necesitamos elegir las *palabras claves*, *descriptores* o *términos de búsqueda*, los cuales deben ser distintivos del problema de estudio y se extraen de la idea o tema y del planteamiento del problema. Este último requiere de algunas lecturas preliminares para afinarse y completarse. Los expertos también nos pueden ayudar a seleccionar tales palabras. Para elegirlas se recomienda: escribir un título preliminar del estudio y seleccionar las dos o tres palabras que capten la idea central, extraer los términos del planteamiento o utilizar los que los autores más destacados en el campo de nuestro estudio suelan emplear en sus planteamientos e hipótesis. En la mayoría de los artículos de revistas es común incluir los términos claves al inicio o al final. La revisión de la literatura es analizar y discernir si la teoría existente y la investigación anterior sugieren una respuesta (aunque sea parcial) a la pregunta o las preguntas de investigación; o bien, provee una dirección a seguir dentro del planteamiento de nuestro estudio. Asimismo, nos podemos encontrar que los estudios antecedentes presentan falta de consistencia o claridad, debilidades en el método (en sus diseños, muestras, instrumentos para recolectar datos, etc.), aplicaciones que no han podido implementarse correctamente o que han mostrado problemas (Mertens, 2005). En cada caso varía la estrategia que habremos de utilizar para *construir* y *organizar nuestro marco teórico*. Ver **Tabla A.9.**

**Tabla A.9. Grado de desarrollo del conocimiento por la revisión de la literatura**

- Que existe una teoría completamente desarrollada, con abundante evidencia empírica y que se aplica a nuestro problema de investigación.
- Que hay varias teorías que se aplican a nuestro problema de investigación.
- Que hay *piezas y trozos* de teoría con cierto respaldo empírico, que sugieren variables potencialmente importantes y que se aplican a nuestro problema de investigación (pueden ser generalizaciones empíricas e hipótesis con apoyo de algunos estudios).
- Que hay descubrimientos interesantes, pero parciales, sin llegar a ajustarse a una teoría.
- Que sólo existen guías aún no estudiadas e ideas vagamente relacionadas con el problema de investigación.

Fuente: Henández-Sampieri et al. (2010) con adaptación propia

En el proceso de investigación siempre es conveniente efectuar la revisión de la literatura y presentarlade una manera organizada (llámese marco teórico, marco de referencia, conocimiento disponible o de cualquier otro modo), y aunque nuestra investigación puede centrarse en un objetivo de evaluación o medición muy específico (por ejemplo, un estudio que solamente pretenda medir variables particulares, como el caso de un censo demográfico en una determinada comunidad donde se medirían: nivel socioeconómico, nivel educativo, edad, género, tamaño de la

familia, etc.), es recomendable revisar lo que se ha hecho antes (cómo se han realizado en esa comunidad los censos demográficos anteriores o, si no hay antecedentes en ella, cómo se han efectuado en comunidades similares; qué problemas se tuvieron, cómo se resolvieron, qué información relevante fue excluida, etc.). Esto ayudará a concebir un estudio mejor y más completo. El papel del marco teórico resulta fundamental antes y después de recolectar los datos. Ver **Tabla A.10.**

**Tabla A.10. Papel del marco teórico en el proceso de investigación**

| Antes de recolectar los datos, nos ayuda a: | Después de recolectar los datos, nos ayuda a: |
|---|---|
| • Aprender más acerca de la historia, origen y alcance del problema de investigación. | • Explicar diferencias y similitudes entre nuestros resulta- dos y el conocimiento existente. |
| • Conocer qué métodos se han aplicado exitosa o errónea- mente para estudiar el problema específico o problemas relacionados. | • Analizar formas de cómo podemos interpretar los datos. |
| • Saber qué respuestas existen actualmente para las preguntas de investigación. | • Ubicar nuestros resultados y conclusiones dentro del conocimiento existente. |
| • Identificar variables que requieren ser medidas y observa- das, además de cómo han sido medidas y observadas. | • Construir teoría y explicaciones. |
| • Decidir cuál es la mejor manera de recolectar los datos que necesitamos y dónde obtenerlos. | • Desarrollar nuevas preguntas de investigación e hipótesis. |
| • Resolver cómo pueden analizarse los datos.<br>• Refinar el planteamiento y sugerir hipótesis.<br>• Justificar la importancia del estudio. | --- |

Fuente: Henández-Sampieri et al. (2010) con adaptación propia

*La construcción del marco teórico, se centra en el problema de investigación* que nos ocupa sin divagar en otros temas ajenos al estudio. *Un buen marco teórico* no es aquel que contiene muchas páginas, sino que trata con profundidad únicamente los aspectos relacionados con el problema, y que vincula de manera *congruente* (lógica) *y coherente (bien interconectado) los* conceptos y las proposiciones existentes en estudios anteriores. *Construir el marco teórico no significa sólo reunir información, sino también ligarla e interpretarla* (en ello la redacción y la narrativa son importantes, porque las partes que lo integren deben estar enlazadas y no debe saltarse de una idea a otra).

## *¿Cómo organizar el marco teórico?*

Una vez extraída y recopilada la información que nos interesa de las referencias pertinentes para nuestro problema de investigación, podremos empezar a *elaborar el marco teórico*, el cual se basará en la integración de la información recopilada.

Un paso previo consiste en *ordenar la información recopilada* de acuerdo con uno o varios criterios lógicos y adecuados al tema de la investigación. Algunas veces se ordena cronológicamente; otras, por subtemas o por teorías, etc. Por ejemplo, si se utilizaron documentos en archivos y carpetas (en la computadora) para recopilar la información, se ordenan de acuerdo con el criterio que se haya definido. De hecho, hay quien trabaja siguiendo un método propio de organización. En definitiva, lo que importa es que éste resulte eficaz. Ver **Tabla A.11.**

**Tabla A.11. Métodos para organizar la información del marco teórico**

| Método del mapeo |
|---|
| • Este método implica elaborar un mapa conceptual y, con base en éste, profundizar en la revisión de la literatura y el desarrollo del marco teórico.<br>• Como todo mapa conceptual, su claridad y estructura dependen de que seleccionemos los términos adecuados, lo que a su vez se relaciona con un planteamiento enfocado. |
| **Método por índices** |
| • La experiencia demuestra que otra manera rápida y eficaz de construir un marco teórico consiste en desarrollar, en primer lugar, un índice tentativo de éste, global o general, e irlo afinando hasta que sea sumamente específico, para posteriormente colocar la información (referencias) en el lugar correspondiente dentro del esquema.<br>• A esta operación puede denominársele *vertebrar* el marco o perspectiva teórica (generar la columna vertebral de ésta).<br>• Por otra parte, es importante insistir en que el marco teórico *no es un tratado de todo aquello que tenga relación con el tema global o general de la investigación*, sino que se debe limitar a los antecedentes del planteamiento específico del estudio. Si éste se refiere a la tipología de consumidores en internet en adultos de un cierto perfil, la literatura que se revise y se incluya deberá tener relación con el tema en particular; no sería práctico incluir apartados como: *la historia de internet*, *cómo son las redes sociales en general* o *el efecto del acceso a intenet por niños*, etcétera. |

Fuente: Henández-Sampieri et al. (2010) con adaptación propia

## *Confirmando una revisión adecuada de la literatura*

Se debe tomar en cuenta el planteamiento del problema, el tipo de reporte y políticas a las que entregará el documento y el área en que se encuentre, además del presupuesto. No hay una respuesta exacta ni mucho menos, pero existen sugerencias. Ver **Tabla A.12.**

**Tabla A.12. Referencias sugeridas de acuerdo al documento de**

| Cantidad | Documento |
|---|---|
| 18-29 | • Revista científica de divulgación |
| 30-40 | • Revista científica nacional arbitrada/Tesina de licenciatura |
| 50-79 | • Revista científica nacional/ internacional indizada (50% de menos de 5 años) |
| 80-99 | • Protocolo de investigación y/o tesis de maestría |
| 100-199 | • Protocolo de doctorado (80% de menos de 5 años) |
| +200 | • Tesis doctoral (80% de menos de 5 años) |

Fuente: propia

Deben ser referencias directamente vinculadas con el planteamiento del problema, es decir, se excluyen las fuentes primarias que mencionan indirectamente o de forma periférica el planteamiento, aquellas que no recolectan datos o no se fundamentan en éstos (que son simples opiniones de un individuo) y también las que resultan de trabajos escolares no publicados o no avalados por una institución. Ver **Tabla A.13.**
Una correcta revisión de la literatura así como selección de referencias para integrarlas en el marco o perspectiva teórica, es posible utilizar siguientes criterios en forma de preguntas. Ver **Tabla A.13**

**Tabla A.13. Criterios que confirman una adecuada revisión de literatura**

| |
|---|
| • ¿Acudimos a un par de bancos de datos, ya sea de consulta manual o por computadora? y ¿pedimos referencias por lo menos de cinco años atrás? |
| • ¿Buscamos en directorios, motores de búsqueda y espacios en internet? (por lo menos tres). |
| • ¿Consultamos como mínimo cuatro revistas científicas que suelen tratar el tema de interés? ¿Las consultamos de cinco años atrás a la fecha? |
| • ¿Buscamos en algún lugar donde había tesis y disertaciones sobre el tema de interés? |
| • ¿Buscamos libros sobre el tema en al menos dos buenas bibliotecas físicas o virtuales? |
| • ¿Consultamos con más de una persona que sepa algo del tema? |
| • Si, aparentemente, no descubrimos referencias en bancos de datos, bibliotecas, hemerotecas, videotecas y filmotecas, ¿contactamos a alguna asociación científica del área en la cual se encuentra enmarcado el problema de investigación? |
| • ¿Quién o quiénes son los autores más importantes dentro del campo de estudio? |
| • ¿Qué aspectos y variables se han investigado? |
| • ¿Hay algún investigador que haya estudiado el problema en un contexto similar al nuestro? |
| • ¿Tenemos claro el panorama del conocimiento actual respecto a nuestro planteamiento? |
| • ¿Sabemos cómo se ha conceptualizado nuestro planteamiento? |
| • ¿Generamos un análisis crítico de la literatura disponible?, ¿reconocimos fortalezas |

| |
|---|
| y debilidades de la investigación previa?<br>• ¿La literatura revisada se encuentra libre de juicios, intereses, presiones políticas e institucionales?<br>• ¿El marco teórico establece que nuestro estudio es necesario o importante?<br>• ¿En el marco o perspectiva teóricos queda claro cómo se vincula la investigación previa con nuestro estudio? |

Fuente: Henández-Sampieri et al. (2010) con adaptación propia

# Fuentes de información

La información, se define como el *conjunto de fuentes de conocimientos necesarios para comprender situaciones, tomar decisiones o evaluar las ya tomadas .* Los investigadores deben tener muy claros los objetivos de una investigación, saber qué información reunir , dónde encontrarla y cómo tienen que tratarla. Es nmuy importante saber en que consiste cada una de las fuentes de información porque su utilidad y aplicación no es la misma en las investigaciones. Por ejemplo, no se pueden aplicar *técnicas proyectivas* para hacer una segmentación de mercados, ni las entrevistas son válidas cuando se trata de evaluar el interés que muestran los niños muy pequeños por un juguete nuevo.

## *Por su disponibilidad*

La información puede ya existir cuando surge su necesidad para realizar una investigación o por el contrario, debe crearse para cumplir este cometido. En función de sy disponibilidad se tiuene infomación secundaria y primaria. Ver **Tabla A.14.**

**Tabla A.14. Información primaria y secundaria**

| Informacion primaria |
|---|
| • Es la que el investigador crea expresamente para un estudio concreto. Esta información no existe en el momento en que se plantea la necesidad de utilizarla. Las organizaciones deben entonces, acudir a investigaciones ex profeso para obtenerla, empleando para ello diversas técnicas, como la observación, la pseudocompra, las reuniones de grupo o los métodos experimentales . Otras veces se obtiene la informac mediante diversos tipos de encuestas, que posteriormente serán tratadas con las herramientas estadísticas adecuadfas.<br>• Sin embargo, es poco probable que las necesidades de información sean tan especiales o únicas que no exista previamente ningún tipo de datos o estudios relacionados con la investigación que se desee realizar.<br>• Con mucha freecuencia los investigadores encuentran datos, informes, páginas web o metodologías que les resultan sumamente útiles para sacar adelante sus proyectos. *Esta información que ya existe*, que se encuentra disponible en el momento en que surge la necesidad de su utilización, recibe el nombre de *información secundaria.* |
| **Información secundaria** |
| • La información secundaria puede haber sido creada en el pasado por los |

investigadores o puede haber sido generada por terceros ajenos a ellos. En estos casos se habla, respectivamente, de información *secundaria interna o externa*.

La *importancia* de la información secundaria, al margen de cual sea su soporte resulta evidente, por lo siguiente:

- Un investigador no debería buscar una información si otros ya lo han realizado antes por él.
- La información secundaria ahorra mucho tiempo y esfuerzos, de todo tipo en el desarrollo de las investigaciones
- El costo de adquiriri uinformacion secundaria es inferior al de obtener información primaria

Muchas veces proporciona al investigador la única información que puede emplear, pues este por sus propios medios sería incapaz de recoger todo elcúmulo de datos que precisa para alcanzar susu objetivos.

Fuente: Grande-Esteban y Abascal-Fernández (2017) y Hernández-Sampieri et al., (2010) con adaptación propia

## *Por su tendencia a tratamiento cuantitativo*

La propia naturaleza de a información condiciona las posibilidades de tratamiento estadístico. Por ejemplo, se puede proponer que solamente la información que proporcionan: *las encuestas, las bases de datos y los paneles* sean susceptibles de tratamiento cuantitativo, con la ayuda de técnicas de matemáticas y estadísticas.

Por el contrario, la información que se puede obtener a través de: *grupos de discusión, pseudocompra, observación, técnicas de creatividad oproyectos o entrevistas* no sea posible de tratarla en la inmensa mayoría de los casos con la ayuda d etécnicas estadísticas omatemáticas. ; será interpretada según los conocimentos criterios de los analistas.

Las fuentes de información que proporcioan datos que se pueden tratar con técnicas estadístics o matemáticas reciben el nombre de fuentes de *información cuantitativas*. En caso contrario, se denoinan *técnicas cualitativas.*

## *Por su grado de dinamismo: estáticas y dinámicas*

Existen fuentes de información que la proporcionan *una sola vez en el tiempo (estáticas)* y otras las suministran de *forma periódica y continuada (dinámicas)*. Para analizar determinadas investigaciones, sólo es necesario obteer información en unmomento del tiempo; por ejemplo, para conocer la intención de compra antes de una campaña de mercadotecnia. Otras veces, conviene obtener información en diferentes momentos de tiempo, para hacer comparaciones; por ejemplo, cómo evoluciona la cuota de mercado de un aceite nuevo automotriz, que acaba de salir

## *Por su naturaleza*

Se tienen de varios tipos y cada uno tine características particulares de cómo recoger y analizar. Ver **Tabla A.15.**

**Tabla A.15. Fuentes de información por su naturaleza**

## Fuentes bibliográficas e internet

•Muchas investigaciones se pueden llevar a cabo a partir d einformación que aparece publicada en libroso se pude descargar de internet. Esta informaciòn, debe provenir de fuentes serias como portales de industria, de gobierno así como académicas de diversas bibliotecas tanto de paga como de acceso libre, para serr tomadas en cuenta.Se apoya en los motores de búsqueda de internet (**SEO**. *Search Engine Optimization*).

•Las posibilidades de encontrar cuaquier información son muy altas. Se debe considerar que siempre se está actualizando. Es muy útil formular tanto *palabras* como *frases clave* para irlas aplicando en la búsqueda y hacer descartes de la mismas, con el fin de precisar paulatinamente las búsquedas.

## Observación

**Concepto**

En ocasiones no es posible recoger información proporcionada directa y voluntariamente por las personas. Entonces se presta atención a cómo se comportan. Por ejemplo se puede observar el comportamiento de los supervisiores de la cdena de armado de una empresa automotriz ante el retraso por rezago d eentrega de partes o cómo eligen los compradores la marca de insumos de cómputo que adquieren de los proveedores de agencia. Es una técnica *cualitativa, primaria, personal e indirecta,* También se aplica en investigaciones exploratorias. Consiste en la *atención, examen e interpretación* sin que el investigador intervenga en su desarrollo con objeto de modificar el curso natural de su desenvolvimiento. Se aplica porque en ocasiones no es posible la colaboración de las personas para obtener información, debido a diversas razones:

•Incomodidad del encuestado.

•Las personas puede que sean materialmente incapaces de proporcionar información, como niños pequeños, con evidencias de uso por mordidas a un juguete nuevo por diseño.

•Con frecuencia, las personas no proporcionan información porque no reparan en sus conductas. No todos recurdan a detalle lo que hacen, sien embargo existen técnicas como pegar adhesivos especiales en una revista para que una vez abierta las paginas, estas no vuelvan a pegarse. Esto muestra evidencia de qué paginas le interesó hojear por su contenido.

**Condiciones**

•En ocasiones, no interesa descubrir que se está observando a otras personas. Por ejemplo, un inspector de vendedores en unos grandes almacenes no debe descubrirse como tal, pues de lo contrario aquellos cambiarían de conducta.

La observación recoge información en el momento de consumarse los hechos. Se hace de manera:

•*Formal o estructurada,* con planificación y objetivos previos antes de iniciar la investigación. Un ejemplo es la de observar si un consumidor percibe el cambio de rediseño de un producto y su ubicación dentro de la tienda, buscándolo al no encontrarlo en su lugar habitual. Deben por cierto presentarse condiciones como:

•*Los datos deben ser accesibles*. El observador se interesa por conductas por hechos observables y concretos que esta viendo. Si deseara conocer *motivaciones, actitudes u opiniones,* deberían emplearse otros métodos.

•*La observación se aplicará a conductas supuestamente repetitivas*, *frecuentes o*

*predecibles* Por ejemplo, establecimientos como supermercados, librerías grandes, etc. no tiene sentido hacerlo en situiaciones esporádicas, lento, difícilmente predecible y condicionada por múltiples factores como la decisión de una pareja en una inmobiliaria en la compra de un piso.

•*El suceso debe desarrollarse en un corto período de tiempo*, pued de lo conrario, varían las condiciones del entorno que afectan a las conductas. Es posible *captar condutas impulsivas o rutinarias de forma instantánea*.

o

•*Informal o no estructurada,* limitándose a examinar conductas extrayendo conclusiones generales. No tiene planeación previa, no existen hipótesis ni objetivos concretos de comportamiento. Un ejemplo, es el observar qué intineriario siguen los compradores dentro de un hipermercado, dónde se paran, cómo deambulan para sacar conclusiones de distribución de los artículos a exponer.

**Escenarios**

•Los escenarios para recoger información pueden ser múltiples,sin embargo, se propone:

•En el *entorno natural* de las personas. Por ejemplo, se puede acudir a un supermercado para observar la ruta que sigien los compradores. También se puede observar cómo se comportan las personas en un banco o mientras esperan el autobús, o cómo juegan en una guardería.

•En el *entorno artificial*, creado para estudiar los comportamienbtos o reacciones. Se puede simular unestablecimiento comercial y ofrecer un productp para que sea probado, mientras se filman los hechps con una cámara oculta o a través de un espejo-pared con u sólo sentido de visión hacia adentro (*cámara Gesell*).

La condición de condiciones artificiales presenta ventajas , pro ejemplo, no hay que esperar a que los hechos sucedan, siendo posible adaptar el laboratorio con múltiples cámaras que registren eventos diversosque se presenten durante la prueba, grabarlos y analizarlos de forma controlada.El problema es la conciencia de los actores dentro del laboratorio que sabejn se encuentran en un ambiente controlado, hecho que puede condicionar las acciones y modificar los resultados.

•Por su importancia, merece un trtamiento especial la *observación de residuos* . Los observadores pueden fijarse en qué se deposita en los *contenedores de basura.* Este proceder ha dado origen a una *nueva ciencia*, llamada *garbology (basurología),* de la que se tienen antecedentes de la Universidad de Arizona (1975), la cual permite:

•*Estimar el comportamiento del consumidor de ciertos grupos sociales.* Se estiman perfiles de consumidor.

•*Estimar cuotas de mercado por grupos.* Examinando el número de envases, àuetes, etiquetas, etc. por grupo socioeconómico, se estiman con gran aproximación.

•*Medir la reacción de los consumidores* ante cualquiera delos elementos del mix de mrketing al variar precios, diseño de producto, cambios en la publicidad o la plaza, etc.

•Conocer por lo tanto el impacto de las campañas de marketing directo al revisar, de la basura, la cantidad de correo con publicidad., que no han sido abiertas.

**Medios técnicos**

Los medios técnicos para recoger información por observación, son tres:

- Contar.
- Medir.
- Buscar pistas o retos

Lo cual se refleja en contadore los contadores de tráfico en las carreteras, fotografñias, videos, cámaras ocultas que registran las acciones de los sujetos de estudio.

**Limitaciones**

- La observación *no destaca la motivación* nbi explica el comportamiento del sujeto o consumidor. Se conocen las consecuencias de los procesos delas tomas de decisiones, pero no sus mecanismos.
- En las conductas observadas *influyen variables del entorno*que no siemore son tenidas en cuenta. Por ejemplo, las condiciones de temperatura internas y/o externas de clima de un centro comercial y sus implicaciones de compra por parte de los consumidores.
- La observación es una técnica *estática.* Recoge información de una conducta que se está observando en un momento dado. Esa conducta probablemente sea un hábito o tal vez un comportamiento irrepetible. *El observador nunca lo sabrá.*
- *La percepción de lo que se observa es selectiva.*El observador interpreta lo que ve.. Otra persona pudiera llegar a concluiones diferentes.
- *En la observación, no se seleccionan muestras por procedimientos aleatorios.* Se observan comportamientos de grupos de personas obtenidos por conveniencia, o por detención. Esto significa *que si el muestreo no es aleatorio no se pueden acotar los porcentajes de error.*
- *La extrapolación de conclusiones es difícil.*Esa es una característica más que una limitación, de los metodos cualitativos. En principio sólo son generalizables los resultados de muestras amplias, aleatorias y representativas de una población.Sin embargo, razonando con sentido común, sería poco probable que el observador se hubiera encontrado con uyna población rara. Conla debida cautela se pueden hacer tímidas extrapolaciones.

En caso de que la observación se produzca en un entorno no familiar , se pueden *producir reacciones en las personas que no sean las habituales* en ellas.

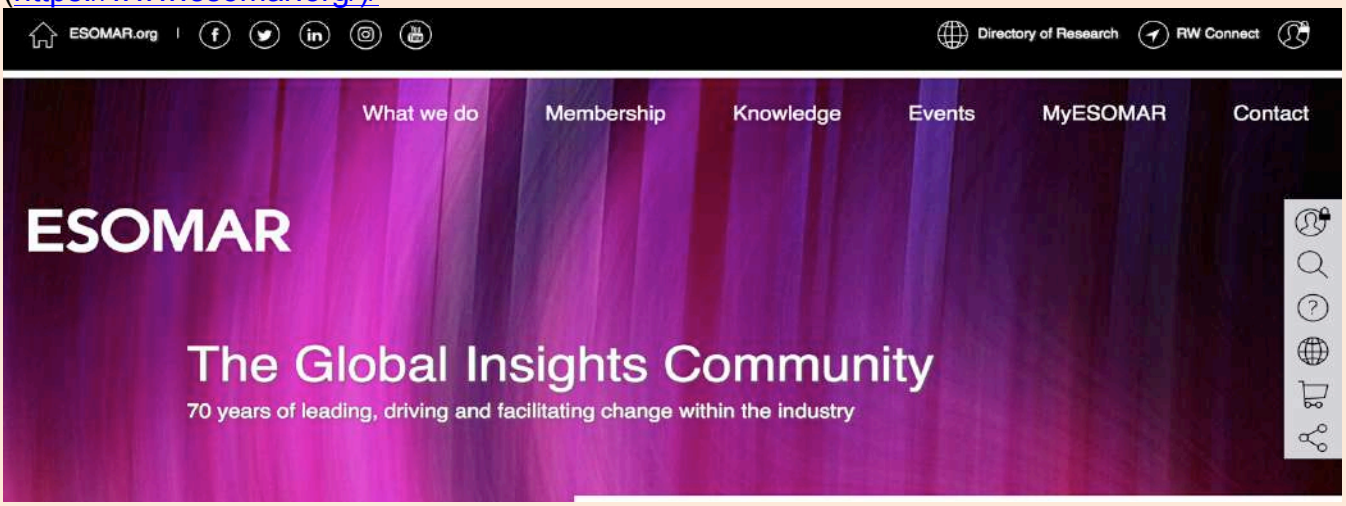
## Pesudocompra (Pseudoadquisición)

Esta técnica consiste en que el investigador actúa como u cliente y observa el comportamiento del vendedor, con objeto de valorar sus conocimientos, actitudes hacia las marcas que vende, sus preferencias, capacitación, competencia y calidad del servicio prestado. Por ejemplo, los evaluadores de la guía Michelin actúan en estos términos. Es una técnica *cualitativa, primaria, estática, personal e indirecta*. También se aplica en investigaciones exploratorias. Es considerada como *la visita qie realizan investigadores a detallistas u otros puntos de venta simulando que se trata de clientes normales que ya sea quieren comprar un bien, producto o servicio, o bien, asesorarse. (Mistery Shopper).* Las personas que la practican no deben de identificarse como investigadores, porque entonces alterarían los comportamientos de las que son investigadas. Ésta es una técnica que permite obtener informaciónsobre los siguientes aspectos:

- *Calidad de servicio* prestado a los clientespor los distintos canales de distribución
- *Grado de conocimiento* que sobre el producto o servicio que ofrece una empresa tienen los distribuidores.
- *Formación de las personas* que atienden al público.

• *Problemas o deficiencias* que puedan observarse respecto al servicio al cliente.
Actitudes de los distribuidores o vendedores con relación a la oferta de la empresa respecto a la oferta de otros competidores, en caso de que se distribuyan productpos de diversos oferentes: cuál se recomienda, qué atributos destacan, etc.
La pseudocompra es un técnica de investigación de mercados y quienes la llevan a cabo deben respetar las normas del código CCI/ESOMAR (https://www.esomar.org/uploads/public/knowledge-and-standards/codes-and-guidelines/ICCESOMAR_Code_Spanish_.pdf)

, especialmente las que se refieren al anonimato del informante. ESOMAR (https://www.esomar.org/)/




AEDEMO (http://www.aedemo.es/aedemo/)

que proporcionan las siguientes directrices para realizarlo:

- *Debe diseñarse y realzarse de forma que no suponga una pérdida de tiempo ni de dinero* ni para el investigador ni para el investigado.
- La investigación debe realizarse de forma que *no se puedan identificar a las personas* o establecimientos investigados.
- La investigación debe realizarse mediante s*imples observaciones, cortas entrevistas o compras de duración normal.*
- La pseudocompra con *fines de control* o de simular incrementos de la demanda, no es investigación de mercados y está prohibida por el código ICC/ESOMAR.
- La pseudocompra también se puede realizar por teléfono. Por ejemplo, se puede preguntar cuánto tiempo tardará en realizar una reparación a su auto sin llevarlo al taller.

## Técnicas de creatividad

Consisten. en reuniones de personas en las que el grupo que participa propone espontáneamente soluciones a problemas o ideas para desarrollar productos nuevos, o para mejorarlos, o para encontrar causas de problemas y buscar posibles soluciones. Por ejemplo, un grupo de empleados de una empresa puede aportar ideas sobre cóno reducir costos o sobre que atributos de producto una aspiradora podría tener para diferenciarlo de la competencia. Constituyen un conjunto de técnicas de obtención de información *cualitativa,primaria, directa, estática y personal,* que consisten en que un grupo de personas proporcionaideas orientadas, muy frecuentemente, a la creación de nuevos productos o a la solución de problemas de diversa índole que pudieran aparecer en las organziaciones . suele aplicarse en investigaciones exploratorias y puede ser:

- *Intuitiva.* Basada en las argumentaciones obtenidas tras un proceso no demasiado formal, de un conjunto de personas.Pr ejemplo el *brainstorming.*
- *Formal.* Basada en juicios razonables que justifiquen las ideas que se proponen . Por ejemplo, la solicitu de ideas de un nuevo tecleado para comoutadora y que se defiendan razinadamente los nuevos atributos propuestos.

Las técnicas de crreatividad requieren la colaboración de un colectivo de personas.

**Limitaciones**

En una organización, la propuesta de ideas por parte de las bajas jerarquías queda condicionada por las de alta jerarquía inhibiendo la originalidad y oportunidad de las nuevas ideas. Se sugiere crear grupos homogéneos de jerarquía organizacional.

**Recomendaciones**

Se sugiere su aplicación para diseño de productos nuevos, proponer mejoras a los existentes, encontrar denominaciones de marca y en general ,para identificar problemas en las organziaciones y proponer soluciones.

## Dinámica de grupos

**Concepto**

También llamada *grupo de discusión.* Esta técnica consiste en que un conjunto de personas debate y opina sobre un tema fijado de antemano , bajo la dirección de un moderador. Por ejemplo, in conjunto de consumidores que debata sobre la utilidad de un nuevop proyector de imágenes. Es una técnica de recogida de información *cualitativa, primaria, estática, personal y directa,* qu esuele aplicarse en investigaciones de naturaleza exploratoria. Consiste en la reunión de grupos de personas , bajo un moderador, que conversan entre ellas y él, sobre un tema prevoiamente convenido., con obejeto de encontrar una solución a un problema o prporcionar información sobre aquel , sugerir un producto o servicio, encontrar causas de ciertos fenómenos, etc. Port ejemplo, se puede solicitar una dinámica d grupos para resolver el problema de qué hacer con la basura de una ciudad.

**Condiciones**

- *Prevismente haber definido los objetivos* a alcanzar en la investigación y que sean conocidos por todos los integrantes del grupo
- *Este debe ser homogéneo y adecuado*, porque posea conocimientos y experiencias par aproporcionar información útil orienbtada a la determinación de los objetivos de la investigación.
- El grupo debe tener una *dimension adecuada*, *normalmente de siete personas,* que permita la conversacion entre los integrantes.
- *El grupo debe ser fácilmente controlable* por el moderador o dorector de la reunión.

Ls reuniones de grupo se pueden grabar . La interpretación se hace basándose en lo que se dice, cómo se dice, qué actitudes se adoptan, la rapidez de las respuestas., etc.

**Utllidad**

- *Conocer el léxico* de los grupos para conseguir comunicaciones eficaces o para preparar cuestionarios comprensibles.
- *Aproximarse a los problemas o familiarizarse con ellos*. Por ejemplo, , proqué los trabajos rechazarían cierto tipo del iderazgo, o consumo de ciertos edulcorantes, porué hay más demanda de turismo médico, etc.
- Generar ideas de nuevos productos. Los consumidores como agente activo de innovación.
- *Desarrollar test de concepto.*Conocer que creen los sujetos de estudio, como los consumidores sobre qué es un producto y sus atributos. Se debe detectar el contraste de lo que piensa el usuario vs.el que lo diseña.
- *Identificar atributos delos productos.* Los consumidores pueden proporcionar a los fabricantes ideas sobre el uso de losproductos o sobre sus atributos. Por ejemplo, suavizantes para ropa que además de dejarla tersa, le abonaban un mejor aroma como consecuencia no prevista.
- *Evaluar la conveniencia de hacer test de mercado*. Cuando las empresas ya han creado sus productos suelen realizar pruebas de mercado, seleccionando áreas geográficas limitadas, donde ensayan precios, canales de distribución y de comunicación.Estas pruebas son caras y

llevan tiempo. Sin embargo, ésta técnica puede servir para clarificar el precio que los consumidores estan dispuestos a pagar, cómo debería diseñarse la comunicación y que tipos de eestablecimientos consideran más convenientes para comprar el producto antes de decidir realizar el test de mercado.

•*Hacer previsionbes del éxito de nuevos productos.* En un grupo se puede detectar el interés que suscite un producto o servicio. En las reuniones se podría detectar la utilidad prevista, ventajas, inconvenientes, etc.

•*Conocer causas de la disminución de ventas* ante situaciones incomprensibles a casos donde por ejemplo, existe evidencia de la fuerza de ventas por impulsarlas, su precio competitivo, calidd de manufactura, etc. El grupo podría evidenciar debilidsdes del producto que explicaran la situación.

•*Conocer el uso que se hace de los productos.* Así, fabricantes de galletas de dulce con baja cntidad de azíucar, se darían cuenta que son mayormente adquiridas por adultos mayores que por niños de edad preescolar

•*Evaluar la competitividad de los productos* basada en la percepción de la calidad de los componentes o prestaciones. Así, componentes de plástico resitente contra metálicos en una podadora, no son un diferencial al cliente si percibe que su aplicscción y rendimiento es el mismo.

•*Evaluar envases.* Los concuidores pueden ser jueces ideales para evaluar prestaciones de ergonomía, estética, diseño, tamaño, capacidad de carga, transporte y conserva, etc.

•*Evaluar precios.* No se requiere realizar cálculos complejos más que solictar al consumidor el precio por el que estaría dispuesto a pagar estando consciente de sus prestaciones y atributos. Por adición se llega a calcular un precio demercado razonable y que puede ser aceptado por los consumidores.

•*Pretest publicitario antes de una campaña de publicidad.* Las campañas sobn caras,. Antes de lanzar un anuncio a los medios, un grupo de personas podría, tras ser visto, opinar sobre su poder de comunicación, credibilidad, aspectos estéticos, etc.

**Diseño**

•*El moderador.* Generalmente se acude a sociólogos y sicólogos, profesonales en esta práctica, considerando:

- •Que debe *ser aceptado* por el grupo
- •Ser sensible a nuevas ideas que surjan del grupo.
- •*Tener buena memoria para los nombres* ya que colaboran mejor
- •Debe ser neutral tanto en el planteamiento de objetivos como en la solución de conflictos de la investigación
- •Debe mantener la conversación del grupo y evitar desvíos

**Características del grupo**

•*Debe ser incentivado* desde el *punto de vista económico* por el tiempo que dedican , normalmente como compensación a su esfuerzo traducido en bonos o cheque de compra en tiendas designadas así como desde el *punto de vista emocional* al animarlos a continuar y concluir.

•*No debe ser numeroso* ya que en caso contrario se vuelve *incontrolable* y existe tendencia a

*dilución de la participación* de algunos miembros,. Generalmente deben ser siete personas para que no se formen grupos o bandos del mismo tamaño.

•Los integrantes deben tener *conocimientos y experiencia* afines a la materia

•*Debe evitarse la formación de conflictos en el grupo.* Una norma general es que compartan características básicas.

**Tipología de participantes**

Se presentan por lo general:

•*El colaborador*. Es el idóneo para participar

•El listo. Bien enterado, valiosos, trta de sobresalir e imponer. Problemático.

•*El agresivo.* Neurótico, beligerante. Debe evitarse que ataque al grupo.

•*El manipulador.* Dominador del moderador, hostil. Se debe neutralizar.

•*El indolente.* Pasivo, poco colaborativo, interesante. Se debe animar a participar.

•*El tímido.*Reacio, con dudas. Se debe animar a participar.

•*El narcisista.*Siente ser superior al resto, presuntuoso, problemático.

•*El terco.* Con ideas fijas, inflexible, difícil.Pr*oblemático.*

•*El contradictorio.* Actitud cambiante, extremosos, confuso. Resaltar lo favorable.

**Valoración**

**Ventajas** notables, son:

•*Rapidez.* Las sesiones duran de 2 a 3 horas. En pocos días se realizan y se obtienen resultados. El problema es localizar e integrar al grupo. Las empresa e institutos de estudios de mercado normalmente tienen grupos localizados para trabajar con ellos diversas temáticas.

•*Flexibilidad.* Técnica muy simple en la que lo más importante es actuar con sentido común. No existen normas muy rígidas y el moderador emplea su iniciativa e imaginación para lograrlas de la mejor manera.

•*Empatía*. El moderador ideal reduce la distancia entre los participantes

•*Dinamismo.* Basado en la *sinergia* ya que se producen más ideas en grupo que en lo individual. Produce el efecto *bola de nieve,* ya que existen ideas que inspiran a los demás.Se producen encadenamientos de ideas. Con un moderador hábil los participantes se sentirán cómodos y darán información espontánea y de calidad.

**Desventajas**

•*Subjetivismo.* Por la interpretación de las opiniones. No existen manuales o tratados de cómo interpretar.

•*Artificialidad.* Es una técnica orquestada, poco natural. Es posible que una persona no manifieste sus propias opiniones, sino las que se espera de su condiciój, o no se atreva a contradecir al resto de un grupo que opina al contrario de ella. También es posible que por diversas razones como la timidez, se reserve la participación de ideas valiosas.

•*Parcialidad.* El moderador u otra persona u otra persona puede cautivar a algún asistente (o no) y hacerle responder , en un sentido no esperado de él en otrs condiciones. Incñuso es posible que sea el moderador quien se vea infuenciado pro algún asistente de la sesión.

## Entrevistas

**Concepto**

Son encuentros don personas para recoger informaciòn relativa a sus comportamientos, opiniones y actitudes. por ejemplo, se puede eentrevistar a un especialista para que dé su opinión sobre el desarrollo tecnológico en el futuro próximo.
Es una técnica *cualitativa, primaria, estática, personal y directa* que suele aplicarse en investigaciones de naturaleza exploratoria. Una entrevista no es más una **conversación entre dos personas,** frente a frente, para intercambiar información, ideas, opiniones o sentimientos. La entrevista , como técnica cualitativa de información, persigue propósitos bien definidos ; es más que una simple conversación . Puede ser de varias clases , en fucnión de la rigidez con que se siga el guión que la conduce.

- La entrevista *estructurada* se caracteriza por que el entrevistador reqaliza exclusivamente las preguntas que figuran en un guión
- Es *semistructurada* cuando existye la libertad para que el eentrevistador introduzca ciertas preguntas según quien sea entrevistado y en función del desarrollo de la entrevista.
- En la *entrevista en produndidad*, se perfila un guión general que no se ciñe a preguntas concretas.

La elección entre los distintos tipos de entrevista dependerá de los objetivos que se persigan y de quien sea el que responde. Si se tratara de analizar un problema concreto , que exige una muestra amplia y opiniones inequívocas, se puede acudir a una entrevista estructurada. Si la muestra fuera reducida e interesara valorar las respuestas infdividuales, las entrevistas semiestructurada y estructurada pueden ser las más adecuadas. Cuando se trate de obtener información proporcionada por personas relevantes , como científicos, artistas, literatos, gerentes, supervisores, etc., lo aconsejable es la utilización de la *entrevista en profundidad.*

**Proceso de entrevista en profundidad**
**El guión.**
Una entrevista no se improvisa. El entrevistador debe de preparar un guióny planificar cómo hará las preguntas. La forma de preguntar depende de tres factotres:

- *La longitud de la entrevista.* Obviamente, si la entrevista fuera corta no se pueden hacer preguntas largas. Se piden respuestas concretas, no divagaciones.
- *La naturaleza del a pregunta.* Si la informacion que se desea obtener es compleja, las preguntas deben formularse con alta precisión y riqueza de matices, para que se comprendan bien y no den lugar a equívocos.
- *La naturaleza dela investigación.* Si fuera *exploratoria* la entevista sigue un guión poco estructurado. Por el contrario, en una investigación *descriptiva,* el guión debe apoyarse en preguntas meditadas y estructuradas, si bien es cierto que la entrevista no suele emplearse en investigaciones de esta naturaleza, en beneficio del uso de cuestionarios que aunque guardan alguna relación con las entrevistas, no son propiamente lo mismo.

**Fases**
**Introduccción**
La entrevista comienza con una fase introductoria , cuyo objetivo es proporcionar información a la persona que va a ser entrevistasa para conseguir mejor su colaboración. Se le informa sobre:

- El objetivo de la entrevista.
- El uso que se hará de la información que proporcione.

•Lo que se espera del entrevistado a lo largo de la misma.
En la fase introductoria, en definitiva, lo que se pretende es informar al entrevistado el objetivo de la misma para que colabore sin recelos y proporcione información fiable.

**Desarrollo**
Finalizada la fase introductoria, el entrevistador comienza a preguntar. Para ello debe trasladar los objetivos de la investigación a preguntas. El entrevisytador debe tener en cuenta los siguientes aspectos ala hacer preguntas

•*No debe se demasiado directo.* Cpnviene que utilice rodeos. Las preguntas demasiado directas pureden proporcionar respuestas poco fiables, pued muchas personas puede que no respondan o, si lo hacen, sea en los términos sociales aceptados, que no tienen por qué coincidir con sus opiniones.

•*No debe optarse por preguntas abiertas.* Las preguntas cerradas procedentes de guines estructurados recogen opiniones solamente, acuerdos o desacuerdos con afirmaciones que se leen. Las preguntas abiertas proporcionan respuestas más elaboradas e información más rica, relativa a puntos de vista, opiniones, motivaciones, actitudes y sentimientos.

•*Uso de antecedentes.* Edto es una pregunta que se ha formulado con anterioridad y que sirve para profundizar en un tema o reconduir la entrevista. Permite relacionar preguntas , siguiendo un proceso continuoy lógico: Relacionando dos preguntas, el entrevistador puede conocer si el ent¡revistado es coherente en sus respuestas, o por el contrario, responde de forma errática.

•*Deben evitasrse las preguntas que puedan dar la impresión de que el entrevistador espera una respuesta* determinada. Algunas llevan implícita la contestación. Por ejemplo, no debe preguntarse: *¿Qué opina usted son¡bre el problema que causan las drogas,* porque la respuesta es evidente y el enetrevistado seguramente repetirá opiniones socialmente pactadas y tópicas. Es mejor preguntar: *¿Qué soluciones sugeririía Usted para atajar el problema de las drogas?*

•*Deben evitarse preguntas con carga emocional,* obviando las palabras que puedan enfatizar este sentido. Por ejemplo, no debería hablarse de los *pobres migrantes o de los indefensos niños.* Deben vigilarse terminos con carga: *afectiva, moral, espiritual o ideológica.* De lo contrario, el entrevistador seguramente sesgara las respuestas, pues actúa como un juez,

•*La entrevista debe seguir un orden secuencial.* Debe tener un principio , un intermedio y un final. Deben ordenarse las preguntas de forma que si el entrevistado se niega a seguir respondiendo en un momento determinado, se miniminice la pérdida de información. Por ejemplo, se podría preguntar:

•¿Qué hace en su tiempo libre?

•¿De esas actividades , cuál es la más importante para Usted?

•¿Va mucho al cine?

•¿Le gustan las películas de aventuras?

El interés de las respuestas a estas preguntas decrece a medida que se avanza en su formulación.

**Factores que dificlutan su desarrollo**

•*El tiempo.*Una entrevista puede durar varias hotas . Los entrevistadod pueden manifestar reticecncias a respondere cuando se les informa de su duración , que podrian considerar

excesiva.

•*El nivel cultural y la edad.* A medida que decrece o se incrementa la edad , aumentan las dificultades para recoger información mediante entrevistas. Las personas muy mayores o muy jóvenes no son los mejores candidatos para hacerles entrevistas. Cuando el nivel cultural es bajo, también aumentan las dificultades para la realización de la entrevista.

•*Confusión* por influencia del entrevistador. Los juicios o las formas de hacer las preguntas pueden frenar o falsear las respuestas. Por ejemplo, un entrevistador podría decir que *los cereales para el desayuno son sólo para niños.* Un adulyto puede que no reconociera su consumo ante un enntrevistador que estudiara los hábitos de consumo.

**Cuándo deben usarse**

Los investigadores proceden a entrevistar a las personas en las siguientes situaciones:

•*Cuando es preciso obtener una información muy compleja,* que no puede recogerese a través de preguntas simples como las que suelen figurar en los cuestionarios. Ppor ejemplo, cuando se trata de que un consumidor explique todas las fases que sigue para compra de una computadora: nacimiento de la necesidad, la recogida de información externa, influencia de los grupos, actitudes ante las marcas, reglas de decisión, reducción de la disonancia, etc. *En este caso, no sirven las encuestas,* porquelas respuestas poseen muchos matices que una encuesta es incapaz de recoger.

•*Cuando las personas deban proporcionar información confiencial* o delicada relativa a su vida, creencias, actos, preferencias, etc. Por ejemplo, la entrevista puede resultar muy adecuada para obtener información depersonas con SIDA, que hayan sufrido alguna agresión sexual , que hayan sido toxicómanaso hayan ejercido la prostitución, haber estado en prisión , etc.

•*Cuando se desee obtener información de profesionales,* como empresarios, artistas, políticos, etc. que puedan prioporcionar opiniones y juicios muy ricos con relación a su actividad. En estos casos, los cuestionarios resultan un freno para obtener todo el potencial de información que estas personas pueden proporcionar. Es mejor dejarlas hablar libremente, bajo un guión previamente diseñado.

•También se pueden utilizar las entrevistas como *fase previa para la elaboración de cuestionarios*, pues las personas entrevistadas pueden ir definiendo, a partir de sus respuestas, los contenidos o las preguntas que convenga formular.

## Técnicas proyectivas

Son instrumentos que descubren aspectos inconscientes de las personas a través de asociaciones de palabras, o posubles explicaciones de conductas o hechos. Por ejemplo, se podría pedir a una persona que dijera si la marca de computadoras Dell fuera un animal, qué sería: un pez o un pájaro. Un analista especializado en este tipo de técnicas podría valorar la imagen que tiene esa persona de esta marca de computadora, de acuerdo con el significado simbólico de los animales en nuestra cultura occidental. Son técnicas *cualitativas, primarias, estáticas, personales e indirectas.* Es un instrumento que permite conocer aspectos inconscientes de las percepciones, motivaciones, actitudes y conductas. Son técnicas cualitativas muy útiles en investigaciones exploratorias. Creadas por sicólogos, empezaron a utilizarse en el marketing, durante los años 50s del siglo pasado debido a que muichos aspectos del comportamiento del consunidor no son conscientes. Existe un caso muy ilustrativo de aplicación de ésta técnica. Cuando hace años se introdujeron en el mercado del EUA los alimentos congelados, se preguntó a las amas de casa sobre su uso. Muchas de

ellas tras probarlos, manifestaron su deseo de no consumirlos. La argumentación consciente porporcionada fue que no eran tan buenos como los frescos. Sin embargo, tras aplicar técnicas proyectivas los investigadores encontraron los verdaderos motivos del rechazo. Las amas de casa habían sido educadas para ser buenas madres y esposas excelentes de manera tradicional. Los alimentos ongelados suponían ahorro de tiempo. Para ellas ese ahorro de tiempo era una *fuente de remordimientos*, pues pensaban que por dedicar menos tiempo a cocinar para su familia la atendían peor.

**Clases de técnicas y aplicaciones**

Las técnicas proyectivas de mayor utilidad, como investigación de mercados, son:

**Asociación**

- Consiste en leer una lista de palabras y la persona que participa en la prueba debe responder con aquellas que espontáneamente se le ocurran .
- La asociación de palabras es muy útil en marketing, especialmente para buscar *denominaciones de marca, personalidad de marca e identificar atributos de productos.* En la práctica, el test consiste en pronunciar una palabra *y valorar sólo aquella respuesta que se proporcione en tres segundos como máximo*. Si el tiempo fuera mayor significaría ignorancia, confusión o falta de ideas claras, incluso en el subconsciente.

- **Construcción**

Consiste en inventar historias o diálogos a partir de estímulos visuales, como fotos o dibujos. La técnica más conocida es el **TAT** (Test de Apercepción Temática). En él se muestran fotos o dibujos de productos, anuncios, etc. Las personas sometidas a eesta prueba pueden aportar información muy valiosa sobre actitudes hacia los productos, *imagen de empresa y factores que afectan a la conducta de los consumidores.* Por ejemplo, se puede mostrar una foto de una persona que esté mirando unas gafas de sol.Ante esta foto, las personas construyen sus propias historias. Una podría decir quese está fijando en el precio, siendo por tanto, una persona sensible a él. Una persona derrotista, pesimista podría decir que necesita unas gafas oscuras para disimular la picadura de una abeja en el párpado. Otra podría pensar que el observador pronto se marchará de vacaciones y por eso está mirando unas gafas, que deberá comprar.

**Complementación**

- Consiste en completar frases sin terminar. El investigador comienza una afirmación que no termina, o muestra un dibujo o foto y la persona bajo prueba debe acabar las frases que se le proporcionan o explicar o justificar lo que ve en las fotos o dibujos.*Esta técnica es muy útil para detectar motivaciones o actitudes.*

Si se formulara la frase: *los hombres que se compran un auto eléctrico…*caben diversas respuestas, que permiten detectar motivaciones o actitudes hacia ellos , que pudiern ser:

- *Se preocupan por economía*
- *Lo hacen porque recorren muchos kilómetros al año*
- *No tendrán problemas con tiempo húmedo*
- *Disfrutarán durante años de su coche*

Estas respuestas a una frase incompleta muestran motivos de uso y actitudes. Aparecen la economía la duración, la facilidad de mantenimiento, la resistencia a condiciones atmosféricas

hostiles o la ecología.

**Expresión**

•Las técnicas de expresión se parecen mucho a las de construcción. En ellas se pide a las personas que representen el rol de otra o que identifiquen a otras personas o entes con *seres animadas o inanimados.* Tamién sirve para medir *imágenes o percepciones.* Existen diversas técnicas, siendo las más importantes:

•*El juego de papeles (role playing).* Se pide a las personas que actúen en lugar de otras, que se pongan en su lugar. Así, se pueden conocer sus necesidades, deseos, emociones o motivaciones. Se podría pedir a una persona, un director financiero de una empresa, que actuara como director comercial y proporcionara argumentos para vender un producto. De esta forma podría conocerse cómo lo percibe él.

•En la *personificación* se pide a las personas que participan en la prueba que materialicen en un ser vivo (que animen, en términos antropológicos), empresas, marcas, productos, etc. Por ejemplo, alguien podría decir que la empresa X, si fuera un ser vivo, sería un tiburón o un buitre. Este juicio es muy elocuente sobre la imagen que tiene de la empresa. Si opinara que tal vez fuera un perro pastor alemán estaría proporcionando una imagen positiva de la empresa, por todo lo que se asocia a esta raza de perros: nobleza, fidelidad, protección.

**Valoración**

•Las técnicas proyectivas no se basan en cuestionarios largos o complicados. Bastan unas fotos o dibujos o preguntas muy sencillas para que *las personas proporcionen información de forma muy simple y sin fatiga.*

•En las técnicas proyectivas *no se descubre el fin de la investigación*. Las personas que participan en ellas no pueden adib¡vinar para qué servirála prueba. Responden con naturalidad, sin que sean capaces de introducir intencionalmente sesgos.

•Como toda técnica cualitativa *se interpreta de forma subjetiva.* Las conclusiones pueden variar según quién sea el investigador.

•*Como todo resultado cualitativo no se debe generalizar*, porque procede de muestras pequeñas y generalmente , de conveniencia, no aleatorias.

Mostramos así, un cuadro resumen

**Técnicas proyectivas y su utilidad en marketing**

| Técnica | Concepto | Utilidad en la investigación de mercados |
|---|---|---|
| Asociación | •Pronunciar palabras que construyen una respuesta espontánea a otras que se proporcionan | •Identificar la imagen de marca<br>•Buscar denominaciones de marca<br>•Identificar atributos<br>•Posicionar productos o marcas |
| Construcción | •Crear historias basadas en un estímulo | •Proponer usos de los productos<br>•Valorar la utilidad de los productos<br>•Medir actitudes de los productos<br>•Identificar factores que afectan o explican la conducta de los |

| | | |
|---|---|---|
| | | consumidores<br>• Identificar motivaciones de consumo |
| Complementación | • Completar frases sin terminar | • Medir actitudes<br>• Identificar motivaciones de compra o de rechazo a la compra<br>• Descubrir usos delos productos<br>• Valorar atributos de productos, marcas, personas, organizaciones |
| Expresión | • Representar el rol de otra persona<br>Identificar personas o entes con seres animados o inanimados | • Medir actitudes<br>• Medir percepciones<br>• Medir imágenes |

## Bases de datos

Son conjuntos de informaciónes reacionadas y accesibles conforme a algún criterio. Por ejemplo, las generadas por instituciones de gobierno como INEGI en México sobre el rubro de empleo, ingreso per cápita, actividad preponderante , etc.

Si los términos son vagos y generales obtendremos una consulta con muchas referencias e infor-mación que *no* es pertinente para nuestro planteamiento. En este sentido, las bases de referencias funcionan como los *disparadores o motores de búsquedas* (Google, Yahoo, Altavista, etcétera).

Por ejemplo, si hacemos una consulta con palabras como *escuela*, *educación*, *comunicación*, *empresas* o *personalidad* aparecerán miles de referencias y nos *perderemos en un mundo de información*.

Entonces, los términos de búsqueda deben ser precisos, por lo que si nuestro planteamiento es concreto, la consulta tendrá mayor enfoque y sentido y nos llevará a referencias apropiadas. Asimismo, nuestra búsqueda deberá hacerse con palabras en español y en inglés, porque gran cantidad de fuentes primarias se encuentran en este idioma.

Al acudir a una base de datos, sólo nos interesan las referencias que se relacionen estrechamente con el problema específico a investigar. Por ejemplo, si pretendemos analizar la relación entre el clima organizacional y la satisfacción laboral, ¿cómo encontraremos las fuentes primarias que en verdad tienen que ver con el problema de estudio que nos incumbe? Primero, con la revisión de una base de datos apropiada. Si nuestro tema trata sobre clima organizacional y satisfacción laboral, *no* consultaríamos una base de referencias sobre cuestiones de química como Chemical Abstracts ni una base de datos con referencias de la historia del arte, sino una base de información con fuentes primarias respecto a la materia de estudio, tal es el caso de Wiley InterScience, Communication Abstracts y ABI/INFORM (bases de datos correctas para nuestra investigación). Si vamos a comparar diferentes métodos educativos por medio de un experimento, debemos acudir a la base de referencias adecuada: ERIC (Education Resources Information Center). En español también hay algunas bases, como *Latindex* y *Redalyc*, para diversas ciencias y disciplinas; *bvs*, ciencias de la salud; *ENFISPO*, enfermería, etcétera).

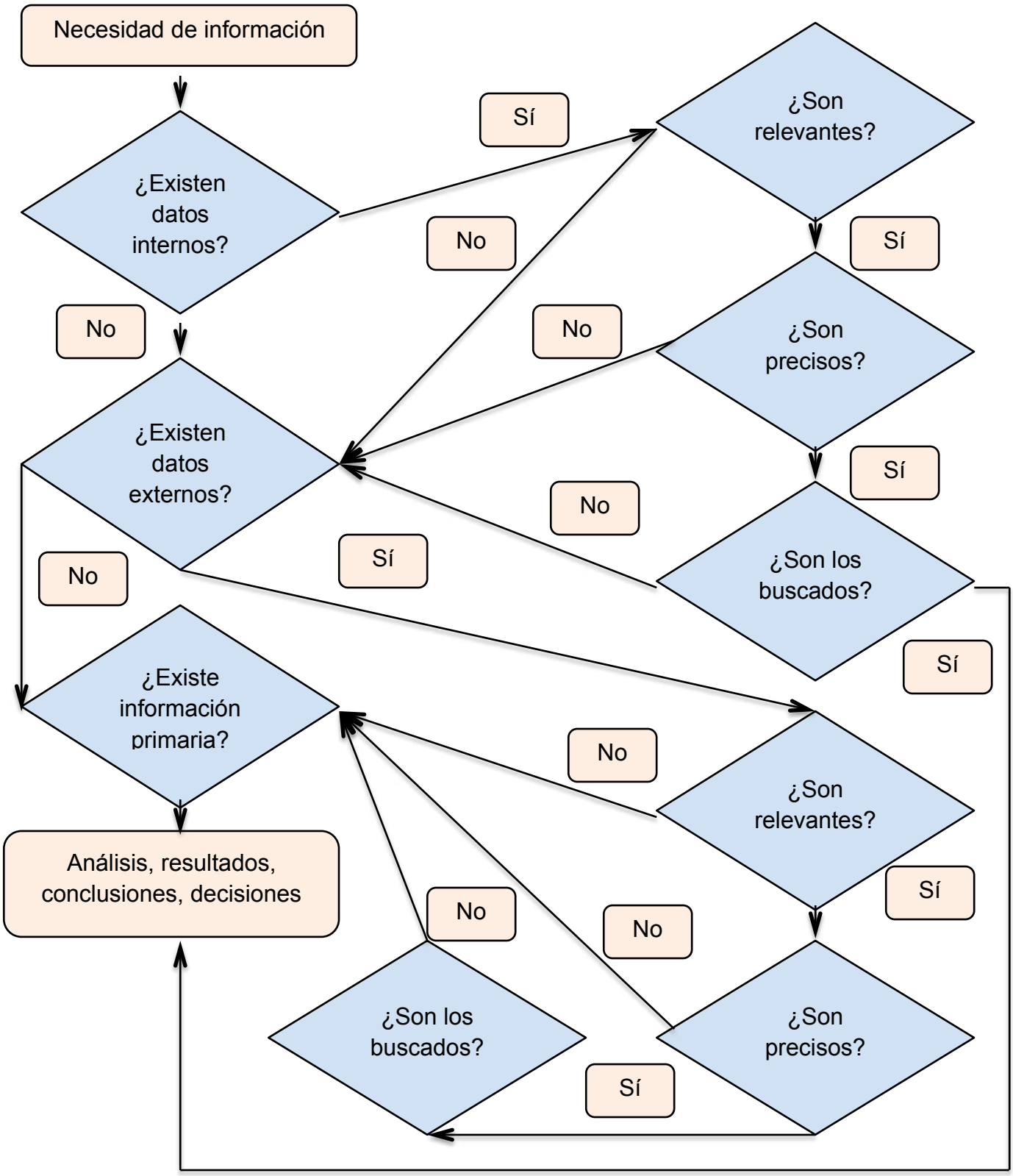
Una vez elegida la base de datos que emplearemos, procedemos a consultar el *catálogo de*

| |
|---|
| *temas, conceptos y términos* (*thesaurus*) respectivo, que contiene un diccionario o vocabulario en el cual podemos hallar un listado de palabras para realizar la búsqueda. Del catálogo debemos seleccionar las palabras o conceptos *claves* que le proporcionen dirección a la consulta. También podemos hacer una *búsqueda avanzada* con esos términos, utilizando los operadores del *sistema booleano*: *and* (en español *y* ), *or* (en español *o* ) y *not* (en español *no* ). Con los descriptores y las preposiciones estableceremos los límites de la consulta al banco o la base de referencias.<br>La búsqueda nos proporcionará un listado de referencias vinculadas a las palabras clave (dicho de otra manera, el listado que obtengamos dependerá de estos términos llamados descriptores, los cuales escogemos del diccionario o simplemente utilizamos los que están incluidos en el planteamiento). Por ejemplo, si nuestro interés se centra en *procedimientos quirúrgicos para el cáncer de próstata en ancianos* y vamos a revisar en la base de referencias MEDLINE_1997-2008 (para Medicina), si seleccionamos las palabras o descriptores *cáncer próstata*, el resultado de la consulta será una lista de todas las referencias bibliográficas que estén en tal base y que se relacionen con dichos términos (enfermedad). Si la búsqueda la hicimos el 28 de enero de 2009 se obtienen 39 643 referencias (que son demasiadas, por lo que tenemos que utilizar más descriptores o incrementar nuestra precisión). Al agregarle el término *anciano* el resultado fue de 14 282 referencias (todavía muchas). Y al agregarle cirugía (porque realmente nuestro estudio se centra en ello), el número es mucho más manejable, 132 fuentes primarias. Desde luego, las búsquedas avanzadas pueden acotarse por fechas (por ejemplo, últimos tres años, de 2005 a 2010, de 2000 a 2009). |
| **Paneles** |
| Son muestras estables de consumidores o de establecimientos comerciales que proporcionan información periódica sobre sus actos de consumo y hábitos de exposición a medios de comiunicación. Tratándose de establecimientos comerciales, la información se refiere al comportamiento de las ventas, existencias, promociones. |
| **Encuesta** |
| Consisten en un conjunto articulado y coherente de preguntas que se formulan a las personas sobre la base de una batería de preguntas en torno a un tema específico. Se puede realizar por correo, frente a frente, en la calle emn establecimientos comerciales o en el domicilio de las personas, po rtelefono u online |

Fuente: Grande-Esteban y Abascal-Fernández (2017) y Hernández-Sampieri et al., (2010) con adaptación propia

Estas fuentes de información, como se aprecia, son muy diferentes entre sí y su capacidad para proporcionar conocimiento con el que llevar a cabo investigaciones es muy dispar. Estas fuentes se pueden clasificar atendiendo a criterios muy diversos. al mercado. En este caso, se emplea un panel de consumidores, como observatorio, que proporciona información mensual sobre el comportamiento observado. Algunas técnicas son por su naturaleza estáticas y otras dinámicas. Ver **Figura A.5.**

**Figura A.5. Selección de fuentes de información**




Fuente: Grande-Esteban y Abascal-Fernández (2017) con adaptación propia

## *Por su grado de contacto*

Algunas técnicas de recogida de infoamción suponen mucho contacto con las personas; por ejemplo, las entrevistas personales, las dinámicas de grupo, las encuestas a domicilio o las tñécnicas proyectivas. Otras veces el contacto es nulo, como ocurre con las encuestas por correo, la obsrvación o cuando se emplean las bases de datos. Atendiendo al grado de contacto con las personas, las técnicas pueden clasificarse en *personales* e *impersonales.*

## *Por su grado de colaboración*

Las personas que colaboran en las investigaciones oueden ser plenamente conscientes de a información que proporcionan, de su utilidad y previsible uso. Esto ocurre, por ejemplo, cuando se manifiestan intenciones de compra a un encuestador. Otra s veces las personas que colaboran saben que proporcionan información, pero no llegan a saber exactamente qué se hará con ella, *no intuyen su utildad.*Esto ocurre, por ejemplo, cuando a una persona se le pide que relacione marcas de cámaras fotográficas con animales. Cuando las personas conocen o intuyen la utilidad de la información que aportan se habla de *técnicas directas.* Cuando ello no es posible, se trata de *técnicas indirectas.*

## *Por su  proceso de investigación*

Probablemente, el criterio más adecuado que en teoría podría encontrarse para clasificar las fuentes de informacion, sea relativo a la función que desempeñan para llevar a cabo las investigaciones, ya que:

Las investigaciones exploratorias, suelen realizarse cuando no se tiene n muchos conocimientos previos sobre la realidad que se va a investigar. En estos casos se suele acudir a obtener *información secundaria* que sirva para centrarse sibre aquella. En este sentido son muy útiles las *dinámicas de grupo, la observación o las técnicas proyectivas, las consultas bibliográficas o la compra y explotación de bases de datos.* No puede defenderse que las investigaciones exploratorias deban desarrollarse siempre con fuentes de *información secundaria*, por ejemplo, a partir de una información de base de datos. Se puede acudir a una *dinámica de grupos*, que es una técnica de recolección de datos de *informacion primaria* y además, cualitativa. Una vez base de datos puede ser información primaria par aquien la compra, pero las empresas que las crean para su uso, tienen que considerarla como *información secundaria interna*. Esa misma base de datos puede servir a unas empreesas para hacer investigaciones descriptivas y otras podrían comprarla para hacer investigaciones exploratorias.

No hay que obsesionarse con el intento de clasificar o encasillar las fuentes de información aplicando un criterio u otro Lo que debe hacerse es conocer bien las fuentes, sus características y su utilidady saber cuando deben aplicarse.

La **Tabla A.16** muestra ls características que se han mencionado en este apartado para un conjunto de fiuentes de información considerando su naturaleza.

**Tabla A.16. Características de las fuentes de información**

| Naturaleza | Criterios de clasificación delas fuentes de información | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Disponibilidad | | Tratamiento de datos | | Dinamismo | | Contacto | | Colaboración | | Utilidad En la Inevstigación | |
| Fuente de información | P | S | Q | C | E | D | Pe | I | Di | In | Ex | Co |
| Bibliografía, web | N | S | N | S | - | - | - | - | - | - | S | N |
| Observación | S | N | N | S | S | N | N | S | N | S | S | N |
| Pseudocompra | S | N | N | S | S | N | S | N | N | S | S | N |
| Creatividad | S | N | N | S | S | N | S | N | S | N | S | N |
| Dinámicas de grupos | S | N | N | S | S | N | S | N | S | N | S | N |
| Entrevistas | S | N | N | S | S | N | S | N | S | N | S | N |
| Técnicas proyectivas | S | N | N | S | S | N | S | N | N | S | S | N |
| Bases de datos | N | S | S | N | N | S | N | S | - | - | S | S |
| Paneles | N | S | S | N | N | S | S | N | S | N | S | S |
| Encuestas | S | N | S | N | S | S | S | S | S | N | S | S |

Notas:
P. Primaria; S. Secundaria; Q. Cuantitativa; C. Cualitativa Pe. Persona; I. Impersonal; Di. Directa; In. ; Ex. Exploratoria; Co. Cioncluyente; N. No; S. Sí
Fuente: Grande-Esteban y Abascal-Fernández (2017) con adaptación propia

# Características de las fuentes de información en las ciencias económico-administrativas

La elección de la fuente de información para realizar una investigación es una tarea compleja. Un mismo problema puede requerir el uso de fuentes diversas. El investigador debe tner muy claro el objetivo de la investigación. Por ejemplo, ante una reducción del potencial innovador de la organización, se uede pensar en la conveniencia de difundoir una encuesta entre el área de mercadotecnia y el de investigación y desarrollo., para conocer los motivos. También, se podría pensar en estudiar la evoñución del precio, la legislación en materia de estímulos a la innovación, las facilidades para el acceso a fondos extraordinarios y otros.

Para conocer los motivos por los cuales los consumidores prefieren los autos eléctricos a los de etanol, podría difundirse una encuesta en la población objetivo. La obtención de información a través de una encuesta y su posterior análisis resultan caros, Tal vez las causas del problema que se desea conocer se puedan identificar estudiando sólo aspectos cualitativos. En este caso, la explicación podría ser una mayro sensibilidad sobre la conservación medioambiental. También, es posible que

los aspectos cualitativos que puedan identificarse deban de ser contrastados con un análisis empírico posterior.
*Un principio básico de investigación es que cualquier estudio debe comenzar por una fase cualitativa,* que permita aproximarse al problema y conocerlo. En caso de que ya sea conocido, la fase cualitativa le ayudará, sin duda, a definirlo aún con más precisión y desarrollarlo en mejores condiciones. Una vez realizada la fase cualitatuiva se puede pasar, si fuera necesario, a una fase *cuantitativa,* en la que se analizarán los datos, generalmente primariaos, con técnicas estadísticas y matemáticas. Muy frecuentemente, los estudios cuantitativos refrendan las coonclusiones a las que se llega tras realizar estudios cualitativos. Por lo atnato, a fin de decidior qué fuentes deben emplearse hay que tener en cuenta:

- Si la investigación es *exploratoria* debe cpmenzarse con el análisis de fuentes cualotativas o secundarias
- Si la investigación es *descriptiva* también debe realizarse una fase cualitativa, para centrar bien los objetivos y diseñar la investigación.Posteriormente, se dearrollará la fase cuantitativa.

# Fuentes descriptivas de información

Los datos cualitativos consisten en *descripciones de situaciones, sucesos, personas interacciones y conductas observadas; relatos directos de experiencias personales, actitudes, creencias, pensamientos; pasajes completos de documentos, correspondencia, informes y casos históricos.*Ver **Tabla A.17.**

**Tabla A.17. Fuentes de información descriptiva**

| Primarias |
|---|
| Se aplican para desarrollar investigaciones cuando surge su necesidad. Esta afirmación debe matizarse, pues la bibliografía o páginas web son secundarias. En general, no obstante, se puede afirmar que las fuentes cualitativas son globalmente, primarias. |
| Estáticas |
| Se emplean para obtener información adecuada a los objetivos concretos de una investigaciónen un momento dado. Generalmente, los resultados que se obtienen no suelen tener una validez permanente. |
| Personales |
| Requieren la presencia o colaboración de personas para poder ser aplicadas. |
| Adecuadas para investigaciones exploratorias |
| Permiten aproximarse a problemas y comprenderlos, pero los resultados no suelen ser extrapolables porque se emplean muestras pequeñas y poco representativas |
| Excepciones |
| Por ser cualitativas, son contadas las excepeciones en las que es posible un *tratamiento de datos* que requiera la utilización de *técnicas estadísticas No* |

| |
|---|
| *paramétricas* |
| Uso en las primeras fases |
| Se emplea en las primeras fases del diseño de un estudio y sirven, sobre todo, para generar hipótesis que podrán ser verifoicadas en un posterior estudio cuantittivo |

Fuente: Grande-Esteban y Abascal-Fernández (2017) con adaptación propia

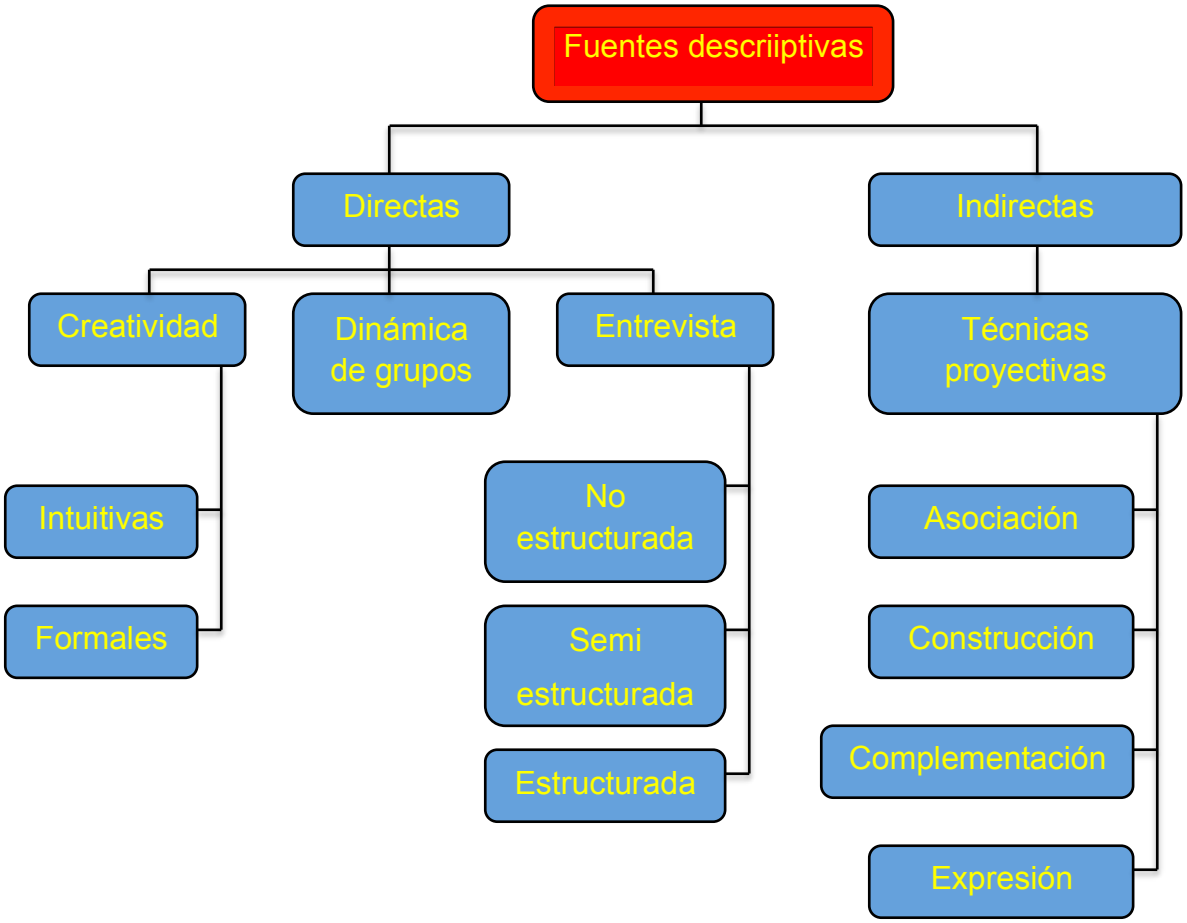
Un elemento diferenciador importante de las diversas fuentes cualitativas, radica en la clase de colaboración que proporcionan las personas. Pueden identificarse dos grupos de técnicas.El primero lo constituyen, las que requieren *colaboración activa de las personas* conformándolo: *las técnicas de creatividad, las dinámicas de grupos y las entrevistas.* El segundo grupo lo conforman las técnicas en las que las personas proporcionan información sin ser conscientes de ello. Lo constituyen: *la observación, la pseudocompra y las técnicas proyectivas de asociación, complementación, expresión y construcción.* Las investigaciones, deben comenzar en principio con recogida de *información cualitativa* empleando para ello técnicas o fuentes específicas. Ver **Tabla A.18 y Figura A.6.**

**Tabla A.18. Importancia de las fuentes descriptivas**

| Importancia de las fuentes descriptivas |
|---|
| • *Desarrollar investigaciones de naturaleza exploratoria*, como obtener información previa sobre un determinado campo del que no se tiene ningún conocimiento para que el investigador se familiarice con él. |
| • *Desarrollar investigaciones explicativas* a partir de actitudes, motivaciones o creencias. Una cadena de hipermercados podría encontrar que se incrementan las ventas de equipo de ejercicio y disminuyen las de aceite. Una investigación cualitativa podría poner de relieve que la población se encuentra preocupada por la relación de consumo de grasas y aceites y la salud cardiovascular y compensarla con la adquisición de compra de equipo casero de gimnasio.<br>• *Evaluar actividades de marketing,* como el diseñoi de productos, su percepción, la fijación de presios, el diseño de los envases o de anuncios publicitarios. En una investigación cualiotativa, se puede percibir que un precio es caro, que un producto posee atributos no valorados, que un envase es poco funcional o que un anuncio resulta desagradable, o no se entiende.<br>• *Vonocer la terminología de los usuarios y/o consumidores,* y en generaL, su comportamiento. Antes de diseñar un cuestionari, debería llevarse a cabo un estudio cualitativopara conocer el lenguaje del público al que se dirige y proceder a una redacción asequible y adecuada, para su correcta comorensión. No es lo mismo dirigirse a personas cultas que de escasa cultura, mayores o niños. Cada población tienee un lenguaje propio, que debe ser respetado y recogido en el cuestionario. |

Fuente: Grande-Esteban y Abascal-Fernández (2017) con adaptación propia

**Figura A.6. Fuentes descriptiva**




Fuente: Grande-Esteban y Abascal-Fernández (2017) y Hernández-Sampieri et al., (2010) con adaptación propia

# El Cuestionario

Es un conjunto congruente coherente de preguntas diseñado para obtener información necesaria para poder realizar una investigación. Sus funciones son:

- Dado que su diseño considera: ordenación, estructura, y aspecto para ser acertado, el cuestionaro es una herramienta que contribuye a que los sujetos proporcionen información requerida.
- Homogeiniza la obtención de información, dado que los encuestados responden a una misma batería de preguntas, ya que está diseñado en su formulación para aplicarse a todos por igual.
- Traslada el objetivo de la investigación a preguntas concretas respondidas por el sujeto encuestado, por lo que es un instrumento eficaz para que el tratamiento de los datos sea rápido y efectivo particularmente en sistemas informáticos de soporte.

# Etapas y proceso de diseño

Sus etapas para realizarlo, son mostrados en la **Tabla A.19.**

**Tabla A.19. Etapas para realizar un cuestionario de investigación**.

**Información a requerir**

Se debe tomar en cuenta los objetivos de la investigación y las hipótesis que se hayan formulado , pues son los determinantes de la naturaleza de la información que se requiere recopilar. Aunque paarece obvio su planteamiento, es relativamente fácil que se produzca una distorsión entre quien decide los objetivos de la investigación, quien la prepara y quien la implementa. Por ejemplo:

- Una investigación cuyp pbjetivo es analizar la viabilidad de un servicio nuevo, debe de medir actitudes , preferencias y/o valores.
- Una investigación que se oriente a revisardeterminantes de posicionamientp de marca, debe cuestionar sobre los atributos percibidos, su valoración y plantear cómo realizarlo; puntuándolos o asociándolos a las marcas.

Una investigación se diseña *globalmente.* Se fijan objetivos de conocimiento,que determinan qué información debe de recopilarse. Paralelamente, debe seleccionar la técnica adecuada para tratar la información. Cada tipo de datos exige una técnica de análisis y cada estudio por lo tanto, debe aplicar las técnicas idóneas a su objeto y sujeto de estudio. La información recogida debe ser capaz de permitir aceptar o rechazar las hipotesis o conjeturas formuladas. Recuerde que una investigación puede estar en el *continuum: exploratorio-descriptivo-correlativo-explicativo.*

**Tipo de información a recoger**

**Hechos o comportamientos**

Objetiva e inequívocamente contrastables y verificables. Por ejemplo: cúánto se compra de un producto, quién compra, dónde, etc.
Las siguientes preguntas, registran hechos:

| Pregunta | Si | No |
|---|---|---|
| *¿Tiene Usted tarjeta de crédito?* | | |
| *¿Usa Usted su tarjeta de crédito?* | | |
| *¿Utiliza Usted los cajeros automáticos?* | | |

**Conocimientos**

Una investigación puede encaminarse a valorar la nototiedad de una organización, empresa, marca, etc. si las personas conocen el producto y/o servicio, etc.

**Opiniones o juicios**

Esto es, por ejemplo: ¿qué opinan los empresarios de una política de innovación ?, ¿cómo valoran los usuarios la implementación de una tecnología híbrida?, ¿cómo percibe que es la cadena de suministro ebn cuanto a su efectividad?, etc.Las siguientes preguntas recogen opiniones sobre motivos de queja de establecimientos comerciales.
*Indique el grado de importancia de los siguientes motivos de queja* , donde:
M.Mucho; B.Bastante; P.Poco. N.Nada

| Pregunta | M | B | P | N |
|---|---|---|---|---|
| No encontró lo que buscaba | | | | |

| | | | | |
|---|---|---|---|---|
| El establecimiento es caro | | | | |
| Lo atendieron mal | | | | |
| Las reclamaciones no se atienden satisfactoriamente | | | | |

## Actitudes o predisposiciones de ánimo

La información que se busca con los estudios de actitudes se refiere a lo que se piensa acerca de imagen , valoración de las cosas, y la razón de las conductas. Por ejemplo, se puede preguntar qué se piensa sobre las residencias para la tercera edad, sobre los autos elèctricos no contaminantes, las campañas antiaborto, etc.

Las preguntas que aparecen a continuación, pueden ser válidas para conocer las actitudes de la población ante nuevos productos y el grado de racionalidad que demuestra su adquisición. Por ejemplo:

*Usted ve con frecuencia en los anuncios productos nuevos como hornos de microondas, lectores de memorias sólida, televisores internet, etc. Con relación a esos productos nuevos, marque con una X todas las opciones que quiera.*

| Pregunta | Marque con una X |
|---|---|
| Con los que tengo ahora, me es suficiente | |
| Los compraré cuando reponga los que ahora tengo | |
| Antes de comprarlos, creo que es importante preguntar a otras personas que ya los tienen | |

## Motivos o explicación de conductas concretas

Podría preguntarse por qué se compran vehículos híbridos o por qué se consumen productos congelados, etc. Por ejemplo: *Indique por qué compra Usted su refresco en envase de 1 litro. Marque con X la7s opción/es que desee.*

Porque:

| Pregunta | Marque con una X |
|---|---|
| Es más fácil de almacenar en el refrigerador | |
| Es la que suele estar de oferta en el supermercado | |
| El envase es muy resistente | |
| El diseño es fácilmente de cargar | |

## Posibles conductas futuras

Se podría preguntar si consumiría, o no, un producto preparado a base de algas marinas; si se estaría dispuesto a ingresar, llegado el momento, en una residencia para la tercera edad, por qué partido político votará, etc.Cruzando las clases de investigaciones con la naturaleza de la información, se puyeden conseguir las siguientes asociaciones:

- Una *investigación exploratoria* debe recoger conocimientos, opiniones, actitudes o posibles conductas futuras.
- Las *investigaciones* de tipo descriptivo o causal deben recoger resultados, conductas, hechos o motivos.

*Si sobre el papel no es demasiado difícil hacer estas distinciones, en la práctica puede resultar complicado diferenciar una actitud de una opinión* y predefinir el tipo de información que debe recogerse de cada investigación. Por ejemplo, una persona puede tener una actitud favorable hacia los motores de etanol porque son económicos en consumo y menos contaminantes. Es posible que incluso lo adquiera y cambie a la misma opción de vehículo cuando lo requiera. *Sus opiniones están*

*condicionadas por sus actitudes,* que a su vez influirán, presumiblemente, en cionductas futuras. En el cuestionario siguiente, algunas preguntas recogen actitudes, otras opiniones e incluso hechos.

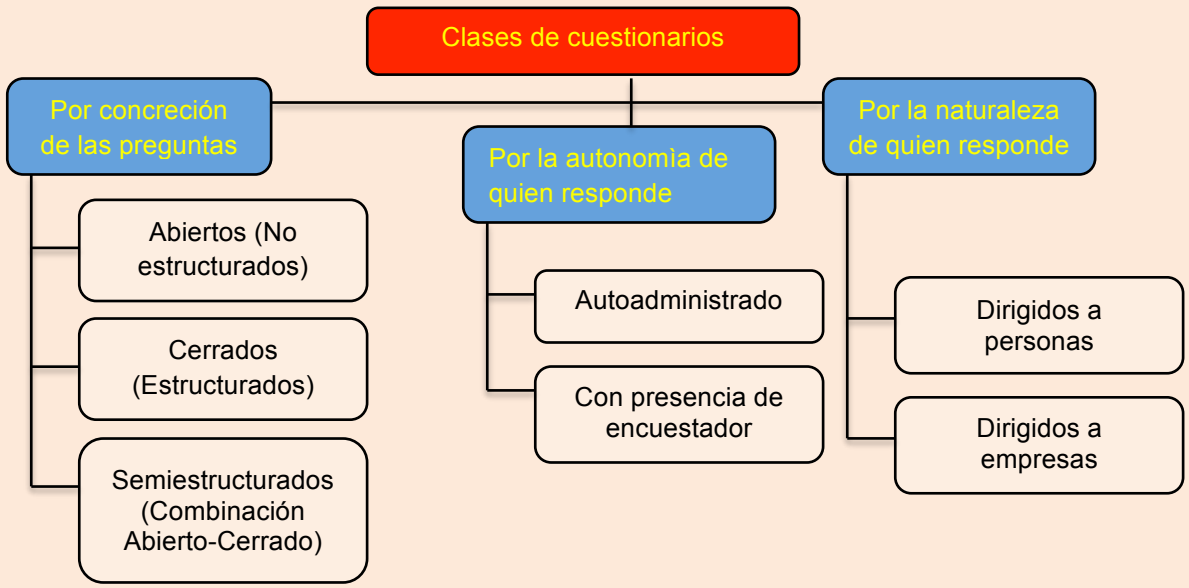
*Con relación a la venta por correo electrónico, Usted cree…(MARQUE CON x todas las alternativas que requiera)*

| Pregunta | Marque con una X |
|---|---|
| Esta forma de venta no es fiable. Es fácil que lo engañen | |
| Es mejor ver físicamente los productos en tienda | |
| Sólo compro lo que veo y toco | |
| Es un sistema cómodo para comprar | |
| Permite conocer más los productos | |
| Ofrece marcas desconocidad que me hacen desconfiar | |
| No existe asesoría ni me aconsejan | |
| En caso de devolución, no me queda claro cómo proceder | |
| Ofrecen cosas interesantes | |
| En caso de una garantía, no me queda claro cómo proceder | |

## Decidir la clase de cuestionario

Esta fase es de gran relevancia ya que el investigador tiene oportunidad de abordar al sujeto de estudio en distintas modalidades, al sujeto de estudio. Es así que se le permitirá desde dar amplias respuestas a la pregunta planteadfa como a decidir de diversas alternativas y niveles de percepción así como el detectar si son individuales o de percepión corporativa. No olvide que cada una de las clases de cuestionarios demanda un enfoque y atención diferenciados por lo que deberña guardar todas las cautelas para que su diseño sea útil basado en su precisión. Así, se tienen:

- Por concrecion de preguntas
- Autonomía de quien responde
- Por la naturaleza de quien responde




## Cuestionarios abiertos y cerrados

La variedad de cuestionarios que se pueden elaborar es muy amplia. Una clasificación muy operativa, se deriva del *grado de concreción* de las preguntas. En función de la naturaleza de la investgación, debe optarse por un tipo de cuestionario u otro:

•Los cuestionarios se llaman *no estructurados o abiertos* cuando todas las preguntas están abiertas. Se emplean cuando se van a desarrollar investigaciones *exploratorias.* Los investigadores tienen conocimientos escasos, ambiguos, vagos, sobre lo que van a estudiar y no son capaces de dar contenidop a los cuestionarios. Son útiles para recoger información que se empleará para diseñar cuestionarios con preguntas cerradas o para que los investigadores se familiaricen con el tema que deben investiga. Se utilizan entrevistas a domicilio, dinámicas de gruposy entrevistas en profundidad. Los cuestionarios no estructurados se administran a muestras pequeñas y poco representativas; no son susceptibles de ser tratados estadísticamente y la información que proporcionan es básicamente cualitativa.

•Los cuestionarios son *semiestructurados* cuando se combinan preguntas cerradas y abiertas. Se aplicn cuando se desea una mayor variedad de respuestas o cuando éstas, no pueden preverse. A veces, los investigadores no conocen la totalidad de respuestas que pueden proporcionar las personas , porque es muy difícil acotarlas todas. , o porque no se tiene el conocimiento suficiente. En estos casos, se cierran las preguntas que puedan y se deja la posibilidad de que las personas añadan alguna respuesta . Dse emplean en entrevistas personales y, ocasionalmente, en entrevistas por teléfono.

.

•Cuestionarios *estructurados* o *cerrados* cuando todas las preguntas están cerradas. Se utilizan en investigaciones de tipo *descriptivo o causal, descriptivas y causales.* Por ejemplo, para medir la imagen de empresas, detectar un posicionamiento de producto, segmentar un mercado con variables demograficas, etc. **estos cuestionarios emplean escalas** y se analizan con técnicas cuantitativas univariantesy/o multivariantes. Generalmente, las personas que los cumplimentan han sido seleccionadas aplicando alguna técnica de muestreo. Se emplean cuando se va a obtener información porporcionada por más de **50** personas dadas las exigencias de cumplimiento de técnicas estadísticasd e comprobaciòn de hipótesis, que requieren tamaños muestrales mínimos. Para utilizar cuestionarios estructurados, es necesario que puedan preverse las respuestas con bastante exactitud. Se emplean en encuestas en la calle, casa, establecimientos, por teléfonoo cuando deban cumplimentarse por el propio encuestado. Es fácil de responder . Es el tipo ideal de custionario *autoadministrado.* Debe estar probado y depurado tras una serie de *pruebas piloto* . Se procura que su redaccion sea sencilla, inequívoca y que las escalas y categorías se adapten al nivel cultural y de inteligencia de las personas que proporcionarán la información. Por ejemplo, las escalas *mult-iítem* con más de **5**

Categorías son muy complicadas para las personas mayores o muy jóvenes. Ocurre lo mismo con las métricas contínuas.

Lo más frecuente es que los cuestionarios sean estructurados o semiestructurados. Los cuestionarios los pueden cumplimentar las mismas personas que proporcionan

la infomación. Otras veces, se leen las preguntas y los encuestadores o entrevistadores marcan las respuestas.

Al primer tipo de cuestionarios se les llama *autoadministrados*. Se emplean en encuestas postales y en paneles. Se entregan personalmente o se envían por correo a quienes deban responderlos. Tienen el atractivo de que no hace falta una red de entrevistadores para recoger la información. Sin embargo, no es sencillo obtener la información puntualmente.

Estos cuestionarios tienen que estar perfectamente diseñados; correcta y claramente redactados. También deben ser muy fáciles de cumplimentar. Las normas que deben seguirse para su diseño, son:

- Debe enviarse una carta que explique lo que se espera del encuestado. Se explicará la finalidad dela investigación.
- Es conveniente que la carta esté personalizada.
- Conviene explicar ómo ha sido seleccionada la persona.
- Debe explicarse que el cuestionario es anónimo.
- Deben proporcionarse instrucciones claras para su cumplimentación.
- En lo posible deben utilizarse preguntas cerradas.
- Las preguntas más convenientes son las que muestran un grado de aceptación (o no) del indicador.
- El diseño y presentación debe ser atractivo, pues así aumenta la tasa de respuestas.
- **La experiencia aconseja que debe estar entre 30 a 40 preguntas**

El ejemplo opuesto son las *entrevistas vía telefónica*, Deben tener un diseño especial y ser cortos, pues resulta difícil mantener a una persona al teléfono durante largo plazo. Lo ideal es que sean respondiso en *menos de **10** minutos.* La introducción a la entrevista debe hacerse rápida y eficazmente para conseguir la colaboración. La voz del encuestador es muy importante. Las frases no deben ser largas para facilitar la comprensión. Es muy importante seguir un orden en la formulación de las preguntas, para no confundir al encuestado.

**Clases de cuestionarios según el grado de concreción de las preguntas**

<table>
<tr><th>Cuestionarios Abiertos</th><th>Cuestionarios Semicerrados</th><th>Cuestionarios cerrados</th></tr>
<tr><td>Su fin es obtener información cualitativa o diseñar custionarios que más adelante se cerrarán</td><td colspan="2">Su fin es medir hechos, actitudes, conductas, preferencias, etc.</td></tr>
<tr><td colspan="2">Se emplean en investigaciones exploratorias</td><td>Se emplean en investigaciones descriptivas y causales</td></tr>
<tr><td>No contienen escalas</td><td colspan="2">Contienen escalas</td></tr>
<tr><td>No se emplean técnicas de muestreo</td><td colspan="2">Se emplean técnicas de muestreo</td></tr>
<tr><td>Se administran con la presencia física de un investigador</td><td colspan="2">Pueden administrarse con presencia física de un investigador o ser autoadministardos</td></tr>
</table>

| | | |
|---|---|---|
| Se emplean en entrevistas y en dinámicas de grupos | Se emplean con fuentes primarias estáticas: encuestas personales, telefónicas y correo | Se emplean con fuentes primarias estáticas: encuestas personales, telefónicas, correo así como con fuentes dinámicas (páneles y encuestas ómnibus) |
| La información se analiza cualitativamente | | La información se analiza cuantitativamente |
| No tiene sentido inferir resultados | Se pueden generalizar los resultados | |

**Cuestionarios dirigidos a personas o empresas**

- Cuestionarios dirigidos a personas físicas. Se diseñan siguiendo las normas generales expuestas en este capitulo.
- Cuestionarios para empresas. En ele caso de las empresas son los encuestados quienes opinan en calidad de representantes de estas. Las personas que poroporcionan información pueden ser muy diversas. Por ejemplo, el director de compras de una gran empresa tendrá un perfil muy distinto al responsable de estas en una Pyme. Las compras serán diferentes en volumen y frecuencia. Por eelo se deduce que siendo tan distintas las empresas, resulta imposible diseñar un cuestionario que sea válido para todas. Sin embargo, pueden seguirse normas generales parta la elaboración de cuestionarios de empresas, tales como:
  - Debe dejarse una cabecera para identificar la empresa en funcion de sus características.
  - Conviene emplear también preguntas abiertas, dada la heterogeneidad de empresas y respondientes y su alta cualificación.
  - Las personas que responden a las encuestas suelen tener una elevada formación. No debe tenerse reparo en formular preguntas complejas.

**Dar contenido a las preguntas**

El cuestionario debe estar cuidadosamente elaborado en contenido y forma para escoger la información necesaria y llevar a cabo correctamente la investigacion que se desarrolla. Además, debe estar redactado de forma que facilite su cujmplimentación y permita obtener una tasa elevada de respuesta. En su diseño ha de definirse la codificación delas preguntas más adecuada para su tratamiento posterior.

**Clases de preguntas**

Se tienen las siguientes:

| Criterio de clasificación | Clases de preguntas |
|---|---|
| Libertad de elección de respuestas | • Abiertas<br>• Cerradas<br>• Semicerradas |
| Cantidad de respuestas y su relación | • Dicotómicas<br>• Respuesta múltiple<br>• Baterías de preguntas |

| | |
|---|---|
| Fines delas preguntas | • Filtro<br>• Control |
| Forma de realizarse | • Directas<br>• Indirectas |

**Por la libertad de elección de respuestas**

• **Preguntas abiertas**. En elllas el encuestado responde libremente. Este tipo de preguntas se emplea en las fases previas a la elaboración de un cuestionario cerrado. Sirve para encontrar posibles actitudes, motivaciones, opiniones, etc.
También se emplean en entrevistas en las que las personas por sus características, pueden aportar una información muy rica en matices. Las respuestas son más espontáneas y .seguramente, más valiosas que las proporcionadas por preguntas con alternativas de respuestas cerradas, pero su análisis es más complicado y costoso. Las preguntas abiertas suelen utilizarse para preparar cuestionarios cerrados y elegir las modalidades de respuesta*. Sin embargo, las preguntas abiertas no son recomendables con personas de nivel cultural o intelectual bajo*. Por ejemplo, la preparación de un cuestionario definitivo sobre la innovación, se podrían distribuir algunos custionarios previos, con una pregunta abierta entre otras, en la que se preguntara *¿qué ventajas encuentra Usted con la práctica de la innovación?*. Tras ele examen de los cuestionarios, el investigador podría encontrar respuestas como: ventaja competitiva, rapidez de fabriocación, ahorro en costos, etc. que le servirían para diseñar un cuestionario con preguntas cerradas. *Un problema que puede enfrentar es que existan demasiadas alternativas*, ante un número excesivo de posibles respuestas.

• **Preguntas cerradas.** En ellas las alternativas de respuesta están limitadas. El encuestado marca una, o varias, de las alternativas que se le ofrecen. Las preguntas cerradas son útiles en cuestionarios que han de responderse con rapidez o por teléfono. No hacen pensar mucho y resultan interesantes cuando no es necesario profundizar demasiado o sólo se trata de recoger hechos, aunque también sirven para medir actitudes o motivaciones. *Son las más adecuadas para personas con nivel cultural bajo, escasa memoria o con dificultades de comprensión.* Por ser menos fatigosas de responder que las preguntas abiertas pueden formar parte de cuestionarios más largos. Por ejemplo, los administradores de la información basada en tecnolgía (IT. *Information Technologies*) podrían contestar alguna de las siguientes modalidades cerradas, de respuestas.

| Pregunta *:¿Cuántas veces respalda la información de la empresa a la semana?* | Marque con una X |
|---|---|
| Una | |
| Dos o tres | |
| Cuatro o cinco | |
| Más de cinco | |

Su codificación es inmediata y su tratamiento estadístico es más sencillo. No

obstante, debe prestarse especial atención a la elección de las, modalidades de respuesta, de forma que se contemplen todas las posibilidades y todas tengan un número elevado de posibles respuestas. Por ejemplo, una mala elección de modalidades de respuesta, sería:

| Pregunta :*¿Cuántas veces respalda la información de la empresa a la semana?* | Marque con una X |
|---|---|
| Una | |
| Entre dos y diez | |
| Más de diez | |

Porque la mayoría de los administradores de IT, se encontrarían en la segunda modalidad. La pregunta no aportaría información relevante.

•**Preguntas semicerradas.** A través de ellas se proporcionan alternativas de respuestas cerradas y también se deja la posibilidad de que el encuestado responda libremente a aalguna de las preguntas. Constituyen un híbrido entre las preguntas abiertas y cerradas, con sus ventajas e incob¡venientes. Se utiliza cuando se sabe que existen unas pocas modalidades que contienen la mayor parte de las respuestas, pero se desea una informacion exhaustiva. Así, se obtiene esta con un coste menor que el correspondiente a las preguntas abiertas. Por ejemplo, se podría preguntar:

| Pregunta : *¿Qué tipo de certificación tiene su empresa?* | Marque con una X |
|---|---|
| Seguridad e higiene | |
| Procesos de manufactura | |
| Calidad | |
| Otros (especificar) | |

**Por la cantidad de respuestas y su relación**

•**Preguntas dicotómicas.** En ellas sólo existen dos alternativas de respuestas, las mutuamente exluyentes. Por ejemplo: se adopta un proceso de innovación o no se adopta; se tiene administración del conocimiento o no se tien. Las preguntas dicotómicas son fáciles de responder. El encuestado proporciona la información rápidaente y con pocas posibilidades de sesgo, siempre que estén correctamente formuladas. Una pregunta dicotómica no debería sustituir a otras preguntas más complejas o encademnadas. Por ejemplo: no debería preguntarse : *Conteste SÍ o NO ¿Su empresa adopta procesos de innovación así como de seguridad e higiene para lograr una mejor integración de la calidad?*. Esta pregunta debería desglosarse en tres, una por cada línea.

•Preguntas con múltiples respuestas. Ante una pregunta o afirmación se proporcionan varoias alternativas de respuesta. Por ejemplo (se podria contestar a todas las que fuera posible)

| Pregunta : *¿Cuál de las modalidades de innovación que realiza su empresa es la* | Marque con una X |
|---|---|

| *que le representa mayor productividad?* | |
|---|---|
| 1.Organizacional | |
| 2.Social | |
| 3.Del modelo de negocios | |
| 4.Producto/Servicio | |
| 5.Proceso | |
| 6.Tecnología | |
| 7.Mercadotecnia | |

La pregunta con múltiples respuestas está cerrada; es inequívoca, fácil de responder y fácil de codificar (1,2,3,4,5,6,7). Sin embargo, no es adecuada para estudiar *motivaciones.* Es adecuada para conocer *hechos, preferencias y opiniones.*
Puede suceder que el enúmero de alternativas de respuesta sea demadiado corto o elevado. En el primer caso, tal vez no se recoja toda la información que pudiera tener interés. Si fueran muchas las alternativas los respondientes podrían aburrirse. Además, es posible que esta pregunta pueda tener múltiples respuestas, porque una empresa podría contestar más modalidades de las que se reflejan en la tabla. en estos casos la codificación es complicada.En tal caso, cada alternativa se convierte en dicotómica para poder tabularla correctamente.

**Preguntas en batería**
Se llama así a una serie de preguntas, relacionadas con un mismo tema, que se integran y complementan. Por ejemplo:

| Pregunta : Indique qué marca de computadora su empresa prefiere adquirir A-B-C-D-E-F-G-H… | Marque con una X |
|---|---|
| Indique dónde la adquiere | |
| -Proveedor especializado | |
| -Representante de ventas de la marca | |
| -Tienda de electrónicos | |
| Indique el propósito de su uso | |
| -Apoyo a ventas | |
| -Apoyo a diseño e ingeniería | |
| -Apoyo a la operación y el mantenimiento | |
| Indique sobre personal a capacitar en su uso inicial | |
| -Entre tres y cinco personas | |
| -Entre seis y diez personas | |
| -Más de 11 personas | |

**Por los fines de las respuestas**
•**Preguntas filtro.** Así se denomina a una serie de preguntas dispuestas de forma que se subordinen unas a otras. Una respuesta afirmativa o negativa a una pregunta excluye o requiere de la siguiente,. Porf ejemplo, las preguntas en batería que acaban de exponerse sólo deberian ser respondidas por personas que devclararan usar las computadoras. En alguna parte del cuestionario tendría que figurar una pregunta dicotómica tal como:

| Pregunta | Marque con una X | |
|---|---|---|
| | SÍ | NO |
| ¿Usted utiliza su computadora en la organización por más de 6 horas? | | |

Esta pregunta es un filtro. Los usuarios que cumplan responderán a la bateria de preguntas anotada. Quienes no lo sean, serán conducidos a otra parte del cuestionario.

**Preguntas de control**

Su misión es verificar la coherencia entre las respuestas proporcionadas por el encuestado.Una persona que nose encuentre certificada y que así lo constate, en el cuestionario, no puede responder posteriormente que aplica la certificació en particularidades requeridas. Las preguntas de control son útiles para eliminar cuestionarios por incoherencia.

**Por la forma de realizarse**

•**Las preguntas directas** se formulan para obtener respuestas directas, sin rodeos. La pregunta dorecta- y dicotómica- es *¿tiene Usted certificado de educación superior?* Sólo cabe decir SÍ o NO. Las respuestas a las preguntas directas pueden cerrarse en los cuestionarios.

•**Las preguntas indirectas** se formulan asi porque de lo contrario sería más difícil obtener una respuesta. Se aplican principios de técnicas proyectivas. Por ejemplo, se podría preguntar *¿qué piensa Usted de las personas que tienen creencias religiosas?* La respuesta popdría permitir conocer si la persona encuestada las tiene, o al menos cómo las valora. Una pregunta diorecta como *¿tiene Usted creencias religiosas?* Posiblemente daría una respuesta cuya credibilidad sería discutiuble, pues es demasiado íntima para ser contestada a un extraño.

Cerrar las preguntas indirectas es difícil o incómodo, aunque es posible. Las preguntas indirecvtas generan respuestas muy ricas en matices, que deberian recogerse en preguntas abiertas. Las preguntas indirectas, por lo general, no se emplean en cuestionarios, sino en la aplicaci`ón de técnicas cualitativas proyectivas.

**Características de las preguntas**

| Clase de pregunta | Características |
|---|---|
| Pregunta abierta | •Recoge información con un mínimo de indicacionespara el encuestado<br>•Es útil cuando la gama de respuestas es muy amplia<br>•Es más costosa de codificar, tabular y analizar<br>•Los encuestadores deben registrar las respuestas de forma textual<br>•Los resultados dependen mucho de la calidad de las entrevistas<br>•Es fácil de introducir sesgos al escribir las respuestas |
| Pregunta dicotómica | •Es muy fácil de preguntar y de responder<br>•Debe de segurarse que sólo existan dos alternativas. |

| | |
|---|---|
| | No cabe el : No sabe/ No contesta<br>•Muchos tipos de informacion sólo admiten preguntas dicotómicas: Hombre/Mujer; Rural/Urbano; Compra/No Compra.<br>•Es más difícil introducir sesgos que con otros tipos de preguntas |
| Preguntas cerradas | •Son fáciles de preguntar y tabular<br>•Suelen ser más sencillas de preguntar que las preguntas abiertas<br>•Hay que estar seguro de incluir todas las alternativas posibles |

**Cómo preguntar**

La calidad de una investigación se encuentra condicionada, en gran medida, por lo acertado del diseño de las preguntas. Si estas están mal definidas, son ambiguas o no responden a los objetivos de la investigación, por muy bueno que sea el analista y por muchas tñecnicas de análisis de datos que conozca, llegará a resultados nulos. La investigación debe partir siempre de una fase cualitativa, en la que se trata de definir qué se va a preguntar y cómo se van a medir las respuestas. Preparar un cuestionari no es tan fácil com parece,por lo que se sugiere:

•*No deben emplearse nunca palabras ambiguas*. Por ejemplo:

**Relación de palabras ambiguas**

| Palabras | Por què es ambigua |
|---|---|
| Todo<br>Siempre | Pueden ser *obvias: ¿todas las aves tienen dos patas?; ¿siempre debe avanzar con la luz del semáforo en verde?*<br>Pueden ser *imprecisas, vagas: ¿hace Usted todo lo posible para encontrar un buen precio cuando compra?; ¿está siempre de acuerdo con todo?* |
| Malo | La gente no suele criticar en los cuestionarios |
| País | Posee connotaciones políticaa. Se puede entender como región, Estado, nación, etc. |
| Diario | Puede entenderse como de Lunes a Viernes o de Lunes a Domingo |
| Comida | Se puede entender como alimento, almuerzo, o comida de mediodía |
| Poco, Mucho, Demasido | Puede ser poco para algunos pero mucho o demasiado para otros. Por ejemplo, tomar 5 copas de vino diarias. |

•**Las preguntas deben ser concretas,** si se requieren respuestas concretas. Por ejemplo, se puede preguntar a un fumador : *¿cree Usted que fuma demasiado?* Esta es una pregunta que no es concreta. Para algunas personas demasiado podrían ser 10 cigarros; para otras 20 o 30, o tal vez más. El término *demasiado* es subjetivo. Es mejor preguntar

| | |
|---|---|
| Pregunta : *¿Cuántos cigarros fuma Usted al día?* | Marque con una X |

| | |
|---|---|
| Menos de 5 al día | |
| Entre 5 y 10 | |
| Entre 10 y 20 | |
| Más de un paquete diario | |

| Pregunta : *¿Cuál de las modalidades de innovación que realiza su empresa es la que le representa mayor productividad?* | Marque con una X |
|---|---|
| 1.Organizacional | |
| 2.Social | |
| 3.Del modelo de negocios | |
| 4.Producto/Servicio | |
| 5.Proceso | |
| 6.Tecnología | |
| 7.Mercadotecnia | |

•**Deben evitarse los cálculos.** Las personas encuestadas no deben hacer cálculos, especialmente cuando los cuestionarios son largos, por diversas razones; porque se cansan, se pueden confuindir o su memoria y/o actitud no son buenas.

•**Las preguntas deben de ser breves,** porue así se facilita su comprensión y se reduce el cansancio de los encuestados.

•**Use vocabilario comprensible,** no rebuscado y complejo

•**No deben hacerse preguntas que contengan implícitamente la respuesta .**Si se deseara hacer un estudio sobre la visión de los conductores acerca de la influencia de los motores de los vehículos y la contaminación, la pregunta:

| *Pregunta: Considerando que los motores de combustión interna son demasiado contaminantes y que la legislación en medioambiente será cada vez más estricta, opine sobre la conveniencia de fomentar la producción de* | Acertada | Poco Acertada | Nada Acertada |
|---|---|---|---|
| Motores de gasolina con plomo | | | |
| Motores de gasolina sin plomo | | | |
| Motores diesel | | | |

*Está mal formulada, pues contiene la respuesta*

•**Las preguntas deben ser neutrales.** No deben incorporar juicios, opiniones o valoraciones. Por ejemplo, no estaría bien preguntar: *¿qué le parecen a Usted la rapidez y amabilidad con que le atienden en el establecimiento X?*, es mejor: *¿qué opina Usted del trato que le procuran en el establecimiento X?*

•**Las preguntas se formularán en términos inequívocos** para garantizar su interpretación. El encuestado no tiene qué averiguar qué es lo que se le quiere preguntar. Por ejeemplo, el Ayuntamiento de una ciudad podría estar considerando

la conveniencia de convertor el casco antiguo de la ciudad en peatonal y preguntar a la población: *¿cree Usted que podría resultar comveniente estudiar la posibilidad de que, bajo ciertas circunstancias, se cerrara discrecionalmente y parcialmente al tráfico, la parte antigua d ela ciudad para cierto tipo de vehículos que pudieran tener algunas características considerada eespeciuales?* Una persona a quien se le formjulara un apregunta así, no se enteraría de nada. Si puensara una posible respuesta podría proporcionar multiples argumentos a una pregunta que ella debería esforzarse ( y cómo) en interpretar. La pregunta debería ser:
*¿Aprueba Usted que se cierre definitivamente al tráfico la parte vieja de la ciudad?*
Con posibles respuestas: SI, NO ,Indiferente.

•**Conviene facilitar la memoria**. El encuestado no debe esforzarse en recordar; es cansado o puede que tenga mala memoria. Por ejemplo, si se deseara saber cuánto pago un ama de casa por un litri de aceite en su última compra no debería preguntarse directamente el precio. Salvo que sea una persona especialmente sensible a él no lo recordará. Los consumidores no recordamos exactamente los precios, pero sí intervalos dentro de los cuales se desenvuelven. La pregunta sobre el precio del aceite, podría ser

| Pregunta: ¿Cuánto pago por un litro de aceite? | Marque con una X |
|---|---|
| Entre 30 y 40 pesos m.n. | |
| Entre 41 y 50 pesos m.n. | |
| Más de 51 pesos m.n | |

Si se quisiera saber cuántas veces ha ido una persona al cine al año, no se preguntaría: *¿cuántas veces ha ido Usted al cine en el último año?*; es mejor preguntar: *¿cuántas veces, promedio, va Usted al cine por semana?*, de contestar dos, se deduciría que es persona ve el cine unas 100 veces al año.

•**No deben hacerse preguntas embarazosas** o que puedan dañar la sensibilidad del encuestado. Los estados emocionales deberñan evitarse. En los EUA, se realizó una investigación de mercado en 1959, para definir el perfil de compradores de dos marcas de autos: Chevrolet y Ford. Entre otras preguntas del cuestionario figuraba una relativa a las tendencias sexuales del encuenstado. En una primera fase, fueron muchas las personas que se negaron a contestar el cuestionario en su totalidad. Finalmente, se tuvo que desapoarecer la pregunta. Esta anécdota pone de relievela necesidad de desarrollar fases cualitativas para diseñar los cuestionarios y la conveniencia de jhacer algunas pruebas previas a su distribución masiva.
Desde una perspectiva menos Freudiana, algunas preguntas son también delicadas, como por ejemplo, la relativa a los ingresos de las personas o su edad. No debe preguntarse directamente la renta. Es mas fácil obtener una respuesta cuando se formula la pregunta :

| Pregunta: *Por favor, indique dentro de qué tramo se enucntran sus ingresos anuales* | Marque con una X |
|---|---|
| Menos de 100, 000 pesos m.n. | |
| Entre 101,000 y 200,000 pesos m.n. | |

| | |
|---|---|
| Más de 201,000 pesos m.n | |

Algo semejante sucede con una pregunta relativa a la edad de la persona encuestada. Es más fáciñ obtenrr una respuesta si se pregunta la edad por intervalos que si se pide directamente

## Estructurar el cuestionario

Un cuestionario, consta de tres partes: *introducción, cuerpo, identificadores, orden preguntas y de cuestionario*

**Introducción**

- Se deberá identificar a quien haya encargado la investigación, salvo que esto no lo desee.
- Se exploicará el objetivo de la investigación.
- Se explicará cómo se seleccionó a la persona que responde, a fin de evitar recelos y conseguir la colaboración.
- Si el cuestionario se cumplimentara cara a cara, el encuestador hará una valoración de la idoneidad de la persona seleccionada para ser encuestada, si sospechara que su colaboración fuera dudosa.

**Cuerpo**

O llamado también, *conjunto de preguntas* que se van a formular y que se puede subdividir en los siguuientes *tipos*:

- De aproximación, que suelen ser muy sencillas, poco comprometidas y que *generan interés por colaborar.* Por ejemplo: *¿con qué frecuencia convoca a su comité de investigación y desarrollo?*
- Figuran al principio del cuestionario. A medida que se avanza en el mismo, se puede aumentar el grtado de complejidad de su formulación o contenido. Si las preguntas delicadas se colocan a principio, la posibilidad de una negativa a colaborar es mayor.

**Transición**

- Se formulan antes de las preguntas complejas. Por ejemplo: *Describa en este momento cómo se desarrolla la transmisión de conocimiento en su organización.*
- Se deben colocar entre las preguntas de aproximación y las complejas.

**Complejas**

- Requieren de mucha atención, interés y juicio por parte de quien responde. Por ejemplo: *Proponga Usted un nuevo proceso de innovación de mercadotecnia a su organización, que le permita incrementar sus ventas en un 2% mensual.*

**Identificadores del encuestado**

- Son variables generalmente nominales que permiten reconoicer las características demográficas o sociodemagráficas de quienes proporcionan la información y que *son fundamentales para formar grupos o identificar segmentos homogéneos.* Deben ir al final del cuestionario
- Una vez respondido, , la persona que lo haya hecho comprenderá los objetivos de

la investigación y entenderá mejor por qué se le piden datos que la caractericen. Tras finalizar el levantamiento de información, es más fácil que se proporcionen los identificadores relativos a edad, renta, estudios y actividad laboral.

**Orden de las preguntas**

Esto no debe de afectar a las respuestas. *esta es probablemente la norma de redacción de cuestionarios más difícil de aplicar.* Lo deseable es que las respuestas emitidas por los encuestados NO dependan del orden en que se formulan. *cada pregunta debería ser independiente* de las anteriores. En una investigación sobre tecnología aplicada, podrían hacerse estas dos preguntas, entre otras:

- *El gobierno estatal va a invertir para impulsar la tecnología transgénica en semillas resistentes a plagas, a fin de garantizar la productividad y garantizar la viabilidad alimentaria, así como los empleos de la región Indiique si está de acuerdo.*
- *Considerando que existen dudas sobre la seguridad de la tecnología transgénic, indique alternativas sustitutivas de combatir las plagas para garantizar la productividad y garantizar la viabilidad alimentaria, así como los empleos de la región.*Como se observa, el orden de presentación, tiende a condicionar la respuesta.

**El orden del cuestionario**

- Este deberá ser congruente, es decir, lógico. Las preguntas deben realizarse por temas afines y en orden de dificultad creciente. En una investigación sobre el gasto de tecnología en una pyme, deberán hacerse las preguntas por naturaleza de gastos.
- Por ejemplo: insumos, logística, procesos, investigación y desarrollo, etc. y no mezclaros.
- Por ejemplo, cada pregunta, debe conducir a la siguiente. En el caso de identificar de los poseedor de un equipo tecnológico en uso específica.

| Pregunta | Nombre |
|---|---|
| Indique la marca | |
| Indique el modelo | |
| ¿Cántos años tiene de uso? | Marque con una X |
| Menos de uno | |
| Entre uno y dos | |
| Entre dos y cuatro | |
| Más de cuatro | |
| El mantenimiento, ¿quién lo realiza? | Marque con una X |
| Su empresa | |
| Servicio de agencia | |
| Servicio de un tercero | |
| El desemèño de su equipo es | Marque con una X |
| Satisfactorio | |
| Poco satisfecho | |
| Nada satisfecho | |

| Cuando termine la vida útil de su equipo | Marque con una X |
|---|---|
| Comprará otro de la misma marca | |
| No creo repetir la marca | |
| Compraré de otra marca | |

## Prueba y valoración del cuestionario

Una vez que se ha diseñado el cuestionario, deberá responder:

- ¿Responde el cuestionario a los objetivos de la investigación?
- ¿Son necesarias todas las preguntas?
- ¿Podrá el encuestado responder todas las preguntas?
- ¿Querrán los encuestados contestar a todas las preguntas?
- ¿Es fluído?, ¿cansa?
- ¿Es de una extensión razonable?
- ¿La secuencia de preguntas es correcta?
- ¿Se han incluído transiciones e introducciones?

## Técnicas para realizar preguntas difíciles

*Cuando se rata de preguntas delicadas, es difícil obtener respuestas sinceras, si las preguntas se hicieran directamente*.

- Por ejemplo, si se preguntara a algún trabajador de procesos de alto involucramiento de su cuerpo si ha consumido drogas, es muy probable que diga que no, aunque la evidencia muestre lo contrario
- Si se desarrollara una investigación para conocer la cantidad de trabajadores del sector, con las mismas características, que hubieren consumido alguna droga, no sería necesario conocer las respuestas individuales, sino *una estimación del porcentaje.* En casos como este se puede acudir a la técnica de **respuestas aleatorizadas,** que garantice el anonimato de la respuesta del entrevistado y asegura su colboración. Para aplicar esta técnica, se formulan dos pregintas simultáneamente. Una delicada y una intrascendende:

**a**.Acaba su número de SS en cinco.

**b**.¿Ha consumido drogas en alguna ocasión?

Antes de responder, se le proporciona al encuestado el medio de elegir entre ambas preguntas

- Por ejemplo, *lanzando un dado*. Si saliera un uno o un dos, deberá responder ña pregunta **a.** En caso contrario, deberá responder a la **b.**

El investigador no conoce el resultado del lanzamiento y no sabe a qué pregunta responden los encuestados. Sólo conoce una respuesta: SÍ o NO. La probabilidad de responder a la pregunta **a** es conocida como **p0;** en este caso 2/6. La de responder a la pregunta **b** es **1-p0,** en este caso 1-(2/6), es decir 4/6. La probabilidad de SÍ a la pregunta **a** es conocida como pa=1/10, suponiendo que sean 10 las posibles terminaciones del SS, todas con igual probabilidad. Sea **p,** la proporción de *Síes* obtenidos en la encuesta. La estimación de la proporción de Síes a la pregunta **b**, se calcula:

$$(p-p0pA)/(1-p0)$$

Por ejemplo, en una encuesta realizada a 1000 trabajadores en la que 300 han

respondido SI a las preguntas (a la que corresponda tras el lanzamiento del dado), la proporción de trabajadores del sector que se drogan, es:

**((30/1000) –(2/6)(1/10))/ (1-2/6)= 0.4**

• Otra técnica para hacer preguntas delicadas, es proporcionar al entrevistado unos cartones con las respuestas codificada, de forma que el entrevistado responda a cada pregunta únicamente con su código. Como el entrevistador no tiene a la vista estos códigos, no identifica la respuesta.

## ¿Responden bien los encuestados?

Un cuestionario puede estar muy bien diseñado y el encuestador ser excelente. No obstante, *siempre existe la duda sobre la veracidad de la información* que se obtiene.

**Fuentes de incertidumbre**

• **Quien responde tiene dificultad para expresarse o para comprender, el cuestionario.** Puede ser por razones culturales o intelectuales. En general, los cuestionarios con *preguntas cerradas* son los más aconsejables para personas que tengan estas características en grado medio o bajo. Cuando se tratade *preguntas abiertas* es más difícil obtener respuestas, porque obligan a un ejercicio mental. A veces el número de ítems o de categorías en cada uno de ellos puede resultar excesivo, especialmente con personas mayores o muy jóvenes.

• **Quien responde, tiene mala memoria.** Este problema aparece con frecuencia cuando se trata de personas de edad, pero también puede aparecer aisladamente en cualquier otra. Para conseguir información fiable, se puede acudir a listas u otros elementos visuales (fotos, catálogos, etc.), que ayuden a recordar. También se puede invitar a que la persona escriba un diario.

• **Quien responde puede ser reacio a contestar.**Las razones pueden ser de diversa naturaleza. Pueden ser *inconscientes o irracionales* pues tal vez el encuestado no pueda dar argumentos sobre su negativa a responder. Otras veces, pueden surgir *barreras sociales o de inadmisibilidad.* Por ejemplo, una persona que sea alcohólica, será muy reacia a informar sobre la cantidad de alcohol que consume.

Otras veces las personas tienen a dar respuestas socialmente aceptadas, aunque internamente piensenlo contrario. Es posible que contesten en un sentido por educación o para acabar la entrevista o encuesta cuanto antes.

## El cuestionario piloto

Una vez que se haya diseñado el cuestionario, este debe ser sometido a un grupo de personas para efectuar una prueba. En un primer resultado de diseño, es posible que no se acierte con aspectos semánticos en las preguntas. Es decir, *la redacción del cuestionario puede no ser del todo correcta o que no se comprenda bien.* Es posible también, que algunas preguntas importantes no se hayan incluído o no estén bien matizadas, o que haya un exceso de preguntas y algunas no sean significativas.

Los defectos de contenido o forma que pudieran aparecer en el cuestionario, se detectan mediante pruebas piloto dirigidas a pequeños grupos. Una vez subsanados los errores o perfeccionado el cuestionario, se podrá dirigir a la totalidad de las personas que se considere oportun que deben responderlo. *De esta manera, se*

*evita tener que repetir la investigación* por haber difundido un custionario confuso o erróneo.
Así, la prueba piloto del cuestionario persigue:

- Eliminar ambiguedades
- Eliminar preguntas superfluas
- Añadir al cuestionario preguntas relevantes
- Simplificar preguntas difíciles
- Cambiar el orden de las preguntas para agilizar el flujo de respuestas
- Corregir la redacción
- Eliminar faltas de ortografía, puntuación, sintaxis y/o gramática
- Comprobar que los códigos para grabar los datos más adelante sean correctos

En la prueba piloto, se mide la consistencia interna del cuestionario a través del coeficiente a de Cronbach o el coeficiente 0. **Se medirá la validez de la escala y si fuera necesario, se eliminan o añaden ítems o categorías**. L as pruebas piloto se repiten las veces necesarias hasta conseguir la mayor validez del cuestionario. Cuando se haya conseguido se reproduce y distribuye.

Fuente: Grande-Esteban y Abascal-Fernández (2017) y Hernández-Sampieri et al., (2010) con adaptación propia

# Medición. Escalas y sus propiedades

El concepto de *medición* puede entenderse como *la forma de obtener símbolos para representar propiedades de las personas , cosas, sucesos o ideas, según la cual, esos símbolos guardan una relación entre ellos dela misma forma que los entes que representan* (Green et al. 1988). La medición también puede entenderse como la asignación de números a objetos para representar sus propiedades o las relaciones entre ellos. Par auna mayor comprensión de lo que representa las escalas de medición, debe tenerse en cuenta lo mostrado en la **Tabla A.20.**

**Tabla A.20. Propiedades de los númeors**

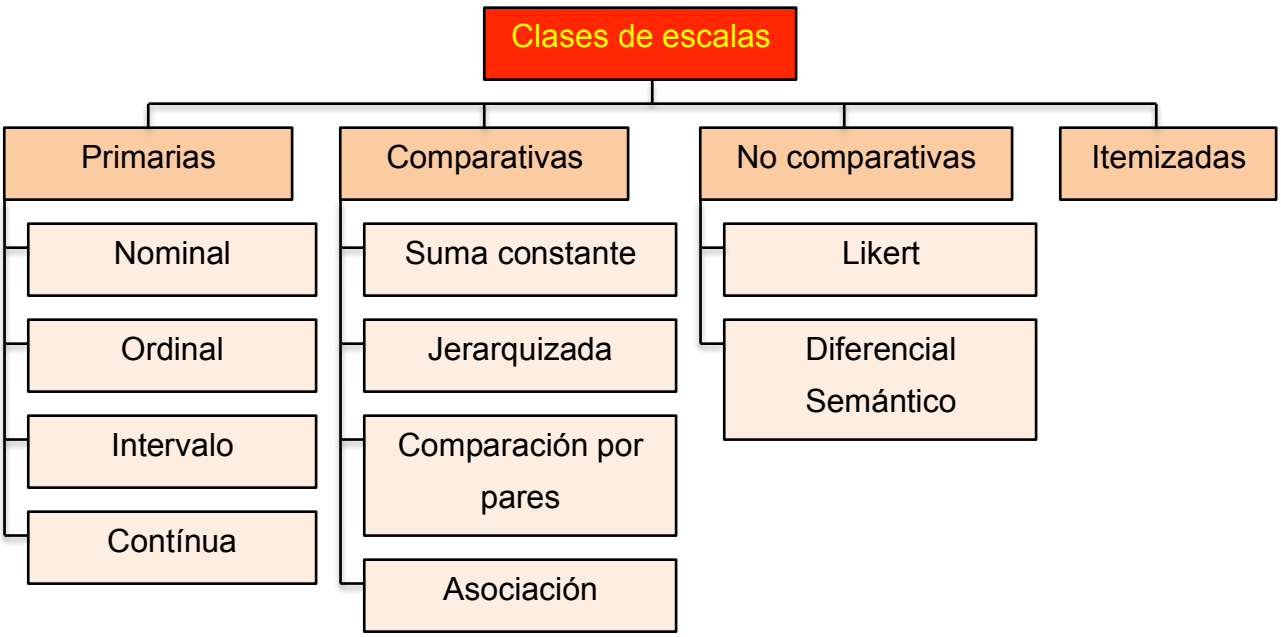
| Propiedad | Descripción |
|---|---|
| Orden | O secuencia numérica ordenada . Por ejemplo un kilogramo, dos kilogramos, tres kilogramos, etc. |
| Distancia | O diferencias entre números, sin que se preste atención al origen. Por ejemplo, entre 15 grados Celsius y 20 grados Celsius existen 5 grados Celsius. |
| Origen | Las series pueden tener un único origen determinado por el cero. por ejemplo, la edad, la renta o el peso son magnitudes cuyo origen es el cero. |

Fuente: Grande-Esteban y Abascal-Fernández (2017) con adaptación propia

# Clases de Escalas

Se tienen básicamente: primarias, comparativas, no comparativas, itamizadas las cuales tienen subdivisiones que pueden ser consideradas para diferentes necesidades del investigador, como se muestra en la **Figura A.7.**

**Figura A.7. Clases de escalas**




Fuente: Grande-Esteban y Abascal-Fernández (2017) y Hernández-Sampieri et al., (2010) con adaptación propia

## *Escalas Primarias*

La clasificación más elemental de escalas es la que diferencia entre nominales, ordinales, métricas de intervalo y métricas contínuas, Ver **Tabla A.21.**

**Tabla A.21. Escalas primarias**

| **Escalas nominales** |
|---|
| Son las escalas más irrestricTas de todas. No tienen orden, ni distancia, ni origen. Se emplean para medir aspectos cualitativos. Los números identifican y clasifican objetos mutuamente excluyentes. Ejemplos muy ilustrativos son las referencias de nombre, cargo, etiquetas de tallas, los números de las camisetas de los deportistas, etc. Son identificadores que se dan de forma arbitraria: pueden ser nombresm números, letras o códigios alfanuméricos. Se emplean generalmente para identificar a los individuos (personas, empresas, organismos, establecimientos comerciales, etc.), en los cuestionarios. Los individuos de una misma clase tienen el mismo número. En un cuestionario se podría llamar 1 a los homnres y 2 a las mujeres.. Si se prestara atención a la situación laboral de las personas, 1 podría indicar cuadro superior; 2, indicaría cuadro intermedio; 3, funcionario; 4, pernsionado; 5, desempleado, etc. Esta escala no es proporcional, es decir, 2 no es el doble de 1 ni 10 es 2x5. Los números son sólo códigos de identificación.<br>**Características**<br>• Sirven para identificar obketos<br>• Los valores numéricos son nombres de categorías. |

- Los números no tienen magnitud
- Los números no miden ninguna relación

**Ejemplos**

- Género
- Profesiones
- Usuario/No Usuario
- Tipos de establecimiento
- Niveles de renta
- Nivel de estudios
- Tallas de ropa o calzado
- etc.

**Aplicaciones**

- Identificación de cuestionarios
- Variables ilustrativas en análisis factoriales
- Estudios de segmentación
- Identificación de perfiles de consumidores

**Tratamiento estadístico**

- No se pueden hacer operaciones matemáticas
- Análisis cualitativo
- Tablas cruzadas
- Frecuencias
- Tablas de porcentajes simples y acumulados
- Gráficos
- Análisis factorial de correspondencias múltiples
- Elementos ilustrativos e análisis de componentes principales o factorial de correspondencias

## Escalas ordinales

Las escalas ordinales sólo indican posiciones relativas. Sólo poseen la propiedad de orden, pero no las de distancia ni origen. Entre los números no existe ninguna relación de proporcionalidad. La diferencia entre ellos no mide ninguna magnitud que tenga sentido. Es muy utilizada esta escala para medir preferencias, calidad, etc. Por ejemplo, un consumidor de productos automotrices podría, con base a su percepcion de calidad, calificar con 1 a Nissan, 2 Ford, 3 Kia, etc.

En resumen:

**Características**

- Establecen un orden de preferencias
- Los valores numéricos muestran un orden o secuencia.
- Los intervalos no miden magnitudes
- Es una forma fácil de recoger información sobre cualquier tipo de ítems

**Ejemplos**

- Preferencias por marcas

**Aplicaciones**

- Estudios depreferencias
- Similitudes
- Disimilitudes
- Posicionamiento

**Tratamiento estadístico**

- Métodos no paramétricos
- Correlaciones
- Análisis de varianza por rangos
- Escalas multidimensionales
- Percentiles
- Medianas
- Tablas cruzadas

## Escalas métricas de intervalo

En éstas, los números sirven para medir las diferencias entre los objetos, de forma que las distancias en la escala representan las mismas distancias en valoración de esos objetos, como sucede con las temperaturas. Estas esclas miden la distancia entre dos números. Estas escalas tienen las propiedades de *orden y distancia, pero el cero es arbitrario.* La escala métrica de intervalo mantiene un orden y las diferencias de valores tienen sentido, cosa que no sucede con las escalas ordinales.No poseen la propiedad de origen, porque el cero es arbitrario.

La diferencia entre 1 y 3 es la mismo que 3 y 5, pero 3 no es el triple de 1. La diferencia entre 5, 10, 15, 20, 25 grados Celsius, es de cinco grados entre cada valor consecutivo. Llamamos 0 grados Celsius a la temperatura de congelación del agua y 100 grados Celsius a la ebullición, eso es la escla centígrada. Obsevemos que en la escla Reamur , el agua se congela a 0 grados Celsius , pero hierve a 80 grados Celsius.. En la escala Farenheit la temperatura de congelación es a los 32 grados Celsius y la de ebullición es de 212 grados Celsius.No tiene sentido calcular relaciones; 20 grados Celsius no significa que esté 20 veces mejor que a 1 grado Celsius. Las operaciones que se pueden realizar son, además de las escalas reseñadas, medias aritméticas, desviaciones, correlaciones, etc.

**Caracerísticas**

- Miden distancia entre objetos
- Los valores numéricos muestran orden y magnitud
- Los intervalos son iguales
- El cero no es origen
- El cero es arbitrario

**Ejemplos**

- Grados de acuerdo
- Temperatura

**Aplicaciones**

- Estudios de actitudes
- Estudios de opinión

**Tratamiento estadístico**

| |
|---|
| •Correlaciones<br>•Contrastes de hipótesis<br>•Análisis de la varianza<br>•Análisis factorial |
| **Escalas métricss contínuas** |
| Tienen las propiedades de orden, distancia y origen. Se asignan valoraciones a algún objeto de forma que esa valoración tenga las propiedades de origen, distancia y orden. Por ejemplo, se podría pedir que se valorara de 0 a 5 la calidad de un producto, entendiendo que 0 es la peor calidad y 5 la mejor. Se pueden valorar estímulos , objetos o atributos concretos. Se pueden valorar atributos de una marca de ropa como calidad, diseño, precio y elegancia o la calidad de servicio de un hotel, restaurante, etc. Estas escalas poseen todas las propiedades de las escalas nominales, ordinales y de intervalo..<br>**Características**<br>•Sirven para hacer comparaciones abolutas de magnitudes<br>•Los valores numéricos muestran origen. Orden y distancia<br>•Los intervalos son iguales<br>•El cero indica ausencia de magnitid<br>**Ejemplos**<br>•Edad; •Peso<br>•Renta<br>•Ventas<br>•Cuotas de mercado<br>**Aplicaciones**<br>•Estudios de imagen<br>•Estudios de segmentación<br>•Estudios de posicionamiento<br>**Tratamiento estadístico**<br>•Permiten la utilización de cualquier técnica estadística o matemática |

Fuente: Grande-Esteban y Abascal-Fernández (2017) y Hernández-Sampieri et al., (2010) con adaptación propia

## *Escalas comparativas*

En el apartado anterior se han expuesto escalas por las propiedades de los números empleados: orden, distancia y origen. Sin embargo, debe profundizarse más. Existen escalas a aplicar que son *comparativas* y *no comparativas*. Las primeras permiten hacer comparaciones entre objetos (marcas, modelos, productos, entidades, etc.) que se estudin. Se muestran las más frecuentes en la **Tabla A.22**

**Tabla A.22. Escalas comparativas**

**Escalas de suma constante**

Consiste en repartir una puntuación entre ciertos atributos determinados. Por ejemplo, se podrían repartir 100 puntos entre tres atributos de computadoras:

| | Diseño | Calidad | Batería | Total |
|---|---|---|---|---|
| Lenovo | 30 | 45 | 25 | 100 |
| Toshiba | 20 | 50 | 30 | 100 |
| Dell | 50 | 25 | 25 | 100 |

**Escalas jerarquizadas**

El respondente establece un orden de preferencias, por ejemplo: *las marcas de aparatos de smartphones que más me gustan son: Iphone, Samsung, Motorolla, Alcatel, Huawei.*

**Escalas comparación por pares**

Las personas que responden tienen que elegir uno de los dos objetos que secuencialmente se le presentan. Por ejemplo, se le pueden presentar dos marcas de motociletas: Suzuki y Honda; otro par puede ser: Honda y BMW. Si prefiere Suzuki a Honda y Honda a BMW, se concluye que su orden de preferencias es: Suzuki, Honda y BMW, por aplicación de propiedad transitiva.

**Escalas de asociación**

A través de ellas se trata de relacionar atributos de diversos conjuntos. Una escala de asociación, es:

| Las motos | Suzuki | Honda | BMW |
|---|---|---|---|
| Potencia | X | | |
| Rápidas | | | X |
| Estabilidad | | X | |

Cuando se está de acuerdo con la asociación, se marca la casilla correspondiente . Las escalas de asociación proporcionan, cuando se tabulan, *tablas de frecuencia* de asociación entre los atributos y los objetos (en este caso marcas de motos), por ejemplo:

| Las motos | Suzuki | Honda | BMW |
|---|---|---|---|
| Potencia | 100 | 124 | 235 |
| Rápidas | 256 | 345 | 321 |
| Estabilidad | 156 | 450 | 125 |

Fuente: Grande-Esteban y Abascal-Fernández (2017) y Hernández-Sampieri et al., (2010) con adaptación propia

## *Escalas no comparativas*

Las escalas no comparativas más conocidas son las de Likert y los diferenciales semánticos. Ver **Tabla A.23.**

**Tabla A.23. Escalas no comparativas**

**Escala de Likert**

Consiste en formular proposiciones relativas a una serie de atributos de un objeto y que el entrevistado exprese su grado de acuerdo o desacuerdo en una escala de varias categorías , que pueden ser 3, 5, 7, 9 u 11. Cada categoría se puntúa con un número que se le asigna. Por ejemplo, ante la pregunta: *¿Cuál es su percepción sobre los motores de los autos eléctricos? (Marque con una X)*

*Escala de Likert de cinco categorías*

| Percepción | Muy de acuerdo (5) | De acuerdo (4) | Indiferente (3) | Desacuerdo (2) | Muy en desacuerdo (1) |
|---|---|---|---|---|---|
| Sencillos de mantener | | | | | |
| Potentes | | | | | |
| Económico | | | | | |

*Escala de Likert de 3 categorías*

| Percepción | De acuerdo (3) | Indiferente (2) | En desacuerdo (1) |
|---|---|---|---|
| Sencillos de mantener | | | |
| Potentes | | | |
| Económicos | | | |

**Escala diferencial semántico**

El criterio no viene definido por una proposición, para la que se indican los grados de acuerdo o desacuerdo, sino por categorías opuestas de una dimensión. En el ejemplo que figura a continuación se consideran algunas dimensiones de unos motores, como el diseño, el consumo, la fiabilidad, el ruido, etc. Quien responda a la escala, debe manifestar una tendencia hacia las categorías opuestas de la dimensión considerada. Por ejemplo:

| *Los motores Tesla (Marque con una X)* | | | | | | |
|---|---|---|---|---|---|---|
| Tienen un diseño avanzado | | | | | | Su diseño es anticuado |

| | | | | | | |
|---|---|---|---|---|---|---|
| Consumen gran cantidad de energía eléctrica | | | | | | Consumen poca cantidad de energía eléctrica |
| Se averían poco | | | | | | Se averían con frecuencia |
| Son ruidosos | | | | | | No son ruidosos |
| Su mecánica es simple | | | | | | Su mecánica es complicada |
| Contaminan poco | | | | | | Contaminan mucho |

El encuestado se acerca hacia la *zona* con la que esté más de acuerdo. Habitualmente, los atributos desfavorables se sitúan unas veces a la izquierda y otras veces a la derecha. Se hace así para evitar que algunas personas respondan siempre a la izquierda o a la derecha sin leer las afirmaciones. A veces es difícil encontrar la definición de los extremos.

Algunos d elos extremos de la escala son muy sencillos, porque existen *antónimos,* por ejemplo, *bueno-malo*, *claro-oscuro.* Tras veces, es más difícil precisar; por ejemplo, los *antónimos de potente, ergonómico, aerodinámico* hay que definirlos mediante redacciones exprofeso.

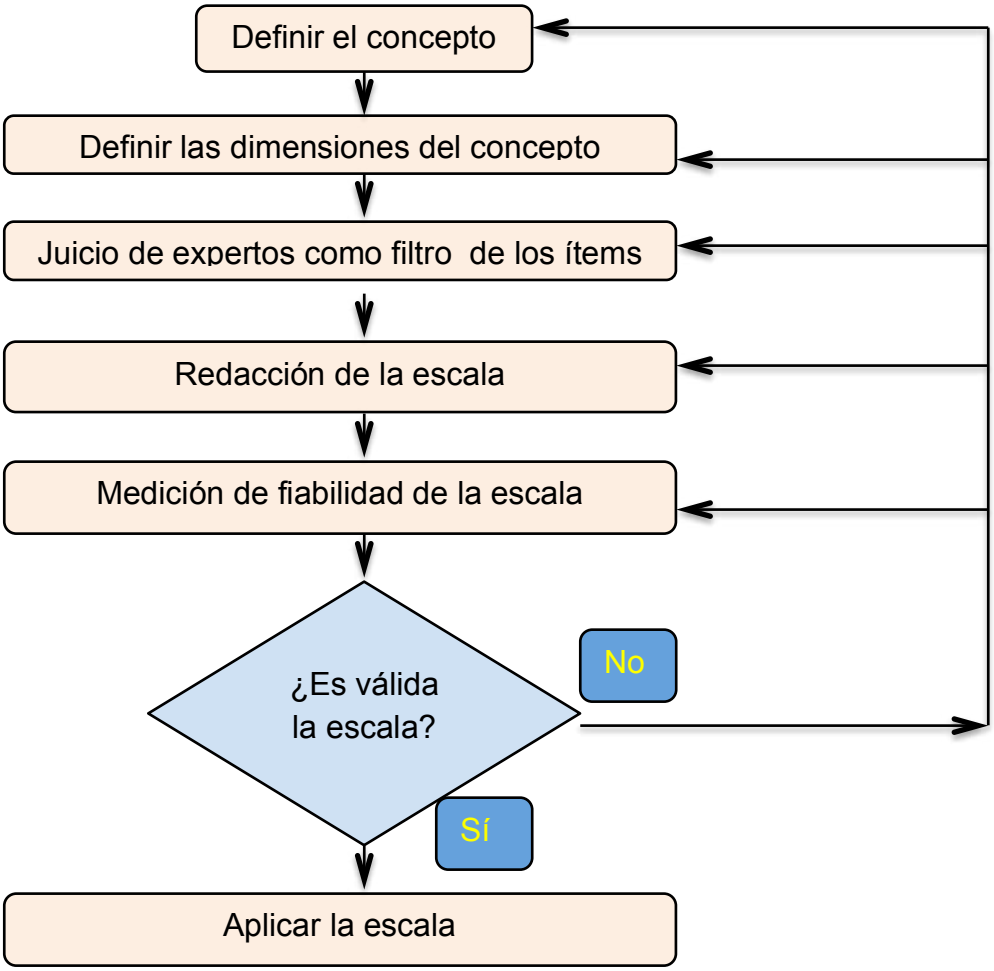
Cada casilla es codificada con un número, que oscila entre el mácimo (siete en este caso, aunque 9 y 11 son frecuentes), para la proposición más favorable.

Fuente: Grande-Esteban y Abascal-Fernández (2017) y Hernández-Sampieri et al., (2010) con adaptación propia

## *Escalas Itemizadas*

Este tipo de escala no comparativa se caracteriza porque contiene varias afirmaciones o proposiciones, relativas a un objeto, que guardan relación entre síy que deben ser asociadas. Tienen varios ítems porque miden conceptos complejos. Ina escala mide un concepto; por ejemplo, la calidad. Pero los conceptos pueden tener varias *dimensiones.* La calidad se mide a través de elementos tangibles, como: el aspecto de las personas, de las instalaciones o de los materiales que se emplean. Otras dimensiones de la calidad son: la capacidad de respuesta, la empatía, la seguridad, la fiabilidad o la cortesía. Una escala multi-ítem recoge las distontas dimensiones del concepto que se pretende medir con la escala. Las razones que aconsejan emplear escalas itemizadas son la *fiabilidad* de las mediciones y la *precisión* cuando los conceptos son complejos . Ver **Figura A.8.**

**Figura A.8. Metodología para construir escalas**




Fuente: Grande-Esteban y Abascal-Fernández (2017) con adaptación propia

Vea la **Tabla A.24.**

**Tabla A.24. Componentes básicos en la construcción de escalas**

| El concepto |
|---|
| •En primer lugar deben definirse una serie de características u objetos para medir, que en nuestro idioma se denomina *concepto.*Su findamento se basa en una extensa investigación cualitativa/cuantitativa , revisión bibliográfica, o en estudios empíricos. Pofr ejemplo, la innovación, el *customer knowledge management,* la musicología, etc.<br>•En segundo lugar, deben definirse las *dimensiones del concepto.*Tratándose de la calidad, algunas de sus dimensiones *pudieran ser* la tangibilidad, la seguridad, la foabilidad, la comunicación, la empatía y la cortesía. *Es posible que el número de dimensiones sea excesivo*. Un *juicio de expertos* permitiría filtrar las dimensiones propuestas. Una vez definidas las dimensiones del concepto, se redactan los ítems o indicadores, es decir, las preguntas. Por ejemplo, para medir la tendencia de la fiabilidad de los consumidores, se pueden utilizar los siguientes ítemsante los cuales |

se pide al encuestado que indique su *grado, nivel* de acuerdo o de desacuerdo con las siguientes afirmaciones:
**1.**Generalmente compro la misma marca siempre.
**2.**Una vez que elijo una marca, tiendo a segur con ellla y me olvido de las demás.
**3.**Una vez que elijo una marca, detesto los cambios.
**4.**Si me gusta una marca, raramente cambio oara probar algo distinto.
**5.**Aun cuando un mismo producto esté disponible en varias marcas, tiendo a comprar la misma.

•En tercer lugar, *las preguntas se someten a un proceso de filtramiento*, en el que un grupo de personas opinará sobre sonveniencia o adecuación. Puede que existan ítems que no tenga muchp sentido incluir, que sean irrelevantes o que no se entiendan. Un *juicio de expertos* clasifica los ítems por su afinidad. El grupo lo forman personas con conocimientos y criterios sobre la materia que se van a estudiar para poder clasificar los ítems en grupos homogéneos. *En la propuesta de escala anterior, el grupo de expertos ha decidido eliminar el primer ítem por considerarlo redundante y es a su juicio (aquí es recomendable justificarlo con un análisis factorial exploratorio, visto más adelante), igual que el segundo*, con lo cual, la escala queda así:
**1.**Una vez que elijo una marca, tiendo a segur con ellla y me olvido de las demás.
**2.**Una vez que elijo una marca, detesto los cambios.
**3.**Si me gusta una marca, raramente cambio oara probar algo distinto.
**4.**Aun cuando un mismo producto esté disponible en varias marcas, tiendo a comprar la misma.

•Después, *se redacta la escala de forma definitiva*, se aplica y se miden la fiabilidad y su validez, es decir, su capacidad de medición del concepto en toda circunstancia. En caso afirmativo, la escala se aceptay en caso contrario, se vuelve a las fases previas.

## Las categorías

Para diseñar las escalas, deben tenerse en cuenta:

•**El número de categorías de la escala**. Si es métrica contínua tienen uno, el valor que asignba el encuestado. Generalmente las escalas de Likert consideran cinco categorías, aunque es frecuente encontrar estudios con siete, nueve e incluso trece. Los diferenciales semánticos, suelen tener entre siete y nueve categorías. *Cuanto mayor sea el nñumero de categorías, más recisa será la escala*, en principio. *Cuando el número de categorías es amplio, cosa que sucede cuando se dispone de cinco o más, los códigos asignados a las categorías tienen las propiedades de los números, es decir, la escala se convierte en métrica.* Sucede sin embargo, que algunos grupos de población no son capaces de desenvolverse con demasiadas categorías. Dificultades de percepción derivadas de la edad reducida o elveda , cultura o nivel de inteligencia de los encuestados, pueden aconsejar reducir el número de categortías a tres o cinco como máximo. La práctica aconseja reducir el número de categorías que sean siete +- dos, es decir cinco, siete, nueve.El número de categorías también depende de la forma de recoger la información. Si hay tiempo y calma, por ejemplo en encuestas en establecimientos o por teléfono. También influye el grado de conocimiento que se tenga sobre los ítems que se preguntan. A

mayor conocimiento o interés por responder, mayor es el número de categorías posible.

•**Categorías equilibradas o no.** Una escala se encuentra equilibrada cuando el número de categorías favorables y desfavorables coinciden, como en la siguiente:
*¿Cuál es su percepción sobre los motores de los autos eléctricos? (Marque con una X)*

*Escala de Likert de cinco categorías*

| Percepción | Muy de acuerdo (5) | De acuerdo (4) | Indiferente (3) | Desacuerdo (2) | Muy en desacuerdo (1) |
|---|---|---|---|---|---|
| Sencillos de mantener | | | | | |
| Potentes | | | | | |
| Económico | | | | | |

Ya que *existen dos desacuerdos y dos acuerdos.* En general, es mejor utiliar escalas equilibradas. Sin embargo, si tras una primera prueba se observara que existen muchas más respuestas favorables (o desfavorables), podrían añadirse más categorías en el sentido observado y reducirlas en el que menos respuestas se hayan encontrado.
La escala que aparece a continuación, *está desequilibrada:*
*¿Cuál es su percepción sobre los motores de los autos eléctricos? (Marque con una X)*

*Escala de Likert de cuatro categorías*

| Percepción | Muy de acuerdo (4) | De acuerdo (3) | Indiferente (3) | Desacuerdo (1) |
|---|---|---|---|---|
| Sencillos de mantener | | | | |
| Potentes | | | | |
| Económico | | | | |

•**Número de Categorías pares o impares**. En la siguiente escala de Likert, el número de categorías es impar:
*¿Cuál es su percepción sobre los motores de los autos eléctricos? (Marque con una X)*

*Escala de Likert de cinco categorías*

| Percepción | Muy de acuerdo (5) | De acuerdo (4) | Indiferente (3) | Desacuerdo (2) | Muy en desacuerdo (1) |
|---|---|---|---|---|---|

| | | | | | |
|---|---|---|---|---|---|
| Sencillos de mantener | | | | | |
| Potentes | | | | | |
| Económico | | | | | |

Pero la siguiente es par ya que se ha suprimido la categoría indiferente
*¿Cuál es su percepción sobre los motores de los autos eléctricos? (Marque con una X)*

*Escala de Likert de cuatro categorías*

| Percepción | Muy de acuerdo (4) | De acuerdo (3) | Desacuerdo (2) | Muy en desacuerdo (1) |
|---|---|---|---|---|
| Sencillos de mantener | | | | |
| Potentes | | | | |
| Económico | | | | |

La decisión de un número de categorías par o impar debe tomarse despues de haber hecho alguna prueba con el custionario. Si algunas personas manifiestan la conveniencia de incluir la categoría indiferente así se hará.

• **Respuestas forzadas.** En una escala, las respuestas se fuerzan cuando no se proporcionan opciones intermedias, porque se *elimina la opción indiferente*, o no se contempla el *no se sabe-no contesta*. Cuando no saben contestar tienden a responder *indiferente.* En estops casos , las personas tienden a proporcionar una valoración media quer no es real y que puede distorsionar los resultados. *No sabe-no contesta,* tiene que ser una opción más de respuesta, que puede materializarse en una categoría adicional o sencillamente, en dejar en blanco la respuesta por parte de la persona que responde o recohe la información. El siguiente ejemplo, muestra una escala par y además, forzada.
*¿Cuál es su percepción sobre los motores de los autos eléctricos? (Marque con una X)*

*Escala de Likert de cuatro categorías*

| Percepción | Muy de acuerdo (4) | De acuerdo (3) | Desacuerdo (2) | Muy en desacuerdo (1) |
|---|---|---|---|---|
| Sencillos de mantener | | | | |
| Potentes | | | | |

| | | | | |
|---|---|---|---|---|
| Económico | | | | |

•**Grado de descripción verbal.** Debe prestarse mucha atención a los adjetivos que se empleen. Por ejemplo, cabe esperar respuestas distintas ante estas afirmaciones en una escala de diferencial semántico. Por ejemplo: *Telcel (Marque con una X)*

| *Telcel (Marque con una X)* | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Proporciona un servicio pésimo | | | | | | | | Proporciona un servicio óptimo |
| Proporciona un servcicio malo | | | | | | | | Proporciona un servicio bueno |

Cuando se hacen *afirmaciones rotundas* (servico pésimo-servicio óptimo), las personas tienden a respoinder alrededor de la media. Cuando se es *menos contundente* (servicio malo- servicio bueno) las respuestas se pueden aproximar a los extremos.

Las recomendaciones para redactar las propoisiciones de cada ítem, son:

•El ítem debe expresar sólo una idea o concepto

•Deben emplearse y alternarse proposiciones positivas y negativas *para evitar respuestas carretilla*

•Deben evitarse expresiones vulgares, jerga o de argot.

•La redacción debe adaptarse al nivel cultural o intelectual de las personas que responderán.

Fuente: Grande-Esteban y Abascal-Fernández (2017) y Hernández-Sampieri et al., (2010) con adaptación propia

Así, mostramos un resumen en la **Tabla A.25.**

**Tabla A.25. Valoración de las escalas**

| Escala nominal | |
|---|---|
| Ventajas | Desventajas |
| •Muy fácil de diseñar<br>•Muy fácil de comprender<br>•Ausencia de sesgos en su redacción<br>•No requiere de conocimientos especiales del encuestado<br>•Muy útil en estudios convaroables cualitativas<br>•No fatiga | •Pocas posibilidades de tratamiento cuantitativo<br>•Análisis estadístico muy limitados |
| Escala ordinal | |
| Ventajas | Desventajas |

| | |
|---|---|
| •Muy fácil de diseñar<br>•Muy fácil de comprender<br>•Ausencia de sesgos en su redacción<br>•No fatiga | •Pocas posibilidades de tratamiento estadístico<br>•Puede cansar en cuestionarios largos<br>•Requiere experiencia o conocimeintos para poder jerarquizar<br>•Puede ser difícil de aplicar cuando la lista a ordenar es amplia |
| **Escala intervalo** | |
| Ventajas | Desventajas |
| •Fácil de redactar<br>•Riesgo bajo de introducior sesgos en su redacción<br>•Amplias posibilidades de tratamiento cuantitativo | •Reservadas a personas conocimientos elevados de la materia que investiga<br>•Requiere un nivel cultural medio o alto<br>•Obliga a reflexionar de forma contínua y profunda<br>•Tiende a fatigar |
| **Escala métrica continua** | |
| Ventajas | Desventajas |
| •Fácil de redactar<br>•Máximas posibilidades de tratamiento cuantitativo | •Reservadas a personas conocimientos elevados de la materia que investiga<br>•Requiere un nivel cultural medio o alto<br>•Obliga a reflexionar de forma contínua y profunda<br>•Requiere de una redacción muy precisa<br>•Puede haber dificultades de comprensión<br>•Tiende a fatigar |
| **Escala suma constante** | |
| Ventajas | Desventajas |
| •Muy fácil de redactar | •Fatiga por aburrimiento<br>•Su comprensión puede ser difícil<br>•Queda reservada a cuestionarios cortos<br>•Sólo responden bien las personas capaces de calcular ágilmente<br>•Reservada a expertos en la temática que se investiga<br>•La tabulción y el análisis dedatos pueden parecer complejos |
| **Escala de asociación** | |
| Ventajas | Desventajas |
| •Muy fácil de redactar | Tratamiento cuantitativo limitado , |

| | |
|---|---|
| •Muy fácil de comprender<br>•No cansa<br>•Las respuestas son muy fiables<br>•No requiere conocimientosespeciales de los encuestados<br>•Se analizan con análisis factorial de correspondencias | contrastes de Chi-cuadrada, CFA |
| Escala de Likert | |
| Ventajas | Desventajas |
| •Muy fácil de redactar<br>•Muy fácil de comprender<br>•Riesgo bajo de sesgos en su redacción<br>•No requiere conocimientos especiales de los encuestados<br>•No fatiga<br>•Amplias posibilidades de tratamiento cuantitativo | •Peligro de desequilibrar la escala<br>•Peligro de forzar la escala<br>•Peligro de decantaciónhacia posiciones intermedias por parte de los encuestados<br>•Tratamiento cuantitatoivo limitado<br>Peligro de incluir pocas/demasiadas categorías<br>•Puede haber dificultades para depurar la escala |
| Escala de diferencial semántico | |
| Ventajas | Desventajas |
| •Amplias posibilidades de tratamiento cuantitativo<br>•Fácil de comprender | •Pocas pistas verbales<br>•Redacción muy escueta<br>•Algunos de los extremos puede ser ambiguo o discutible<br>•Es difícil de radactar sin introducir sesgos |

Fuente: Grande-Esteban y Abascal-Fernández (2017) y Hernández-Sampieri et al., (2010) con adaptación propia

# Confiabilidad de las escalas

Una escala es un instrumento de de medida, como lo es un termómetro o una cinta métrica. Los instrimentos de medida tienen que cumplir dos requisitos: ser *confiables y válidos. Son confiables cuando siempre miden lo mismo. Son válidos cuando además miden, lo que tienen que medir.* Todos los instrumentos de medida cometen *errores,* que pueden ser de dos clases: *aleatorios y sistemáticos.* Ver **Tabla A.26.**

**Tabla A.26. Clases de error y sus fuentes**

| Error aleatorio |
|---|
| *Los cosntituyen todos aquellos factores que no se pueden controlar* y confunden la medición de un fenómeno. Por ejemplo, el velocímetro de un auto puede tener un error variable e impredecible de hasta un +_ 30%. Esto significa que la velocidad real del auto cuando llegue a los 100 Km/Hr en su aguja podrá tomar un valor comprendido entre 70 y 130 Km/Hr. *El error es variable en cantidad y sentido.* Hace |

que no sea posible saber a qué velocidad real se circula, con peligro de ser denunciado por los radares que controlan la velocidad.
En las investigaciones de ciencias económico-administrativas, los errores aleatorios proceden de las personas que responden o preguntan, o de las situaciones que rodean la recogida de información. *Este error no es constante.* Afecta a las mediciones en forma variable cada vez que se hacen. Todos los instrumentos de medición basados en escalas, generan errores aleatorios en mayor o menor grado. *Cuanto mayor sea el error aleatorio, menor sera la fiabilidad de un instrumento.* Lo deseable es diseñar instrimentos de medición, escalas en nuestro caso, que lo miniminicen. Cuando el error aleatorio es pequeño , por ejemplo el velocímetro lo tiene entre +- 1%, la estimación de la medición se puede considerar precisa.
Así, se tiene en resumen como fuentes de error aleatorio:

- Errores de codificación
- Errores detabulación
- Característcas delas personas que responden, como inteligencia o cultura, salud o estado emocional
- Factores coyunturales, como ruido, distracción, presencia de otras personas, cansancio
- Características de los encuestadores o entrevistadores muy dispares, como simpatía, empatía o *don de gentes*
- Énfasis de las palabras de los encuestadores o entrevistadores

**Error sistemático**

Lo constituyen aquellas fuentes que afectan a las mediciones de forma constante y suponen *factores estables que afectan de la misma forma e intensidad* cada vez que se hacen las mediciones. Por ejemplo, un velocímetro podría tner la aguja torcidad y marcar siempre diez Km/Hr en exceso. Entonces a cualquier velocidad se sabe que la real será 10 Km/Hr menos de los marcados. En tales caos, un conductor sabrìua exactamente a que velocidad circula en todo momento. Le bastaría restar 10 Km/Hr. a lo que marca el velocímetro.
Así, se tiene en resumen como fuentes de error sistemático:

- Ítems mal seleccionados
- Ausencia de claridad de la escala

Errores físicos en la escala, como mala calidad de impresión de la misma, exceso de ítems, diseño poco claro, etc.

- Errores de redacción de los ítems de la escala

Fuente: Grande-Esteban y Abascal-Fernández (2017) y Hernández-Sampieri et al., (2010) con adaptación propia

# Medición de la confiabilidad de las escalas

Se dice que una escala es confiable cuando proporciona *resultados consistentes* cada vez que se realizan ediciones con ella, cuando siempre mide lo mismo. Como las fuentes de errorsistemático afectan siempre de la misma foirma y con la misma intensidad a las mediciones, *la confiabilidad de una escala está condicionada por el error aleatorio.*

Si una escala estuviera *libre de error aleatorio*, su confiabilidad tenderñia a ser perfecta. *Una escala confiable NO significa que sea válida.* Imagine una cinta métrica que fuera *indilatable* de 30 cm. Le llamaramos *metro* , nos proporcionará mediciones confiables. Siempre mediría lo mismo, peroo NO seria válida, porque el concepto de metro es que tiene 100 cm. Si ese metro de 30 cm fuera de goma elástica , nuestras mediciones NO serían NI confiables NI válidas. La longitud dela regla dependerìa de la tensión que se aplicara a los extremos, que podría tener la consideración de aleatoria en función de la persona que la utilizara, su fuerza, su envergadura. Podríamos medir metros de 50 cm., de 60 cm., de 80 cm. Según la fuerza que se aplicara. Esta regla no sería ni confiable ni vñalida. Cada medición sería distinta y ningina de un metro de longitud.

Para medir la confiabilidad de una escala, se pueden emplear varias técnicas. Vea la **Tabla A.27**

**Tabla A.27. Técnicas para medir la confiabilidad de una escala**

| Test-retest |
|---|
| Consiste en suministrar la misma escala a la misma muestra en condiciones similares para que se responda. Para medir la fiabilidad se calcula el *coeficiente de correlación de Pearson* entre las respuestas dadas por la misma muestra de personas a una misma escala en dos momentos separados en el tlempo. Aún así, *no se considera una prueba adecuada para medir la confiabilidad* , debido a:<br>• Las pruebas se repiten pasado un tiempo, produciendo una *maduración* y también puede haber *efectos principales e interactivos*, como sucede con los experimentos. Puede que las personas tengan mayres conocimienntos, experiencia interés o preocupación por el hecho que se investiga.<br>• Las *condiciones del medio* pueden cambiar mucho y las condiciones de la obtención de información en distintos momentos puede que no sean comparables.<br>• *Es posible que nose pueda repetir e test*. Sería el caso de la medición de actitudes ante un producto nuevo, una vez comercializado.<br>• En realidad, mide más la estabilidad que la confiabilidad |
| **Formas alternativas** |
| Para medir la coniabilidad mediante formas alternativas, se procede a diseñar dos escalas equivalentes que se proporcionan para ser respondidas en dos momentos diferentes del tiempo. Además de las limitaciones del test-retest, puede resultar complicado preparar dos escala a la vez. De hecho, ya es bastante complicado la primera. |
| **Método de las mitades** |
| El método de las dos mitades es otra forma de confiabilidad de una escala. Consiste en calcular la correlación entre dos mitades de un conjunto de ítems. Por ejemplo, se distribuye una escala y se calcula la corrrelación entre los ítems pares y los impares, o entre la mitad de muestra y la otra mitad. El coeficiente de *Spearman-Brown* permite calcular la confiabilidad de la escala. Toma valores comprendidos |

entre cero y uno. Su expresión simple es:

$$p_{sp}=2p/(1+p)$$

donde:

***p.*** Es el coeficiente de correlación de Pearson entre las dos mitades consideradas
Se observa que si ***p=0*** , el coeficiente de fiabilidad de Spearman-Brown valdrá cero.

Cuando la correlación se perfecta, ***p***=1, el coeficiente de fiabilidad $p_{sp}$ será máxima y valdrá uno. Este método tiene la ventaja de estimar la confiabilidad en un momento del tiempo. No necesita de dos o más muestras a lo largo del tiempo, como con los métodos antes expuestos.Por ejemplo, una prueba con 100 preguntasp roporciona una correlación de 0.84 entre las preguntas pares e impares. El coeficiente de confiabilidad del test, es:

$$p_{sp}= 2* 0.84 / (1 + 0.84) = 0.91$$

La fórmula de Spearman-Brown, se le denomina *profética*, prorque sirve para predecir coeficientes de confiabilidad según su expresión general, que es:

$$p_{sp} = K\ p/ [1 + (K +1)\ p]$$

Donde:

**K.** Es el número de veces que la longitud del test es mayor o menor que el test original. Si una escala tiene 180 preguntas y un coeficiente de confiabilidad de 0.92, si se reduce a 90, la confiabilidad de la escala será **(K=90/180=0.5)**

$$p_{sp}= K\ p /[1 + (K-1)\ p] = 0.50*0.92/[1+(0.5-1)*0.92]=0.852$$

Si una escala tiene 15 preguntas y su confiabilidad es de 0.65, si tuviera 75 su confiabilidad sería:

**(K= 75/15=5)**

$$p_{sp}= K\ p /[1 + (K-1)\ p]= 5*0.65/[1 + (5-1)*0.65]= 0.902$$

**Ver Apéndice 1.**

## Consistencia interna o coeficiente de alfa

Esta forma de medir la confiabilidad de una escala, es el de medir su *consistencia interna,* a través del instrumento denominado *coeficiente alfa de Cronbach* . Si se parte de la matriz de correlaciones entre ítems, el *coeficiente alfa de Cronbach*se estima a partir de la correlación media entre los ítems:

$$\text{Alfa}= N\ p / [1+ p\ (N-1)]$$

Donde*:*

***p.*** Es el coeficiente de correlación media entre ítems

**N.** Es el número de ítems
El *coefiiente alfa de Cronbach* es una prueba muy utilizada para pronunciarse sobre la confiabilidad de una escala. Mide en un momento del tiempo y sin necesidad de hacer repeticiones, la correlación esperada entre la escala actual y otra forma alternativa. Sin embargo, conviene saber adicionalmente que:

- **El coeficiente se eleva a medida ue se eleva el número de ítems.** Esto significa que, en principio, cuantos más íteks se incluyan, mayor será la consistencia interna y la confiabilidad de la escala. Pero todo tiene un límite.
- **La confiabilidad no crece proporcionalmente con el número de ítems.** Además, cuantos más ítems se tengan en cuenta, más cuesta construir la escala, recoger la información, tratarla, anaizarla, etc.

Aunque no existe un acuerdo sobre cuánto debe valer, el *coeficiente alfa de Cronbach* debe tomar un valor de **0.70** como mínimo, para poder concljuir que existe consistencia interna o que la escala multi-ítem es confiable. (Este concepto es analizado a detalle más adelante)

**El coeficiente de theta**

Este coeficiente mide lo mismo que el de alfga, sólo que en la expresión se cambia la correlación media por el primer valor propio obtenido cuando se aplica la técnica de *análisis de componentes principales.* El valor propio es un indicador de la cantidad de información que contiene una combinación linel de variables

$$\textbf{Theta= (N / N -1) (1-1/ } F_1 \textbf{ )}$$

Donde:

El máximo valor que puede obtener $F_1$ **es N.** Se deduce que cuando el primer factor recoge todos los ítemns, la consistencia interna medida por Theta es 1.

Fuente: Grande-Esteban y Abascal-Fernández (2017) con adaptación propia

# Influencia del diseño de una escala sobre la confiabilidad

Como se muestra en la **Tabla A.28** existe una relación tal que la inluencia de una escala sobre la confiabilidad *depende de la longitud de la escala así como del n´mero de categorías.*

**Tabla A.28. Influencia del diseño de la escala sobre la confiabildad**

| Diseño de la escala | Efecto sobre la confiabilidad |
|---|---|
| Número de ítems | • Aumenta la confiabilidad |
| Dificultad de los ítems | • No existen evidencias en uno u otro sentido |
| Tipo de escala | • No existe a priori, superioridad de una escala multi-itemm sobre otra |
| Número de categorías | • A medida que aumenta, se incrementa la confiabilidad |
| Formato de la escala | • No existen evidencias que los colores, gráficos, etc. afecten la confiabilidad |
| Escalas con categorías neutras (indiferente) | • No existen evidencias sobre su influencia en la confiabilidad |

Fuente. Churchill (1991) con adaptación propia

# Validez de las escalas. Primera aproximación

Una escala es válida cuando mide lo quie tiene que medir. Los voltímetros miden tensiones eléctricas, los amperímetros miden corrientes eléctricas, los termómetros miden miden temperaturas, los manómetros presiones. Una escla para medir innovación y una escala para medir riesgo cada una deben de medir para lo que se les diseñó. *Una escala será tanto más válida cuanto menor sea el error sistematico.* Ver **Tabla A.29.**

**Tabla A.29. Tipos de validez**

| Validez pragmática |
|---|
| Una escala es valida desde el punto de vista pragmático cuando sirve para hacer *predicciones* actuales y futuras. Se habla de *validez concurrente y predictiva* (analizada a detalle más adelante). Un test de inteligencia permite hacer predicciones sobre la capacidad de intelectual de una persona del presente. Es así que es un instrumento que tiene *validez concurrente.* Un test de actitudes ante el trabajo, sirve para hacer predicciones sobre el éxito que pueda alcanzar en su vida. Este tiene *validez predictiva.* |
| **Validez de contenido** |
| *Considera si los ítems de la escala son adecuados para hacer las mediciones.* Por ejemplo, un test de inteligencia basado en la resoluciòn de cálculo integral, no es adecuado por su contenido para medir la inteligencia de un niño. Una persona con poca información matemática puntuará peor en un test de operaciones aritméticas que otra con conocimientos avanzados de matemáticas. El contenido de una escala se determina buscando ítems adecuados para hacerlas mediciones, por ejemplo, mediante revisiones de literatura , en otras escalas similares o usando el sentido común. La escala inicialmente propuesta, será sometido a *juicio de expertos* para su filtraje. Posteriormente, se puede afinar más a base de calcular la consistencia interna quitando y añadiendo ítems. (Analizada a detalle más adelante). |
| **Validez de concepto** |
| Este tipo de validez guarda relación con lo que está midiendo realmente la escala. Se dice que una escala posee *validez convergente* cuando mide lo mismo que escalas que miden el mismo concepto. Una escala posee *validez discriminante* cuando sus mediciones no guardan ninguna relación con escalas miden conceptos distintos. Por ejemplo, una escala elaborada para medir fidelidad a la marca debería proporcionar correlaciones elevadas con otras escalas alternativas que tambien miden fidelidad. Las correlaciones con otras escalas que no miden fidelidad, por ejemplo, etnocentrismo o riesgo, deberían ser nulas o no significativas. (analizada a detalle más adelante |
| **Validez nomológica** |
| Se dice que una escal posee este tipo de validez cuando los resultados y conclusionesa los que se llega tras su aplicación, son coherentes con las teorías. (Analizada a detalle más adelante) |

Fuente: Grande-Esteban y Abascal-Fernández (2017) y Hernández-Sampieri et al., (2010) con adaptación propia

# Criterios de valoración de las escalas

La **Tabla A.30.** resume los criterios a fin de apoyar a los investigadores en formación ya que es frecunte que acudan a examinar estudios previos para elaborar sus escalas. En dicha tabla, se muestran los aspectos para considerar antes deutilizar una escala, como:

- La coherencia de la escla, avalada por fundamentos teóricos o validaciones empíricas
- Pruebas piloto desarrollados pro los investigadires para definir los ítems
- Pruebas estadísticas realizadas con la muestra
- Tipo de muestreo adoptado
- Correlación media entre los ítems, estimada a aprtir de la matriz de correlaciones
- Valor mínimo del *coeficiente alfa de Cronbach*
- Número de factores encontrados en caso de emplear análisis de componentes principales para validar la escala
- Valor mínimo de la correlación entre ítems y tiempo mínimo transcurrido en caso de aplicar test-retest
- Influencia del entorno de las mediciones
- Validez convergente
- Validez discriminante

Según estos parámetros, las escalas pueden ser: *excelentes, buenas regulares, mínimas o sin calificar.*

**Tabla A.30. Criterios para la valorización de escalas**

| Criterio de evaluación | Calificada como: | | | | |
|---|---|---|---|---|---|
| | Excelente | Buena | Regular | Mínima | Sin calificar |
| Desarrollos teóricos o empíricos sobre la escala | Refleja investigaciones de calidad en ese campo de conocimiento y validaciones empiricas | Presenta revisiones de literatura o validaciones empíricas | Revisa mas de una fuente | Revisa una fuente | No se basa en nada. La escala es *ad hoc* |
| Estudios piloto para definir los ítems | Más de 250 ítems inicialmente acompañados de varios estudios piloto para depurarlos | De 100 a 250 ítems y más de dos pruebas piloto | Entre 50 y 100 ítems y más de dos pruebas piloto | Se eleiminan algunos ítems. Se realiza una prueba piloto | No se eliminan ítems. No se realiza prueba piloto |
| Estadística descriptiva, contrastes de hipótesis | Máxima para la muestra y submuestras | Máxima para la muestra y parcial para algunas submuestras | Máxima para algunas submuestras | Sólo para la muestra | No se realiza |
| Muestreo | Aleatorio | Aleatorio | Aleatorio/ Conveniencia | Varias muestras de conveniencia | Una muestra de conveniencia |
| Correlación media entre | Mínimo **0.30** | Entre **0.2-0.30** | Entre **0.10-0.19** | Por debajo de **0.10** | No se calcula |

| | | | | | |
|---|---|---|---|---|---|
| ítems | | | | | |
| Coeficiente alfa de Cronbach | **0.80** o más | Entre **0.7** y **0.79** | Entre **0.6** y **0.69** | Menor de **0.6** | No se calcula |
| ACP | Un solo factor | Un solo factor | Un solo factor | Varios ítems en un solo factor | No se calcula |
| Test-retest | Correlación mayor que **0.5** al menos un año después | Correlación mayr que **0.4** entre 3 y 12 meses después | Correlación mayor que **0.3** entre 1 y 3 meses después | Correlación mayor que **0.2** eal menos 1 mes después | No se calcula |
| Influencias del entorno | Tres o más estidios que demuesten independencia | Dos estudios que demuestren independencia | Un estudio que demuestre independencia | Algunois estudios muestran independencia, pero otros no | No se aportan pruebas |
| Validez discriminante | Significativamente diferente de cuatro o más medidas no afines | Significativamente diferente de dos o tres medidas no afines | Significativamente diferente de una medida no afin | Diferente de una medida correlacionada | No se calcula |
| Validez convergente | Correlaciones significativamente elevadoas con dos o más medidas afines | Correlaciones significativamente elevadoas con dos o más medidas afines | Correlaciones significativamente elevadoas con dos o más medidas afines | Correlaciones significativamente elevadoas con una medida afin | No se calcula |

Fuente: Baerden et al. (1993) y Robinson, et al. (1991) con adaptación propia

# Fuentes aleatorias de información

Recibe este nombre porque la información que proporcionan se puede *generalizar,* debido a que se obtiene a partir de datos proporcionados por muestras aleatorias , amplias y representativas, que son analizadas mediante la aplicación de técnicas estadísticas y matemáticas. Ver **Tabla A.31.**

**Tabla A.31. Fuentes aleatorias de información y sus características**

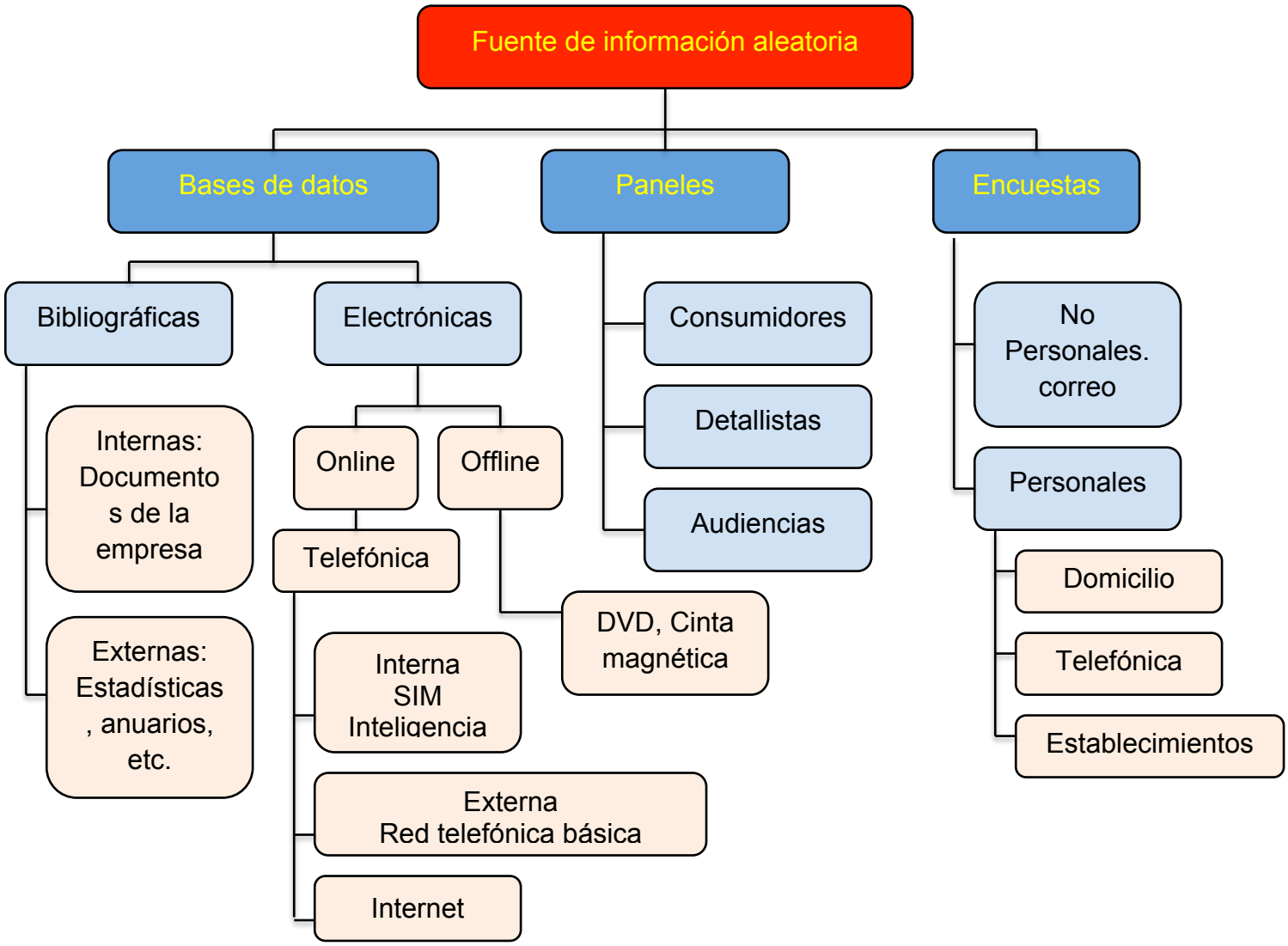
| Secundarias o primarias |
|---|
| Algunas de las fuentes de información aleatorias pueden considerarse *secundarias*porque ua existen en el momento de ser necesaria la información que contienen. *Esto ocurre con las bases de datos o con los paneles*. Sin embargo, las *encuestas* son por naturaleza *primarias*, porque se difunden en el momento en que se hace necesaria la obtención del a información que proporcionarán. |
| **Estáticas o dinámicas** |
| Una base de datos contiene información que se actualiza con determinada frecuencia. Si una empresa crea su propia base de datos, esta fuente de información es *dinámica.* Si la base se compra a un tercero sin acordar de la actualización de la información, la base de datos o el panel *son fuentes de datos estáticas.* Las encuestas son, salvo excepciones, *estáticas*. Los *códigos deontológicos* prohíben encuestar a la misma muestra en dos momentos distanciados en el tiempo salvo que dé su consentimiento la primera vez que se entra en contacto con ella. |
| **Personales o impersonales** |
| Ciertas fuentes secundarias en algún momento requieren contactos con personas para crear la información , o necesitan de la presencia o de la colaboración de personas para poder ser |

| |
|---|
| aplicadas. Esto ocurre con los *paneles,* en general y con la mayoría de las encuestas. Otras veces no es necesario el contacto personal, como ocurre con las encuestas por correo. |
| Para investigaciones exploratorias y descriptivas |
| Las fuentes cuantitatvas puede servir para realizar investIgaciones *exploratorias o descriptivas.* Una *base de datos* o los resultados del análisis de información que contiene un panel de consumidores se pueden comprar para hacer un estudio exploratorio. Sin embargo, lo más frecuente es que estas fuentes de infromación sirvan para desarrollar i*nvestigaciones descriptivas,* especialmente las *descriptivas,* que pueden servir para medir los resultados de decisiones tomadas. , medir imagen posicionamientos, segmentar el mercado, etc. Una *encuesta* nunca se utiliza para desarrollar investigaciones exploratorias . Par ello están las fuentes cualitativas. |

Fuente: Grande-Esteban y Abascal-Fernández (2017) y Hernández-Sampieri et al., (2010) con adaptación propia

Ver **Figura A.9**

**Figura A.9. Fuente de información aleatoria**




Fuente: Grande-Esteban y Abascal-Fernández (2017) con adaptación propia

Así, tenemos la siguiente descripción de las fuentes de información aleatorias más relevantes y utilizadas. Ver **Tabla A.32.**

**Tabla A.32. Descripción de las fuentes de información aleatoria**

| Bases de datos |
|---|
| Es una colección de informaciones relacionadas y accesibles conforme algún criterio de selección. Pueden ser de distinta naturaleza basada en : *quién las crea y qué soporte se utiliza..*<br>• Desde el primer punto, pueden ser *externas o internas* Las *internas* las crean los propios usuarios, como ocurre en una empresa que construya un archivo de clientes. Las *externas* las desarrollan las organizaciones para su uso privado, o el de terceros , o ambas cosas a la vez. Por ejemplo *Bloomberg* tiene bases de datos de información de empresas que puede ser compradapor terceros para ahacer *consultas*, El instituto Nacional de Estadística y Geografia de México (INEGI) que recoge datos de utilidad para elaborar sus publicaciones de interés general, y también los proporciona a los investigadores a título gratuito u oneroso.<br>•Desde el punto de vitsa de soporte, las bases de datos pueden ser *bibliográficas o electrónicas.* Las primeras, se basan en el papel y son publicaciones convencionales en forma de libros, caidas ya en desuso. Las *electrónicas* tienen soporte magnético o de descargra ddirecta de la nube. Se dice también que son *offline, o sin conexión directa con el proveedor.* Existen otras basadas en la conexión telefónica convencional y reciben el nombre de bases de datos *online.* Para ello, se habrá registrado como usuario, pagando una cantidad de dinero en concepto de alta y también pagará por cada consulta realizada o por tiempo de conexión. Internet actualmente ofrece el acceso a información con características de uso muy cómodas para quien investiga. Finalmente, debe decirse que los criterios de clasificación expuestos se pueden combinar. Así, existen bases de datos *internas* o *externas*, que a su vez pueden ser *bibliográficas o electrónicas.*<br><br>**Motivos para crear bases de datos**<br>•Existe mucha información importante para las empresas que no aparecen en ninguna fuente secundaria, coo por ejemplo, las relativas a sus *relaciones comerciales.* Las fuentes secundarias o son *rígidas*, es decir, no ser diseñadas a la medida de las necesidades de las empresas, o son *estáticas,* pues proporcionan datos referidos a un momento de tiempo determinado como las estadísticas. Otas veces, las *fuentes primarias* son instrumentos adecuados para obtener información, porque sea muy compleja o abundante. Por este motivo, muchas empresas crean sus propias bases de datos inernas, que se construyen a medida de las necesidades que surgen.<br>•Los intereses de las organizaciones por crear su información y acceso a la misma, se basan en esencia a: estructura compleja (niveles jerárquicos, procesos burocráticos, etc.), las fuentes diversas (consumidores, proveedores, distribuidosres, etc), las combinaciones dinámicas de producto/servicio que ofrecen (de manera muy rápida), la tendencia a desvirtuar la información o falsearla y/o que las fuentes de información se tengan contenidos en diveros medios desde escritos (libros contables, |

directorios) hasta electrónicos (CD, la nube).

**Usos potenciales**

•Son las bases de datos internas, el cimiento para: la predicción de consumo, facilita la comunicación o promoción,mide la difusión de productos, conocer el perfil de sus consumidores, examina tendencias, identificar determinantes demográficos de la demanda (sexo, edad, ocupación ,etc.), identificar a los clientes, priveedores, distribuidores principales, descubrir nuevos nichos, descubrir hábitos de compra, origen geográfico, fidelidad a la marca, preferencias, contro de existencias, minimizar riesgo de marcado erróneo, separación de infirmaciuón regional, analizar el efecto del marketing mix por cada uno de sus componentes, etc. Aún así, aspectos importantes como las motivaciones de los consumidores aún no se pueden analizar. Sin embargo, esto ha cimentado las bases para la práctica del *geomarketing,* que se define como: *la investigación de mercados y atención a los consumidores que emplea bases de datos que contienen información relativa a sus características demográficas y socioeconómicas vinculadas a una localización geográfica.* Las bases de datos así diseñadas, son fáciles de obtener, manipular, portar , proyectar , facilitan las tareas de segmentación.

**Diseño**

•Las bases de datos se componen de *registros*, formados por *campos*. Por *campo* de una base de datos se entiende un criterio, característica de la información o cualidad para acceder a su contenido. *Registro* es un conjunt de campos mutuamente relacionados con la información que se maneja. El diseño, debe permitir *añadir* datos ebn cualquier momento y poder editarlos, para corregirlos o eliminarlos. La base debe permitir el acceso a los datos para su consulta. Las bases de datos deben permitir clasificar la información, con el uso de operadores *and, or.* Debe ser posible extraer o exportar datos para su análisis con base a criterios.

| Puesto | Funciones generales | Finciones específicas |
|---|---|---|
| Responsable de información | •Definir la naturaleza de la información que debe contar la base<br>•Revisar periódicamente el trabajo de recogida de datos | •Valorar qué tipo de decisiones toman<br>•Decidir qué clase de información hace falta para tomar las decisiones y con qué frecuencia<br>•Buscar dónde se puede conseguir la información |
| Documentalista | •Alimentar la base de datos | •Seleccionar información<br>•Evaluar información<br>•Extraer datos relevantes<br>•Aplicar dónde se puede conseguir la información |
| Especialista en informática | •Programar la base de datos | •Diseño del sistema<br>•Posibilitar el cruce de la |

| | | |
|---|---|---|
| | •Configurar el software | información<br>•Depurar la información<br>•Actualizar la base<br>•Incorporar campos nuevos<br>•Eliminar la información no interesante o anticuada |
| Operadores | •Manejar físicamente la base | •Atender y orientar a los usuarios para realizar consultas |

**Limitaciones**

•Por parte de los empleados un instrumento de control
•Celos profesionales
•Directivos que no le tengan confianza
•La falta de cultura sobre le uso de la información
•Aspectos éticos contra el mal uso por asociaciones de información tendenciosas

**Cookies**

•Son archivos que se almacenan en los computadores de personas que navegan eninternet . Estos archivos identifican la computadora del usuario, sirven para diferenciarlos y actuar de forma diferente con ellosen función del compoirtamiento observado.

•Se idearon para crear cestas de la compra virtuales, en las que los consumidores van depositando aquello que desean comprar. Se instalan mientras se navega por internet por los sitios web que visita el consumidor o por terceros relacionados con ese sitio. De esta forma se puede conocer la actividad de los consumidores , como el lugar donde se accede, tipo de dispositivo, páginas más visitadas, productos más consultados, tiempo dedicado, productos/servicios adquiridos, número de clics, etc.

•Las cookies se identifican con sitios web. Los consumidores introducen información que también los identifica con datos personales. Las cookies prmiten que el servidor sepa quién se ha conectado, quién está autorizado para pernitir el acceso a lo que ofrece la tienda o plataforma virtual.

•Son extremadamente útiles para hacer seguimientos de los consumidores en sitios web y obtener estadística de uso y para crear perfiles de consumidores, anónimos, y ofrecerles productos que podrían ser de su interés y también para diseñar campañas de comunicación adecuadas a los perfiles de consumidores encontrados.

•Contrariamente a lo que algunas personas creen, las *cookies* no son virus, gusanos, troyanos, ni generan correos no deseados. Son solamente datos, sin capacidad de alterar una computadora.No pueden ni borrar ni leer la información almacenada por un equipo informático. Crean perfiles de usuario que son anónimos y además se pueden borrar, en cualquier caso el consumidor tiene libertad para que estos archivos que lo identifican puedan ser creados o no.

- El proceso de creación es sencillo. El navegador pregunta al consumidor si las acepta o no, o cuales aceptará. Se puede se selectivo con los sitios web , rechazar las *cookies* de terceros, aceptar las que sen temporales que se borran al apagar el equipo, Los navegadores incorporan *gestores de cookies.*
- El titular de las *cookies* recuerda a los usuarios que su uso está sujetoa su aprobación durante la instalación o actualización del navegador. La aceptación puede ser revocada. Los titulares ofrecen información sobre la política de *cookies.*

Desde una perspictiva ética hay que tener presente que la informacion de las *cookies* es totalmente anónima. En cualquier caso, existe legislación al respecto, que varía según los países.

**Clases de cookies**

- Se tienen de varios tipos, pro ejemplo, las creadas por el *propietario DEL SITIO* (como Amazon), las que son *permanente,* las de tipo *variable* y las que se borran al apagar la computadora. Las *cookies*, se envían al servidor que las definió o a otros dominios llamadas *cookies de terceros.* Estas últimas son, básicamente, de tres tipos: *sociales, mediadoras del tráfico y publicitarias.* Por ejemplo, un banco pudiera diferenciar sus *cookies de terceros:*
- *Cookies de redes sociales.* El banco podría utilizar cookies de Facebook, Twitter, Linkedin y Google para que el consumidor pueda compartir contenidos en estas redes sociales.
- *Cookies mediadoras de tráfico de portales.* Un banco podría utilizar *cookies* de Google Analytiics, Nielsen Netratings, entre otras,para recopilar datos estadísticos generados por la actividad de los clientes del bamco.
- *Cookies publicitarias.* El bnaco podría utilizar *cookies almacenadas* por terceras empresas que crean campañas publicitarias para el mismo banco.

**Utlidad de las cookies**
**Para el consumidor:**

- Recordar el usuario y contraseña en una webpara acceder con mayor rapidez
- Recordar preferencias de navegación y personalización en la web
- Recordar el carrito de compra completo antes de hacer el pago y el pedido
- Facilitar las compras futuras por ser usuario registrado , a través de formularios
- Compartir información en redes sociales

**Para las empresas:**

- Hacer seguimientos de usuarios: qué compran, cuánto ganan, con qué frecuencia visitan al sitio, gasto medio por compra, productos examinados y no comprados, tiempo medio de permanencia en el sitio, número de clics realizados.
- Obtener perfiles de consumidor y evaluar su potencial
- Segmemtar el mercado siguiendo criterios sociodemográficos y comportamentales: edad, género, localización geográfica, volumen de compras, gasto, frecuencia, repetición, fidelidad observada, etc.
- Diseñar campañas publicitarias y de promoción a medida de los segmentos identificados. Las empresas publicitarias pueden hacer seguimientos de los

consumidores en muchos sitios web e identificar perfiles útiles para segmentar el mercado y orientar con mayor precisión las campañas de comunicación.

- Ahorrar en investigación de mercados, se reducen los costes de obtención de información, el tiempo en conseguirla y en analizarla.
- Obtener estadísticas en caso de ser un *webmaster*

**Riesgos**

- *Pérdida de anonimato*. Por algun razón del consumidor podría ser identificado con datos personales: nombre, dirección postal, datos bancarios, etc.
- *Identificación inexacta.* Los consumidores pueden utilizar más de un navegador, como: Google, Firefox, Explorer, etc. más de un computador o mas de una cuebnta. Las *cookies* en este caso, identofocan varios grupos de combinaciones de equipo, navegadores y cuentas.
- *Robo o falsificación de cookies.* Edtas se enc`ían en dos sentidos , entre computador y servidor. Alguien podría interceptar información sensible que identificra con datos personales al consumidor , sus datos financieros, etc. Un atacante podríafalsificar las *cookies*, co funestas consecuencias, como cargar a otro consumidor una compra, un cargo, falsificar un precio o una dirección de compra por ejemplo.

## Panel de consumidores

Los paneles son muestras aleatorias, estratificadas, representativas de consumidores o establecimientos detallistas que proporcionan información periódicasobre determinadas variables o comportamientos. Cosntituyen *fuentes de información aleatoria, secundaria, dinámica personal y directa.* Se pueden distinguir *paneles de consumidores, paneles detallistas y paneles de audiencias.* El panel es por naturaleza un instrumento *dinámico.*

| Paneles de consumidores | Paneles de detallistas | Paneles de audiencias |
|---|---|---|
| • Estructura de gasto de los hogares | • Comportamiento de ventas y existencias de productos y marcas | • Atención a medios de comunicación |
| • Tamaño del mercado y tendencias observadas | • Importncia de los mercados y tipos de establecimientos | • Estructura de kas audiencias por edad, sexo, franjas horarias, días, etc. |
| • Característ cas de los consumidores | • Estructura de mercado de las marcas y formatos | • Audiencia de los diferentes soportes |
| • Precios de los productos | • Porcentajes de establecimientos que tienen el producto genérico y sus marcos en | • Cracterísticas de la audiencia |

| | | |
|---|---|---|
| | stock | |
| •Concurrencia de las compras | •Posicionamiento de marcas propias y de la competencia | •Evolución de anuncios |

**Metodología para la creación de los paneles de consumidores**

•Se requiere en principio definir *qué información será necesaria,* cuál será el tamaño de la muestra, donde y cómo se localizará, cómo se conseguirá su colaboración y durante cuánto tiempo.

•Una vez definido el universo o población se hace necesaria su cuantificación para determinar el tamaño de la muestra que de él se extraerá.

•Una vez identificado el universo y sus características así como el tamaño de la población, debe extraerse una *muestra aleatoria estratificada.*

•Una vez cuantificada y *perfilada la muestra, es necesario entrar en contacto con ella.*

•A los panelistas debe explicárseles la frecuencia de recogida de información es necesario elegir el soporte que se empleará para recogerla.

•Una vez constituído el panel, debe conseguirse *cooperación* y cómo se ke compensará

•Debe de informarse el horizonte temporal de la colaboración

•Mantener la representatividad a través de incentivos.,

•Los institutos que desarrollan paneles venden a sus clientes la explotación de los datos. Cuando diseñan el panel defonen el *contenido de la información.* Que servirá de base para vender sus informes Estos pueden estar estandarizados y limitarse a datos y análisis acordados de antemano ci nlos clientes o, por el contraio, *informes a medida, específicos, puntiuales* que pueden ser solicitados en un momento dado.

| Tipo de diseño del panel | |
|---|---|
| Fases de diseño | Decisiones a tomar |
| •Diseño inicial del panel | •Definir la ionformación que se desea obtener y analizar<br>•Definición del universo o población que puede proporcionar la información<br>•Determinación de la muestra necesaria<br>•Selección de la muestra<br>•Localización de la muestra<br>•Conseguir la cooperación de los panelistas<br>•Prefijar la duración del panel y de los panelistas |
| •Definir la frecuencia de a toma de información | •Dependiendo de la natiraleza de los productos y de la urgencia en disponer de la inforación para la toma de decisiones , la frecuencia de obtención de la información puede ser *diaria*, como las audiencias de TV, *semanal o mensual*, como ocurre con los alimentos , o trimestral , como sucede con la ropa |

| | |
|---|---|
| •Decidir cuál será el soporte de la información | •Físico como un diario, o electrónico a través de internet o e-mail |
| •Mantener la representatividad del panel | •Sustituir a los panelistas ppor otros de similares características cuando abandonan el panel |
| •Analizar los datos | •Determinar qué técnicas estadísticas se emplearían Decidir qué programas estadñisticos serán los más convenientes |
| •Presentar informes | •*Estándar.*Informes periódicos que se ofrecen conforme a un diseño prestablecido , acordado que resume los principales resultados en forma d etablas y/o gráficos.<br>•*Especiales.* Informes a medida de los deseos de los clientes que se presenta y exponen personalmente |

**Clases de paneles**

•Un panel de consumidores, lo constituye una muestra estable de personas que proporciojna información sobre sus actos de consumo y otras variabes de comportamiento. Ésta técnica apareció en los años 20 del siglo pasado ; en 19389 se utilizó para estudiar las reacciones de los lectores de la revista *Woman´s Home Companion;* en ese mismo año se diseño otro estudio para estudiar la conducta de compra y marcas. En los años 40, para estudiar conducta de votos,así como pruebs de producto en alimentos. A partir de los 50, se usaron para las audiencias de radio y fidelidad de marca. Es a partir d los 60, donde se aplica en múltiples estudios y necesidades.

| Clases de paneles | | |
|---|---|---|
| Por unidades muestrales | •Personas o grupos que proporcionan la información | •Paneles de hogares<br>•Paneles de individuos |
| | •Productos específicos | •Panel de automóvil<br>•Panel de bebés<br>•Panel de perfumería<br>•Panel de textil y calzado<br>•Panel de productos farmacéuticos |
| Sistema de toma de información | •Postales | •La información se envía por corre |
| | •*Home audit* | •Una persona recoge la información a domicilio |
| | •*Dustbin check* | •Los panelistas guardan en una bolsa etiquetas, pruebas de compra o envases, códigos de barras, etc., que posteriormente es recogida por la empresa que realiza las investigaciones |
| Especiales | •Pruebas de | •Los integrantes pueden ser |

| | | |
|---|---|---|
| | producto | periódicamente encuestados para realizar pruebas de productos |
| | •Con dispositivo | •La muestra de panelistas se controla mediante tarjetas especiales que vinculan la compra a las características del consumidor |
| | •Q-link | •Se contacta con panelistas para conocer patrones de compra y opiniones. |

**Paneles de consumidores a consultar sugeridos:**

•**Kantar WorldPanel (https://www.kantarworldpanel.com/mx )**

•**Nielsen:** http://www.nielsen.com/es/es/solutions/capabilities/consumer-panels.html

**Problemas**

La técnica presenta siempre *sesgos,* cuyas fuentes, son: el tamaño de la familia (difícil ubicar personas que viven solas), identificación de los respondientes (quié es el que toma las decisiones de compra), edad del ama de casa (ciertas edades son más colaborativas, como de 25-35, ya que mayores son más difíciles), quien es el propietario de la vivienda (normalmente colaboran mejor los que son propietarios que los que pagan renta),sensibilidad en el precio, renta (los extremos normalmente no colaboran igual que los de clase media), nivel cultural (a mayor educación mayor tendencia a colaborar), la naturaleza de los productos (si son o no de marca, si son o no lujosos, si son o no de diseño, etc. tienden a ser más o menos recordados) , por el tipo de compra (son más recordados los que se planean compras por su utilidad y costo que los que son productos de consumo diario o impulsivo como una golosina)

La aparición de *sesgos* tiene influencias importantes. Obsérve que el perfi de un panelista puede provenir de ser de clase media, jove, propietario de su vivienda.y otros no cumplan con esta condición. Esto significa que los resultados de los paneles deben ser ponderados , teniendo en cuenta el peso de los grupos socioeconómicos que proporcionan la información. En todo caso, se intenta que la muestra sea representativa y para ello se encentiva a la colaboración de los panelistas.

**Utilidad**

Los panelistas en general, *proporcionarán información de forma continua*, que deben registrarse en un diario.Se registran productos, marcas, precios, cantidades adquiridas, tamaños de envases y lugar de compra, precio pagado , establecimiento, etc.

| Utilidad de los paneles de consumidores | |
|---|---|
| • Estudios que se pueden realizar | • Información qu proporciona el estudio a partir de los datos del apnel |
| • Comportamiento del consumidor: los hábitos de compra | • Qué se compra en el hohar<br>• Quién compra<br>• Cuándo se compra<br>• Cuánto se compra<br>• Dónde se compra<br>• A qué precio se compra<br>• etc. |
| • Conocimiento de la ventas y de los mercados | • Ventas en el país<br>• Ventas en la región, localidad<br>• Ventas según los canales de distribución<br>• Ventas por cadenas de establecimientos<br>• etc. |

| | |
|---|---|
| •Segmentación de mercados | •Por marcas<br>•Por formatos<br>•Por variedades<br>•Por sabores<br>•Por envases<br>•etc. |
| •Características de los consumidores | •Edad del ama de casa<br>•Clase social<br>•Tamaño de la familia<br>•Presencia de niños<br>•Clase social<br>•Profesión<br>•Poder adquisitivo<br>•Hábitat<br>•etc. |
| •Tamaño del mercado tendencias de los productos | •Evolución de ventas por productos<br>•Evolución de ventas por mercados<br>•Difusión de productos nuevos<br>•Transferencias entre marcas<br>•Predicción de éxito de productos nuevos<br>•etc. |
| •Efectos del as políticas de marketing | •Efectos de las alteraciones de los precios<br>•Efectos de las modificaciones de envases<br>•Fidelidad a las marcas<br>•Eficacia publicitaria<br>•etc. |
| •Comportamiento del consumidor con relación al estableciiento | •Fidelidad al establecimiento<br>•Número de establecimientos donde compra<br>•Frecuencia de compra en el establecimiento<br>•Importe de la compra por establecimiento<br>•etc. |

**Grupos testigo**

•*Es posible que las empresas no busquen información permanente y siempre idéntica.* En estas situiaciones se emplean este tipo de grupos. un *grupo testigo* es una muestra de consumidores qu aceptan colaborar en encuestas.

•Con frecuencia son antiguos panelistas o pueden haber sido reculutados a través de anuncios en prensa o internet. Se acude a ellos para *recoger información que no tenga carácter repetitivo.*

•Para la formación del grupo se aplican los *principios de representatividad de las*

*muestras.* Los grupos pueden servir para realizar estudios *cuantitativos o cualitativos.* Con los grupos testigo, puede suceder algo parecido a lo que ocurre con los paneles en lo relacionado a su duración: *las personas pueden abandonar.*Sin embargo, son fáciles de hacer, son fiables y su costo, aún es ventajoso frente a otras alternativas.

**Valoración**

- Se puede decir que la información recogida en diario *tiene carácter objetivo* porque no surgen los sesgos que pueden introducir los encuestadores o entrevistadores.
- Los paneles se degradan con el paso del tiempo. Los panelistas abandonan , porque se aburren o porque cambian sus condiciones.
- En general, sólo proporciona *información cuantitativa.* No permiten conocer creencias, opiniones, preferencias, motivaciones, actitudes.
- Pueden resultar caros, motivo por el cual, las organizaciones como institutos o centros de investigación los llevan a cabo.
- Los paneles *se degradan* con el paso de tiempo. Los panelistas abandonan por que se aburren o porque cambian sus condiciones.
- Pueden proporcionar *información sesgada*

## Paneles de detallistas

**Concepto y calidad**

Las empresas que producen bienes, los distribuyen en multitud de establecimientos comerciales. Los productos que salen de una factoría no están, necesariamente vendidos. Pueden permanecer en almacenes de mayoristas o minoristas o en establecimientos comerciales a la espera de ser vendidos. Un panel de detallistas es, en definitiva, una muestra permanente representativa de establecimientos a partir de la cual, las empresas pueden conocer la difusión de sus productos, rotaciones, valor de sus existencias, etc. Así, el panel se forma debido a:

- *Las empresas necesitan conocer* sus ventas, la rotación de sus productos, la penetración de sus marcas en los mercados y sus cuotas de participación, o los tipos de establecimientos donde se venden
- *Es necesario hacer una valoración de las existencias* iniciales y finales en cada período de referencia. Las empresas no pueden valorar las existencias que se encuentran repartidas en multitud de establecimientos o almacenes, pues desconocen su número, características y localización.

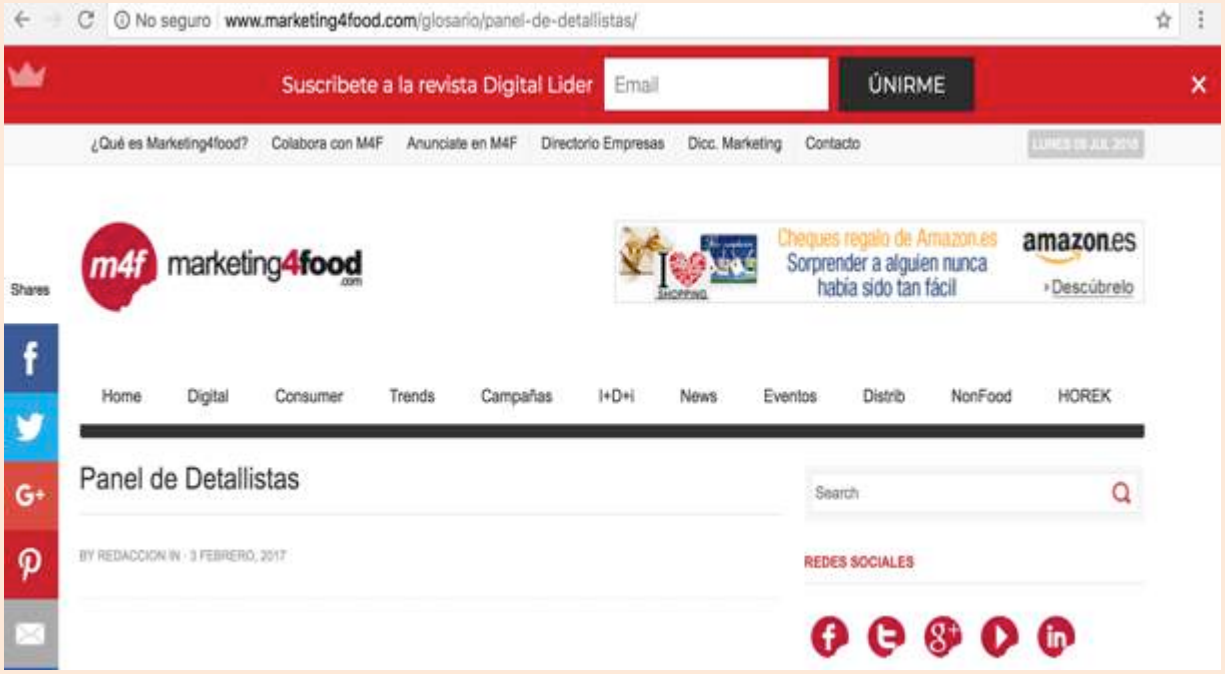
*Las empresas necesitan informacion con carácter periódico.* Si los objetivos de la información reseñados fueran esporádicos, tal vez las empresas pudieran conseguirlos por su cuenta. Sin embargo, estas necesidades de datos son muy frecuentes, su volumen es considerable y tienen que conseguirlos con mucha frecuencia.

**Principales consultoras**

**Paneles de detallista a consultar sugeridos:**

**ACNielsen:** **https://www.marketingdirecto.com/diccionario-marketing-publicidad-comunicacion-nuevas-tecnologias/panel-de-detallistas**

**Marketig4Food**
**http://www.marketing4food.com/glosario/panel-de-detallistas/**




**Operativa del panel**

•Una vez seleccionada la muestra de establecimientos, el instituto investigador se pone en contacto con ellos para solicitar su colaboración y permitirles que le dejen tomar la información que solicite (la fórmula para conocer las ventas es de contabilidad elementral):

**Ei + Compras = Ventas + Ef**

•O sea, las existencias iniciales más las compras, son iguales a las ventas más las existencias finales.

**Potencialidades**

•*Identificación del producto.* Marca, tamaño y tipo de envase. Por ejemplo: Refresco X sin cafeína en lata de 330 c.c.

•*Ventas.* Hacia el consumidor en términos monetarios absolutos y en cuota de mercado. Por ejemplo, Refresco X podría haber vendido 1'000,000 de usd en su versión Light en la zona metropolitana de la ciudad capital. Ello podría suponer un 75% de las ventas de refrescos bajos en calorías.

•*Ventas en unidades físicas* en términos absolutos y relativos. El volumen de ventas del mencionado producto fue de 40'000,000 de latas de 330 c.c.y en términos relativos, un 68.54%

•*Compras efectuadas por detallistas en volumen.* Por ejemplo, 43'567,890 latas de 330 c.c.

•*Existencias iniciales y finales*. Las iniciales fueron por ejemplo: 2'345,546. Si se consideraran las ventas , perno no las existencias finales, es:

**2'345,546+ 43'567,890= 40,000,000+Ef**

**Ef=5'913,436 unidades**

•*Duración de los stocks* suponiendo un ritmo de consumo como el actual. Si se consumen 40'000,000 de latas a mes y quedan 5'913,436 y aplicando una sencila regla de tres, pone de relieve que las existencias se agotarán en cinco días aproximadamente.

•*Porcentaje de establecimientos que ha trabajado con el producto.* Por ejemplo un 67%. Ese peso se puede expresar en porcentaje sobre ventas. Se podría decir que las ventas de Rfresco X Light en tiendas tradicionales, en usd, han supuesto un 2.34% del total.

•*Porcentaje de establecimientos que no tenían existencias* en el momento de recoger la información

•*Porcentaje de establecimientos a los que se suministraon los productos* durate el periodo de estudio

•*Stock medio del producto y ventas medias* por tipo de establecimiento y por zona

•*Precio medio unitario del producto* por tiepo de establecimiento y por zona

•*Cuota de mercado* en la zona y respecto a otros mercados incluyendo el local

•Promociones desarrolladas durante el período de estudio y su descripción, rebaja en el precio, sorteos, regalos, cantidad de producto gratis, tipo de establecimiento,etc.

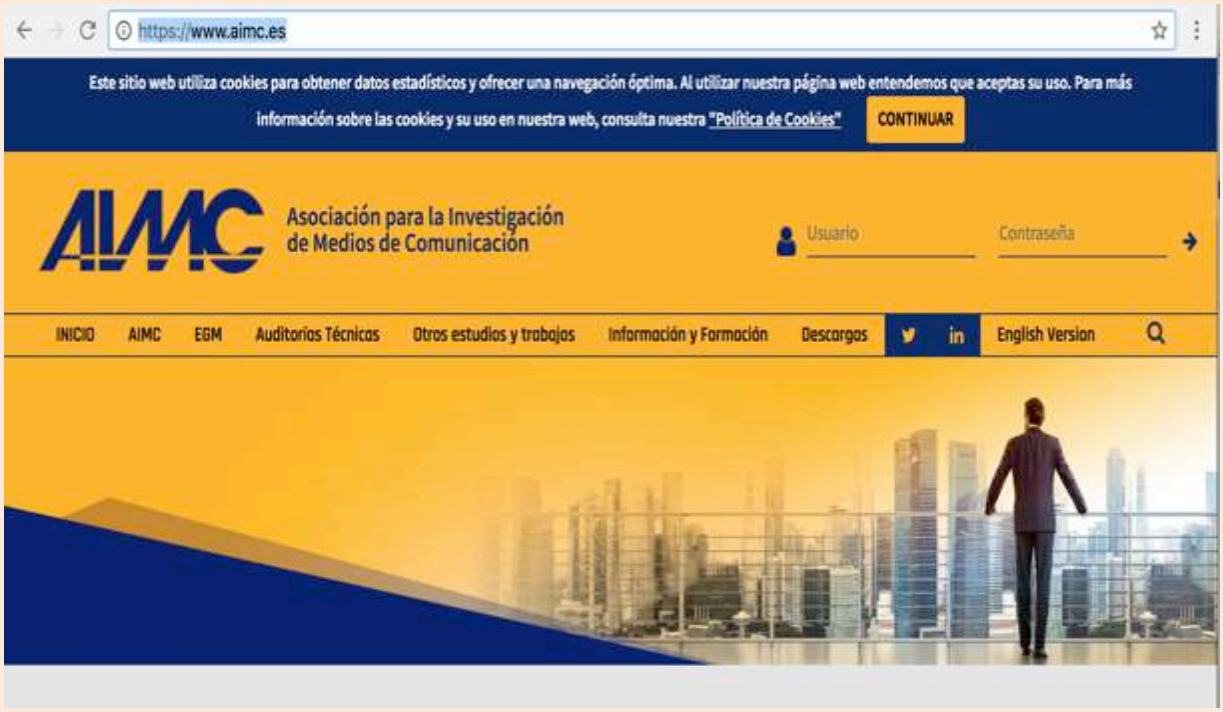
## Paneles de audiencia

•Las empresas necesitan diseñar campañas de comunicación eficaces.Esto significa que deben impactar al mayor número de personas de su población objetivo con el menor costo posible. Un fabricante de relojes de lujo considerados caros, no haría anuncios en TV, sino mñas bien, en revistas especializadas de las que consume su mercado objetivo. Normalmente, las empresas conocen sobradamente de estas

ideas, pero no suelen ser capaces de desarrollar campañas de comunicación y simultáneamente comprobar las audiencias y sus perfiles. Las agencias publicitarias necesitan conocer audiencias delos medios de difusióno los lectores de prensa. A partir de esas informacionesse pueden diseñar campañas de comunicación con muchas posibilidades de ser eficaces. Las empresas más significativas realcionadas con el estudio de audiencias, son:

- **Kantarmedia: https://www.kantarmedia.com/es**

- o la **Asociacion para la investigación de medios de comunicación: https://www.aimc.es/**



## La encuesta

**Concepto**

Es la técnica de recogida de información que consiste en la formulación de preguntas a las personas que deban responderlas sobre la base de un *cuestionario.*

**Ventajas**

•*Estandarización.*Significa que en el uso de encuestas, se harán las mismas preguntas a todos loes elementos de la muestra., lo que implica *homogeneidad* de la información. Todo el mudo contesta a las mismas preguntas

•*Faciliidad de administración, porque las encuestas están redactadas.* El encuestador se encuentra con un texto qie se limita a leer y no tiene que explicar ni interpretar nada. El encuestado, en potsal o e-maill, se limita a leer preguntas y responderlas.

•*Facilidad de tratamiento de datos*. Las encuestas poseen némeros y códigos que identifican las respuestas, es decir, se puede codificar la encuesta o convertoir a números aspectos cualitativos. Esta característica facilita grabar los datos en computadoras para poder tratarlos posteriormente con programas informáticos.

•*Obtención de información adicional.* Al encuestar a las personas en sus domicilios o establecimientos, se puede conseguir información relativa a su aspecto o hábitat.. Por ejemplo, se aprecia su nivel cultural o de inteligencia, cómo es su casa, etc.

•*Posibilidad de hacer estudios parciales.*Debido a que las encuestas poseen identificadores de las personas que responden, es posible hacer estudios imponiendo condiciones. Por ejemplo, se puede estudiar el comportameinto de consumidores de videojuegos atendiendo su edad, sexo, residencia, nivel socieoeconómico, etc. La utiización de encuestas es fundamental para: ahcer estudios de segmentación, imagen, posicionamiento, obtención de tipología de consumidores, etc.

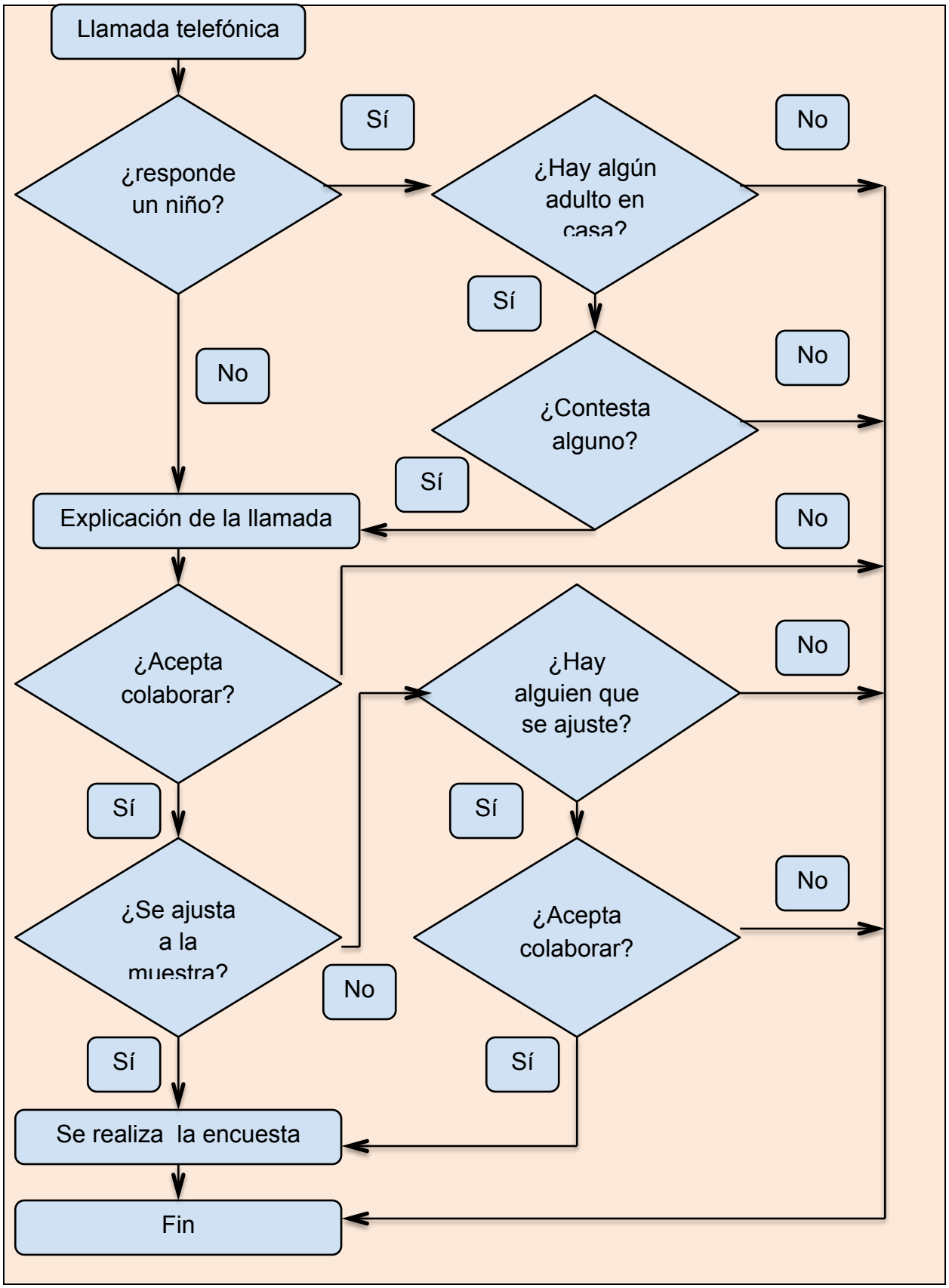
**Clases de encuestas**

•*Personales.* Consisten en un encuentro entre dos personas en el que una de ellas obtiene información proporcionada por la otra sibre la base de un cuestionario (*estructurado o no estructiurado*).El encuentro puede realizarse en diversos lugares. Las preguntas están fijadas aunque el encuestadpr puede aportr también las suyas o adaptarlas al encuestado Se tienen de varias clases, como:

•*A domicilio.* En los años 30 era muy utilizada , es costosa y ahora está en desuso

•*En establecimientos.* Cuando se obtiene información en tiendas, supermercados, grandes alñmacenes, etc. el cuestionario *suele ser estructurado.*, aunque es frecuente que se deje libertad de respuesta al encuestado, al menos en algunas preguntas.

•*Telefónica*. Cuando se realiza la obtención de información se efectúa mediante una conversación por teléfono entre el entrevistador y la persona seleccionada., sobre la base de un cuestionario, generalmente estructurado. En ocasiones, para mejorar el proceso se emplea un computador (**CATI.** *Computer Aided Telephone Interview*) que facilita las tareas de recogida, control y tratamiento de información como: analizar nuevos productos, conocer la estructura de la población, pruebas de eficacia publicitaria, actitudes post-compra, etc. Ver diagraa de proceso de llamada telefónica.


Llamada telefónica
¿responde un niño?
Sí
¿Hay algún adulto en casa?
No
Sí
¿Contesta alguno?
No
No
Sí
Explicación de la llamada
No
¿Acepta colaborar?
¿Hay alguien que se ajuste?
No
Sí
Sí
¿Se ajusta a la muestra?
¿Acepta colaborar?
No
No
Sí
Sí
Se realiza la encuesta
Fin

*Directrices de la encuesta telefónica*

•El encuestador debe dejar *muy claro de que se trata la investigación* y que no se trata de vender nadaen ese momento ni en el futuro

•Debe grantizarse *el anonimato* a los respondientes

•Debe infoirmarse a los respondientes que pueden *verificar la garantía de la empresa investigadora*

•Debe explicarse clara y convincentemente de *por qué se está llamando a la persona* Las llamadas se deberán hacer en *horas razonables.*. Generalmente , entre las 14 y las 16 o entre las 20 y 22 horas. Fuera de este horario, las personas pueden estar ausentes. Despues de las 22, están descansando.

•Finalmente, existen encuestas personales llamadas *omnibus* que consiste en una variedad de encuesta realizada normalmente por un instituto, que la realiza periódicamente con un cuestionario multitemático.

•Cabe señalar, que un problema de primer orden de las grandes encuestas, es el costo. Muchas empresas no pueden soportarlo. Además, a la gran mayoría no les hace falta suinistrar grandes cuestionarios con muchas preguntas a recoger la información . Les basta con fprmular dos otres preguntas para lograr su cometido. Es en esta circunstancia que una encuesta ómnibus apoyada por otras, les es de utilidad ya que comparten costosde levantamiento. Puede servir para realizar varias investigaciones simultáneas siempre que la muestra responda al contenido diverso del cuestionarioy que sus dimensiones sean suficientes. Una encueesta omnibus puede servir para conocer aspectos, como:

•Características básicas de los ercados , como el grado de consumo y compra de productos y marcas , el cpnocimiento de las marcas, y hñabitos de consumo.

•Las características de colectivos o ciertos grupos de consumidores, según sus condiciones socioeconómicas, o sociodemográficas o ambas.

•Actividades culturales, deportivas, de trabajo y uso de tiempo para entretenimiento

•Conocer opiniones sobre temas diversos como: educación, política, sociedad, etc.

**Levantando una encuesta omnibus**

•Es necesario definir la población de la que desea obtener información, que puede ser de tres clases: *personas de 14 años en adelante, hogares o cabezas de familia.*

•Definir el ámbito geográfico en el que se obtendrá la información

•Se debe calcular, a continuacón, el tamaño de la mues, mediante los procesos sugeridos en la creación de paneles.

•La muestra debe ser *aleatoria y estratificada* y su tamaño oscila entre 300 y 800 muestras.

•Se deben practicar criterios de estratificación socioeconómica y geográfica. Se puede partir de las regiones de Nielsen

•La muestra se localiza fisicamente a través de rutas aleatorias en cada lugar de muestra respetando las cuotas establecidas.

•Las encuestas no personales se hacen por correo . Las encuestas postales convencionales, se llaman así porque la recogida d einformación se efectúa mediante un cuestionario que se envía a una persona , que lo cumplimenta y lo devuelve por

correo.

•Junto a la encuesta se debe anexar una carta de preesentación, las insttrucciones para llenarla y un sobre pagado para devolver el cuestionario.

Es frecuente recompensar al encuestado , generalmente en forma de obsequio cuando devuelva la encuesta.

**Criterios para selección de una encuesta adecuada**

Una duda que asalta al investigador es decidir qué tipo de encuesta conviene utilizar. Es necesario basarse en criterios que consideren: *las características de la investigación, de la técnica, de la muestra, de las tareas de muestreo y de los encuestadores.*

**Características de la investigación**

•*Complejidad de la investigación.* Si es compleja, no se recomienda vía telefónica.No debe durar más de 10 a 15 min. Buen pronóstico har en domicilio, porque el entrevistador puede dar explicaciones.

•*Control de encuestadores.* Estos pueden responder indebidamente el cuestionario; se debe supervisar a los encuestadores. El sistema CATI puede grabar su desarrollo. En encuestas a domicilio, se puede verificar si el encuestador asistió. Contartación de agencia encuestadora.

•*Volimen de información.* Se debe planificar la toma de muestra representativa, ya que las encuestas largas requieren tiempo y serían costosas ahcerlas por teléfono, vía postal o domicilio.

•*Costo de la investigación.*Las encuestas personales tienen costos diferentes , son más económicas vía postal, más caras las telefónicas y las más costosas son las hechas a domicilio.

•*Facilidades para recoger la información.* Las másdesgastantes son las hechas a domicilio ya que hay que savar distancias.Las telefónicas y postales son intrmedias y las realizadas en establecimientos son las más económicas en esfuerzo.

•Facilidad para tratar la información. Las asistidad por cati son más fácilies de almacenar. El resto de las modalidades se debe codificar y guardar en sistemas informaticos.

•*Posibilidad de estudios de comportamiento del consumidor*. Todas las modalidades sirven, aunque la telefónica es limitada en tiempo, la postal y hecha a domicilio permie preguntas de profundidad y en estudios dinámicos, la conveniente es la ómnibus.

•*Posibilidad de obtener tendencias de comportamiento*. Todas las modalidades permiten verificar tendencias con menor o mayor profundidad.

•*Urgencia de la investigación.* En telefónica son horas, en domicilio semanas y en postal meses.

•*Temporalidad de la investigación.* Se debe tomar en cuenta el tiempo en el que se levantan las encuestas, por ejemplo pascua, navidad o vacaciones de verano, se tienen diversos comportamientos de los sujetos de estudio, con tendencias a sesgos.

•*Utilidad para la gestión.* Se sugieren muestras aleatorias. Se dice que la realizada

en establecimientos directos es la única que no lo cumple ya que los resultados no son generaliables a la población y son de carácter orientativo..Si se requiere controlar cuota de mercado, una empresa debiera contratar servicios de compañía encuestadora a través de paneles o encuestas ómnibus.

•*Zona geográfica del estudio* En zonas muy amplias, la encuesta a domicilio es lenta y cara; la postal y telefónica son más convenientes y menos costosas. La postal en zonas de baja densidad muy dispersas es la versiónmás recomendable.

Carcaterísticas y posibilidades del tipo de encuesta que se aplica

•*Conocimiento del hábitat.*En la encuesta a domicilio, el encuestador puede proporcionar información muy valiosa sobre l vivienda. O qué aspecto tiene la persona que encuesta . Esta informacion cualitativa complementaria puede ser muy valiosa. Por correo o teléfono es difícil obtener esta información.

•*Posibilidad de empleo de material auxiliar.*En las encuestas a domicilio, se puedne mostrar fotos, bocetos, maquetas,catálogos, etc. Muy util en la creacion de prototipos para nuevos productos. Si se trata de nuevos servicios se pueden presentar diversos tipos para que el encuestado los prdene de acuerdo a su preferencia.muy difícil de hacer telefonica o postalmente.

•*Posibilidad de formular preguntas diversas.* En ocasiones, los encuestadores tienen que formular preguntas alternativas, porque los encuestados no entienden bien las que se hacen. Esto no puede hacerse por correo y si mediante encuestas persdonales. Se corre el riesgo de que en la formulación de las preguntas el encuestador introduzca sesgos.Esto no ocurre con CATI.

•*Posibilidad de inferencias de conclusiones.* Siempre que se trate de muestras aleatorias, los resultados se pueden generalizar con cierto grado de error y confianza. Solo la encuesta en establecimientos no es aleatoria y no se pueden generalizar los resultados.

•*Participación.* El porcentaje d epersonas que responde es variable. En encuestas postales ronda el 10%, con encuestas telefónicas de 50 al 60% . Estos porcentajes deben tenerse en cuenta porque las muestras efectivas ,requiereb muestras teóricas más amplias. Para incentivar la participación se sugiere estimularlas con presentes o dinero.

.

**Características de la muestra**

•*Anonimato percibido.* Responder una encuesta puede percibirse como una intromisión a la intimidad delas peersonas. La obtención de respuestas será tanto más fácil cuanto más alta sea la sensación de anonimato. En ewste sentido, la encuesta por correo es la más anónima. A domicilio la menos anónima. Por teléfomo, la sensación de anonimato es intermedioa.

•*Aspectos sociodemográfico.* Las caracteristicas de la población de la que se extraerá la muestra afectan la elección del tipo de encuesta. Se debe tener en cuenta factores factores como la facilidad de su localización (a través de censos), indicadores demográficos (género, edad, zona de residencia) y socioeconómicos (renta, profesión, etc.). Gente culta responderá rápido; gente iletrada será más lento y quizá será personal.

•*Control de coherencia.* Al respoder los encuestados, suele haber problemas de

detección de coherencia , cuando la encuesta es a domicilio o por teléfono. En encuestas postales nada se puede hacer.Las encuestas por CATI, detectan incoherencias en las respuestas. En este caso el entrevistador recibe instrucciones por pantalla y salta a las preguntas convenientes cuando obtiene respuestas a preguntas filtro.

•*Control de proceso de investigación*. Se requiere de supervisión continua del proceso, haciendo preguntas recurrentes como: porcentaje de muestra encuestado, cantidad de encuestas aplicadas por género, edad, etc. CATI es la única técnica que no requiere de conteo físico ya que sus algoritmos permiten detener la encuesta cuando se reponden las preguntas de control..

•*Deseo de colaboración.* En la encuesta a domicilio, si se ecxponen bien los objetivos de la investigación, las personas no suelen negarse a colaborar. En este sentido también influye mucho la confianza que proporcionan los encuestadores. Son muy importantes su aspecto personal y el modo de comportarse como tal. Por correo es más difícil conseguir la colaboración. En los establecimientos, la gente suele tener prisa y no les gusta pararse a responder.

•*Influencia de terceros.* Cuando se envía una encuesta por correo para ser contestad ay devuelta, no se tiene la copmpleta seguridad de que haya contestado la persona seleccionada o de que no haya sido influenciada. Es posible incluso, que las respuestas hayan sido *aprobadas* por un *equipo de personas*. En encuestas personales a domicilio, en establecimientos o por teléfono no hay influencia de terceros.

•*Información delicada,* Ante un encuestador que pregunta cuestiones delicadas como preferenciassecuales, ideologíaspoliticas, ingresos, etc. una persona se puede sentir cohibida. Por teléfono es menor y postal es nula.

**Tareas de muestreo**

•*Accesibiidad.* Depende del tipo de encuesta a usar. Por teléfono es muy fácil entrar en contacto con la muestra, pues basta con marcar un número y repetir las veces necesarias la llamada en caso de no encontrar respuesta.También permite encuestar a persona sdifícilmente accesibles o que viven en lugares lejanos , o cuando hace mal tiempo. Obviamente sólo puede emplearse con personas que tengan teléfono y eso es una limitante. CATI genera automáticamente los números de teléfono y simplifica el proceso de muestreo. La accesibilidad de la meustra en la encuesta a domicilio es menor , pues requiere tiempo y desplazamientos. Por correo hace falta conocer los nombres y domiciñios delos detinatarios para acceder a ella.

•*Facilidad de muestreo.* Mediante la encuesta telefónica convencional o CATI es fácil de planificar el muestreo. Los números de teléfono se asocian a barrios, ciudades o provincias. Se pued eprogramar el n úmero de los teléfonos para que su selección proporcione muestras estratificadas y mantenga la representatividad de las zonas. En encuestas a domicilio o por correo, la planificación del muestreo y la localización de las muestras es más difícil.

•*Posibilidad de influencia de resultados.* Sólo las muestra aleatorias proporcionan resultados generalizables. Sólo la encuesta en establecimientos recoge muestras no aleatorias de información.

•*Representatividad.* Cuando se trata de una encuesta postal, puede que la muestra de respiondientes no sea representativa., debido a que el mayor porcentajes de respuestas lo aporta la clase media. En paísis poco desarrollados, la encuesta telefónica sería proveniente de personas de case alta.

**Características de los encuestadores**

•*Flexibilidad para formular preguntas.* Los encuestadores deben formular preguntas a la muestra a veces de forma *rígida o flexible*, es decir, preguntan sobre un cuestionario *cerrado* y marcan las respuestas o formulan las preguntas en términos menos concretos. En establecimientos ser rápido y no entretener demasiado a las personas. Los cuestionarios serán cerrados. A domicilio, se puden hacer preguntas *abiertas.* La encuesta postal no es flexible. L atelefónica, en especial la cati y a domicilio da opción a la flexibilidad de los encuestadores.

•*Habilidad de los encuestadores.* El encuestador que se aplica a domicilio o por teléfono para obtener informaicón, requiere de habilidad especial para hacer las preguntas y brindar explicaciones. Si una empresa dudara de esta habilidad del encuestador, se le reomienda hacerla vía postal.

•*Posibilidad de introducir sesgos.* Los errores que pueden aparecer durante la captura de información, son parcialmente debida a los encuestadores, orque leen mal, se equivocan al marcar las respuestas o confunden a los encuestados. En toda encuesta personal aparecen seshgos debido a los encuestadores , en maypr o menor grado, cosa que no ocurre en las postales..

**Utilización conjunta de encuestas**

Los diversos tipos de encuestas tienen sus limitaciones . ninguna de ellas puede decirse que sean netamente superior a las demás. *En la práctica es frecuente combinarlas*. Se puede entregar un cuestionario en mano en el momento oportuno en una visita concertada. Posteriormente, se puede llamar por teléfono para concertar una cita y recogerlo. Este procedimiento se emplea cuando los cuestionarios son largos y necesitan de explicaciones para llenarlos. Tambien, se puede llamar por teléfono para anunciar el envío de in cuestionario postal y conseguir así, una mayor tasa de respuesta. Después, se vuelve a llamar para recordar que se pasarápara recoger el cuestionario. Actuando en estos términos , se pueden conseguir tasas de respuesta de un 90%, con un porcentaje de cuestionarios útiles de un 75%.

Los investigadores pueden llamar porteléfono para concertar una cita y realizar una encuesta . de esta forma se consigue una tasa elevada de respuesta.

Tratandose de una encuesta postal se puede llamar a las personas que no han respondido para insistir en su colaboración.

Tras realizar una entrevista personal se puede llamar por teléfono para verificar la información proporcionada.

| Cracterísticas de las encuestas | | | | | | |
|---|---|---|---|---|---|---|
| Caracteríticas | Domicilio | Establecimientos | Teléfono | CATI | Postal | Omnibus |
| Complejidad recoger info | Mucha | Poca | Poca | Poca | Moderada | Moderada |
| Control encuestador | Moderado | Alta | Alta | Máxima | No existe | No existe |
| Costo | Alto | Moderada | Moderada | Moderada | Baja | Moderada |
| Facilidades recoger info | Moderado | Moderada | Moderada | Moderada | Poca | Moderada |
| Facilidades tratar info | Media | Media | Alta | Máxima | Media | Media |
| Comportamiento consumidor | Buena | Buena | Buena | Buena | Buena | Máxima |
| Tendencias | Nula | Nula | Nula | Nula | Nula | Muy buena |
| Urgencia | Baja | Moderada | Alta | Máxima | Muy baja | Muy baja |
| Gestión | Muy buena | Moderada | Muy buena | Muy buena | Buena | Muy buena |
| Amplitud geográfica | Baja | Alta | Muy alta | Muy alta | Muy alta | Muy alta |
| Hábitat encuestado | Muy alta | Baja | Nula | Nula | Nula | Nula |
| Material auxiiar | Muy alta | Moderada | Nula | | | |
| Preguntas variadas | Máxima | Muy alta | Muy alta | Máxima | Nula | Nula |
| Inferencia | Máxima | Nula | Máxima | Máxima | Máxima | Máxima |
| Tiempo recoger info | Mucho | Moderado | Corta | Corta | Mucho | Alta |
| Anonimato | Mínimo | Mínimo | Moderada | Moderada | Máxima | Máxima |
| Coherencia | Bajo | Baja | Baja | Máximo | No existe | No existe |
| Colaboración | Bajo | Baja | Moderada | Moderada | Moderada | Alto |
| Influencia terceros | Nula | Nula | Nula | Nula | Alta | Moderada |
| Info delicada | Difícil | Nula | Poco difícil | Poco difícil | Fácil | Fácil |
| Participación | Alta | Moderada | Moderada | Moderada | Baja | Alta |
| Confiabilidad | Alta | Dudosa | Alta | Alta | Moderada | Alta |
| Accesabilidad a muestra | Difícil | Fácil | Fácil | Muy fácil | Difícil | Difícil |
| Muestreo facilidad | Difícil | No se usa | Fácil | Muy fácil | Difícil | Difícil |
| Posibilidad Inferencia | Alta | Nula | Alta | Alta | Alta | Alta |

| | | | | | | |
|---|---|---|---|---|---|---|
| resultados | | | | | | |
| Representatividad muestra | Alta | Nula | Moderada | Moderada | Moderada | Moderada |
| Flexibilidad preguntas | Máxima | Moderado | Baja | Máxima | Nula | Nula |
| Habilidad encuestador | Máxima | Normal | Alta | Alta | No existe | No existe |
| Sesgos | Moderada a baja | Elevado | Elevada | Moderada a baja | No existe | No existe |

## La encuesta online

El desarrollo de la informática, en general, y del comercio electrónico, en particular, han propiciado la aparición de una modealidad de encuesta que proporciona información a través de cuestioanriosque se distribuyen y cumplimentan online.
Este proceder empieza a ser utilizado con cierta intensidad y tiene *ventajas* y *limitaciones:*

**Ventajas**

•Por una parte, *no pueden aparecer sesgos* derivados de la intervención de los encuestadores, muchas veces inevitables, sencillamente porque en la encuesta online no existen encuestadores. Estos pueden condicionar o derivar las respuestas de las personas que responden y ser una fuente de error aleatorio cuando se aplican escalas o, simplemente, cuando se recoge información. Los cuestionarios son autoadministrados por los propios respondientes.
Por otra parte, *la información se consigue rápidamente*. Los cuestionarios se distribuyen de forma prácticamente instantánea cuando los servidores no están saturados y las personas también puede responder rápidamente. Además la información proporcionada alimenta bases de datos que pueden descargarse en formatos habituales como: TXT, Excel o SPSS, por ejemplo y ser analizadas, sin más preámbulos. Las pesadas tareas de grabación de datos, configuración de bases, etiquetado de variables, depuración de la información, etc. son mucho más sencillas ahora y algunas tareas , como la grabación de datos , avcabarán siendo un recuerdo sin duda. La encuesta online, *facilita el tratamiento de datos.*
Finalmente, *la encueta online se puede distribuir en cualquier lugar* del planeta donde se pueda conectar un dispositivo a internet ya sea PC, laptop, tablet, smartphoine. Basta con que exista cobertura o acceso a redes informáticas. Esta característica resulta particularmente atractiva cuando ha de obtenerse información en áreas geográficas muy extensas, con poblaciones muy diseminadas o cuando la densidad de las poblaciones es reducida. En realidad, la encuesta online es la heredera de la encuesta postal y ademas con muchas ventajas adicionales sobre ella, como las mencionadas anteriormente.

**Liimitaciones**

•*Dificultad de acceso a la población*. Las empresas, sencillamente, no tienen todos los correos electrónicos porque no hau un censo tan completo. Existen bases de datos de teléfonos fijos, pocas de móviles , *pero no de correos electrónicos* , que

pueden anularse o crearse con mucha facilidad. . Una persona puede poseer varias cuentas de correo. También hay que tener presente la protección de datospuede constituir un freno a la elaboración de bases de mails, que además y por los motivos mencionados, pueden ser muy volátiles.

•Un problema adicional es que el correo referido a la encuesta sea considerado como *spam* y que los filtors de la computadora lo eliminen.

•Algunas bases de mails son relativamente sencillas de cosntruir y bastante estables. Por ejemplo, con trabajo y paciencia es posible armar una base de datos que contenga los emails de los profesores en un campus universitario inspirada en los datos que presentan en web, las universidades del país. Dicha base pòdría ser del interés del mundo editorial o para las organizaciones que entrar en contacto con los investigadores y quisieran formalizar colaboraciones en materia de docencia e investigación.

•También, se podrían crear bases de datos para grupos profesionales a través de sus respctivos colegios o directorios ya publicados. En cualquier caso, no hay que olvidar que se puede vulnerar la intimidad de las personas y que existe una Ley de Protección de Datos que las ampara. Las empresas pueden pedir a sus clientes que se registren. Muchas no permiten las compras ni otras actividades si no existe un registro previo, n el que deben proporcionarse datos personales de identificación, localización, contacto e e-mail, aparte de contraseñas de acceso , por seguridad del propio consumidor.

•La propia actividad de los consumidores crea el censo de clientes que se enriquece a lo largo del tiempo, con el cual se puede entrar en contacto y recoger información sobre el ehistorial de compras, intereses por nuevos productos, grado de satisfacción ,calidad percibida, actitudes, valoraciones, etc. y cualquier otra informaciónsusceptible de ser recogida a través del *método de la encuesta.*Las *cookies* desempeñan un papel fundamental en este sentido.

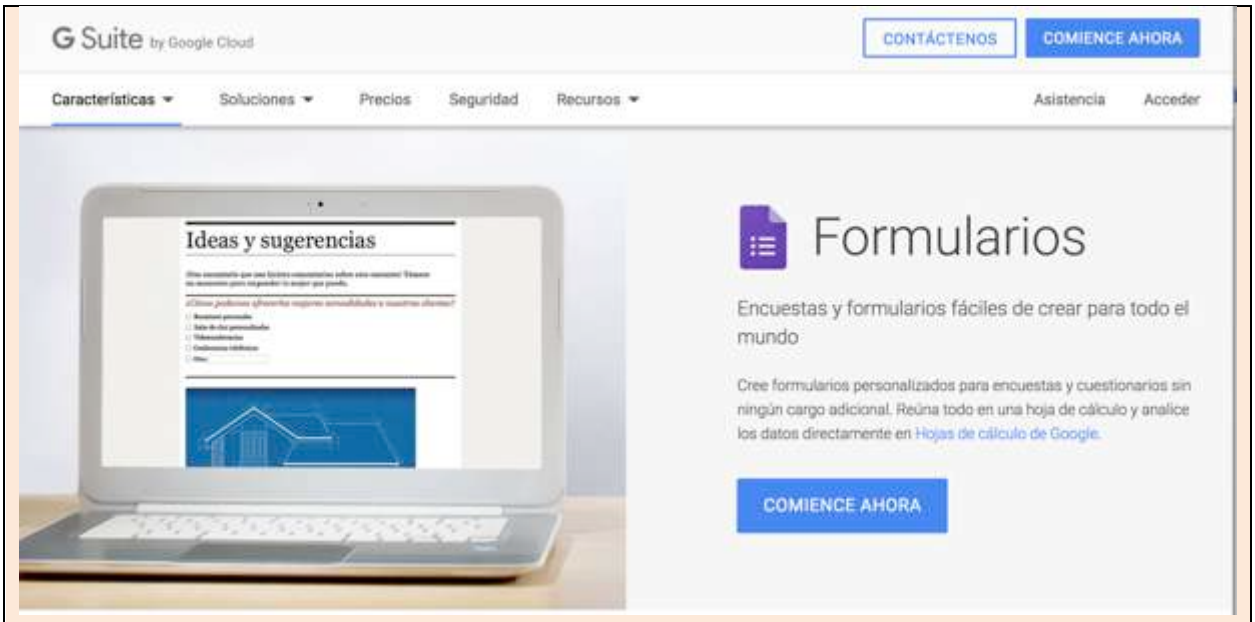
•Lo que ocurre es que el censo de los consumidorres *se crea a posteriori.* Lo ideal sería disponer a *priori* de los mail de poblaciones enteras para hacer sondeos sobre nuevos productos, interés que despiertan, identificar segmentos de mercado, etc.

•Otra limitación lo constituye el *control de personas* con las que se contacta. Se envían cuestionarios online a personas deyterminadas, pero esos cuestionarios pueden ser respondidos por otras, o biuenb, influir en las respuestas.. En este sentido, los peligros que acechan a la encuesta online so nsimilares a los ya mencionadospatra la encuesta postal.

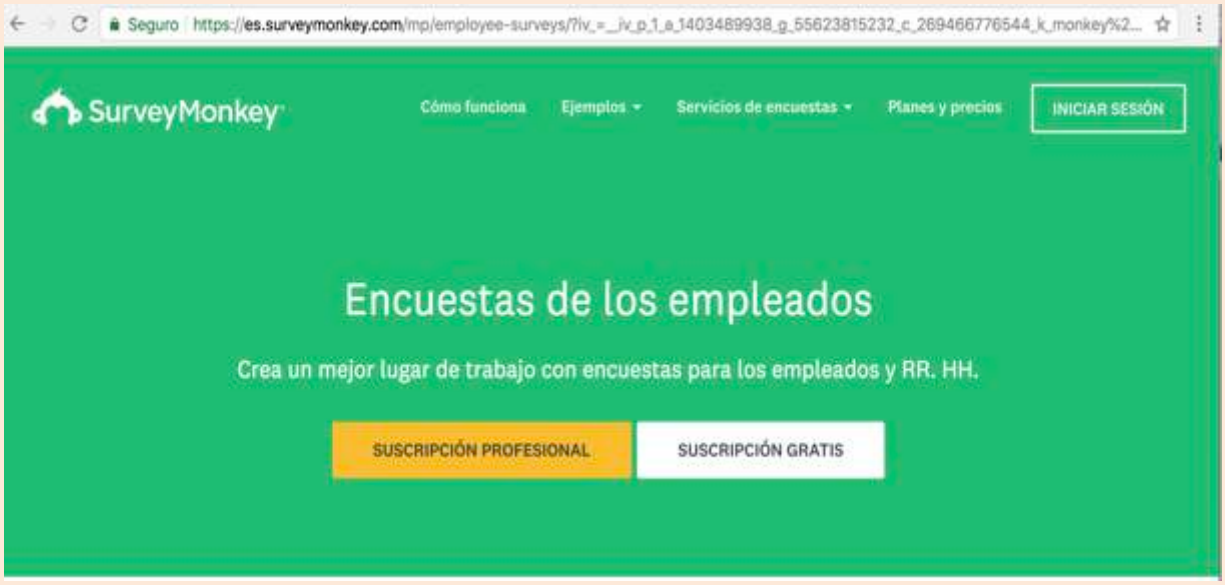
•No obstante, a pesar de todo lo expuesto, la encuesta online se vislumbra como un método de obtención de información que, junto a la telefónica, serán los dominantes en la investigación de mercados a medio plazo.

Software más representativo:

**Google Formats** **https://gsuite.google.com/signup/basic/welcome**




**Survey Monkey** [https://es.surveymonkey.com/](https://es.surveymonkey.com/)




Fuente: Grande-Esteban y Abascal-Fernández (2017) y Hernández-Sampieri et al., (2010) con adaptación propia

# Formulación de hipótesis

Son explicaciones tentativas del fenómeno investigado que se formulan como proposiciones. No todas las investigaciones cuantitativas plantean hipótesis. El hecho de que formulemos o no hipótesis depende de un factor esencial: el alcance inicial del estudio. Las investigaciones cuantitativas que formulan hipótesis son aquellas cuyo planteamiento define que su alcance será correlacional o explicativo, o las que

tienen un alcance descriptivo, pero que intentan pronosticar una cifra o un hecho. Ver **Tabla A.33.**

**Tabla A.33. Forumulación de hipótesis y sus alcances**

| Alcande de la investigación | Formulación de hipótesis |
|---|---|
| Exploratorio | •No se formulan hipótesis |
| Descriptivo | •Sólo se formulan hipótesis cuando se pronostica un hecho o dato |
| Correlacional | •Se formulan hipótesis correlacionales |
| Explicativo | •Se formulan hipótesis causales |

Fuente: Hernández-Sampieri et al., (2010) con adaptación propia

El *enfoque cualitativo*, por lo regular, no formulan hipótesis antes de recolectar datos (aunque no siempre es el caso). Su naturaleza es más bien inducir las hipótesis por medio de la recolección y el análisis de los datos. En una investigación podemos tener una, dos o varias hipótesis. las hipótesis son proposiciones tentativas acerca de las relaciones entre dos o más variables, y se apoyan en conocimientos organizados y sistematizados. Una vez que se prueba una hipótesis, ésta tiene un impacto en el conocimiento disponible, que puede modificarse y por consiguiente, pueden surgir nuevas hipótesis.

En este punto es necesario definir una *variable* como una propiedad que puede fluctuar y cuya variación es susceptible de medirse u observarse y tienen un papel preponderante en el planteamiento de las hipótesis, una vez que es revisada la literatura . Existe, pues, una relación muy estrecha entre el planteamiento del problema, la revisión de la literatura, las hipótesis y las variables que la conforman. La revisión inicial de la literatura hecha para familiarizarnos con el problema de estudio nos lleva a plantearlo, después ampliamos la revisión de la literatura y afinamos o el planteamiento con la descripción de las variables, con las que derivamos las hipótesis. Al formular las hipótesis volvemos a evaluar nuestro planteamiento del problema, objetivos y preguntas de investigación.Ver **Tabla A.34.**

**Tabla A.34. Características de una hipótesis**

| Carácterísticas de las hipótesis |
|---|
| •La hipótesis debe referirse a una situación *real* ya que sólo pueden someterse a prueba en un universo y un contexto bien definidos. En ocasiones, en la misma hipótesis se hace explícita esa realidad y otras veces la realidad se define por medio de explicaciones que acompañan a la hipótesis, contextualizando más sobre la realidad<br>•Las variables o términos de la hipótesis deben ser comprensibles, precisos y lo más concretos posible. Términos vagos o confusos no tienen cabida en una hipótesis<br>•La relación entre variables propuesta por una hipótesis debe ser clara y verosímil (lógica). |

•Los términos o variables de la hipótesis deben ser observables y medibles, así como la relación planteada entre ellos, o sea, tener referentes en la realidad.

•Las hipótesis deben estar relacionadas con técnicas disponibles para probarlas.

## Utiidad de las hipótesis

•Son las guías de una investigación en el enfoque cuantitativo. Formularlas nos ayuda a saber lo que tratamos de buscar, de probar. Proporcionan orden y lógica al estudio. Son como los objetivos de un plan administrativo: las sugerencias formuladas en las hipótesis pueden ser soluciones a los problemas de investigación. Si lo son o no, efectivamente es la tarea del estudio (Selltiz *et al*., 1980.)

•Tienen una función descriptiva y explicativa, según sea el caso. Cada vez que una hipótesis recibe evidencia empírica en su favor o en su contra, nos dice algo acerca del fenómeno con el que se asocia o hace referencia. Si la evidencia es en favor, la información sobre el fenómeno se incrementa; y aun si la evidencia es en contra, descubrimos algo acerca del fenómeno que no sabíamos antes.

•Prueban teorías. Cuando varias hipótesis de una teoría reciben evidencia positiva, la teoría va haciéndose más robusta; y cuanto más evidencia haya en favor de aquéllas, más evidencia habrá en favor de ésta.

•Sugieren teorías. Diversas hipótesis no están asociadas con teoría alguna; pero llega a suceder que como resultado de la prueba de una hipótesis, se pueda construir una teoría o las bases para ésta. Lo anterior no es muy frecuente, pero ha llegado a ocurrir.

## Operacionalización de variables en la hipótesis

•Para que el investigador, sus colegas, los usuarios del estudio y, en general, cualquier persona que lea la investigación le den el mismo significado a los términos o variables incluidas en las hipótesis, es común que un mismo concepto se emplee de maneras distintas

•Asegurarnos de que las variables pueden ser medidas, observadas, evaluadas o inferidas, es decir que de ellas se pueden obtener datos de la realidad.

•Confrontar nuestra investigación con otras similares. Si tenemos definidas nuestras variables, podemos comparar nuestras definiciones con las de otros estudios para saber “si hablamos de lo mismo”. Si la comparación es positiva, confrontaremos los resultados de nuestra investigación con los resultados de las demás.

•Evaluar más adecuadamente los resultados de nuestra investigación, porque las variables, y no sólo las hipótesis, se contextualizan.

•En conclusión, sin definición de las variables no hay investigación. Las variables deben ser definidas de dos formas:

•*Conceptual.* Una definición conceptual trata a la variable con otros términos. Es lo que nos indica la revisión de la literatura

•*Operacional.* Es basado en la perspectiva o punto de vista en el que se abordará el estudio. Debe especifica qué actividades u operaciones deben realizarse para medir una variable. Una definición operacional nos dice que para recoger datos respecto de una variable, hay que hacer esto y esto otro, además articula los procesos o acciones de un concepto que son necesarios para identificar ejemplos de éste. Así, la definición operacional de la variable *temperatura* sería el termómetro o innovación

la que se define en el Manual de Oslo (OCDE,2005). De aquí, lo valioso de conocer escalas previas de medición.

Cuando dispone de varias opciones para definir operacionalmente una variable, debe elegir la que proporcione mayor información sobre la misma, capte mejor su esencia, se adecue más a su contexto y sea más precisa. O bien, una mezcla de tales alternativas.. Los criterios para evaluar una definición operacional son básicamente :

•*Adecuación al contexto,* capacidad para captar los componentes de la variable de interés, confiabilidad y validez. "

•*Recolección de los datos cuantitativos.* Una correcta selección de las definiciones operacionales disponibles o la creación de la propia definición operacional se encuentran muy relacionadas con una adecuada revisión de la literatura.

•Cuando ésta ha sido cuidadosa, se tiene una gama más amplia de definiciones operacionales para elegir o más ideas para desarrollar una nueva.

•Asimismo, al contar con estas definiciones, el tránsito a la elección del o los instrumentos para recabar los datos es muy rápido, sólo debemos considerar que se adapten al diseño y a la muestra del estudio.

•En la investigación comúnmente se tienen diversas variables y, por tanto, se formularán varias definiciones conceptuales y operacionales.

•Algunas variables no requieren que su definición conceptual se mencione en el reporte de investigación, porque ésta es relativamente obvia y compartida. El mismo título de la variable la define; por ejemplo, *actividad del sujeto* y *edad*.

•Prácticamente todas las variables requieren una definición operacional para ser evaluadas de manera empírica, aun cuando en el estudio no se formulen hipótesis. Siempre que se tengan variables, se deben definir operacionalmente.

## Hipótesis de investigación

Éstas se definen como proposiciones tentativas acerca de las posibles relaciones entre dos o más variables, y deben cumplir con los requisitos mencionados. Se les suele simbolizar como Hi o H1, H2, H3, etc. (cuando son varias), y también se les denomina *hipótesis de trabajo*. Estas pueden ser:

•*Descriptivas de un valor o dato pronosticado.*

Se utilizan a veces en estudios descriptivos, para intentar predecir un dato o valor en una o más variables que se van a medir u observar. Pero cabe comentar que no en todas las investigaciones descriptivas se formulan hipótesis de esta clase o que sean afirmaciones más generales. Ejemplo*: la ansiedad en los gerentes de producción de componentes electrónicos, es del 90%.*

•*Correlacionales.*

Especifican las relaciones entre dos o más variables. Por ejemplo: *la motivación de logro se encuentra vinculada con la satisfacción laboral y la moral en el trabajo* Sin embargo, las hipótesis correlacionales no sólo pueden establecer que dos o más variables se encuentran vinculadas, sino también cómo están asociadas. Alcanzan el nivel predictivo, parcialmente explicativo y qué dirección existe entre ellas. Por ejemplo: *A mayor autoestima, habrá menor temor al éxito.*

•*Diferencia de grupos.*

Estas hipótesis se formulan en investigaciones cuya finalidad es comparar grupos.

Por ejemplo: *¿es más eficaz un comercial televisivo en blanco y negro que uno en colores?*
Cuando el investigador no tiene bases para presuponer en favor de qué grupo será la diferencia, formula una *hipótesis simple de diferencia de grupos* (como la anterior). Y cuando sí tiene bases, establece una hipótesis direccional de diferencia de grupos (por ejemplo: *los gerentes que administran tecnología, atribuyen más importancia a el conocimiento de la innovación ,que los gerentes de manufactura*). Esto último, por lo común, sucede cuando la hipótesis se deriva de una teoría o estudios antecedentes, o bien, el investigador está bastante familiarizado con el problema de estudio. Esta clase de hipótesis llega a abarcar dos, tres o más grupos.
Algunos investigadores consideran a las hipótesis de diferencia de grupos como un tipo de hipótesis correlacional, porque en última instancia relacionan dos o más variables.

•*Causales*

Este tipo de hipótesis no solamente afirma la o las relaciones entre dos o más variables y la manera en que se manifiestan, sino que además propone un *sentido de entendimiento de las relaciones*. Tal sentido puede ser más o menos completo, esto depende del número de variables que se incluyan, pero todas estas hipótesis establecen relaciones de causa-efecto. Se simbolizan como *X—Y* ; y las hipótesis causales, se representan como:

$$X \xrightarrow{\text{Influye en o causa}} Y$$

*Correlación y causalidad son conceptos asociados, pero distintos*. *Si dos variables están correlacionadas, ello no necesariamente implica que una será causa de la otra.* Suponga que una empresa aporta un servicio vía App que se vende poco y decide mejorarlo en su diseño, lo hace y lanza una campaña de mercadotecnia digital para anunciarlo en internet. Después, se observa un aumento en las ventas del servicio. Los ejecutivos de la empresa pueden decir que el lanzamiento de la campaña de mercadotecnia digital está relacionado con el incremento de las ventas; pero si no se demuestra la causalidad, no es posible asegurar que la campaña haya provocado tal incremento. Quizá la campaña de mercadotecnia digital sea la causa del aumento de las descargas de la App, pero tal vez la causa sea en sí la mejora del servicio vía App, una excelente estrategia de distribución en redes sociales u otro factor, o bien, todas pueden ser causas.
Si cada vez que se obtiene una correlación se supusiera causalidad, ello equivaldría a decir que cada vez que se observa a una persona con saco y corbata y a un joven estudiante juntos, se supusiera que él es su profesor, cuando puede ser su padre, su pariente o cualquier otra persona de vestimenta formal, que por azar se colocó muy cerca del joven.
*Para establecer causalidad antes debe haberse demostrado correlación, pero además la causa debe ocurrir antes que el efecto.* Asimismo, los cambios en la causa tienen que provocar cambios en el efecto. Al hablar de hipótesis, a las *supuestas causas* se les conoce como *variables independientes* y a los *efectos* como *variables dependientes*. Únicamente es posible hablar de variables

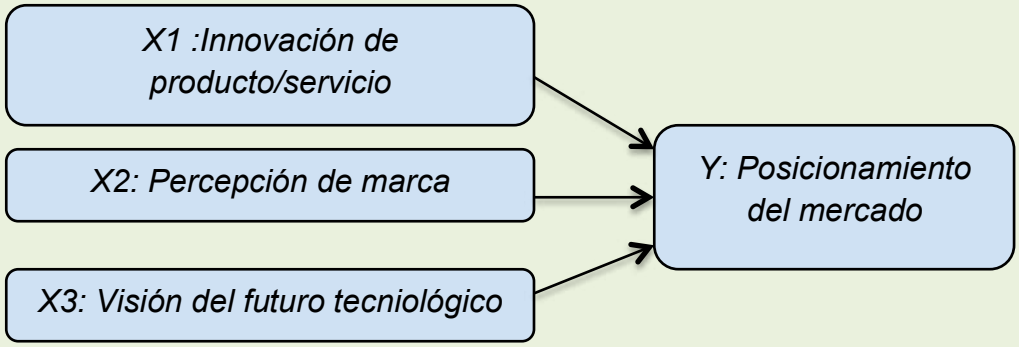
independientes y dependientes cuando se *formulan hipótesis causales o hipótesis de la diferencia de grupos,* siempre y cuando en estas últimas se explique cuál es la causa de la diferencia supuesta en la hipótesis.

**Tipos de hipótesis**

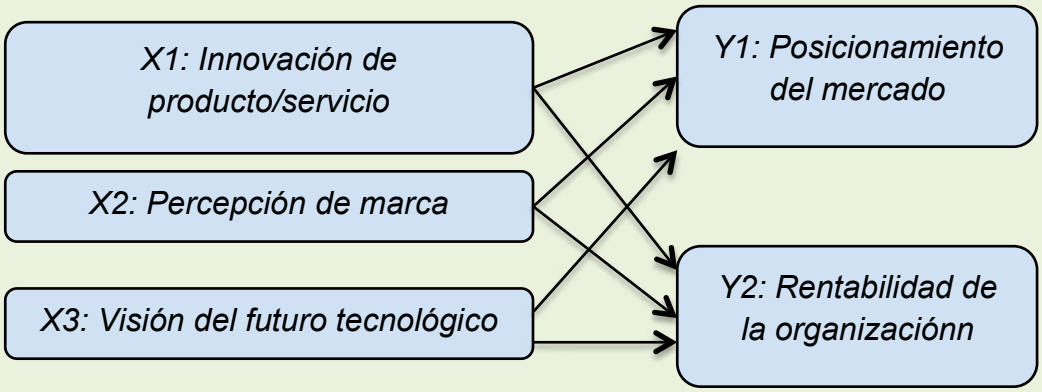
Hipótesis causales

•*Hipótesis causales bivariadas.* En éstas se plantea una relación entre una variable independiente y una variable dependiente. Por ejemplo: *La percepción de que otra producto de una compañía similar al que hemos utilizado es similar en calidad, prestaciones y diseño, nos provoca mayor atracción hacia dicho producto*

•*Hipótesis causales multivariadas.* Plantean una relación entre diversas variables independientes y una dependiente, o una independiente y varias dependientes, o diversas variables independientes y varias dependientes. Por ejemplo: *La innovación de producto/servicio así como la percepción de la marca sometida a una visión del futuro tecnológico, determinan el posicionamiento de mercado .*

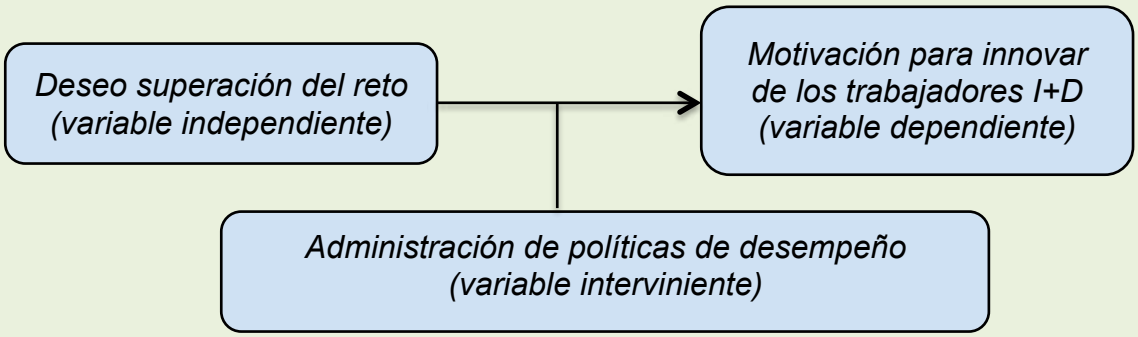



Otro ejemplo: *La innovación de producto/servicio así como la percepción de la marca sometida a una visión del futuro tecnológico, determinan tanto el posicionamiento de mercado, así como mayor rentabilidad de la organizacón*

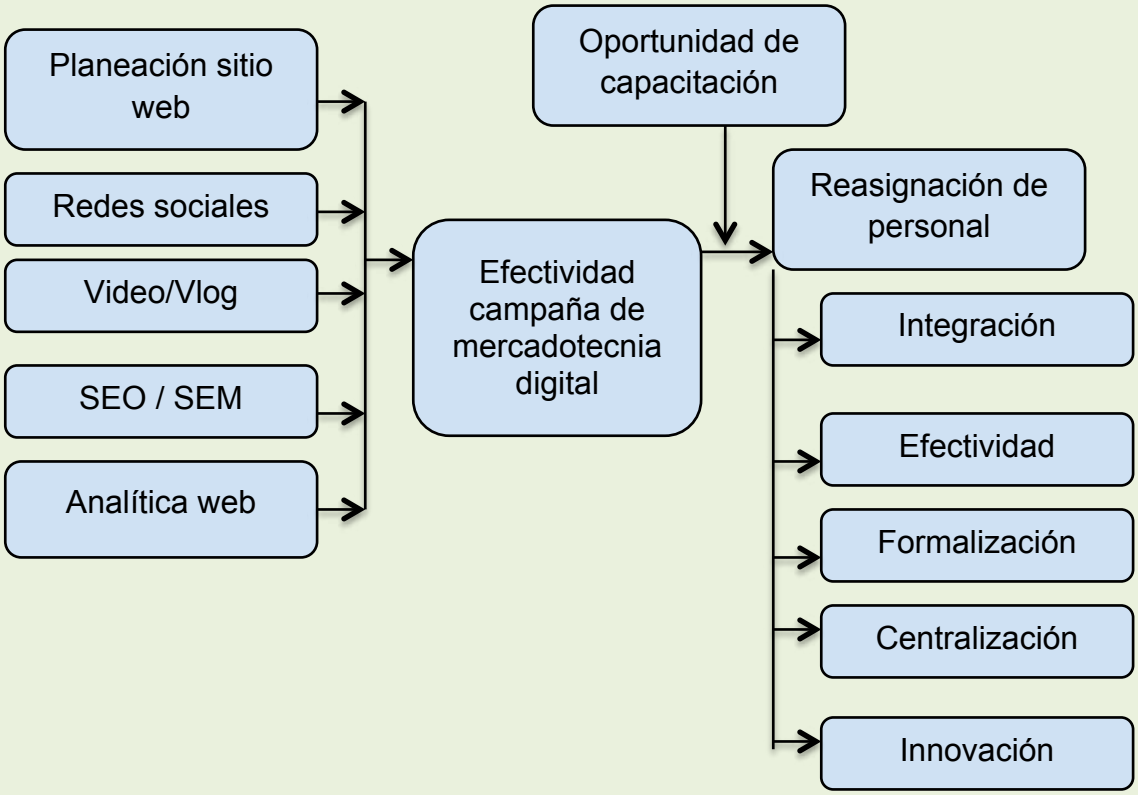



Las hipótesis multivariadas pueden plantear otro tipo de relaciones causales, en donde ciertas variables intervienen modificando la relación *(hipótesis con*

*presencia de variables intervinientes*). Por ejemplo: *El deseo de superación del reto aumenta la motivación para innovar de los trabajadores de I+D, cuando se administra de acuerdo con políticas de desempeño*




Es posible que haya estructuras causales de variables más complejas que resulta difícil expresar en una sola hipótesis, porque las variables se relacionan entre sí de distintas maneras. Entonces se plantean las relaciones causales en dos o más hipótesis, o de forma gráfica




H1: *La planeaciòn sitio web incrementa la efectividad de la campaña de mercadotecnia digital.*
H2: *Las redes sociales web decrementan la efectividad de la campaña de mercadotecnia digital.*
H3*: El Video/ Vlog, SEO/SEM y la analítica web incrementan la efectividad de la campaña de mercadotecnia digital.*

H4: *La efectividad de la campaña de mercadotecnia digital influye en la oportunidad de capacitación.*
H5: *La oportunidad de capacitación mediatiza la vinculación entre la efectividad de la campaña de mercadotecnia digital y la reasignación de personal*
H6: *La reasignación de personal afecta la integración, la efectividad organizacional, la formalización, la centralización y la innovación*

Cuando las hipótesis causales se someten al análisis estadístico, *se evalúa la influencia de cada variable independiente (causa) en la dependiente (efecto), y la influencia conjunta de todas las variables independientes en la dependiente o dependientes.*

## Hipótesis nulas

Son, en cierto modo, el reverso de las hipótesis de investigación. También constituyen proposiciones acerca de la relación entre variables, sólo que sirven para refutar o negar lo que afirma la hipótesis de investigación.
Si la hipótesis de investigación propone:
Hi: *Los jóvenes le atribuyen más importancia al atractivo de diseño a los Smartphone que las jóvenes*
La hipótesis nula, postularía:
Ho: *Los jóvenes NO le atribuyen más importancia al atractivo de diseño a los Smartphone que las jóvenes*
Debido a que este tipo de hipótesis resulta la contrapartida de la hipótesis de investigación, hay prácticamente tantas clases de hipótesis nulas como de investigación. Es decir, la clasificación de hipótesis nulas es similar a la tipología de las hipótesis de investigación: hipótesis nulas descriptivas de un valor o dato pronosticado, hipótesis que niegan o contradicen la relación entre dos o más variables, hipótesis que niegan que haya diferencia entre grupos que se comparan, e hipótesis que niegan la relación de causalidad entre dos o más variables (en todas sus formas). Las hipótesis nulas se simbolizan así: Ho.

## Hipótesis alternativas

Son posibilidades alternas ante las hipótesis de investigación y nula: ofrecen otra descripción o explicación distinta de las que proporcionan estos tipos de hipótesis. Si la hipótesis de investigación establece: *este grupo tiene capacitación tipo alfa*, la nula afirmará: *este grupo no tiene capacitación tipo alfa*, y podrían formularse una o más hipótesis alternativas: *este grupo tiene capacitación tipo beta*, *gama, theta.* Cada una constituye una descripción distinta de las que proporcionan las hipótesis de investigación y nula. Se simbolizan como *Ha* y sólo pueden formularse cuando efectivamente hay otras posibilidades, además de las hipótesis de investigación y nula. *De no ser así, no deben establecerse.* Por ejemplo:
H1: *El nuevo software CRM de negocios electrónicos, obtendrá una 50 a 60 % de aceptación en los clientes*
H0: *El nuevo software CRM de negocios electrónicos, NO obtendrá un 50 a 60 % de aceptación en los clientes*
Ha: *El nuevo software CRM de negocios electrónicos, obtendrá más de un 60 % de aceptación en los clientes*
Ha: *El nuevo software CRM de negocios electrónicos, obtendrá menos de un 50 %*

*de aceptación en los clientes*
Otro ejemplo
Hi: *Los jóvenes le atribuyen más importancia al atractivo de diseño a los Smartphone que las jóvenes*
Ho: *Los jóvenes NO le atribuyen más importancia al atractivo de diseño a los Smartphone que las jóvenes*
Ha: *Los jóvenes le atribuyen menos importancia al atractivo de diseño a los Smartphone que las jóvenes*
En este último ejemplo de los jóvenes, si la *hipótesis nula* hubiera sido formulada de la siguiente manera:
Ho: *Los jóvenes no le atribuyen más importancia o le atribuyen menos importancia al atractivo de diseño a los Smartphone que las jóvenes*
No habría posibilidad de formular una *hipótesis alternativa*, puesto que las *hipótesis de investigación y nula* abarcan todas las posibilidades.
Las hipótesis alternativas, como puede verse, constituyen otras hipótesis de investigación adicionales a la hipótesis de investigación original.

Fuente: Hernández-Sampieri et al., (2010) con adaptación propia

## *¿Qué y cuántas hipótesis deben formularse en una investigación ?*

Al respecto no hay reglas universales, ni siquiera consenso entre los investigadores. Se puede leer en un artículo de alguna revista científica un reporte de investigación donde sólo se establezca la hipótesis de investigación; y, en otra, encontrar un artículo donde únicamente se plantea la hipótesis nula. Un artículo en una tercera revista, en el cual se puedan encontrar solamente las hipótesis de investigación y nula, pero no las alternativas. En una cuarta publicación otro artículo que contenga la hipótesis de investigación y las alternativas. Y otro más donde aparezcan hipótesis de investigación, nulas y alter- nativas. Esta situación es similar en los reportes presentados por un investigador o una empresa. Lo mismo ocurre en tesis y disertaciones doctorales, estudios de divulgación popular, reportes de investigaciones gubernamentales, libros y otras formas para presentar estudios de muy diversos tipos. En estudios que contienen análisis de datos cuantitativos, la opción más común es incluir la o las hipótesis de investigación únicamente.
Algunos investigadores sólo enuncian una hipótesis nula o de investigación presuponiendo que quien lea su reporte deducirá la hipótesis contraria.
Se recomienda que aunque exclusivamente se incluyan las hipótesis de investigación, *todas se tengan presentes*, no sólo al plantearlas, sino durante todo el estudio. Esto ayuda a que el investigador siempre esté alerta ante todas las posibles descripciones y explicaciones del fenómeno considerado; así podrá tener un panorama más completo de lo que analiza.
Sobre la cantidad, cada investigación es diferente. Algunas contienen gran variedad de hipótesis porque el problema de investigación es complejo (por ejemplo, pretenden relacionar 15 o más variables), mientras que otras contienen una o dos hipótesis. Todo depende del estudio que habrá de llevarse a cabo. La calidad de una investigación no necesariamente está relacionada con el número de hipótesis que contenga. En este sentido, se debe tener el número de hipótesis necesarias para guiar el estudio, y ni una más ni una menos

En una misma investigación es posible establecer todos los tipos de hipótesis, porque el problema de investigación así lo requiere.
Los estudios que se inician y concluyen como descriptivos, formularán si pronostican un dato (hipótesis descriptivas); los correlacionales podrán establecer hipótesis descriptivas de estimación, correlacionales y de diferencia de grupos (cuando éstas no expliquen la causa que provoca la diferencia); por su parte, los explicativos podrán incluir hipótesis descriptivas de pronóstico, correlacionales, de diferencia de grupos y causales. No debemos olvidar que una investigación puede abordar parte del problema de forma descriptiva y parte explicativa. Aunque debemos señalar que los estudios descriptivos no suelen contener hipótesis, y ello se debe a que en ocasiones es difícil precisar el valor que se puede manifestar en una variable.
Los tipos de estudio que no establecen hipótesis son los exploratorios. No puede presuponerse (afirmando) algo que apenas va a explorarse. Sería como si antes de una primera cita con una persona totalmente desconocida del género opuesto, tratáramos de conjeturar qué tan simpática es, qué intereses y valores tiene, etc. Ni siquiera podríamos anticipar qué tan atractiva nos va a resultar, y tal vez en una primera cita nos dejemos llevar por nuestra imaginación; pero en la investigación esto no debe ocurrir. Si se nos proporciona más información (lugares a donde le agrada ir, ocupación, religión, nivel socioeconómico, tipo de música que le gusta y grupos de los que es miembro), podemos plantearnos hipótesis en mayor medida, aunque nos basemos en estereotipos. Y si nos dieran información muy personal e íntima sobre ella, podríamos sugerir hipótesis acerca de qué clase de relación vamos a establecer con esa persona y por qué (explicaciones tentativas).

## *¿Qué y cuántas hipótesis deben formularse en una*

Como se ha venido mencionando a lo largo de este capítulo, las hipótesis del proceso cuantitativo se someten a prueba o escrutinio empírico para determinar si son apoyadas o refutadas, de acuerdo con lo que el investigador observa. De hecho, para esto se formulan en la tradición deductiva. Ahora bien, en realidad no podemos probar que una hipótesis sea verdadera o falsa, sino argumentar que fue apoyada o no de acuerdo con ciertos datos obtenidos en una investigación particular. Desde el punto de vista técnico, no se acepta una hipótesis a través de un estudio, sino que se aporta evidencia en su favor o en su contra. Cuantas más investigaciones apoyen una hipótesis, más credibilidad tendrá; y, por supuesto, será válida para el contexto (lugar, tiempo y participantes u objetos) en que se comprobó. Al menos lo es probabilísticamente.
las hipótesis, en el enfoque cuantitativo, se someten a prueba en la “realidad” cuando se aplica un diseño de investigación, se recolectan datos con uno o varios instrumentos de medición, y se analizan e interpretan esos mismos datos. Las hipótesis constituyen instrumentos muy poderosos para el avance del conocimiento, puesto que aunque sean formuladas por el ser humano, pueden ser sometidas a prueba y demostrarse como probablemente correctas o incorrectas, sin que interfieran los valores y las creencias del individuo. No siempre los datos apoyan las hipótesis. Pero el hecho de que éstos no aporten evidencia en favor de las hipótesis planteadas de ningún modo significa que la investigación carezca de utilidad. Claro que a todos nos agrada que lo que suponemos concuerde con nuestra realidad.

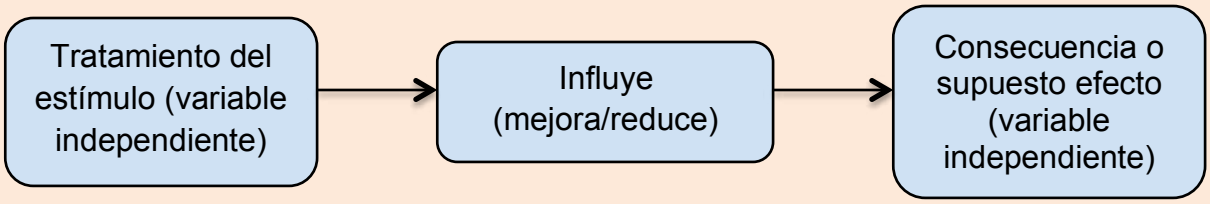
Se recomienda finalmente, mencionar en su reporte, artículo o tesis, *la presentación de los descubrimientos y todos los resultados relevantes, incluyendo aquellos que contradigan las hipótesis,* ofreciendo una discusión que de realce a su investigación validando, comprobando o haciendo lo contrario, dentro de su campo de estudio.

# Diseñando la investigación

Plan o estrategia que se desarrolla para obtener la información que se requiere en una investigación. Sus propósitos son responder las preguntas de investigación, cumplir los objetivos del estudio, someter hipótesis a prueba. En una misma investigación, pueden incluirse dos o más diseños de distintos tipos (diseños múltiples).Ver **Tabla A.35**

**Tabla A.35. Diseño de la investigación**

| Diseño de investigación experimental |
|---|
| •Experimentales (que administran estímulos o tratamientos), y presentan de requisitos:<br>•La manipulación intencional de una o más variables independientes, expresada en grados que para asegurar resultados debe de consultar experimentos antecedentes; evaluar la manipulación antes de realizar el experimento; incluir puntos de verificación durante el experimento<br>•Debe medir el efecto que la variable independiente tiene en la variable dependiente. No hay una recomendación de cuantas variables manejar el experimento, ingresando y suprimiendo según el avance del mismo.<br>•Todo experimento debe cumplir es el *control o la validez interna* de la situación experimental.<br><br>Se dividen los experimentos, en:<br>•*Pre-experimentos.* Con grado de control mínimo<br>•*Cuasi-experimentos.* Implican grupos intactos. manipulan deliberadamente, al menos, una variable independiente para observar su efecto y relación con una o más variables dependientes, sólo que difieren de los experimentos puros" en el grado de seguridad o confiabilidad que pueda tenerse sobre la equivalencia inicial de los grupos. En los diseños cuasiexperimentales los sujetos no se asignan al azar a los grupos ni se emparejan, sino que dichos grupos ya están formados antes del experimento: son grupos intactos (la razón por la que surgen y la manera como se formaron es independiente o aparte del experimento).<br>•*Expermientos puros.* Que implican: manipulación intencional de variables independientes, medición de variables (dependientes), control y validez, dos o más grupos de comparación, participantes asignados al azar. Es típico el uso de: *posprueba y grupo de control; preprueba-posprueba y grupo de control; grupos de Solomon; series cronológicas simples; diseños factoriales*<br><br>•Funcionan con la lógica |




•Los diseños experimentales se utilizan cuando el investigador pretende establecer el posible efecto de una causa que se manipula. Desde luego, hay ocasiones en que no podemos o no debemos experimentar directamente en sujetos y situaciones, como poner deliberadamente en crisis a la organización para verificar cómo responden los gerentes en la toma de decisiones ate situaciones poco éticas y/o situaciones ; no se puede experimenta en el pasado. Siempre deberá garantizarse el control en la cual se manipulan, de manera intencional, una o más variables independientes (causas) para analizar las consecuencias de tal manipulación sobre una o más variables dependientes (efectos).

El *control* en un experimento logra la validez interna y se alcanza mediante:

•Varios grupos de comparación (dos como mínimo). Es necesario que en un experimento se tengan, por lo menos, dos grupos que comparar. En primer término, porque si nada más se tiene un grupo no es posible saber con certeza si influyeron las fuentes de invalidación interna o no. Por ejemplo, si mediante un experimento buscamos probar la hipótesis: *a mayor información científica sobre una clase nueva de tecnología menor prejuicio hacia esta clase*.

Suponga que el nivel de prejuicio antes del estímulo o tratamiento era alto, y después del estímulo, bajo. Quizás el tratamiento sea la causa del cambio, pero tal vez también ocurrió lo siguiente:

•Que la primera prueba de prejuicio sensibilizara a los sujetos participantes y que influyera en sus respuestas a la segunda prueba. Así, las personas crearon conciencia de lo negativo de ser prejuiciosas al responder a la primera prueba (*administración de prueba*).

•Que los individuos seleccionados se agotaran durante el experimento y sus respuestas a la segunda prueba fueran *a la ligera* (*maduración*).

•Que durante el experimento se salieron sujetos prejuiciosos o parte importante de ellos (*mortalidad experimental*).

•O bien otras razones. Y si no se hubiera observado un cambio en el nivel de prejuicio entre la primera prueba (antes del programa) y la segunda (después del programa), esto significaría que la ex- posición al programa no tiene efectos, aunque también podría ocurrir que el grupo seleccionado es muy prejuicioso y tal vez el programa sí tiene efectos en personas con niveles comunes de prejuicio. Asimismo, si el cambio es negativo (mayor nivel de prejuicio en la segunda medición que en la primera), se podría suponer que el programa incrementa el prejuicio, pero supongamos que haya ocurrido un suceso durante el experimento que generó momentáneamente prejuicios hacia esa clase social (una violación en la localidad a cargo de un individuo de esa clase), pero después los participantes *regresa- ron* a su nivel de prejuicio normal (regresión). Incluso podría haber otras explicaciones.

Con un solo grupo no estaríamos seguros de que los resultados se debieran al estímulo experimental o a otras razones. Siempre quedará la duda. Los *experimentos* con un grupo se basan en sospechas o en lo que *aparentemente es*", pero carecen de fundamentos. Al tener un único grupo se corre el riesgo de seleccionar sujetos atípicos (los más inteligentes al experimentar con métodos de enseñanza, los trabajadores más motivados al experimentar con programas de incentivos, los consumidores más críticos, las parejas de novios más integradas, etc.) y de que intervengan la historia, la maduración, y demás fuentes de invalidación interna, sin que el experimentador se dé cuenta.

• *Equivalencia de los grupos en todo*, excepto en la manipulación de la o las variables independientes. Sin embargo, para tener control no basta con dos o más grupos, sino que éstos deben ser similares en todo, menos en la manipulación de la o las variables independientes. El control implica que todo permanece constante, salvo tal manipulación o intervención. Si entre los grupos que conforman el experimento todo es similar o equivalente, excepto la manipulación de la variable independiente, las diferencias entre los grupos pueden atribuirse a ella y no a otros factores (entre los cuales están las fuentes de invalidación interna).

• *Equivalencia inicial*. Implica que los grupos son similares entre sí al momento de iniciarse el experimento. Se logra por adoptar técnicas variadas como:

• Asignación al azar. técnica de control muy difundida para asegurar la equivalencia inicial al ser asigna- dos aleatoriamente los sujetos a los grupos del experimento.

• *Técnica del apareo.* Consiste en igualar a los grupos en relación con alguna variable específica, que puede influir de modo decisivo en la variable dependiente.

• Los experimentos deben buscar siempre validez interna (resultados confiables) y validez externa (posibilidad de generalizar los resultados de un experimento a situaciones no experimentales, así como a otras personas y poblaciones.

• El experimento deberá fincarse en un contexto de campo , o sea, en una situación más real o natural en la que el investigador manipula una o más variables. Un ejemplo es el Experimento en que el efecto de todas o casi todas las variables independientes influyentes no concernientes al problema de investigación se mantiene reducido lo más posible.

Pasos generales para realizar de un experimento

| Paso | Descripción |
|---|---|
| 1 | • Decidir cuántas variables independientes y dependientes deberán incluirse en el experimento. No necesariamente el mejor experimento es el que incluye el mayor número de variables; deben incluirse las variables que sean necesarias para probar las hipótesis, alcanzar los objetivos y responder las preguntas de investigación. |
| 2 | • Elegir los niveles o modalidades de manipulación de las variables independientes y traducirlos en tratamientos experimentales. |
| 3 | • Desarrollar el instrumento o instrumentos para medir la(s) variable(s) dependiente(s). |
| 4 | • Seleccionar para el experimento una muestra de personas que posean el |

| | |
|---|---|
| | perfil que nos interesa. |
| 5 | •Reclutar a los participantes del experimento. Esto implica tener contacto con ellos, darles las explicaciones necesarias, obtener su consentimiento e indicarles lugar, día, hora y persona con quien deben presentarse. Siempre es conveniente darles el máximo de facilidades para que acudan al experimento (si se les puede brindar transporte en caso de que sea necesario, proporcionarles un mapa con las indicaciones precisas, etc.).<br>•También hay que darles cartas (a ellos o alguna institución a la que pertenezcan para facilitar su participación en el experimento; por ejemplo, en escuelas a los directivos, maestros y padres de familia), llamarles por teléfono el día anterior a la realización del experimento para recordarles su participación.<br>•Las personas deben encontrar motivante su participación en el experimento. Por tanto, resulta muy conveniente darles algún regalo atractivo (a veces simbólico). Por ejemplo, a amas de casa, una canasta de productos básicos; a ejecutivos o gerentes, una canasta con dos o tres artículos; a estudiantes, créditos escolares, etc., además de expedirles una carta de agradecimiento. |
| 6 | •Seleccionar el diseño experimental o cuasiexperimental apropiado para nuestras hipótesis, objetivos y preguntas de investigación. |
| 7 | •Planear cómo va a manejar a los participantes del experimento. Es decir, elaborar una ruta crítica de qué van a hacer las personas desde que llegan al lugar del experimento hasta que se re iran. |

**No experimental**

Son estudios que se realizan sin la manipulación deliberada de variables y en los que sólo se observan los fenómenos en su ambiente natural para después analizarlos. No se genera ninguna situación, sino que se observan situaciones ya existentes, no provocadas intencionalmente en la investigación por quien la realiza. En la investigación no experimental las variables independientes ocurren y no es posible manipularlas, no se tiene control directo sobre dichas variables ni se puede influir sobre ellas, porque ya sucedieron, al igual que sus efectos.

La investigación experimental tiene alcances iniciales y finales *correlacionales y explicativos*. La investigación *no experimental es sistemática y empírica* en la que las variables independientes no se manipulan porque ya han sucedido. Las inferencias sobre las relaciones entre variables se realizan sin intervención o influencia directa, y dichas relaciones se observan tal como se han dado en su contexto natural. Son de tipo:

•*Transeccional,* que recolectan los datos en un único momento, siendo de tipo tipo:

•*Exploratorios.* Comlenzan a conocer una variable o un conjunto de variables, una comunidad, un contexto, un evento, una situación. Se trata de una exploración inicial en un momento específico. Por lo general, se aplican a problemas de investigación nuevos o poco conocidos, además constituyen el preámbulo de otros diseños (no experimentales y experimentales).

•*Descriptivos.* Indagan la incidencia de las modalidades, categorías o niveles de

una o más variables en una población, son estudios puramente descriptivos.
Cabe destacar aquí, que una *encuesta,* sería una investigación no experimental transversal o transeccional descriptiva o de correlación-causal, ya que a veces tienen los propósitos de unos u otros diseños y a veces de ambos. Generalmente utilizan cuestionarios que se aplican en diferentes contextos (aplicados en entrevistas *cara a cara*, mediante *correo electrónico o postal*, en grupo).

•*Correlaciones-causales.* Estos diseños describen relaciones entre dos o más categorías, conceptos o variables en un momento determinado. A veces, únicamente en términos correlacionales, otras en función de la relación causa- efecto(causales).

•*Longitudinales* o evolutivos. Analizan cambios a través del tiempo. Son de tipo:

•*Diseño de tendencia* (*trend*), son aquellos que analizan cambios a través del tiempo (en categorías, conceptos, variables o sus relaciones), dentro de alguna población en general. Su característica distintiva es que la atención se centra en la población.

•*De análisis evolutivo de grupos* (*cohort evolution*), examinan cambios a través del tiempo en subpoblaciones o grupos específicos. Su atención son las *cohortes* o grupos de individuos vinculados de alguna manera o identificados por una característica común, generalmente la edad o la época El *diseños de tendencia y de evolución de grupo*, ambos monitorean cambios en una población o subpoblación a través del tiempo, usando una serie de muestras que abarcan a diferentes participantes en cada ocasión, pero en los primeros la población es la misma y en los segundos se toma como universo a los sobrevivientes de la población.

•*Diseños de panel.* son similares a las dos clases de diseños vistas anteriormente, sólo que los mismos participantes son medidos u observados en todos los tiempos o momentos.

•*Estudios de caso.* Son estudios que al utilizar los procesos de investigación cuantitativa, cualitativa o mixta; analizan profundamente una unidad para responder al planteamiento del problema, probar hipótesis y desarrollar alguna teoría. En ocasiones utilizan la experimentación, es decir, se constituyen en estudios preexperimentales. Otras veces se fundamentan en un diseño no experimental (transversal o longitudinal) y en ciertas situaciones se convierten en estudios cualitativos, al emplear métodos cualitativos. Asimismo, pueden valerse de las diferentes herramientas de la investigación mixta. la elección entre la muestra probabilística y la no probabilística se determina con base en el planteamiento del problema, las hipótesis, el diseño de investigación y el alcance de sus contribuciones. Las muestras probabilísticas tienen muchas ventajas, quizá la principal sea que puede medirse el tamaño del error en nuestras predicciones. Se dice incluso que el principal objetivo en el diseño de una muestra probabilística es reducir al mínimo este error, al que se le llama error estándar

**Tipos de estudio, hipótesis y diseño de Investigación**

| Estudio | Hipótesis | Posibles diseños |
|---|---|---|
| Exploratorio | •No se establecen, lo que se puede formular son conjeturas iniciales | •Transeccional descriptivo<br>•Preexperimental |
| Descriptivo | •Descriptiva | •Preexperimental<br>•Transeccional descriptivo |
| Correlacional | •Diferencia de grupos sin atribuir causalidad | •Cuasiexperimental<br>•Transeccional correlacional<br>•Longitudinal (no experimental) |
| | Correlacional | •Cuasiexperimental<br>•Transeccional correlacional<br>•Longitudinal (no experimental) |
| Explicativo | •Diferencia de grupos atribuyendo causalidad | •Experimental<br>•Cuasiexperimental, longitudinal y transeccional causal (cuando hay bases para inferir causalidad, un mínimo de control y análisis estadísticos apro- piados para analizar relaciones causales) |
| | •Causales | •Experimental<br>•Cuasiexperimental, longitudinal y transeccional causal (cuando hay bases para inferir causalidad, un mínimo de control y análisis estadísticos apro- piados para analizar relaciones causales) |

Fuente: Hernández-Sampieri et al., (2010) con adaptación propia

# Muestreo cuantitativo

En la mayoría de las situaciones sí realizamos el estudio en una muestra. Sólo cuando queremos realizar un censo debemos incluir en el estudio a todos los casos (personas, animales, plantas, objetos) del universo o la población. Así, la muestra es *un subgrupo de la población del cual se recolectan los datos y debe ser representativo de ésta, para generalizar características, hipótesis a fin de construir teorías que expliquen a la población o fenómeno, mediante técnicas apropiadas.* Por lo tanto, una *población* es el conjunto de todos los casos que concuerdan con una serie de especificaciones. Ver **Tabla A.36.**

**Tabla A.36. Tipos de muestra cuantitativa**

| Muestra probabilística |
|---|
| •Subgrupo de la población en el que todos los elementos de ésta tienen la misma posibilidad de ser elegidos . todos los elementos de la población tienen la misma posibilidad de ser escogidos y se obtienen definiendo las características de la población y el tamaño de la muestra, y por medio de una selección aleatoria o mecánica de las unidades de análisis. Imagínese el procedimiento para obtener el número premiado en un sorteo de lotería.<br>Este número se va formando en el momento del sorteo. En las loterías tradicionales, a partir de las esferas con un dígito que se extraen (después de revolverlas mecánicamente) hasta formar el número, de manera que todos los números tienen la misma probabilidad de ser elegidos. Las muestras probabilísticas son esenciales en los diseños de investigación transeccionales, tanto descriptivos como correlacionales-causales (las encuestas de opinión o *surveys*, por ejemplo), donde se pretende hacer estimaciones de variables en la población. Estas variables se miden y se analizan con pruebas estadísticas en una muestra, donde se presupone que ésta es probabilística y todos los elementos de la población tienen una misma probabilidad de ser elegidos. Las unidades o elementos muestrales tendrán valores muy parecidos a los de la población, de manera que las mediciones en el subconjunto nos darán estimados precisos del conjunto mayor. La precisión de dichos estimados depende del error en el muestreo, que es posible calcular.<br>Para hacer una muestra probabilística son necesarios dos procedimientos:<br>•Calcular un tamaño de muestra que sea representativo de la población. Uso de software estadístico **SPSS, MATLAB, STATS**, etc.<br>•Seleccionar los elementos muestrales (casos) de manera que al inicio todos tengan la misma posibilidad de ser elegidos. Este consiste en seleccionar los elementos muestrales de manera que al inicio todos tengan la misma posibilidad de ser elegidos. Es decir, cómo y de dónde vamos a elegir los casos. Esto se comentará más adelante. todos los casos del universo tienen al inicio la misma probabilidad de ser seleccionados.<br>•Debe preguntarse: dado que una población es de *N tamaño*, ¿cuál es el menor número de unidades muestrales (personas, organizaciones, capítulos de telenovelas, etc.) que necesito para conformar una muestra (*n*) que me asegure un determinado nivel de error estándar, digamos menor de 0.01?<br>La respuesta a esta pregunta busca encontrar una muestra que sea representativa del universo o población con cierta posibilidad de error (se pretende minimizar) y nivel de confianza (maximizar), así como probabilidad.<br>Pues bien, algo similar hace el investigador al definir un posible nivel de error en la representatividad de su muestra. Los niveles de error más comunes que suelen fijarse en la investigación son de 5 y 1% (en ciencias sociales el más usual es el primero).<br>El porcentaje estimado de la muestra es la probabilidad de ocurrencia del fenómeno (representatividad de la muestra *versus* no representatividad, la cual se estima sobre marcos de muestreo previos o se define, la certeza total siempre es igual a uno, las posibilidades a partir de esto son ***p*** de que sí ocurra y ***q*** de que no ocurra<br>(***p* + *q* = 1**). Cuando no tenemos marcos de muestreo previo, usamos un porcentaje |

estimado de 50% (que es la opción por *default* que nos brinda **SPSS,STATS, MATLAB**, es decir, asumimos que ***p*** y ***q*** serán de 50%, y que resulta lo más común, particularmente cuando seleccionamos por vez primera una muestra en una población).

Finalmente, el nivel deseado de confianza es el complemento del error máximo aceptable (porcen- taje de *acertar en la representatividad de la muestra*). Si el error elegido fue de 5%, el nivel deseado de confianza será de 95%. Una vez más los niveles más comunes son de 95 y 99%.

La selección entre la muestra probabilística y la no probabilística se determina con base en el planteamiento del problema, las hipótesis, el diseño de investigación y el alcance de sus contribuciones. Las muestras probabilísticas tienen muchas ventajas, quizá la principal sea que puede medirse el tamaño del error en nuestras predicciones. Se dice incluso que el principal objetivo en el diseño de una muestra probabilística es reducir al mínimo este error, al que se le llama error estándar

Las muestras probabilísticas son esenciales en los diseños de investigación *transeccionales, tanto descriptivos como correlacionales-causales* (las encuestas de opinión o *surveys*, por ejemplo), donde se pretende hacer estimaciones de variables en la población. Estas variables se miden y se analizan con pruebas estadísticas en una muestra, donde se presupone que ésta es probabilística y todos los elementos de la población tienen una misma probabilidad de ser elegidos. Las unidades o elementos muestrales tendrán valores muy parecidos a los de la población, de manera que las mediciones en el subconjunto nos darán estimados precisos del conjunto mayor. La precisión de dichos estimados depende del error en el muestreo, que es posible calcular. Ver **Apéndice 2.**

| Muestra probabilística estratificada |
|---|
| Muestreo en el que la población se divide en segmentos y se selecciona una muestra para cada segmento. En ocasiones el interés del investigador es comparar sus resultados entre segmentos, grupos o nichos de la población, porque así lo señala el planteamiento del problema. Por ejemplo, efectuar comparaciones por género (entre hombres y mujeres), si la selección de la muestra es aleatoria, tendremos unidades o elementos de ambos géneros, no hay problema, la muestra reflejará a la población. A veces, nos interesan grupos que constituyen minorías de la población o universo y entonces si la muestra es aleatoria simple, resultará muy difícil determinar qué elementos o casos de tales grupos serán seleccionados. Quizá tengamos 300 católicos y dos o tres de otras religiones. Entonces es cuando preferimos obtener una *muestra probabilística estratificada* (el nombre nos dice que será probabilística y que se considerarán segmentos o grupos de la población, o lo que es igual: estratos). La estratificación aumenta la precisión de la muestra e implica el uso deliberado de diferentes tamaños de muestra para cada estrato, a fin de lograr reducir la varianza de cada unidad de la media muestral. |
| Muestreo probabilísticos por racimos |
| Muestreo en el que las unidades de análisis se encuentran encapsula- das en determinados lugares físicos. Son sinónimos de *clusters* o conglomerados, En algunos casos en que el investigador se ve limitado por recursos financieros, por |

tiempo, por distancias geográficas o por una combinación de éstos y otros obstáculos, se recurre al *muestreo por racimos* o *clusters*. En este tipo de muestreo se reducen costos, tiempo y energía, al considerar que muchas veces las unidades de análisis se encuentran encapsuladas o encerradas en determinados lugares físicos o geográficos, a los que se denomina *racimos***.** Muestrear por racimos implica diferenciar entre la unidad de análisis y la unidad muestral. La unidad de análisis indica quiénes van a ser medidos, o sea, los participantes o casos a quienes en última instancia vamos a aplicar el instrumento de medición.

El muestreo por racimos supone una selección en dos o más etapas, todas con procedimientos probabilísticos. En la primera, se seleccionan los racimos, siguiendo los pasos ya señalados de una muestra probabilística simple o estratificada. En las fases subsecuentes y dentro de estos racimos, se seleccionan los casos que van a medirse. Para ello se hace una selección que asegure que todos los elementos del racimo tienen la misma probabilidad de ser elegidos. Implica diferenciar entre la unidad de análisis y la unidad muestral. La unidad de análisis indica quiénes van a ser medidos, o sea, los participantes o casos a quienes en última instancia vamos a aplicar el instrumento de medición. La unidad muestral (en este tipo de muestra) se refiere al racimo por medio del cual se logra el acceso a la unidad de análisis. El muestreo por racimos supone una selección en dos o más etapas, todas con procedimientos probabilísticos. En la primera, se seleccionan los racimos, siguiendo los pasos ya señalados de una muestra probabilística simple o estratificada. En las fases subsecuentes y dentro de estos racimos, se seleccionan los casos que van a medirse. Para ello se hace una selección que asegure que todos los elementos del racimo tienen la misma probabilidad de ser elegidos. En ocasiones se combinan tipos de muestreo, por ejemplo: una muestra probabilística estratifica- da y por racimos, pero siempre se utiliza una selección aleatoria que garantiza que al inicio del procedimiento todos los elementos de la población tienen la misma probabilidad de ser elegidos para integrar la muestra. Los tipos de muestra dependen de dos cosas: del tamaño de la muestra y del procedimiento de selección. De lo primero hemos hablado con detalle, de lo segundo trataremos ahora. Se determina el tamaño de la muestra *n*, pero ¿cómo seleccionar los elementos muestrales? (ya sean casos o racimos). Las unidades de análisis o los elementos muestrales se eligen siempre aleatoriamente para asegurarnos de que cada elemento tenga la misma probabilidad de ser elegido. Los tipos de muestra dependen de dos cosas: del tamaño de la muestra y del procedimiento de selección.

**Procedimientos de selección de la muestra**

| Tómbola |
| --- |
| Muy simple pero muy rápido, consiste en numerar todos los elementos muestrales de la población, del uno al número *N*. Después se hacen fichas o papeles, uno por cada elemento, se revuelven en una caja y se van sacando *n* número de fichas, según el tamaño de la muestra. Los números elegidos al azar conformarán la muestra. Cuando nuestro muestreo es estratificado, se sigue el procedimiento anterior, pero por cada estrato. |

## Números random aleatorios

Una excelente alternativa para generar números aleatorios se encuentra en el programa **SPSS, MATLAB, STATS**, o cualquier software estadístico que contiene un subrutina como Montecarlo, para ello y evita el uso de la tabla de números aleatorios. El programa nos pide que le indiquemos ¿cuántos números aleatorios? (requerimos), entonces tecleamos el tamaño de muestra; se elige elegimos la opción: *Establecer límite superior e inferior.* y nos solicita que establezcamos el límite inferior (que siempre será uno, el primer caso de la población, pues la muestra se extrae de ésta) y el límite superior (el último número de la población, que es el tamaño de la población). Y tecleamos *Calcular* y nos genera automáticamente los números. Vemos contra nuestro listado a quién o a qué corresponde cada número y estos números son los casos que pasarían a integrar la muestra.

## Selección sistemática de elementos muestrales

Este procedimiento de selección es muy útil e implica elegir dentro de una población *N* un número *n* de elementos a partir de un intervalo *K*. Este último (*K* ) es un intervalo que se va a determinar por el tamaño de la población y el tamaño de la muestra. De manera que tenemos que $K = N/n$, en donde *K* = un intervalo de selección sistemática, *N* = la población y *n* = la muestra. Supongamos que se quiere hacer un estudio que pretende medir la calidad de la atención en los servicios proporcionados por los médicos y las enfermeras de un hospital. Para tal efecto consideremos que los investigadores consiguen grabaciones de todos los servicios efectuados durante un periodo determinado.[5] Supongamos que se hayan filmado 1 548 servicios (*N* ). Con este dato se procede a determinar qué número de servicios necesitamos analizar para generalizar a toda la población nuestros resultados. Con STATS® determinamos que se necesitan 307.9 (308) servicios para evaluar (con un error máximo de 5%, nivel de confianza de 95% y un porcentaje estimado de 50% para la muestra [$p = 0.5$]). Si necesitamos una muestra de $n = 308$ episodios de servicio filmados, se utiliza para la selección el intervalo *K*, donde:

$$K=(N/n)= 1548/308=5.0259= 5$$

La selección sistemática de elementos muestrales $1/K$ se puede utilizar al elegir los elementos de *n* para cada estrato o para cada racimo. La regla de probabilidad, según la cual cada elemento de la población debe tener idéntica probabilidad de ser elegido, se cumple al empezar la selección de $1/K$ al azar. Siguiendo nuestro ejemplo, no comenzamos a elegir de los 1 548 episodios, el 1, 6, 11, 16..., sino que procuramos que el inicio sea determinado por el azar. Así, en este caso, podemos tirar unos dados y si en sus caras muestran 1, 6, 9, iniciaremos en el servicio 169, y seguiremos 174, 179, 184, 189... $1/K$... y volveremos a empezar por los primeros si es necesario. Este procedimiento de selección es poco complicado y tiene varias ventajas: cualquier tipo de estratos en una población *X* se verán reflejados en la muestra. Asimismo, la selección sistemática logra una muestra proporcionada, ya que, por ejemplo, tenemos que

el procedimiento de selección 1/*K* nos dará una muestra con nombres que inician con las letras del abecedario, en forma proporcional a la letra inicial de los nombres de la población.

## Listados y otros narcos muestrales

Es un marco de referencia que nos permite identificar físicamente los elementos de la población, así como la posibilidad de enumerarlos y seleccionar los elementos muestrales. Las *muestras probabilísticas* requieren la determinación del tamaño de la muestra y de un proceso de selección aleatoria que asegure que todos los elementos de la población tengan la misma probabilidad de ser elegidos. Todo esto lo hemos visto, aunque nos falta exponer sobre algo esencial que precede a la selección de una muestra: el marco muestral. Éste constituye un marco de referencia que nos permita identificar físicamente los elementos de la población, la posibilidad de enumerarlos y, por ende, de proceder a la selección de los elementos muestrales (los casos de la muestra). Normalmente se trata de un listado existente o una lista que es necesario confeccionar *ad hoc*, con los casos de la población. Los listados existentes sobre una población son variados: guías telefónicas, listas de miembros de las asociaciones, directorios especializados, listas oficiales de escuelas de la zona, bases de datos de los alumnos de una universidad o de los clientes de una empresa, registros médicos, catastros, nóminas de una organización, etc. En todo caso hay que tener en cuenta lo completo de una lista, su exactitud, su veracidad, su calidad y su nivel de cobertura en relación con el problema a investigar y la población que va a medirse, ya que todos estos aspectos influyen en la selección de la muestra.

Los listados existentes sobre una población son variados: guías telefónicas, listas de miembros de las asociaciones, directorios especializados, listas oficiales de escuelas de la zona, bases de datos de los alumnos de una universidad o de los clientes de una empresa, registros médicos, catastros, nóminas de una organización, etc. En todo caso hay que tener en cuenta lo completo de una lista, su exactitud, su veracidad, su calidad y su nivel de cobertura en relación con el problema a investigar y la población que va a medirse, ya que todos estos aspectos influyen en la selección de la muestra.

Con frecuencia es necesario construir listas *ad hoc*, a partir de las cuales se elegirán los elementos que constituirán las unidades de análisis en una determinada investigación. No siempre existen listas que permitan identificar a nuestra población. Entonces, será necesario recurrir a otros marcos de referencia que contengan descripciones del material, las organizaciones o los casos que serán seleccionados como unidades de análisis. Ejemplos de algunos de estos marcos de referencia son los archivos, los mapas y los archivos electrónicos de periódicos en la web. De cada una de estas instancias daremos ejemplos con más detalles y recomendaremos soluciones para algunos problemas comunes en el muestreo.

## Archivos

Cuando en el investigador accede a archivos donde no hay una lista elaborada de los sujetos de estudio. Tales archivos constituyen su marco muestral a partir del cual se obtendrá la muestra. Calcula el tamaño de la población, obtiene el

tamaño de la muestra y selecciona sistemáticamente cada elemento 1/*K*, cada solicitud que será analizada.

## Mapas

Los mapas son muy útiles como marco de referencia en muestras de racimos. Por ejemplo, un investigador quiere saber qué motiva a los compradores de las tiendas de autoservicio. A partir de una lista de tiendas de cada cadena competidora marca sobre un mapa de la ciudad, todas las tiendas de auto-servicios, las cuales constituyen una población de racimos, pues en cada tienda seleccionada entrevistará a un número de clientes. El mapa le permite ver la población (tiendas de autoservicio) y su situación geográfica, de manera que elige zonas donde coexistan diferentes tiendas competidoras, para asegurarse de que el consumidor de la zona tenga todas las posibles alternativas. En la actualidad hay mapas de todo tipo: mercadológicos, socioculturales, étnicos, marítimos, entre otros. El *Global Positioning System* (GPS) o Sistema de Posicionamiento Global ya puede ser muy útil para esta clase de muestreo.

## Tamaño óptimo de muestra

Las muestras probabilísticas requieren dos procedimientos básicos:

- La determinación del tamaño de la muestra y
- La selección aleatoria de los elementos muestrales. Precisar adecuadamente el tamaño de la muestra puede tornarse complejo, esto depende del problema de investigación y la población a estudiar.
- El *teorema del límite central* Señala que una muestra de más de cien casos será una muestra con una distribución normal en sus características, lo cual sirve para el propósito de hacer estadística inferencial. Lo óptimo de una muestra depende de cuánto se aproxima su distribución a la distribución de las características de la población. Esta aproximación mejora al incrementarse el tamaño de la muestra. La "normalidad" de la distribución en muestras grandes no obedece a la normalidad de la distribución de una población. La distribución de diversas variables a veces es *normal* y en ocasiones está lejos de serlo. Sin embargo, la distribución de muestras de 100 o más elementos tiende a ser normal y esto sirve para el propósito de hacer estadística inferencial (generalizar de la muestra al universo). A lo anterior se le llama *teorema del límite central.*
- *Distribución normal:* esta distribución en forma de campana se logra general-mente con muestras de 100 o más unidades muestrales, y es útil y necesaria cuando se hacen inferencias de tipo estadístico.

## Muestra no probabilística

- Subgrupo de la población en la que la elección de los elementos no depende de la probabilidad sino de las características de la investigación.
- La elección de los elementos no depende de la probabilidad, sino de causas relacionadas con las características de la investigación o de quien hace la muestra.
- Aquí el procedimiento no es mecánico ni con base en fórmulas de probabilidad, sino que depende del proceso de toma de decisiones de un investigador o de un grupo de investigadores y, desde luego, las muestras seleccionadas obedecen a otros criterios de investigación.

•Elegir entre una muestra probabilística o una no probabilística depende de los objetivos del estudio, del esquema de investigación y de la contribución que se piensa hacer con ella.

•Las muestras no probabilísticas, también llamadas muestras dirigidas, suponen un procedimiento de selección informal. Se utilizan en diversas investigaciones cuantitativas y cualitativas.

•Seleccionan individuos o casos *típicos* sin intentar que sean representativos de una población determinada. Por ello, para fines deductivos-cuantitativos, donde la generalización o extrapolación de resultados hacia la población es una finalidad en sí misma, las muestras dirigidas implican algunas desventajas. La primera es que, al no ser probabilísticas, no es posible calcular con precisión el error estándar, es decir, no podemos calcular con qué nivel de confianza hacemos una estimación. Esto es un grave inconveniente si consideramos que la estadística inferencial se basa en la teoría de la probabilidad, por lo que las pruebas estadísticas en muestras no probabilísticas tienen un valor limitado a la muestra en sí, mas no a la población.

•Los datos no pueden generalizarse a ésta. En las muestras de este tipo, la elección de los casos no depende de que todos tengan la misma probabilidad de ser elegidos, sino de la decisión de un investigador o grupo de personas que recolectan los datos. La única ventaja de una muestra no probabilística) desde la visión cuantitativa) es su utilidad para determinado diseño de estudio que requiere no tanto una *representatividad* de elementos de una población, sino una cuidadosa y controlada elección de casos con ciertas características especificadas previamente en el planteamiento del problema.

•Para el enfoque cualitativo, al no interesar tanto la posibilidad de generalizar los resultados, las muestras no probabilísticas o dirigidas son de gran valor, pues logran obtener los casos (personas, contextos, situaciones) que interesan al investigador y que llegan a ofrecer una gran riqueza para la recolección y el análisis de los datos.

**Muestreo al azar por marcado telefónico**

•Ésta es una técnica que los investigadores utilizan para seleccionar muestras telefónicas. Involucra identificar áreas geográficas (para ser muestreadas al azar) y sus correspondientes códigos telefónicos e intercambios (los primeros dígitos del número telefónico que las identifican).

•Luego, los demás dígitos del número a marcar pueden ser generados al azar de acuerdo con los casos que requerimos para la muestra (*n*). Es posible reconocer qué intercambios son usados de forma primaria para teléfonos residenciales y enfocar el muestreo en ese subgrupo. Asimismo, es muy útil para incluir en muestras a teléfonos celulares o móviles.

Fuente: Hernández-Sampieri et al., (2010) con adaptación propia

Ya sea que se trate de un tipo de muestreo u otro, lo importante es elegir a los informantes (o casos) adecuados, de acuerdo con el planteamiento del problema y lograr el acceso a ellos.

Los estudios exploratorios regularmente emplean muestras dirigidas, aunque podrían usarse muestras probabilísticas. Las investigaciones experimentales, la mayoría de las veces utilizan muestras dirigidas, porque como se comentó, es difícil manejar grupos grandes (debido a ello se ha insistido que, en los experimentos, la validez

externa se consolida mediante la repetición o reproducción del estudio). Los estudios no experimentales descriptivos o correlacionales-causales deben emplear mues- tras probabilísticas si quieren que sus resultados sean generalizados a una población.
Asimismo, en ocasiones la muestra puede ser en varias etapas (*polietápica*). Por ejemplo, primero elegir universidades, luego, escuelas o facultades, después, salones o grupos y finalmente, estudiantes.

# Muestreo cualitativo

La muestra, en el proceso cualitativo, es un grupo de personas, eventos, sucesos, comunidades, etc., sobre el cual se habrán de recolectar los datos, sin que necesariamente sea representativo del universo o población que se estudia . En la a inmersión inicial, la cual nos sumerge en el contexto, a la par recolectamos y analizamos datos (seguramente ya observamos diferentes sucesos, nos compenetramos con la cotidianidad del ambiente, platicamos o entrevistamos a varias personas, tomamos notas, tenemos impresiones, etcétera).
En algún momento de la inmersión inicial o después de ésta, se define la muestra *tentativa*, sujeta a la evolución del proceso inductivo. Las primeras acciones para elegir la *muestra* ocurren desde el planteamiento mismo y cuando seleccionamos el contexto, en el cual esperamos encontrar los casos que nos interesan. En las investigaciones cualitativas nos preguntamos: ¿qué casos nos interesan inicialmente y dónde podemos encontrarlos.
En los estudios cualitativos el tamaño de muestra *no* es importante desde una perspectiva probabilística, *pues el interés del investigador no es generalizar los resultados de su estudio a una población más amplia.* Lo que se busca en la indagación cualitativa es profundidad. Nos conciernen casos (participantes, personas, organizaciones, eventos, animales, hechos, etc.) que nos ayuden a entender el fenómeno de estudio y a responder a las preguntas de investigación. El muestreo adecuado tiene una importancia crucial en la investigación, y la investigación cualitativa no es una excepción. Por esta razón es necesario reflexionar detenidamente sobre cuál es la estrategia de muestreo más pertinente para lograr los objetivos de investigación, tomando en cuenta criterios de rigor, estratégicos, éticos y pragmáticos como se explicará a continuación. Ver **Tabla A.37**

**Tabla A.37. Tipos de muestra cualitativa**

| Factores para determinar el número de casos |
|---|
| •Capacidad operativa de recolección y análisis (*el número de casos que podemos manejar de manera realista y de acuerdo con los recursos que dispongamos*).<br>•El entendimiento del fenómeno (el número de casos que nos permitan responder a las preguntas de investigación, que más adelante se denominará *saturación de categorías*).<br>•La naturaleza del fenómeno bajo análisis (*si los casos son frecuentes y accesibles o no, si el recolectar información sobre éstos lleva relativamente poco o mucho tiempo).* |
| Tamaño de la muestra |
| •En la indagación cualitativa *el tamaño de muestra no se fija a priori* (previamente a |

la recolección de los datos), sino que se establece un tipo de unidad de análisis y a veces se perfila un número relativamente aproximado de casos, pero la muestra final se conoce cuando las unidades que van adicionándose no aportan información o datos novedosos (*saturación de categorías*), aun cuando agreguemos casos extremos .la decisión del número de casos que conformen la muestra es del investigador.

•Las muestras cualitativas no deben ser utilizadas para representar a una población los estudios cualitativos son artesanales; son *trajes hechos a la medida de las circunstancias.* en una investigación cualitativa la muestra puede contener cierto tipo definido de unidades iniciales, pero conforme avanza el estudio se pueden agregar otros tipos de unidades y aun desechar las primeras unidades. Así, analizaría tanto a los protagonistas de las interacciones como a éstas y sus procesos.

•También, se pueden tener unidades cuya naturaleza es diferente. (Por ejemplo, documentos de la época vs. Testigos)

**Sugerencias de tamaño de muestra cualitativa por tipo de estudio**

| Tipo de estudio | Tamaño mínimo de muestra |
|---|---|
| Entrevistas, observaciones, teoría fundamentada | 30 a 50 casos |
| Historia de una Pyme | Toda la Pyme, cada uno de sus miembros |
| Casos de profundidad | 6 a 10 casos |
| Estudio de caso | 1 a 5 casos |
| Grupos de enfoque | De 7 a 10 casos por grupo, cuatro grupos por cierto tipo de población |
| Biografía | Sujeto central vivo y la mayor cantidad de sobrevivientes |

•En el muestreo cualitativo es usual comenzar con la identificación de ambientes propicios, luego de grupos y, finalmente, de individuos. Incluso, la muestra puede ser una sola unidad de análisis (*estudio de caso*). La investigación cualitativa, por sus características, requiere de muestras más flexibles.

Así, se la esencia del muestreo cualitativo, con: *un objetivo central al seleccionar ambientes y casos que nos ayuden a entender con mayor profundidad un fenómeno y aprender de éste. Lo anterior sirve para: entender significados, actores, e información utilizando técnicas de muestreo con un propósito definido y acorde con la evolución de los acontecimientos.*

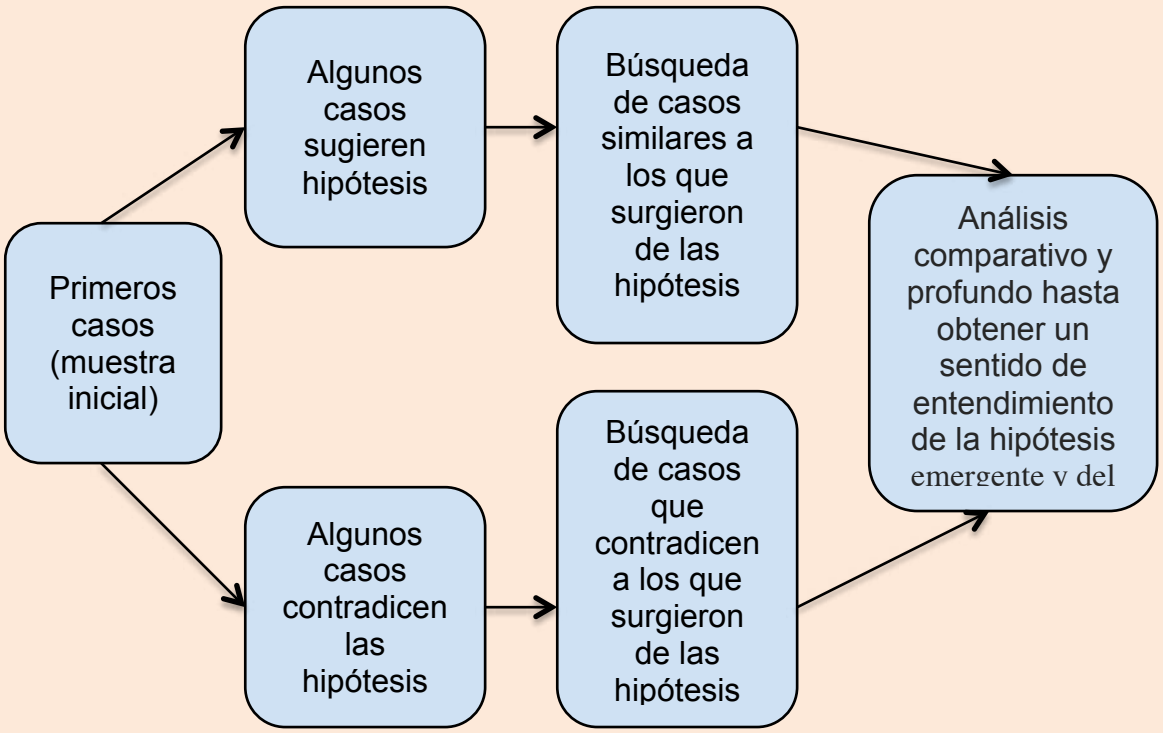
•Los tipos de muestras que suelen utilizarse en las investigaciones son las *no probabilísticas* o *dirigidas*, cuya finalidad no es la generalización en términos de probabilidad. También se les conoce como “guiadas por uno o varios propósitos”, pues la elección de los elementos depende de razones relaciona- das con las características de la investigación. Veamos estas clases de muestras, pero cabe destacar que *no son privativas de los estudios cualitativos*, también llegan a utilizarse en investigaciones cuantitativas, pero se asocian más con los primeros.

| Muestra de participantes voluntarios |
|---|
| Son frecuentes en ciencias sociales y médicas . n estos casos, la elección de los participantes depende de circunstancias muy variadas. A esta clase de muestra también se le puede llamar *autoseleccionada*, ya que las personas se proponen como participantes en el estudio o responden activamente a una invitación. Este tipo de muestras se usa en estudios experimentales de laboratorio, pero también en investigaciones cualitativas, |
| **Muestra de expertos** |
| En ciertos estudios es necesaria la opinión de individuos expertos en un tema. Estas muestras son frecuentes en estudios cualitativos y exploratorios para generar hipótesis más precisas o la materia prima del diseño de cuestionarios.. |
| **Muestra de casos-tipo** |
| También esta muestra se utiliza en estudios cuantitativos exploratorios y en investigaciones de tipo cualitativo, donde el objetivo es la riqueza, profundidad y calidad de la información, no la cantidad ni la estandarización. En estudios con perspectiva fenomenológica, donde el objetivo es analizar los valores, ritos y significados de un determinado grupo social, el uso de muestras tanto de expertos como de casos-tipo es frecuente. |
| **Muestra de casos-tipo** |
| Este tipo de muestra se utiliza mucho en estudios de opinión y de marketing. Por ejemplo, los encuestadores reciben instrucciones de administrar cuestionarios a individuos en un lugar público (un centro comercial, una plaza o una colonia), al hacerlo van conformando o llenando cuotas de acuerdo con la proporción de ciertas variables demográficas en la población. |
| **Muestras más bien orientadas hacia la investigación cualitativa** |
| Es posible considerar otras mues- tras no probabilísticas que, además de las ya señaladas, suelen utilizarse en estudios cualitativos:<br>•*Muestras diversas o de máxima variación*: son utilizadas cuando se busca mostrar distintas perspectivas y representar la complejidad del fenómeno estudiado, o bien, documentar diversidad para localizar diferencias y coincidencias, patrones y particularidades.<br>•*Muestras homogéneas*: al contrario de las muestras diversas, en éstas las unidades a seleccionar poseen un mismo perfil o características, o bien, comparten rasgos similares. Su propósito es centrarse en el tema a investigar o resaltar situaciones, procesos o episodios en un grupo social. Una variante es la llamada *muestras típicas o intensivas*, que eligen casos de un perfil similar, pero que se consideran representativos de un segmento de la población, una comunidad o una cultura (no en un sentido estadístico, sino de prototipo).<br>•*Muestras en cadena o por redes* (*bola de nieve*): se identifican participantes clave y se agregan a la muestra, se les pregunta si conocen a otras personas que puedan proporcionar datos más amplios, y una vez contactados, los incluimos también. *Muestras de casos extremos*: útiles cuando nos interesa evaluar características, situaciones o fenómenos especiales, alejados de la *normalidad.*<br>•*Muestras por oportunidad*: casos que de manera fortuita se presentan ante el investigador, justo cuando éste los necesita. O bien, individuos que requerimos y |

que se reúnen por algún motivo ajeno a la investigación, lo que nos proporciona una oportunidad extraordinaria para reclutarlos.

•*Muestras teóricas o conceptuales*: cuando el investigador necesita entender un concepto o teoría, puede muestrear casos que le ayuden a tal comprensión. Es decir, se eligen las unidades porque poseen uno o varios atributos que contribuyen a desarrollar la teoría.

•*Muestras confirmativas*: la finalidad es adicionar nuevos casos cuando en los ya analizados se suscita alguna controversia o surge información que apunta en diferentes direcciones. Puede suceder que en algunos de los primeros casos surjan hipótesis de trabajo y casos posteriores las contradigan o *no se encuentren tendencias claras*. Entonces, seleccionamos casos similares donde emergieron las hipótesis, pero también casos similares en donde las hipótesis no aplican.




•*Muestras de casos sumamente importantes para el problema analizado*: casos del ambiente que no podemos dejar fuera. En una investigación cualitativa en una empresa, no es conveniente prescindir del presidente(a) o director(a) general. Incluso hay muestras que únicamente consideran casos importantes.

•*Muestras por conveniencia*: simplemente casos disponibles a los cuales tenemos acceso.

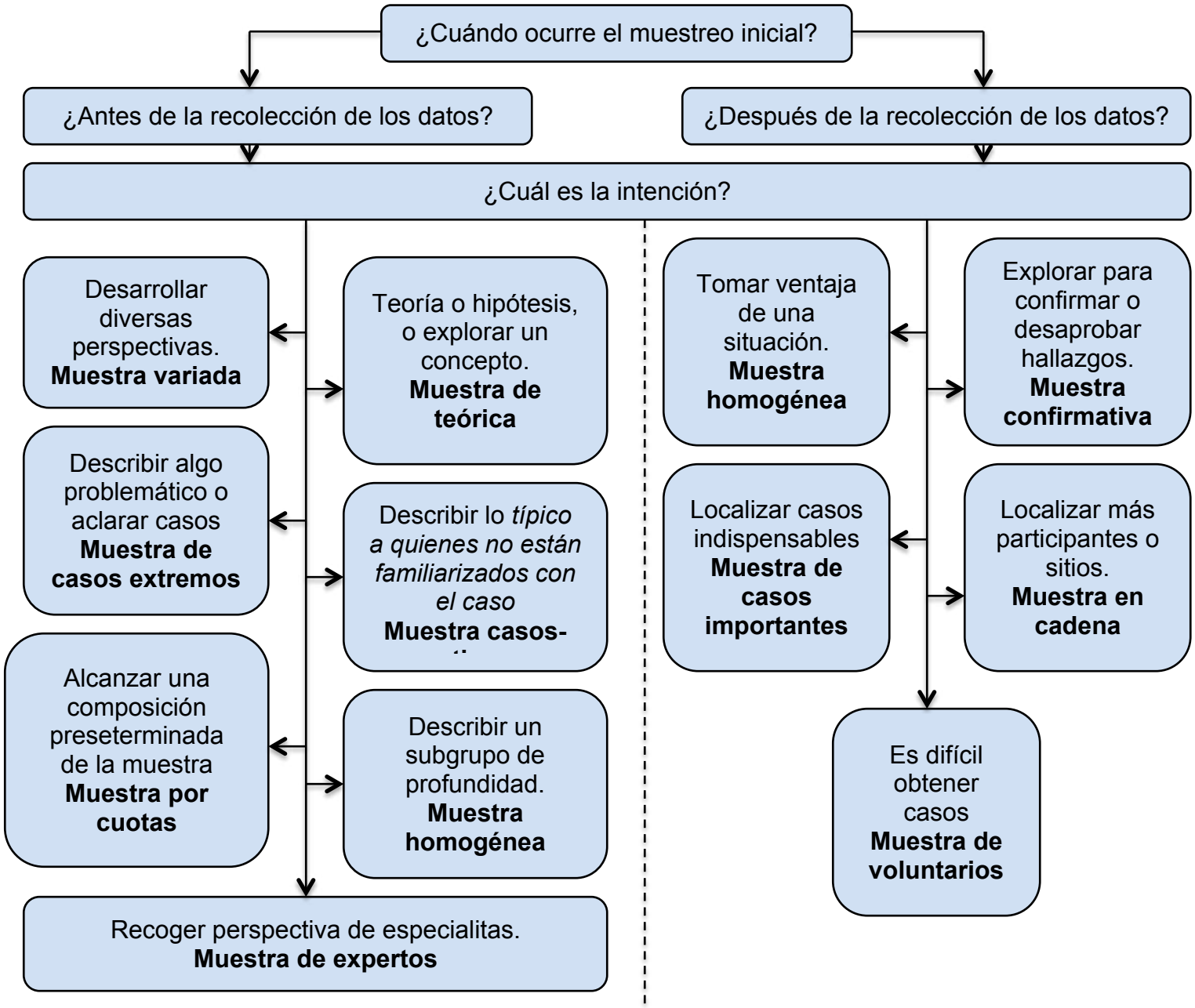
•*Muestras dirigidas.* Son válidas en cuanto a que un determinado diseño de investigación así las requiere; sin embargo, los resultados se aplican nada más a la muestra en sí o a muestras similares en tiempo y lugar (transferencia de resultados), pero esto último con suma precaución. No son generalizables a una población, ni interesa tal extrapolación.

Fuente: Hernández-Sampieri et al., (2010) con adaptación propia

En todo el proceso de inmersión inicial en el campo, inmersión total, elección de las

unidades o casos y de la muestra; debemos tomar en cuenta al planteamiento del problema, el cual constituye el elemento central que guía todo el proceso, pero tales acciones pueden hacer que dicho planteamiento se modifique de acuerdo con la *realidad del estudio* (construida por el investigador, la situación, los participantes y las interacciones entre el primero y estos últimos). El planteamiento siempre estará sujeto a revisión y cambios. Ver **Figura A.10**

**Figura A.10. Esencia de la toma de decisiones para la muestra inicial en estudios cualitativos.**




Fuente: Hernández-Sampieri et al., (2010) con adaptación propia

# CAPÍTULO 1. LA IMPORTANCIA DE REALIZAR MEDICIONES EN RELACIÓN A LAS CIENCIAS ECONÓMICO-ADMINISTRATIVAS

En este capítulo se aclaran el qué es medir, qué es una escala para qué realizarlo, su utilidad en las ciencias económico-administrativas, el planteamiento de escalas con constructos subyacentes con múltiples indicadores así como su importancia, las implicaciones de los conceptos básicos de dimensionalidad, confiabilidad y validez de escala para finalmente, revisar el procesos de las cuatro etapas que las involucran.

## ¿Qué es medir?

De acuerdo al Diccionario Real Academia Española (DRAE,2017), medir es: *Comparar una cantidad con su respectiva unidad, a fin de averiguar cuántas veces es la segunda está contenida en la primera,* es decir, determinar una medida. Medir es una de las principales tareas de la ciencia y de la que es posible afirmar que *si no se mide, no se controla*.

Así, tenemos que la definición de *medida*, como *un conjunto de reglas que permiten asignar números a los objetos observados de tal forma que representen de manera adecuada la cantidad de un determinado atributo que poseen* (Nunally y Berstein, 1994). Para Aldás-Manzano y Maldonado-Guzmán (2015): *en algunos casos, esas reglas son muy obvias, como cuando se trata de medir la altura de un individuo con una cinta métrica. Pero, desgraciadamente, la obviedad no es común en las ciencias sociales. Así, las reglas para determinar el grado de timidez, de etnocentrismo o de inteligencia de un consumidor, o el grado de orientación al mercado de una empresa, no son tan intuitivas*. Para nuestro caso, medir consiste en aplicar reglas que asignan símbolos a objetos para representar numéricamente cantidades de atributos. Medir, incluye la evaluación de números tales que reflejen los diferentes niveles del atributo que se está evaluando (DeVellis, 1991, Haynes et al., 1999, Nunnally y Bernstein, 1994). Se destaca, que en las ciencias económico-administrativas, la mayoría de las veces los *objetos de estudio* son individuos, empresas, etc. y las *reglas* implican una asignación explícita de números, y los *atributos* son rasgos que caracterizan de forma muy particular a cada uno de los *objetos de estudio* que se miden. Lo anterior, quiere decir que por ejemplo, los individuos o empresas *no se miden* pero Sí sus atributos, por ejemplo: el liderazgo o la innovación. En cuanto a *las reglas de medida* requieren un poco más de explicación ya que algunas reglas son *obvias y universales*, como medir la distancia en kilómetros o millas. Sin embargo, las reglas que se emplean para medir los constructos socio-psicológicos no son tan evidentes. Por ejemplo, ¿cuáles son las reglas apropiadas para medir constructos como el liderazgo de los directores, la satisfacción de los clientes la autoestima de los empleados, la satisfacción en el trabajo y/o la autoconfianza del consumidor? Se destaca que *no existen reglas universales* para realizar la medición de tales constructos, pero Sí *el desarrollo de reglas* que son finalmente aceptadas como factor importante para la estandarización y el establecimiento de normas. Así, una medición se dice está estandarizada cuando (a) las reglas de medición son claras, (b) es práctico de aplicar, (c) no es exigente del administrador y/o supervisado, y (d) los resultados no dependen del administrador (Nunnally y Bernstein, 1994). Dicha medición, tiende a producir resultados similares en todas las aplicaciones (es decir, la medición ya es *confiable*) y ofrece puntuaciones que pueden ser fácilmente interpretadas como bajas, medias y altas. Por otro lado, cabe aclarar el *enfoque de la medición* de los atributos también requiere precisarse. Como se ha explicado, realizar mediciones medir a un individuo o empresa no es hacerlo *per se*; sino que se están midiendo sus atributos. Esta distinción es importante porque enfatiza la naturaleza abstracta de la medición en las ciencias económico-administrativas. Es decir, se debe realizar **u***n alto nivel de abstracción* del atributo a medir. En las económico-administrativas, existen una gran variedad de estudios que intentan determinar relaciones entre dos atributos (por ejemplo, competitividad e innovación, o liderazgo con personalidad, etc.). Para evitar la confusión entre los atributos relacionados, la naturaleza exacta del atributo debe ser cuidadosamente determinada y especificada (marco teórico de la investigación, por ejemplo). Además, debe realizarse una evaluación de la factibilidad sobre si el atributo puede medirse en absoluto ya que algunos atributos son tan abstractos que pueden no ser susceptibles de medición (por ejemplo, la clarividencia Nunnally y Bernstein, 1994).

# ¿Qué es una escala de medición?

Para medir, es necesario crear (o utilizar si ya existe) una escala de medida. Una *escala de medida* es un conjunto de ítems, frases, o preguntas que permiten medir el nivel que alcanza una atributo determinado (etnocentrismo, orientación al mercado) *no directamente observable en un objeto* (un consumidor, una empresa). De lo anterior, podemos afirmar que la escala es *un herramental diseñado por el investigador de las ciencias económico-administrativas que representa directamente el modelo conformado por uno o varios constructos basados en factores, variables, dimensiones e indicadores tanto directos como subyacentes, que implementado, le permiten medir el fenómeno objeto de estudio, determinando su magnitud y sentido.*¿Qué ventajas aporta a la investigación la utilización de escalas estandarizadas de medida?. Se resumen en las siguientes (Aldás-Manzano y Maldonado-Guzmán,2015):

- **Objetividad**. Un principio científico básico es que cualquier afirmación efectuada por un investigador, ha de poder ser verificada independientemente por cualquier otro repitiendo el experimento en las mismas circunstancias. Aunque este principio es de limitada aplicación en ciencias sociales, no deja de ser cierto que si, ni siquiera se dispone de un instrumento para medir el atributo, difícilmente se podrán comparar los resultados.
- **Cuantificación.** Los resultados numéricos de medidas estandarizadas tienen la ventaja de que permiten la utilización de técnicas estadísticas avanzadas.
- **Comunicación.** Las medidas objetivas permiten la fácil comunicación de los resultados de las investigaciones entre los distintos científicos.
- **Economía.** Desarrollar una escala adecuadamente es una tarea costosa, pero una vez creadas suponen un gran ahorro de tiempo. Así, por ejemplo, un investigador preparado puede juzgar bastante acertadamente el grado de innovación de un sector manteniendo con su **CEO** una entrevista en profundidad. Valorar esta misma variable en 100 empresas con 100 entrevistas **CEO** le resultará, sin embargo, muchísimo más costoso que administrarles un cuestionario con las preguntas que conformen la escala, por ejemplo, a diseño sugerido por el Manual de Oslo de la **OECD** (2005).

Desafortunadamente, aún y sus evidentes ventajas, no todas las escalas son desarrolladas con el suficiente cuidado en su diseño y aplicación. Se debe tener en claro que una escala *no es un conjunto de preguntas unidas sin más*, sino que su diseño implica necesariamente el cubrir una serie de propiedades adecuadas conocidas como *propiedades psicométricas*, las cuales son básicamente dos: *dimensionalidad, confiabilidad y validez*. Por lo anterior, se hace imperioso al aspirante a investigación que conozca y aplique de manera sistemática un esquema de desarrollo de escalas que le permita generar sus instrumentos de manera adecuada para las medidas que pretenda analizar.

# La utilidad de realizar mediciones en las ciencias económico-administrativas

Dado que las ciencias económico-administrativas tienen un fuerte componente psicológico individual así como social, se hace notar que existen múltiples criterios basados en *propiedades psicométricas y/o de conducta*, que se utilizan para evaluar las mediciones. Los criterios más relevantes, son dependientes de los objetivos de la evaluación y del esfuerzo científico emprendido, de tal manera que las ciencias económico-administrativas se centran en realizar propuestas que incluyen constructos que se describan con escalas de factores, variables y/o indicadores *socio-psicológicas* latentes. Lo anterior, ha generado opiniones y debates incluso muy divergentes, pero una postura que es común a mayoría de los científicos de las disciplinas de económico–administrativas es que los resultados de las mediciones deben ser repetibles y estandarizados ya que son conceptos relacionados.

Por ejemplo, un hallazgo de investigación basado en circunstancias similares y en las mismas mediciones debería replicarse. Este es el principio básico de la repetibilidad, que la medición se realice de forma fiable bajo condiciones de prueba similares. Los procedimientos, por ejemplo de estímulos visuales y/o auditivos para la evaluación de atributos de un producto, a fin desarrollar de la escala incluyen el establecimiento de normas. Cuando estas normas, al ser interpretadas describiendo a una persona como baja, media o alta en un atributo, la medida se considera estandarizada. La estandarización tiene varias ventajas:

- Aunque medimos percepciones que por su propia naturaleza son subjetivas, una medida estandarizada aumenta la objetividad de las ciencias económico administrativas. Cuando entre investigadores verifican de manera independiente una o varias relaciones entre sus constructos, se aumenta la objetividad, dado que las medidas utilizadas son las mismas y están estandarizadas. Si existe desacuerdo en cuanto a la adecuación de las medidas utilizadas para obtener el hallazgo, la objetividad se verá comprometida o en duda. En las ciencias económico-administrativas, a menudo probamos teorías, que sólo son posibles hacerlas en la medida en que los atributos de la teoría (*constructos*) sean adecuadamente medidas. Cuando se acuerdan procedimientos para medir los atributos de interés, se aumenta la objetividad de las pruebas teóricas.
- La estandarización produce resultados numéricos cuantificables que permite la creación de categorías (por ejemplo, baja, media, alta) para análisis matemáticos y estadísticos (**ANOVA**) o para uso como niveles de factor en diseños experimentales. La cuantificación también mejora la comunicación y la generalización de los resultados. El conocimiento se acumula en las ciencias económico-administrativas cuando los investigadores comparan sus resultados con los resultados de estudios previos. Cuando se utilizan las mismas medidas estandarizadas en aplicaciones científicas, los resultados son denominados como *bajos* en el liderazgo o *altos* en la innovación tienen un significado común entre los investigadores. Esto mejora la comunicación de los resultados y la generalización de los hallazgos.
- Finalmente, el desarrollo de mediciones/escalas es una tarea que requiere mucho tiempo. Sin embargo, si una medición ha sido bien desarrollada, ésta recompensa en el futuro. Una vez que se produce la estandarización, la medida y su medición

estarán disponibles para su uso con poco o ningún tiempo invertido debido a las normas acordadas. En el centro mismo de la repetibilidad y la estandarización están las propiedades de medición de fiabilidad y validez.

# Realizando la escala de constructos subyacentes con múltiples elementos

En la literatura de hoy, se observa un número importante de conceptos básicos de la medición, así, es posible afirmar que los investigadores de las ciencias económico-administrativas, enfocan sus esfuerzos en proponer constructos y modelos que expliquen, a través de las mediciones de los atributos de sus objetos de estudio que tienden a ser altamente abstractos y por lo tanto, latentes o subyacentes de origen. *Esto significa que no son directamente determinados y por lo tanto observables y/o cuantificables de forma directa*. Los *constructos subyacentes* son variables en sentido y magnitud de las calificaciones y cambian con el tiempo. Por ejemplo, la necesidad de pertenencia representa un constructo latente o subyacente (es decir, *un atributo perso***nal**). Esto significa que no puede ser observado directamente por un investigador y por lo tanto requiere del diseño de una escala que estime su magnitud y sentido real en un momento dado. Además, el nivel de necesidad de pertenencia de un individuo hacia un grupo puede cambiar con el tiempo. Por ejemplo, Pitonyak (2010) describe como el sentido de pertenencia está en función del grado de soledad, de la sensación de cercanía con otra persona, de la habilidad de relacionarse para hacer amistades, de los niveles de exclusión social que el individuo presenta, entre otros factores subyacentes, los cuales deben ser justificados dentro del marco teórico que se plantee. Por otro lado, algunos constructos son de naturaleza empírica o sin base teórica (*ateórica*). Las encuestas de opinión muchas veces evalúan constructos altamente empíricos. No se está sugiriendo que tales constructos y/o enfoques de medición tengan poco valor en las ciencias económico-administrativas. De hecho, tienen un potencial relevante de utilidad, para el desarrollo del marco teórico. Por ejemplo, una encuesta de opinión (*empírica y deseablemente teórica*) que evalúa el nivel de asimilación tecnológica en una firma y que favorece o no favorece el diseño y desarrollo de productos innovadores bélicos podría revelar que aquellas firmas que son innovadores tempranos, favorecen el control de armas más que aquellos que son innovadores tardíos. Lo anterior es sólo para recalcar la importancia vital del marco teórico como sustento de los constructos de un modelo a fin de poderlo explicar. Se acepta generalmente que las medidas de los constructos teóricos latentes o subyacentes requieren múltiples elementos o afirmaciones para revelar más exactamente los niveles de las variables de dichos constructos; es decir, son escalados (Clark y Watson, 1995, DeVellis, 1991, Haynes, Richard, y Kubany, 1995, Nunnally y Bernstein, 1994).

Así, es posible afirmar que, dado un nivel de un objeto de estudio sobre un atributo que es latente y con tendencia socio-psicológica a ser altamente abstracto no puede medirse directamente en el contexto económico- administrativo, debe construirse una escala. Algunas ocasiones, le será posible inferir un nivel de un constructo socio-psicológico latente a través del comportamiento (por ejemplo, de la compra repetida

de la marca, uno infiere que un individuo cree que la marca es un buen valor para el dinero (*constructo latente*), muchas veces un comportamiento no puede ser indicativo de una construcción latente. Finalmente, se debe reiterar que la escala, y no la indexación, es nuestro enfoque. En la escala, las puntuaciones de los indicadores en la escala son teóricamente impulsadas por la constructo latente o subyacente; es decir, son *un efecto reflectivo (reflected).* Con un índice, los puntajes en los indicadores impulsan la puntuación total del índice; es decir, los indicadores causan el índice construido. Aunque todavía latentes en muchos aspectos, los *indicadores de causa o formativos (formative*) no se consideran escalas porque sus puntuaciones no se reflejan necesariamente en el constructo latente. Un ejemplo frecuentemente usado de *indicadores de causa o formativo (formative*) que resultan en un índice es el nivel de competitividad que se incluye en el premio nacional de calidad de México (PNC,2017). Los indicadores incluyen nivel de liderazgo, generación de valor al consumidor, planeación estratégica, administración del cambio, compromiso social, salud, inclusión, administración del conocimiento y agilidad. Ciertamente, varios de estos indicadores tienen la propiedad latente de no ser directamente observables, sus puntuaciones se consideran el índice de competitividad y no viceversa. Para más información vea Bollen y Lennox (1991), Diamantopoulus y Winklhofer (2001), MacCallum y Browne (1993), y Smith y McCarthy (1995), se hace referencia a las medidas formativa *(formative)* vs. reflectiva (*reflective*).

## Constructo latente.

El fenómeno subyacente que una escala pretende medir se denomina frecuentemente *variable latente.* Por ejemplo, el concepto de la relación Innovación y administración del conocimiento del consumidor, escala desarrollada por Mejía-Trejo, et al. (2016) que muestra tanto a la innovación y su relación con la administración del conocimiento del conocimiento del consumidor como constructos latentes dado que no son directamente observables. Además, de que también la administración del conocimiento del consumidor, se manifiesta a través de otros constructos o variables, como las citadas: **CKM** as a Driver of Innovation (**CKMADI**); (H) **CKM** Support (**CKMS**); (I) **CKM** other Sources of Knowledge (**CKMOSK**); (J) **CKM**, Satisfaction, Experience and Performance (**CKMSEP**). Ver **Tabla 1.1.**

**Tabla 1.1. Modelo innovación-administración del conocimiento del consumidor (Mejía-Trejo, et al. 2016). Constructo: administración del conocimiento del consumidor**

Customer knowledge

| F | Variable | Indicator | Q | Author |
|---|---|---|---|---|
| (G) | 34).- Information from Costumer (IFMC) | Customer is a Resource of NPD ideation; Customer Driven-Innovation (Innovation from Customers). Mutual Innovation. | 78 | Nambisan (2002), Desouza et al. (2007), Gibbert and Probst (2002), Chiu & Fogel (2014) |
| | 35).- Information about the Customer (IABC) | Strategy of close collaboration with customers. Communities of creation. | 79 | Nambisan (2002), Gibbert and Probst (2002) |
| | 36).-Information for Customer (IFRC) | Customer as a User collaborates intensively in the product testing and support. Customer Focused Innovation (Innovation for Customers) | 80 | Nambisan (2002), Desouza et al. (2007) |
| | 37).- Information as a Customer Co-creator (with) (IWIC) | Customer as a Co-creator helps over NPD design and development; Customer Centered Innovation (Innovation with Customers); Prosumerism; Team-Based-CoLearning. Joint Intellectual Property | 81 | Nicolai et al. (2011), Desouza et al. (2007), Gibbert and Probst (2002) |
| | 38).- Negative side effects of Customer Integration (NSEC) | The firm is warned about the dependence on customer's personality (NSEC1) | 82 | Gassmann et al. (2012) |
| | | The firm is warned about the dependence on customer's experience (NSEC2) | 83 | |
| | | The firm is warned about the dependence on customer's point of view (NSEC3) | 84 | |
| | | The firm is warned about to choose the wrong customer (NSEC4) | 85 | |
| | | The firm is warned about the risk to integrate the customer to the company's side (NSEC5) | 86 | |
| (H) | 39).- Knowledge Incentives (KI) | Salary associated with the ability and willingness to share knowledge (KI1) | 87 | Nicolai et al. (2011), OECD (2003) |
| | | Salary determined by willingness to improve skills and upgrade knowledge (KI2) | 88 | |

Customer knowledge

| F | Variable | Indicator | Q | Author |
|---|---|---|---|---|
| | | Tolerance of Failure (KI3) | 89 | Gloet and Samson (2013) |
| | | Rewards and Recognition (KI4) | 90 | |
| | 40).- Knowledge Fluence (KF) | Exchange the knowledge between employees across departments (KF1) | 91 | Nicolai et al. (2011), OECD (2003), Chiu and Fogel (2014) |
| | | Communication among employees and management (KF2) | 92 | |
| | 41).- Knowledge and ICT (KICT) | ICT to support and control the Customer Knowledge Management | 93 | Laudon and Laudon (2012), Mejía-Trejo and Sánchez-Gutiérrez (2013) |
| | | Technical Services (ISOK1) | 94 | |
| | | Engineering Department (ISOK2) | 95 | |
| | 42).- Internal Sources of Knowledge (ISOK) | Research and Design Development (ISOK3) | 96 | Garcia-Murillo and Annabi (2002) |
| | | Production (ISOK4) | 97 | |
| (I) | | Marketing and Sales (ISOK5) | 98 | |
| | | Purchasing and Supply (ISOK6) | 99 | |
| | | Other Employees (ISOK7) | 100 | |
| | | Supplier (ESOK1) | 1 | |
| | 43).- External Sources of Knowledge (ESOK) | Scientist, Universities, Patents, Exhibitions Technological Consultant (ESOK2) | 2 | Garcia-Murillo and Annabi (2002) |
| | | Distributor Agents (ESOK3) | 3 | |
| | | Competitor (ESOK4) | 4 | |
| | 44).- Paradigm (PAR) | If Only We Knew What We Know (KM) as a Customer Retention (PAR1) | 5 | Gibbert and Probst (2002), Garcia-Murillo and Annabi (2002) |
| (J) | | Retention is Cheaper than Acquisition (CRM) as a Customer Satisfaction (PAR2) | 6 | |
| | | If We Only Knew What Our Customer (CKM) Knows as a Customer Experience and Creativity (PAR3) | 7 | |
| | 45).- Performance (PER) | Performance against budget; Customer retention rate. (KM) (PER1) | 8 | |
| | | Performance in terms of customer satisfaction and loyalty (PER2) | 9 | |
| | | Performance against competitors in innovation and growth; Contribution to customer success. (CKM) (PER3) | 10 | |

*Source*: Own.
*Notes*: Factor (F); (A) Innovation Value Added (IVADD); (B) Innovation Income Items (IIIT); (C) Innovation Process (INPROC); (D) Innovation Outcome Items (IOIT); (E) Innovation Performance (IPERF); (F) Innovation Feedback Items (IFEED); (G) CKM as a Driver of Innovation (CKMADI); (H) CKM Support (CKMS); (I) CKM other Sources of Knowledge (CKMOSK); (J) CKM, Satisfaction, Experience And Performance (CKMSEP).

Fuente: Mejia-Trejo et al., 2016

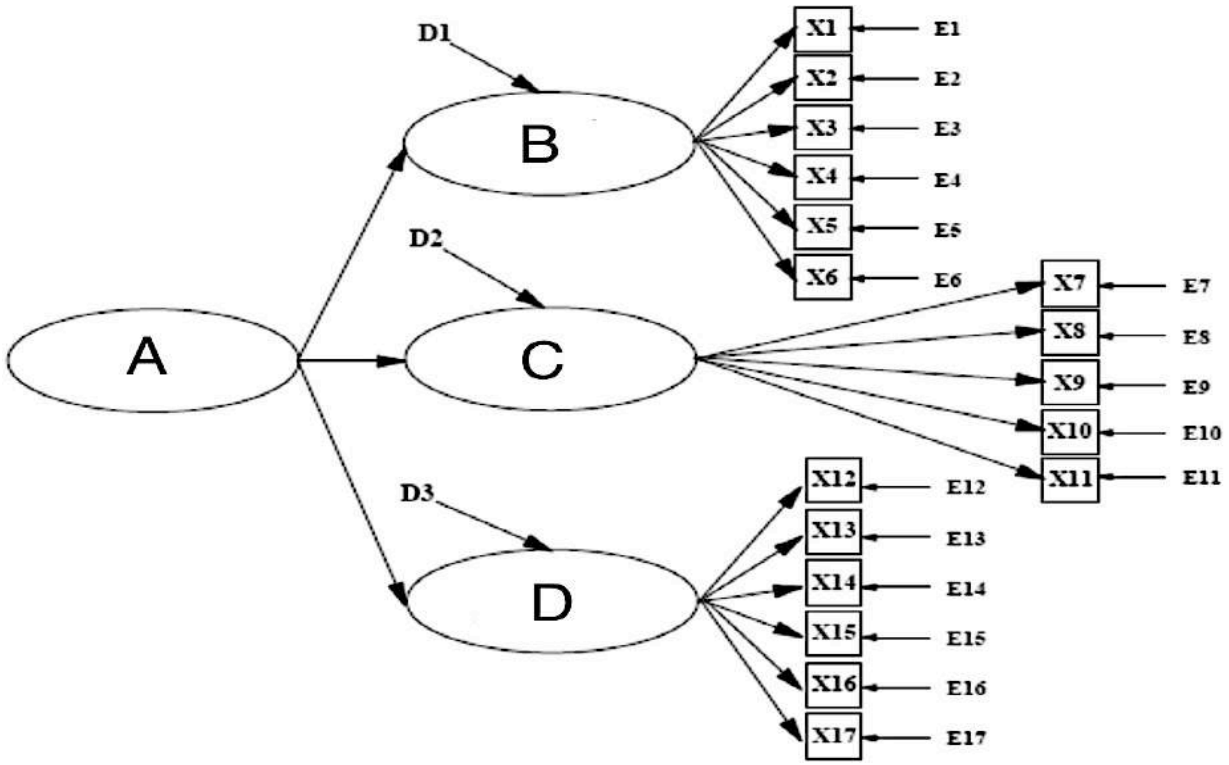
Además la administración del conocimiento del consumidor es también *variable* en cuanto que no es constante, es decir, no está fija en el tiempo. Una escala de medición como la administración del conocimiento del consumidor , buscará estimar la magnitud de esta variable latente en el momento y entorno geográfico actual, lo que denominamos, *puntuación verdadera.* La noción de *constructo latente,* implica que debe existir cierta relación entre ella y los ítems que componen la escala. El *constructo latente* debe ser, si la escala se ha desarrollado correctamente, la causa de la puntuación obtenida en los distintos ítems de la escala. La forma más común de representar las relaciones entre los distintos ítems de una escala y el concepto que miden, son los *diagramas de trayectorias*, facilitando la comprensión de técnicas estadísticas de validación de escalas, como el **CFA.**

# Diagramas de trayectorias

Los diagramas trayectorias son un método para representar las **relaciones causales entre las variables**, y están sujetos a una **serie de convenciones**, tales como:

1. Una flecha (→) entre dos variables, indica que una es la causa de otra, así **A → B** señalaría que **A** *es la causa de* **B**. Si la variable a la que se hace referencia es una *variable latente, ésta se representa en un óvalo o círculos. Si es una variable observable o manifiesta*, esto es directamente medible como son *los indicadores de la escala,* se representa en un *cuadrado***.** Ver **Figura 1.1**.

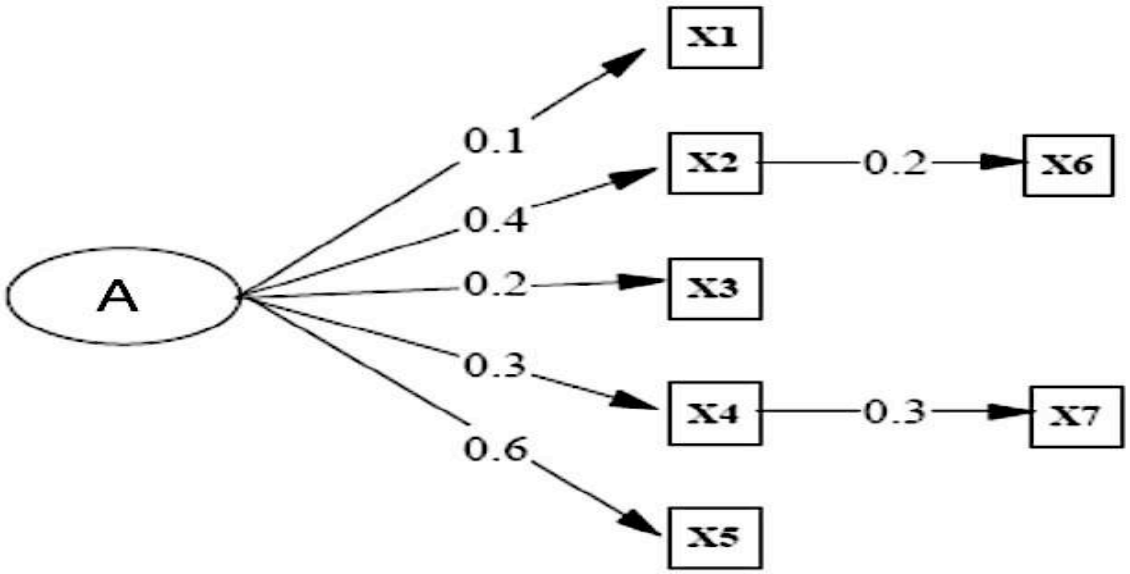
**Figura 1.1. Diagrama de trayectoria conceptual**




Fuente: propia

2. Otra convención de los *diagramas de trayectorias, es la representación del término de error*, que se realiza mediante una *nueva variable causal* (**E** ó **D** en la **Figura 1.1**). El error **E** es el *error de medida* asociado a un *constructo observable* (causado, por ejemplo, por incorrecta interpretación de las preguntas del cuestionario por una empresa, o en su formulación por parte del entrevistador). El **error D** (*disturbance)* hace referencia *al error en la predicción de una variable o constructo latente.*
3. Otro concepto fundamental en los *diagramas de trayectorias*, es el del *coeficiente estandarizado de la relación causal (cargas factoriales en ecuaciones estructurales y análisis factoriales*), que son los números que, una vez estimado el modelo, aparecen sobre las flechas (**Figura 1.2**).

**Figura 1.2. Diagrama hipotético de trayectorias por coeficiente estandarizado de la relación causal**




Fuente: DeVellis (1991)

Estos coeficientes expresan la fuerza de la relación causal entre las variables unidas por la flecha. El hecho de que estos coeficientes se hayan estandarizado, indica que están medidos en la misma escala y son directamente comparables entre ellos. Así, en la **Figura 1.2**, **A** es causa mucho más intensa de $\mathbf{X_5}$ que de $\mathbf{X_1}$.

Existe una útil relación entre los *coeficientes estandarizados y las correlaciones entre las variables manifiestas* (*indicadores de la escala*): *la correlación entre dos variables manifiestas causadas por una misma variable latente, es igual al producto de sus respectivos coeficientes estandarizados*. Por ejemplo, la correlación entre $\mathbf{X_1}$ y $\mathbf{X_5}$ en la **Figura 1.2**, vendría dada por:

$$r_{1\text{-}5}=0.6\times0.1= 0.06$$

Las variables $\mathbf{X_6}$ y $\mathbf{X_7}$ también comparten a **Y** como fuente común y, por ello, se puede calcular la correlación entre ellas a través de la ruta que las une a través

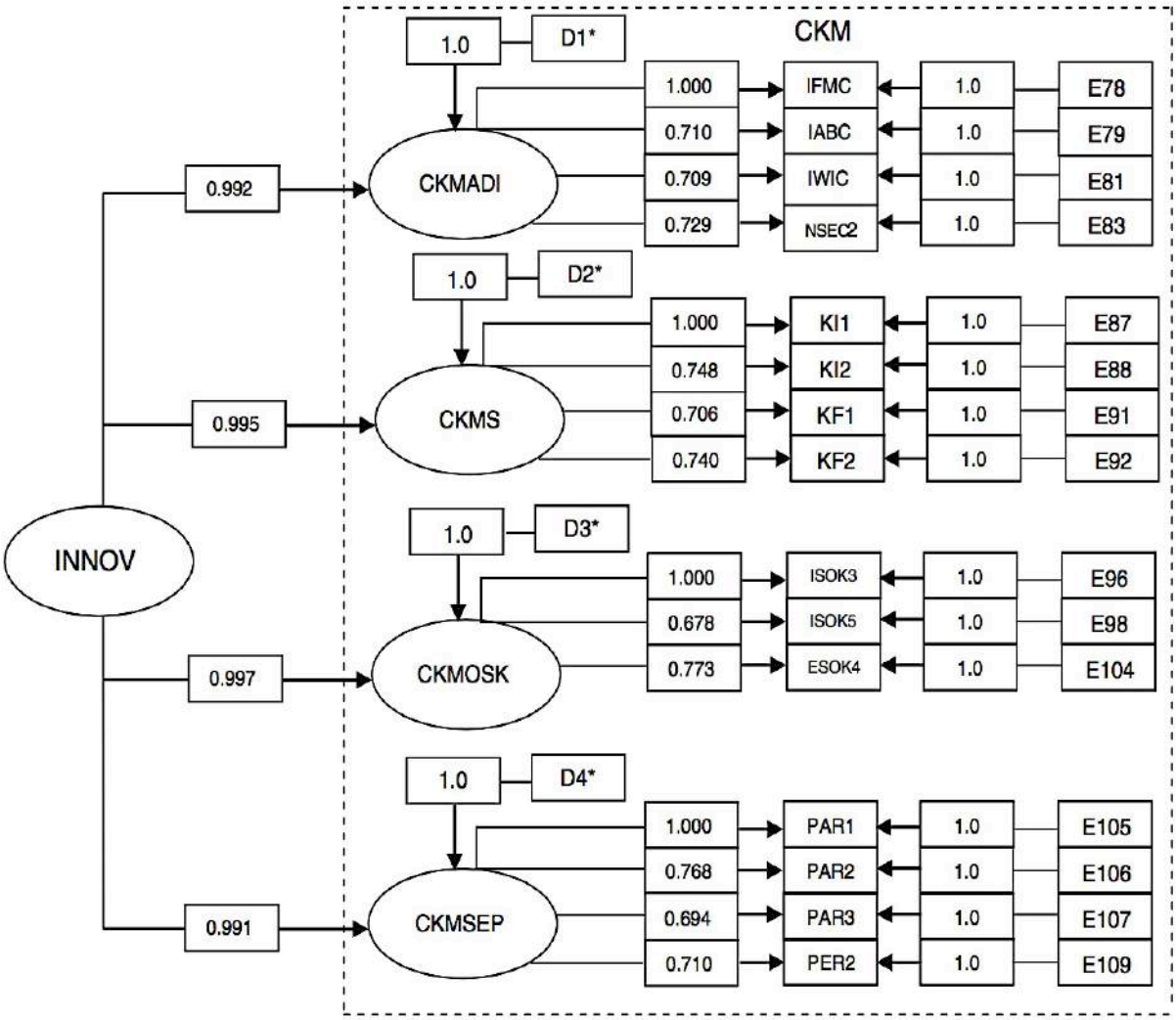
de **Y**:

$$r_{6\text{-}7} = 0.2 \times 0.4 \times 0.3 \times 0.3 = 0.0072$$

Esta relación entre correlación entre las *variables manifiestas y los coeficientes estandarizados* es muy importante, porque servirá de base para estimar los modelos. Nótese que, si aplicamos una escala para medir un *constructo latente*, la información de la que dispondremos serán las puntuaciones de las variables manifiestas (*indicadores de la escala*). Por tanto, se podrá calcular los coeficientes de correlación entre ellas y, a partir de esta información, estimar las *trayectorias (coeficientes estandarizados)* siempre que se cumplan determinados supuestos sobre el *error de medición* que especificaremos a continuación.

Otro ejemplo de mayor detalle de cómo se esquematizan los constructos sin un diagrama de trayectorias, es el presentado en la **Figura 1.3**, relacionando al modelo de innovación administración del conocimiento del consumidor de Mejía-Trejo, et al. (2016).

**Figura 1.3. Diagrama de trayectoria de la relación de la Innovación-administración del conocimiento del consumidor**



Hypothesized second-order factorial model of Customer Knowledge Management influenced by Innovation for SSG. *Notes*: *E*(*n*) – error disturbance; *D*(*n*) – variances of the disturbances. Because the estimation of all higher order factor loadings are typically of interest in second-order models, the variance of the single higher order factor (CKMS) has been constrained to 1.0; note also that the variances of the disturbances (the D's) are designated as freely estimated. Relatedly, their paths are automatically constrained to 1.0 by the program (Byrne, 2006).
Fuente: Mejía-Trejo et al., 2016

# La medición y su error

Es de gran importancia que se forme una idea más clara de cuáles pueden ser las fuentes de error en las mediciones de una variable o *constructo latente mediante una escala* y, posteriormente, especificaremos los supuestos estadísticos que sean necesarios realizar. Así, se consideran *ocho fuentes de variación o fuentes de error* que darán lugar a diferencias en los valores obtenidos en la medición de un concepto (Aldás-Manzano y Maldonado-Guzmán, 2015):

**2.** *Diferencias verdaderas en la característica que se pretende medir.* Que diversos individuos tengan distintas opiniones sobre la variable estudiada, provoca la deseada diferencia entre las puntuaciones.

**3.** *Diferencias verdaderas en otras características estables del individuo que afectan al valor resultante.* Hace referencia a la influencia que variables como, el tamaño de la empresa, pueden ejercer sobre el valor de la variable latente que se mide, digamos, la orientación al mercado.

**4.** *Diferencias debidas a características transitorias del individuo*, como el estado de ánimo, el cansancio o las prisas.

**4.** *Diferencias causadas por factores circunstanciales del entorno,* como que el individuo entrevistado esté sólo o en compañía, en su hogar o en la calle.

**5.** *Diferencias causadas por variaciones en la recolección de datos,* como sesgos introducidos por el encuestador.

**6.** *Diferencias debidas a la muestra de preguntas*, como cambios introducidos en el enunciado de las mismas por el entrevistador.

**7.** *Diferencias debidas a la falta de claridad en el instrumento de medición*, que hacen que el entrevistado no comprenda adecuadamente la pregunta.

**8.** *Diferencias causadas por factores mecánicos,* como falta de espacio para registrar la respuesta, o un error en el registro de la misma por el entrevistador.

La existencia de estas posibles fuentes de variación, obliga a realizar una serie de supuestos para que sea posible asumir que las medidas obtenidas de una variable latente son razonablemente próximas a la realidad. Estos supuestos son los siguientes:

**a.** Sea $\boldsymbol{X_{ij}}$ la respuesta dada por el individuo ***i*** **al ítem** ***j*** de la escala de medida de la *constructo latente Y* (por ejemplo la orientación a la innovación en Mejía-Trejo et al., 2016).

**b.** En el contexto de la *teoría clásica de la medición,* $\boldsymbol{X_{ij}}$ puede considerarse como función de la *puntuación verdadera* que tiene para el *individuo i* la variable medida (la orientación a la innovación para la *empresa i),* que denotaremos $\boldsymbol{m_i}$ y un error $\boldsymbol{e_{ij}}$ provocado por cualquiera de las fuentes de variación señaladas con

anterioridad. Entonces:

$$X_{ij}=m_i+e_{ij}$$

**c.** Pues bien, para poder desarrollar una escala como instrumento de medida del *constructo o variable latente Y,* y poder comprobar sus propiedades psicométricas (*confiabilidad y validez*) mediante los procedimientos que posteriormente expondremos, es necesario asumir los siguientes supuestos:

1. *El error asociado a un ítem determinado es aleatorio, es decir, tiene una media de 0* cuando se considera una muestra amplia de individuos que responden al mismo.
2. *El error de un ítem de la escala no está correlacionado con el de otro ítem.* La única conexión entre los ítems de la escala pasa por la variable latente, nunca a través de ningún término de error. Nótese en la **Figura 1.1** que no hay flechas entre los términos de error.
3. *Los términos de error no están correlacionados en la puntuación verdadera del constructo o variable latente.* Nótese en la **Figura 1.1** que las flechas que salen de la variable latente van hasta los ítems de la escala, nunca llegan al término de error.

# Tendencias en la dimensionamiento de los constructos para las ciencias económico-administrativas

Es un hecho, que las ciencias económico-administrativas han tenido un repunte de crecimiento a partir de los años 1990's, debido a factores como: el desarrollo de software estadístico (**R,SPSS, SAS, BMDP, LISREL, EQS, CALIS, AMOS**, etc.) cada vez más versátil y de fácil uso, disponible en versiones pc y laptop, para el diseño, creación y acceso a bases de datos con contenidos afines así como la aportación cada vez más notable y creciente de investigadores incipientes y con experiencia que aportan al conocimiento de las ciencias económico-administrativas con nuevo modelos, es que *el desarrollo de la escala sigue siendo una industria en crecimiento* (Clark y Watson,1995, p., 309). Esta declaración se aplica a todos los campos relacionados con las ciencias económico-administrativas como los negocios, la mercadotecnia, la competitividad, la administración del conocimiento, la innovación, la contabilidad, sistemas de información de gestión, gestión estratégica, comportamiento organizacional, etc. De hecho, tenemos varios textos que han realizado el esfuerzo por hacer un compilado de las escalas usadas con éxito en la mercadotecnia así como el comportamiento organizacional, tales como: Price y Mueller (1986), Robinson, Shaver y Wrightsman (1991),Bruner y Hensel (1997), Bearden y Netemeyer (1998).

Para las ciencias económico-administrativas, el hecho de crear modelos basados en constructos cada vez más sólidos, hace más evidente, el interés por diseñar, desarrollar y aplicar escalas que faciliten la explicación de problemas donde intervienen factores y variables cada vez más complejos. Lo anterior obliga a comprobarlos de manera objetiva de manera que los modelos aspiren a ser teorías a partir de su la operacionalización de los constructos de interés. Así ocurre que

cuando los constructos son medidos de forma correcta, basados en su *confiabilidad* y se logran mejores bases para la prueba de la teoría. Cabe destacar, que el factor humano, implicado en todos los modelos de las ciencias económico-administrativas obliga a que una escala utilizada una sola vez, requiera ser actualizada o refinada para reflejar y explicar mejor el constructo de interés. Así que no se debe descartar que nuevos artículos y/o tesis que presenten la medición de nuevas escalas, sean más bien actualizaciones derivadas de otras mediciones existentes, a fin de que se sienten con mayor precisión y expliquen más eficientemente los constructos de interés.

Por otro lado, es importante considerar que aún y cuando las tecnologías de información presenten avances notables en el software estadístico que permiten notablemente desarrollos más rápidos las escalas, no necesariamente hacen un mejor desarrollo de la escala. Se sugiere ser muy observante de los procedimientos recomendados que se propongan en los textos afines así como de libros y artículos que han demostrado solidez en el desarrollo de escalas siendo algunos de ellos: (Clark y Watson, 1995, DeVellis, 1991, Haynes et al, 1999, Haynes et al., 1995, Nunnally y Bernstein, 1994, Spector, 1992).

Así, se aprecia que no sólo ha habido una tendencia hacia más escalas, sino que también de los procedimientos utilizados para desarrollar y validar las mismas. Así, se han hecho evidentes pruebas de dimensionalidad de la escala, efectos de los métodos y distribución de la varianza (Bearden y Netemeyer, 1998). También es evidente una preocupación más pronunciada por el contenido y la validez de los elementos en la etapa de desarrollo, así como las consideraciones de longitud de escala (Clark y Watson, 1995; Haynes et al., 1995).

# Marco teórico y validez de los constructos subyacentes

Como se ha explicado hasta el momento, los constructos subyacentes no son directamente observables, por lo que no es posible cuantificarlos directamente, así también las evaluaciones en sus mediciones tienen la posibilidad de ser altamente variables. Se debe hacer la anotación de que existen ya una serie de constructos latentes en las ciencias económico-administrativas, que van desde aquellas que son muy amplias, como los de innovación abierta (Mejía-Trejo, 2017a) hasta constructos más específicos que pueden considerarse subcomponentes de constructos más amplios, como el de modelo de negocios abierto (Mejía-Trejo, 2017b).

Lo anterior, hace notar, que los constructos subyacentes requieren de un tratamiento más bien *efecto* en cuanto al nivel de abstracción y especificidad. Como es de esperar, es aquí donde se hace relevante la importancia de plantear y desarrollar un marco teórico sólido a fin de que las mediciones de los constructos de observación directa y/o subyacente no sean exagerados y por lo tanto, sobredimensionado.

Esto no es nada nuevo, se tienen los trabajos de los clásicos sobre medición y validez, tales como Cronbach y Meehl (1955) y Loevinger (1957) quienes son los pioneros en manifestar de manera muy concreta, la importancia de la teoría en la medición. Cabe destacar, por lo tanto que para que las mediciones de constructos subyacentes tengan la pertinencia, significancia y por lo tanto, relevancia en las

ciencias económico-administrativas, es fundamental y obligatoriamente necesaria, la sustentación del marco teórico.

Es más aún, todo constructo subyacente extremadamente abstracto, pero discutido suficientemente a nivel de marco teórico se dice que es más útil como antecedente o de consecuencia, incorporado en otros constructos subyacentes cuando están incrustados debidamente, en el marco teórico. Así, es posible afirmar que, la importancia de un o varios constructos subyacentes de un modelo que pretende explicar un fenómeno de las ciencias económico-administrativas depende en gran medida de las teorías en las que se basa y formula. En otras palabras y a manera de pregunta, se debe decir *:¿qué predice el constructo latente de interés y/o qué predice el constructo latente?*

Las relaciones entre factores, variables y dimensiones que la explican han sido referidas como una *red nomológica* de una constructo subyacente (Cronbach y Meehl, 1955).

Se debe enfatizar que el marco teórico debe involucrar modelos basados en constructos de factores y variables visibles así como no visibles (subyacentes) de interés, sino también a la *validez* de la medición del constructo, el cual, señala el verdadero impacto del modelo al demostrar su comprobación.

Pero, *¿qué es la validez de constructo?* En términos simples, es una evaluación del grado en que una medición mide realmente el constructo subyacente a la que se destina para medir. Cronbach y Meehl (1955) afirmaron que la demostración de la validez del constructo implica al menos **tres pasos:**

**a.** *Especificar el conjunto de constructos teóricos y sus relaciones basados* (una teoría);

**b.** *Desarrollar métodos para medir los constructos de la teoría* y

**c.** *Probar empíricamente cuán bien los indicadores manifiestos (observables) miden los constructos en la teoría, probando las relaciones hipotéticas entre los constructos de la teoría* (es decir, la *red nomológica*).

Además, la evaluación de la validez de constructo es un proceso continuo. Un estudio que apoya la validez de una estructura no es suficiente para concluir que la medida ha sido validada. Se requieren múltiples pruebas y aplicaciones a lo largo del tiempo, y algunas de ellas pueden requerir un refinamiento del constructo mismo, así como su medición . Según Clark y Watson (1995, p.310): *las mediciones más precisas y eficientes son aquellas que se basan en una validez de constructo establecida… una teoría articulada que está bien soportada por datos empíricos*.

# Importancia de la revisión de la literatura del estado del arte en el marco teórico

El marco teórico es la base para formular y desarrollar una teoría ya que comienza con conceptualizaciones basadas en una revisión exhaustiva de la literatura del estado del arte. Estas revisiones de la literatura sirven a varios propósitos para los investigadores, entre ellos (más en apartado *la importancia de una teoría sólida, revisión de la literatura y su análisis*):

**1.** *Debe alertar de los intentos anteriores delimitando de intereses y teorías en las que el constructo puede resultar útil como variable independiente o dependiente*.

Esto es, debe establecer los límites de los conceptos de factores, variables e indicadores de los modelos, sus fronteras de contenido, antecedentes así como previsión de alcances de los mismos.

2. *Recopilar y revisar de forma rigurosa, escalas previas e intentos de implementación para medir el constructo.* Así, basado en la experiencia e intereses del investigador éste deberá determinar fortalezas, debilidades e incluso oportunidades y amenazas implícitas en la propuesta del modelo, tomando en cuenta incluso, factores como el: político, económico, tecnológico, social, organizacional, ambiental, etc.
3. *La implementación de cualquier escala, debe tomar en cuenta la relación objeto y sujeto de estudio.* En este sentido, es de relevante importancia la metodología empleada para la recopilación de los datos del sujeto de estudio, a fin de evitar en lo posible datos omisos, sesgos, errores, etc. que tiendan a desviar el resultado de la escala.
4. Antes de iniciar el trabajo de crear una escala, se deberá preguntar de manera recursiva e insistente: *¿es necesario crear una nueva escala?* Se debe reconocer que el proceso de generación, implementación y prueba de una nueva escala es un proceso largo y conflictivo El desarrollo y la validación de escala es un proceso largo que implica esfuerzos y costos en el levantamiento de datos.
   Así, la revisión exhaustiva y profunda de la literatura del estado del arte debe ayudar a responderla. Si ya existen previas escalas y mediciones del constructo, el valor de una nueva medida puede ser pequeño en relación con los costos involucrados en el desarrollo. Una nueva escala de medición debe mostrar pertinencia teórica y empírica sobre las escalas y mediciones de modelos existentes del mismo constructo para ser útiles.
5. Para que una nueva escala tenga al menos *validez incremental* sobre las medidas existentes, *debe o hacer una precisión mayor del constructo problema o hacerlo más eficiente* (por ejemplo, que logre niveles notables de optimización para afirmar que es más barato, más corto, más fáciles de usar, de responder, que logra mejores niveles, etc.) que las mediciones existentes.
6. El investigador debe tener en mente que una *revisión exhaustiva de la literatura del estado del arte,* es la búsqueda de una oportunidad que le ayude a evitar la redundancia de desarrollar otra escala para evaluar un constructo ya bien medido.

# Propiedades de una medición: dimensionalidad, confiabilidad y validez

Como se ha visto, el proceso de desarrollo de una escala comienza con una revisión exhaustiva de la literatura en la que se determina una definición teórica sólida del constructo y su dominio, de manera recursiva. Esta definición y descripción del constructo debe incluir lo que está dentro del dominio del constructo y lo que está excluido del dominio del constructo así como su dimensionalidad derivados de la revisión exhaustiva de la literatura del estado del arte existente y más aún, siendo altamente deseable el recopilar la opinión de expertos. En esencia, la definición del constructo y el dominio del contenido determinan la dimensionalidad teórica.

## Dimensionalidad

La dimensionalidad de una medición se refiere a la homogeneidad de los ítems o indicadores. Básicamente, se debe considerar que una medición de tipo unidimensional tiene propiedades estadísticas que demuestran que sus elementos subyacen en un solo constructo o factor. Cuando la medición es multidimensional, los elementos pulsan más de una dimensión o factor.

El dominio de un constructo puede ser hipotético *unidimensional, multidimensional, y/o como un factor de orden superior.* Por lo tanto, la escala (o subescalas/factores) utilizada para operacionalizar el constructo debe reflejar la dimensión hipotética. Dado que la unidimensionalidad de escala (*factor*) es considerada un requisito previo para la confiabilidad y validez, la evaluación de la unidimensionalidad debe ser primordial (Cortina, 1993; Hattie, 1985; Schmitt, 1996).

Al momento, han sido empleados varios procedimientos para verificar la dimensionalidad de una escala (por ejemplo, análisis de ítems o indicadores y análisis factorial exploratorio, entre los más comunes).

Una técnica muy conocida es el *Análisis Factorial Confirmatorio (**CFA**. Confirmatory Factor Analysis)*, en el que se pueden especificar y evaluar varios factores con varios indicadores (y las relaciones entre ellos) sobre los criterios utilizados para evaluar la dimensionalidad, por ejemplo: índices de ajuste, presencia de errores de medición correlacionados y grado de carga cruzada, entre otros. (Anderson y Gerbing, 1988, Clark y Watson, 1995, Floyd y Widaman, 1995, Gerbing y Anderson, 1988, Hattie, 1985, Kumar y Dillon, 1987).

## Confiabilidad

La confiabilidad se refiere a la parte de la medición que se debe a efectos permanentes que persisten de una muestra a otra. Hay dos tipos de confiabilidad basados en la literatura psicométrica, que es referente en las ciencias económico-administrativas

1. *Test-retest (o de estabilidad temporal):* referido a la correlación entre la puntuación de la misma persona en el mismo conjunto de ítems o indicadores pero en dos puntos diferentes en el tiempo. Se refiere a la estabilidad de las respuestas de un participante a lo largo del tiempo. Un coeficiente *test-retest* o *de estabilidad* generalmente se estima por la magnitud de la correlación entre las mismas medidas (y muestra) en diferentes ocasiones de evaluación. Si el coeficiente de estabilidad es bajo en magnitud, sin cambios en el constructo a lo largo del tiempo, la fiabilidad de la medida está en duda. Por lo tanto, la confiabilidad *test-retest* es útil porque ofrece información sobre el grado de confianza que se tiene de que la medición refleja del constructo y que es generalizable para otras ocasiones de evaluación (Haynes et al., 1999). Curiosamente, la fiabilidad *test-retest* no se ha evaluado en el uso o desarrollo de la escala tan frecuentemente como la consistencia interna (Robinson et al., 1991). Al momento, las estimaciones de *test-retest* están disponibles en muy pocas escalas de las ciencias económico-administrativas, por lo que se sugiere utilizarla además de usar otros procedimientos de evaluación de confiabilidad y validez .
2. *Consistencia interna (de evaluación de interrelación entre los indicadores)*: los indicadores o ítems que componen la escala (o subescala) deben mostrar altos

niveles de consistencia interna. Algunos de los criterios comúnmente utilizados para evaluar la consistencia interna son correlaciones individuales corregidas de ítem a total (*individual corrected ítem-to-total correlations*), el promedio de la correlación interítem entre los ítems de la escala **(***the average interítem correlation among escale ítems*) y una serie de fiabilidad de coeficientes (*reliability coefficients*, en Churchill, 1979; Cortina, 1993; DeVellis, 1991; Nunnally y Bernstein, 1994; Robinson et al., 1991)**.**

El coeficiente de confiabilidad de consistencia interna (*internal consistency reliability coefficent*) más ampliamente utilizado, es el coeficiente alfa de Cronbach *(Cronbach's coefficient alpha*) Cronbach (1951).

Si bien, existen algunas reglas generales que se refieren a lo que constituye un nivel aceptable de coeficiente alfa **(α)** la longitud de la escala debe ser considerada. Esto es, a medida que aumenta el número de ítems o indicadores, el alfa **(α)** tenderá a aumentar. Debido a que la parsimonia también es una preocupación en la medición (Clark y Watson, 1995, Cortina, 1993), una pregunta importante es *¿cuántos elementos se necesitan para medir un constructo?* La respuesta a esta pregunta depende parcialmente del dominio y las dimensiones de el constructo. Naturalmente, un constructo con un dominio amplio y múltiples dimensiones requerirá más indicadores para aprovechar adecuadamente la relación dominio/dimensiones que un constructo con un dominio estrecho y de pocas dimensiones. Dado que la mayoría de las escalas son autoadministradas, la fatiga y/o la falta de cooperación de los encuestados deben ser consideradas, por lo que la brevedad de la escala es a menudo ventajosa (Churchill y Peter, 1984; Cortina, 1993; DeVellis, 1991; Nunnally y Bernstein, 1994). Con el advenimiento del modelado de ecuaciones estructurales, otras pruebas de consistencia interna o de estructura interna (o de estabilidad interna) se hicieron disponibles.

La confiabilidad compuesta (*composite reliability* o confiabilidad del constructo), que es similar al coeficiente alfa **(α)**, se puede calcular directamente a partir de la producción de **LISREL, CALIS, EQS o AMOS** (véase Fornell y Larcker, 1981). Una prueba más estricta de la consistencia interna (estructura interna o estabilidad interna) implica evaluar la cantidad de varianza capturada por la medición de un constructo en relación con la cantidad de varianza debida al error de medida, esto es la índices de varianza media extraída (**IVE**/***AVE****. Average Variance Extracted*). Mediante el uso de una combinación de los criterios anteriores, es decir, correlaciones individuales corregidas de ítem a total (*individual corrected ítem-to-total correlations*), el promedio de la correlación interítem entre los ítems de la escala (*the average interítem correlation among escale ítems*), el alfa de Cronbach (*Cronbach's coefficient alpha*), la confiabilidad compuesta (*composite reliability*) y el **AVE,** las escalas pueden desarrollarse de manera eficiente sin sacrificar la consistencia interna.

## *Validez*

La validez de constructo se refiere a qué tan bien una medición mide realmente el constructo. La validez del constructo es el objetivo final en el desarrollo de un instrumento de evaluación y abarca todas las pruebas relativas a una medición

(Haynes et al., 1999). Hasta aquí, se puede observar que todas las pruebas relativas a una medición contribuyen a establecer la validez del constructo, por lo que una medición también debe a priori mostrar su dimensionalidad teórica así como dar evidencia de confiabilidad para ser considerada válida. Como tales, la dimensionalidad y la fiabilidad son necesarias, pero las condiciones pueden ser insuficientes para la validez de constructo. Aún existen desacuerdos en cuanto a la clasificación y los tipos de validez de constructo a aplicar, por lo que se tienen por el momento*: contenido, aparente, convergente, discriminante, relacionado a criterio, nomológica***,** *y la validez de grupo conocido* .

**1.** *La validez de contenido.* Se refiere al contenido de los ítems o indicadores; se han delineado dos subtipos: *el contenido (content validity) y la validez aparente (face validity):*

-El término *validez de contenido* ha sido definido de muchas maneras, con la mayoría de las definiciones enfatizando que los ítems o indicadores de una medida son una muestra apropiada del dominio teórico del constructo (Messick, 1993; Nunnally y Bernstein, 1994). La mayoría de las definiciones son consistentes con lo siguiente:

L*a validez del contenido refleja el grado en que los elementos de un instrumento de evaluación son relevantes y representativos del constructo objetivo para un propósito de evaluación particular* (Haynes et al., 1995, p.238**)** .

Los *elementos* se refieren al contenido de los ítems o indicadores individuales, formatos de respuesta e instrucciones a los encuestados, y *representatividad* se refiere al grado en que los elementos son proporcionales a las facetas (**dominios**) del constructo objetivo y en la medida en que se ha muestreado todo el dominio. Es decir, los ítems deben aparecer consistentes con el dominio teórico del constructo en todos los aspectos, incluyendo formatos de respuesta e instrucciones.

En las escalas de desarrollo que son válidas para el contenido, se recomienda generalmente que se generen una serie de ítems o indicadores que *aprovechen el dominio del constructo*, que los indicadores sean examinados por expertos en la literatura, y que las pruebas piloto sobre muestras de las poblaciones pertinentes para refinar el conjunto de ítems o indicadores (DeVellis, 1991; Robinson et al., 1991). Haynes et al. (1995) ofrecen una excelente descripción de los procedimientos para establecer la del contenido.

**2.** La validez aparente (*face validity*), se ha referido como la simple apariencia de que una medida tiene validez (Kaplan y Saccuzzo, 1997, p.132). Aunque los términos validez aparente (face validity) y validez de contenido (content validity) se han utilizado indistintamente, algunos argumentan que la validez aparente debe estar separada de la validez de contenido (Anastasi y Urbina, 1998; Nevo, 1985). Otros van un poco más lejos e indican que la validez aparente (face validity) vs la validez del contenido (*content validity*) se define en términos de investigador y encuestado. Un instrumento con una gran carga de validez aparente (*face validity*) mejora su uso en situaciones prácticas (entre otras cosas) induciendo la cooperación de los encuestados a través de la facilidad de uso, nivel de lectura adecuado, claridad, facilidad de leer las instrucciones y formatos de respuesta fáciles de usar. Una definición un tanto aceptada de la validez

aparente (*face validity*) implica que un instrumento o prueba, cuando se utiliza: *debe ser práctico, pertinente y relacionado con los propósitos del instrumento [prueba], además de tener validez pragmática o estadística..*es decir, no sólo debe ser válida, sino que también debe parecer válida para los encuestados (Nevo, 1985, p.228). Por lo tanto, la validez aparente (*face validity*) puede estar más preocupada con lo que los encuestados de las poblaciones relevantes infieran con respecto a lo que se está midiendo.

3. *Validez convergente se* refiere al grado en que se relacionan dos medidas diseñadas para medir el mismo constructo. La convergencia se encuentra si las dos medidas diferentes del mismo constructo están altamente correlacionadas.
4. *La validez discriminante* valora el grado en que se relacionan dos medidas diseñadas para medir constructos similares pero conceptualmente diferentes. Una correlación baja a moderada a menudo se considera evidencia de validez discriminante. Las matrices multi-trama-multimetodo (**MTMM***. MultiTrait-MultiMethod Matrices*) se han utilizado con frecuencia para evaluar la validez convergente y discriminante cuando se requieren métodos de medición diferentes; es decir, autoinforme vs. observacional (Campbell y Fiske, 1959).
5. *La validez nomológica* se ha definido como el grado en que las predicciones de una red teórica formal que contiene el concepto bajo escrutinio son confirmadas (Campbell, 1960). Evalúa el grado en el que los constructos que están teóricamente relacionadas lo están también empíricamente (es decir, sus medidas se correlacionan significativamente en la dirección prevista). También existen pautas para establecer validez nomológica, pero también han sido criticadas (Peter, 1981). Al igual que con la consistencia interna y la validación convergente y discriminante, recientemente se han utilizado paquetes de ecuaciones estructurales para evaluar la validez nomológica de las medidas de escala. Varios libros (por ejemplo, Bollen, 1989, Byrne, 2001, Hayduk, 1996, Hoyle, 1995, Schumacker y Lomax, 1996) y artículos (por ejemplo, Anderson y Gerbing, 1988, Bagozzi, Yi y Phillips, 1991, Bentler y Chou , 1987) ilustran técnicas de modelado, criterios evaluativos y guías para lo que constituye validez nomológica.
6. La *validez relacionada con el criterio (criterion-related validity)* varían, y algunas definiciones son similares a las definiciones de otros tipos de validez. Por ejemplo, la validez relacionada con el criterio se ha referido como el grado en que una medida covaría con medidas previamente validadas (*gold-standard*) de los mismos constructos (Haynes et al., 1999). Esta definición es similar a la de la validez convergente. La validez relacionada con el criterio también se ha referido como la relación en que una medida corresponde a otra medida de interés (Kaplan y Saccuzzo, 1997). Algunos afirman que es la misma que la *validez predi*ctiva, o *relación funcional* entre un predictor y un criterio antes, durante o después de aplicar un predictor (Nunnally y Bernstein, 1994)**.** Este enfoque se basa en la relación temporal entre el predictor y su criterio, es decir, la validez *post-dictiva, concurrente y predictiva.* Lo que la mayoría de las definiciones tienen en común es que la validez relacionada con los criterios se evalúa mediante un patrón de relaciones teóricamente especificadas entre una medida y un criterio a menudo denominado coeficiente de validez *(validity coefficient).*
7. *La validez del grupo conocido (known-group validity)* implica la capacidad de la

medida para distinguir de manera fiable entre grupos de personas que deben anotar alto en el rasgo y bajo en el rasgo. Como ejemplo, una persona que es verdaderamente conservadora debe obtener una puntuación significativamente mayor en una escala de conservadurismo que una persona liberal y las personas de ventas en el negocio de inmobiliarias y grandes empresas farmacéuticas deben diferir en sus niveles de orientación al cliente (Saxe y Weitz, 1982). Por lo tanto, las diferencias de puntuación media entre los grupos de un constructo dado pueden utilizarse como evidencia de la validez del grupo conocido. Ver más en Jarvis y Petty (1996).

# El rol del investigador en las ciencias económico-administrativas

Se debe tomar en cuenta, que los investigadores deben mantener un perfil adecuado a sus funciones en la creación de escalas, tal como se suigiere en la **Tabla 1.2.**

**Tabla 1.2. Rol del investigador en la creación de escalas**

| Rol | Descripción |
|---|---|
| **Creativo** | Es fundamental no relegar a un segundo término el ingenio y la creatividad;; la creatividad es absolutamente necesaria, pese a que tiende a utilizar métodos poco aceptados. |
| **Competencia** | Consiste en aplicar cada método en su justa medida y para su ámbito de acción; la metodología científica no puede utilizarse para explicar qué está bien y qué está mal. |
| **Ético** | El investigador no escapa a la influencia que pueda ejercer sobre ciertos colectivos al utilizar determinados métodos, pero debe respetar las normas de comportamiento al desarrollar su estudio |
| **Técnico** | Para realizar con garantías una investigación es necesario que el investigador esté adiestrado en las técnicas de investigación, posea suficiente bagaje metodológico, sea conocedor avanzado de su campo de estudio y domine las técnicas cuantitativas. |
| **Comunicador** | Se puede decir que no existe la investigación que no se comunica, por ello las dos tareas fundamentales son:<br>a) Informar a la comunidad sus líneas de trabajo, con ello se suelen evitar duplicaciones y se incrementa la cooperación entre colegas.<br>b) Informar a los colegas, las empresas y la sociedad de los hallazgos realizados, de su aplicabilidad y de su posible idoneidad para mejorar aspectos concretos. |
| **Imparcialidad** | El investigador necesita cierto grado de aislamiento respecto de su entorno y de inmunidad respecto del sistema burocrático de su organización, tratando de ser lo más neutral y objetivo posible. |
| **Escepticismo** | El científico no es un creyente que trata de validar sus opiniones o presunciones, sino que debe, por principio, dudar de las explicaciones oportunistas, vulgares, demagógicas o que no se han obtenido mediante una metodología científica. |

Fuente: Maldonado-Guzmán,(2016)

## El Cientificismo

Uno de los peligros que acecha al investigador es el entendiendo como tal a toda aquella persona, argumento o propuesta que tiene la pretensión no legítima de tener o aparentar naturaleza científica. En esta definición hay **tres elementos** importantes a considerar (Maldonado-Guzmán, 2016):

- *La pretensión,* que se suele referir a una actitud o un comportamiento de un individuo
- *La ilegitimidad,* que se refiere a no guardar las normas de la comunidad científica o utilizar argumentos, elementos y teorías no aceptadas en ella
- La naturaleza científica del fenómeno o la *calificación profesio***nal** de los sujetos que sustentan un argumento o propuesta.

No obstante, tratar de sustentar la validez universal de esta definición y de sus connotaciones sería ya de por sí una actitud inadecuada, pues hay importantes *divergencias sobre lo que está bien o lo que está mal, sobre lo que es legítimo e ilegítimo, lo que es científico y lo que es cientificista: lo que para unos es ciencia, para otros es cientificismo.*

Los argumentos cientificistas suelen basarse muchas veces en analogías y pueden pretender:

- Ampliar una disciplina a ámbitos no aceptados actualmente.
- Aplicar directamente conocimientos de otras disciplinas sin considerar su naturaleza, objeto y condiciones.
- Dar por válidos argumentos o teorías no adecuados o insuficientemente probados por la comunidad científica.

## Determinación del problema a investigar

La investigación inicia cuando se determina qué problema o fenómeno hay que estudiar y por qué hay que hacerlo, y no es suficiente saber la temática sobre la que se ha de trabajar. También se deberá tener muy claro el porqué de la elección del tema (importancia, idoneidad, oportunidad), para que al momento de elegir un tema éste debe ser , como se muestra en la **Tabla 1.3.**

**Tabla 1.3. Características que debe tener el tema a diseñar una escala de medición**

| Característica | Descripción |
|---|---|
| **Viable** | Significa que el estudio debe ser realizable, y que el problema objeto de análisis es alcanzable; muchas veces puede haber problemas derivados del propio investigador o del tema elegido. |
| **Científico** | El concepto a transmitir es claro: el tema debe tener valor científico;; para que esto ocurra la metodología debe ser rigurosa, adaptarse al objetivo de estudio y ser factible de analizar y aplicar. |
| **Propio** | El problema debe enmarcarse dentro del área de conocimiento en la que se es especialista;; esto no quiere decir que haya que huir de temáticas multidisciplinares, sino que hay que rodearse de un equipo también multidisciplinar. |
| **Preciso** | El problema a investigar debe quedar claramente formulado y |

| | |
|---|---|
| | delimitado; hay que huir de temas ambiguos y genéricos. |
| **Significante** | Aunque es perfectamente posible realizar un estudio sobre un caso aislado, lo ideal es que el tema elegido permita obtener hallazgos susceptibles de generalización o, en su defecto, que representen el comportamiento de un colectivo o un fenómeno para un período de tiempo lo más amplio posible. |
| **Importante** | La investigación está para buscar conocimiento, por lo que el *problema elegido debe representar alguna novedad*, en determinadas ocasiones, comprobar ciertos hechos o comportamientos es una tarea razonable y justificada. |
| **Motivación** | Llama la atención la gran cantidad de investigadores escasamente motivados por su temática; si no hay predisposición y entusiasmo hacia el tema que se estudia, se termina aplicando la *ley del mínimo esfuerzo*, lo que produce detrimento en la calidad del trabajo. |
| **Enfoque** | Hay que destacar la fuerte tendencia de muchos investigadores a no plantear cuestiones sobre las teorías existentes y, así, analizar y describir fenómenos o modelos procedentes de dichas teorías. |

Fuente: Maldonado-Guzmán (2016)

# La medición en las ciencias económico-administrativas

El investigador, deberá siempre tener en cuenta que su aportación en las ciencias económico-administrativas se basará en el diseño pertinente y adecuado de una escala de medición, por lo que se sugiere considere lo mencionado en la **Tabla 1.4.**

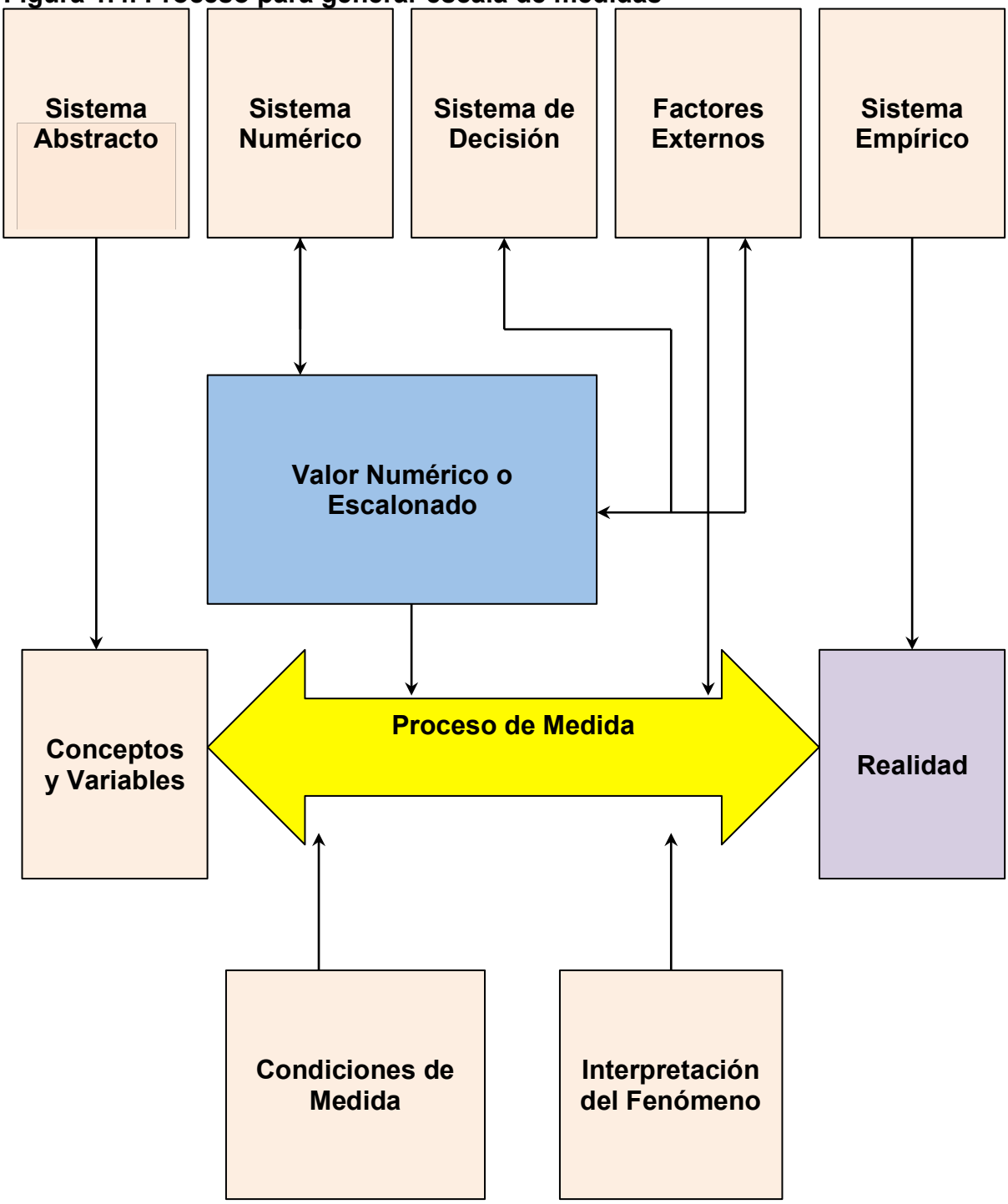
**Tabla 1.4. Puntos basicos a considerar en el diseño de una escala de medición**

| |
|---|
| • Las escalas conforman la parte fundamental del cuestionario |
| • La escala está compuesta por diferentes ítems o preguntas |
| Los enunciados de los ítems deben ser:<br>• Cortos (20 palabras aproximadamente)<br>• Lenguaje simple<br>• Expresen una sola idea<br>• Significado unívoco, sin ambigüedades<br>• Evitar enunciados con negaciones<br>• Misma proporción de elementos en sentido favorable y desfavorable.<br>• Evitar términos absolutos (siempre, todas, nadie...) salvo que el objetivo sea la polarización. |

Fuente: propia

Churchill (1979) afirma que cuando un investigador pone en conexión el *sistema abstracto* con el *sistema empírico* utilizando números para valorar e interpretar los fenómenos objeto de su estudio, *es cuando se produce una medición*. **Ver Figura 1.4.**

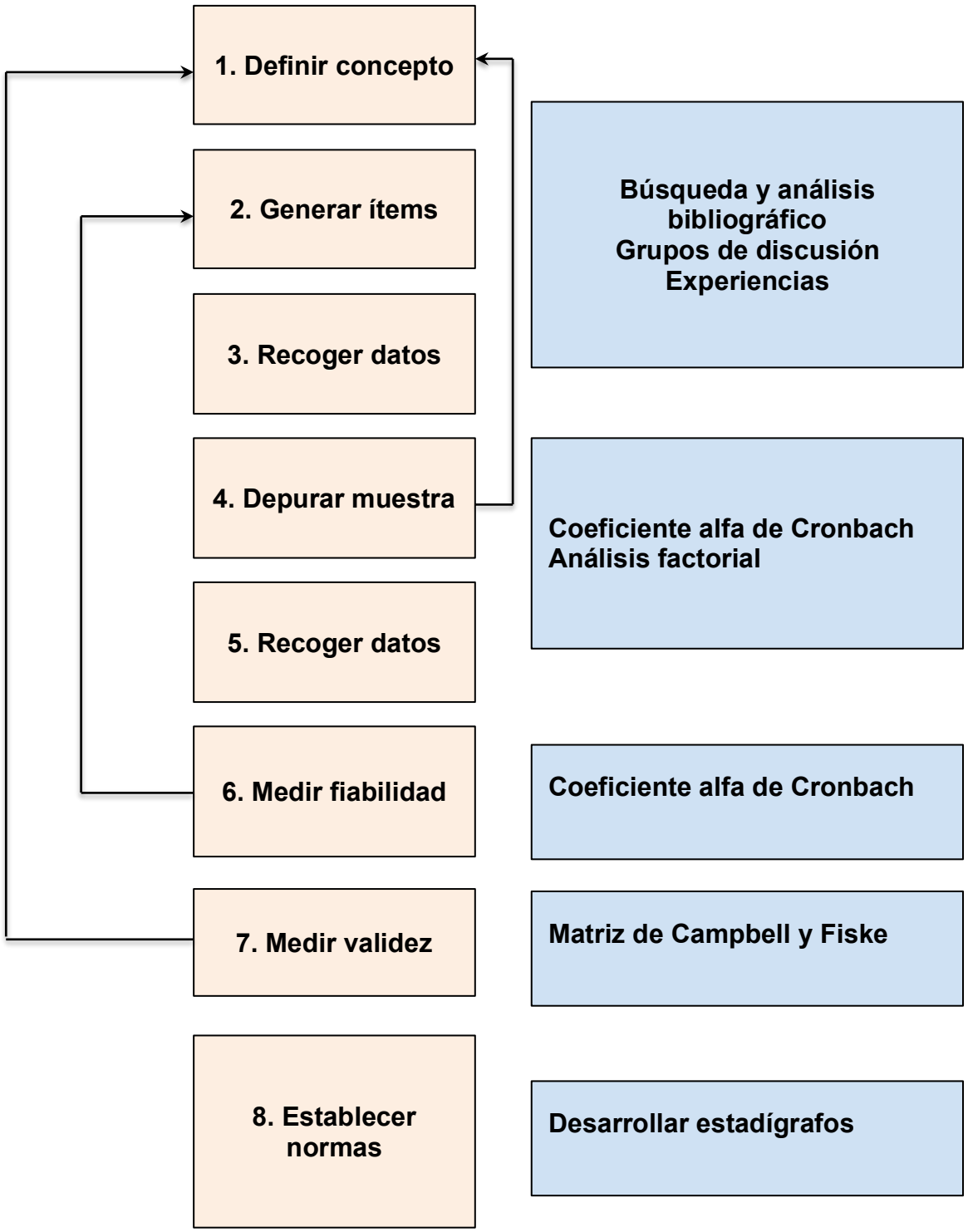
**Figura 1.4. Proceso para generar escala de medidas**




Fuente: Churchill (1979) en Maldonado-Guzmán (2016)

Se han sugerido diversas operativas para crear medidas, siendo la de Churchill (1979) la más aceptada por la comunidad investigadora Ver **Figura 1.5**

**Figura 1.5. Proceso de creación de escalas**




Fuente: Churchill (1979) en Maldonado-Guzmán (2016)

# Concepto y variables: definición y tipos

A manera de repaso, mostramos la **Tabla 1.5.** (Mejía-Trejo, 2017c, Capítulos 2 y 3)

**Tabla 1.5. Concepto y variables: definición y tipos**

Conceptos y variables

- El primer paso para poder medir es desarrollar los conceptos entiendo que concepto es *una representación literaria y abstracta de una realidad o idea.*
- Para que sea completo el concepto es necesario que todas las dimensiones de la mencionada realidad o idea estén perfectamente delimitadas, clarificadas, expresadas en lenguaje consensuado y basadas, siempre que sea posible, en desarrollos anteriores que le sean afines.
- Por otra parte, que el concepto sea una representación literaria no impone que su estructura no sea científica o bien al contrario, al enunciar un concepto el propósito ha de ser científico.
- En muchos casos, los conceptos se construyen a partir de generalizaciones de casos concretos o se limitan a plantear representaciones ideales de algún aspecto de la realidad. En cualquier caso es imprescindible que la construcción de conceptos cumpla al menos con los siguientes criterios:
- El contenido *no ha de ser circular*
- Los elementos que lo definen han de ser precisos y unitarios, por ello debe rechazarse todo concepto que en su formulación utilice expresiones dilógicas.
- La expresión (definición) del concepto ha de señalar los elementos esenciales de lo que se trata de definir.
- Debe tener potencial operacionalizador, es decir, deben tener capacidad para ser puestos a prueba. No se deben considerar aquellos conceptos que por su propia naturaleza no puedan ser comprobados directa o indirectamente.
- Los conceptos pueden ser clasificados en objetivos, subjetivos, científicos y literarios. Los científicos se pueden dividir a su vez en generales (sirven para muchas disciplinas, desarrollando fenómenos o ideas de rápida, fácil y amplia aplicación) y específicos (sólo tienen sentido operativo en un campo concreto). Un ejemplo se muestra en esta tabla:

| | Objetivos | Subjetivos |
|---|---|---|
| Generales | Dimensión<br>Saturación | Calidad/Estilo de vida<br>Liderazgo |
| Específicos | Apalancamiento<br>Participación de mercado | Flexibilidad organizativa<br>Lealtad a la marca |

## Requisitos de variables e índices

Así también, se deberá guardar una clara diferencia entre variables e índices, como se menciona en la siguiente tabla:

| Variable | Índice |
|---|---|
| -Responden a una definición formal.<br>-Relacionarse con el contenido significativo del concepto.<br>-Discriminación del contenido.<br>-Capacidad para ser medida. | -Capacidad definitoria<br>-Debe tener relación con el contenido significativo de la variable.<br>-Internamente consistente.<br>-Ofrecer una valoración específica. |

## Ejemplos de lo anterior

| Conceptos | Variables | Índice (para cada variable) |
|---|---|---|
| Relaciones en el canal de distribución | Conflictos en el canal<br>Control del canal<br>Poder en el canal | Escala de Schul (1987) Escala de Dewey y Welsh (1985) Escala de El-Ansary y Stern (1972) |
| Segmento de mercado | Demográfica Económica Psicográfica Comportamiento | Densidad de población<br>Nivel de ingresos VALS<br>Frecuencia de uso/común |
| Comportamiento de compra | Motivación social<br>Grado de innovación<br>Fidelidad a la marca | Índice de Moschis (1978) Escala de Hawes y Lumpkin (1984)<br>Tasa de recompensa |

## Valor numérico y escalas naturales

Una vez definidas las variables, se debe reconocer cuál es la naturaleza de su valor numérico o como se sabe, las valoraciones que se otorgan a las variables provienen del sistema numérico, que tiene las siguientes características:

- *Universalidad:* el proceso de numeración es estándar y consensuado
- *Operacionalidad:* los números admiten operaciones y manipulaciones
- *No Limitación:* el conjunto de los números es ilimitado, tanto categóricamente como dentro de un intervalo o rango dado. No obstante, la serie de los números se puede truncar a voluntad por el investigador.
- *Identificación:* cada número representa una situación concreta de la cualidad que representa e implica una valoración exclusiva.
- *Ordenación:* la asignación de números a diversos sujetos considerando una cualidad determinada implica una orden natural si dichos sujetos se comparan. Los números poseen una ordenación (0, 1, 2, ..., 9) y la transmiten cuando son asignados a cualidades de los objetos o de los fenómenos.

| |
|---|
| -Distancia: se pueden cuantificar las diferencias entre los números, y tales diferencias tienen sentido matemático.<br>• *Origen:* la medición tiene un origen, cuyo significado es la ausencia de valoración (0). Los números negativos indican la dirección del sistema y de su interpretación. |
| Problemas en la definición de escalas |
| • No todas las mediciones son fáciles de realizar; la mayoría de las variables utilizadas no son directamente observables (calidad de servicio) o no se les puede otorgar una valoración directa (imagen de marca), en estos casos se debe crear un instrumento de medida que permita medir la variable y estimar la naturaleza del concepto de forma fiable y válida.<br>• En las ciencias económico-administrativas, donde el estudio por encuesta está bastante generalizado, se utilizan escalas compuestas como las psicológicas (actitudes) y las sociológicas. En el campo del marketing, Bruner y Hensel (1997) han recopilado más de 800 instrumentos (escalas), y Bearden y Netemeyer (1998) han analizado otros 124. De cualquier forma, un investigador que necesite realizar un estudio puede acudir a las publicaciones o desarrollar su propio instrumento de medida. |

Fuente: Netenmeyer et al. (2002)

# Etapas para el desarrollo de una escala

Existen numerosos artículos y libros sobre cómo desarrollar una escala, tales como Churchill (1979), DeVellis (1991), Spector (1992), Nunnally y Bernstein (1994), Clark y Watson (1995), Haynes et al. (1999) de los que sugieren pasos y procedimientos que varían en base a los objetivos y propósitos de la medición, pero que comparten un conjunto común de pautas para el desarrollo de la escala, siendo brevemente, los mencionados en la **Tabla 1.6**.

**Tabla 1.6. Etapas para la creación de una escala**

| Etapa | Descripción |
|---|---|
| **1 Definición del constructo y dominio de contenido** | Como hemos dicho a lo largo de este capítulo introductorio, la importancia de la teoría en el desarrollo de la escala no puede ser exagerada, y desarrollar y refinar una teoría requiere una revisión bibliográfica exhaustiva. Durante la revisión de la literatura del estado del arte y los procesos de desarrollo de la teoría, se deben enfatizar varias cuestiones: Cuestiones a considerar:<br>**a.**La importancia de una definición clara del constructo, el dominio del contenido y el rol de la teoría y el marco teórico de soporte;<br>**b.**El enfoque de los ítems o indicadores como *efecto reflectivo* (*reflective*) vs. *indicadores formativos (formative).*<br>**c.**Dimensionalidad del constructo: ¿unidimensional, multidimensional o construcción de orden superior? |
| **2 Generación y clasificaci** | Se deben considerar varias cuestiones, como:<br>**a.**Suposiciones teóricas sobre los ítems o indicadores.<br>**b.**Generar indicadores potenciales y determinar el formato de respuesta, es decir: |

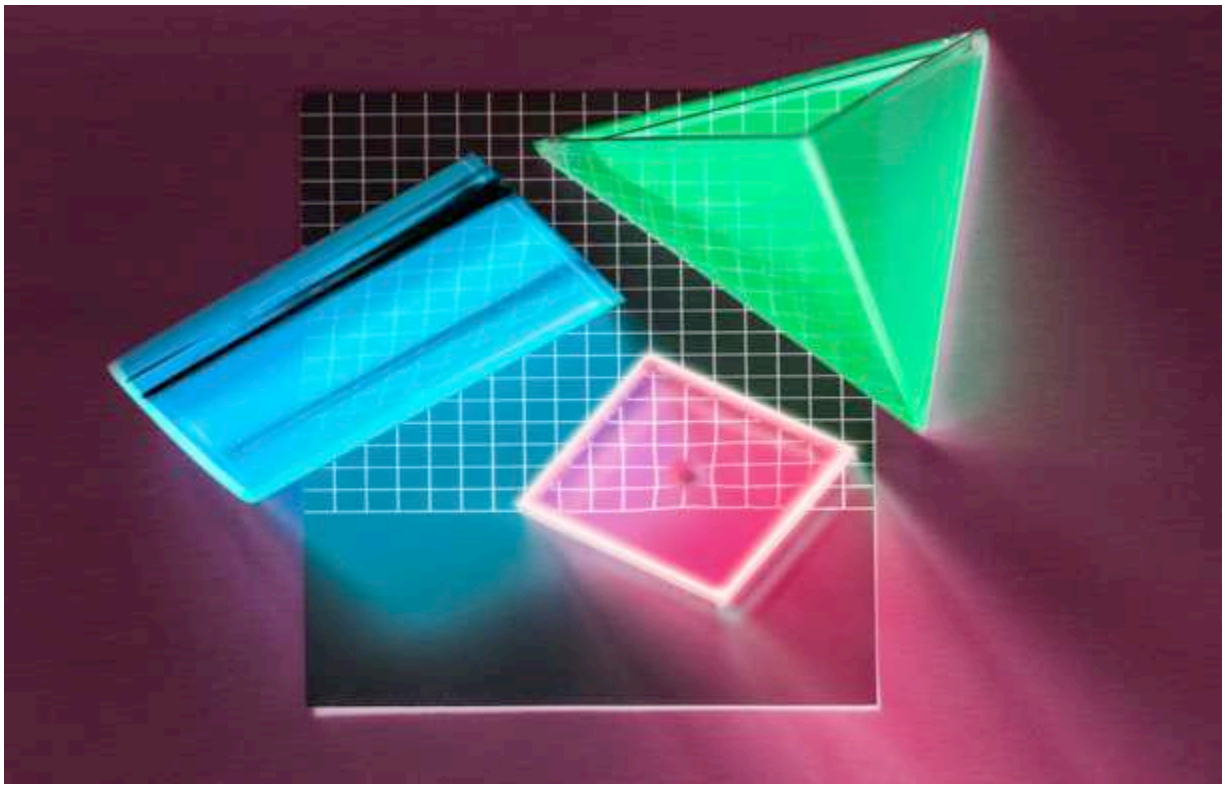
| | |
|---|---|
| **ón de indicadores de medición** | • Cuántos ítems como grupo inicial;<br>• Formato *Dicotómico vs Multicotómico;*<br>• Problemas con las palabras en los indicadores;<br>**c.**El enfoque en la validez de contenido en relación con la dimensionalidad teórica, y<br>**d.**Clasificación de indicadores por académicos, expertos e interesados (validez por contenido (*content validity*) y validez aparente (*face validity*) |
| **3 Diseño de la escala** | Una vez que se ha generado y juzgado un grupo adecuado de ítems, las pruebas empíricas de los ítems en muestras relevantes son el siguiente paso. Entre las cuestiones y procedimientos que deben considerarse se incluyen:<br>**a**.La prueba piloto como procedimiento de reducción-optimización de ítems o indicadores;<br>**b.**El uso de varias muestras de poblaciones relevantes para el desarrollo de la escala;<br>**c**.El diseño de estudios para probar las propiedades solicitadas como ciencias económico-administrativas;<br>**d.** Análisis inicial de ítems o indicadores vía análisis factorial exploratorio (**EFA.** *Exploratory Factor Analysis*);<br>**e.**Análisis inicial de ítems o indicadores y estimados de consistencia interna;<br>**f.**Estimados iniciales de validez; y<br>**g.**Retención de ítems para el siguiente conjunto de estudios. |
| **4 Finalización de la escala** | Se deben utilizar varios estudios para ayudar a finalizar la escala. Muchos de los procedimientos utilizados y cuestiones relacionadas con el perfeccionamiento de la escala también serán aplicables para derivar la forma final de la escala. Esto incluye:<br>**a.**La importancia de varias muestras de las poblaciones pertinentes;<br>**b.**El diseño de los estudios para probar los diferentes tipos de validez;<br>**c.**El análisis de los ítems o indicadores a través de **EFA**:<br>• Consistencia de **EFA** en los resultados entre los pasos 3 al 4;<br>• Derivación de una estructura factorial inicial con teoría y dimensionalidad<br>**d.**Análisis de ítems e indicadores con análisis factorial confirmatorio (***CFA.*** *Confirmatoty Factor Analysis*);<br>• Prueba de la estructura factorial teórica y el modelo especificado;<br>• Estudios de cruce de modelo factorial de invarianza (por ejemplo, análisis multigrupo)<br>**e.**Análisis de ítem o indicadores adicionales vía estimados de consistencia interna.<br>**f.** Estimaciones adicionales de validez;<br>**g.**Establecimiento de normas entre los estudios |

Fuente: Netemeyer et al. (2003)

# Conclusión

Para que un constructo sea valioso, debe tener relevancia teórica y/o práctica para el investigador en ciencias económico-administrativas. Por lo tanto, debe considerarse cuidadosamente lo que predice el constructo de interés y/o lo que predice el constructo de interés. Aquí, la noción de teoría y *conocer* la literatura es muy importante. Además, dada la importancia de la medición en las ciencias sociales, cualquier medida debe ser válida para permitir construir inferencias confiables a partir de estudios empíricos. Dicha validez se basa en cuán bien el constructo subyacente que se mide se basa en teoría. También en este capítulo introductorio, hemos resaltado los conceptos de dimensionalidad, fiabilidad y validez, así como resumimos una serie de pasos para derivar medidas con propiedades adecuadas. El resto de nuestro texto elabora sobre dimensionalidad, fiabilidad y validez, y los cuatro pasos en la construcción de escala.

# CAPÍTULO 2. DIMENSIONALIDAD



Esta primera etapa de la dimensionalidad de los constructos, es una parte importante del proceso de desarrollo de la escala. Es prácticamente imposible desarrollar buenas medidas de un constructo sin conocimiento de su dimensionalidad. Por ejemplo, considere la tarea de desarrollar una escala para medir el liderazgo. ¿Tiene el liderazgo una sola faceta o dimensión (es decir, es unidimensional), o tiene múltiples facetas o dimensiones (es decir, es multidimensional)? Si el liderazgo es un constructo unidimensional, entonces se requiere un sólo número para medirlo. Sin embargo, si se suscribe a la teoría ***MLQ5X*** (*Multifactor Leadership Questionnaire Level 5*) de Avolio y Bass (2004) (Ver Mejía-Trejo, et al. 2014), el liderazgo comprende las siguientes cuatro variables: transfotmacional, transaccional, pasivo y resultados de liderazgo con 12 dimensiones y 45 indicadores. Aún así, se tiene que la dimensionalidad ha llegado a significar cosas muy diferentes a los investigadores sustantivos. El propósito de este capítulo es discutir el significado de dimensionalidad de constructos y su relación con la confiabilidad. Específicamente, abordaremos las siguientes dos preguntas:

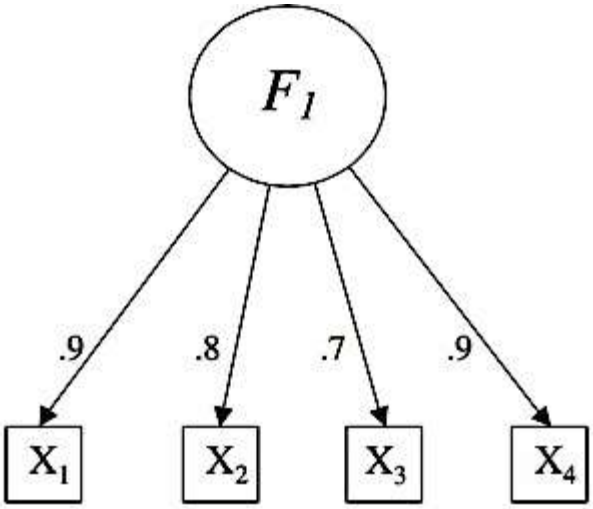
**a.** ¿Qué es la unidimensionalidad?, y

**b.**¿Existe una distinción entre la dimensionalidad de un constructo, la dimensionalidad de un conjunto de elementos manifiestos utilizados para medir el constructo y la dimensionalidad de sólo los elementos manifiestos utilizados para medir el constructo?

# Unidimensionalidad de constructo, de indicadores y de conjunto de indicadores

Observe la **Figura 2.1**, la cual muestra como la variable $F_1$ es la causa del efecto de las variables $X_1, X_2, X_3$ y $X_4$.

**Figura 2.1. Modelo de 1 factor con relación total de las variables**




Fuente: propia

A su vez es presentada la matriz de correlación de la **Tabla 2.1**, de la que de acuerdo con nuestro enfoque de mediciones efecto reflectivo (*reflective*) o causal formativo (*formative*) se puede apreciar de la **Figura 2.1** que $\mathbf{F_1}$ afecta $\mathbf{X_1, X_2, X_3}$ y $\mathbf{X_4}$. También se supone que las variables **X** se pueden medir u observar, mientras que $\mathbf{F_1}$ no se puede medir ni observar.

**Tabla 2.1. Matriz de correlación del Modelo de la Figura 2.1**

| | $F_1$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|---|---|---|---|---|---|
| $F_1$ | 1.00 | | | | |
| $x_1$ | 0.90 | 1.00 | | | |
| $x_2$ | 0.80 | 0.72 | 1.00 | | |
| $x_3$ | 0.70 | 0.63 | 0.56 | 1.00 | |
| $x_4$ | 0.90 | 0.81 | 0.72 | 0.63 | 1.00 |

Fuente: SPSS IBM 21, con datos propios.

Es decir, $\mathbf{F_1}$ típicamente se denomina como constructo no observable, latente o subyacente, y las variables **X** se denominan indicadores, elementos o variables manifiestas del constructo latente o subyacente. La correlación parcial, $\mathbf{r_{12*3}}$, entre

cualquier par de variables (es decir, las variables 1 y 2) después de **remover** o **parcializar** el efecto de una tercera variable (es decir, la variable 3) está dada por la **Ecuación. 2.1.**

**Ecuación 2.1. Correlación parcial $r_{12*3}$**

$$r_{12\bullet 3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1 - r_{13}^2}\ \sqrt{1 - r_{23}^2}}\ .$$

Así, utilizando los valores de la **Tabla 2.1** y la ecuación anterior, la correlación parcial entre $\mathbf{X_1,X_2}$ después de la aparición del efecto de $\mathbf{F_1}$ es igual a:

$$\frac{0.72 - 0.9 \times 0.8}{\sqrt{1 - .9^2}\sqrt{1 - .8^2}} = 0.$$

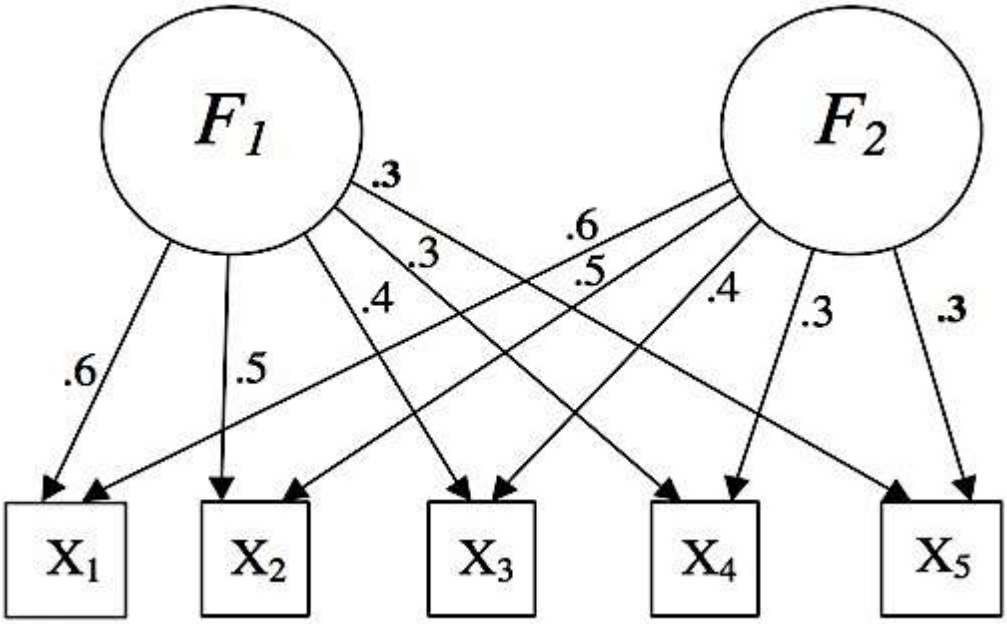
Así, al usar la **Ecuación 2.1,** se puede demostrar fácilmente que las correlaciones parciales entre todos los pares de variables **X** dados en la **Tabla 2.1** son iguales a cero. Es decir, una vez que el efecto de $\mathbf{F_1}$ ha sido removido o parcializado, las correlaciones parciales o las relaciones entre las variables **X** desaparecen. En otras palabras, $\mathbf{F_1}$ es responsable de todas las relaciones entre las variables **X**; por lo tanto, $\mathbf{F_1}$ se conoce como *factor común.* El conjunto de indicadores $\mathbf{X_1,X_2,X_3}$ y $\mathbf{X_4}$ se considera que es unidimensional porque las correlaciones entre ellos, después de haber sido parcializadas para el efecto de un único factor común (es decir, $\mathbf{F_1}$), son iguales a cero. Por lo tanto, un conjunto de elementos se considera unidimensional si las correlaciones entre ellos pueden ser contabilizados por un sólo factor común. (McDonald, 1981 y Hattie, 1985).

Observe que cada uno de los cuatro indicadores de la **Figura 2.1** es una medida de uno y sólo un constructo y, por lo tanto, *cada indicador es unidimensional*. Es decir, se considera que un elemento es unidimensional si es una medida de un solo constructo o factor latente o subyacente. Ahora bien, si se toman múltiples conjuntos de **n** elementos del dominio del constructo y las correlaciones parciales entre cada conjunto son iguales a cero, entonces se dice que el **constructo es unidimensional**. Sin embargo, *es posible que un conjunto de indicadores pueda ser unidimensional, pero el constructo o los indicadores individuales pueden no ser unidimensionales*, circunstancia que se explicará más adelante.

# Multidimensionalidad de constructo, de indicadores y de conjunto de indicadores

Observe la **Figura 2.2**, la cual muestra como la variable $\mathbf{F_1}$ es la causa del efecto de las variables $\mathbf{X_1}$,$\mathbf{X_2}$,$\mathbf{X_3}$, $\mathbf{X_4}$ y $\mathbf{X_5}$. así como la variable $\mathbf{F_2}$ es la causa del efecto de las variables $\mathbf{X_1}$,$\mathbf{X_2}$,$\mathbf{X_3}$, $\mathbf{X_4}$ y $\mathbf{X_5}$.

**Figura 2.2. Modelo de dos factores con relación total de las variables**




Fuente: propia

La **Tabla 2.2** muestra la matriz de correlación para el modelo de dos factores dado en la **Figura 2.2.**

**Tabla 2.2. Matriz de correlación del Modelo de la Figura 2.2**

| | $F_1$ | $F_2$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ |
|---|---|---|---|---|---|---|---|
| $F_1$ | 1.00 | | | | | | |
| $F_2$ | 0.00 | 1.00 | | | | | |
| $x_1$ | 0.60 | 0.60 | 1.00 | | | | |
| $x_2$ | 0.50 | 0.50 | 0.60 | 1.00 | | | |
| $x_3$ | 0.40 | 0.40 | 0.48 | 0.40 | 1.00 | | |
| $x_4$ | 0.30 | 0.30 | 0.36 | 0.30 | 0.24 | 1.00 | |
| $x_5$ | 0.30 | 0.30 | 0.36 | 0.30 | 0.24 | 0.18 | 1.00 |

Fuente: SPSS IBM 21, con datos propios.

En la que las variables **X** se ven afectadas por los factores comunes $\mathbf{F_1}$ **y** $\mathbf{F_2}$. La

correlación parcial entre $\mathbf{X_1}$ y $\mathbf{X_2}$ después de que el efecto de $\mathbf{F_1}$ es parcializado, es igual a:

$$\frac{0.60 - 0.6 \times 0.5}{\sqrt{1 - 0.6^2}\ \sqrt{1 - 0.5^2}} = 0.43.$$

Similarmente, la correlación parcial entre $\mathbf{X_1}$ y $\mathbf{X_2}$ después de que el efecto de $\mathbf{F_2}$ parcializado, es igual a:

$$\frac{0.60 - 0.6 \times 0.5}{\sqrt{1 - 0.6^2}\ \sqrt{1 - 0.5^2}} = 0.43.$$

correlaciones parciales no son iguales a cero. Esto sugiere que las correlaciones entre las variables no pueden ser contabilizadas por un solo factor y, por lo tanto, el conjunto de elementos no es unidimensional. La correlación parcial entre cualquiera de las dos variables 1 y 2 después de parcializar el efecto de las variables 3 y 4 está dada por la **Ecuación 2.2.**

**Ecuación 2.2. Correlación parcial $r_{12*34}$**

$$r_{12\bullet 34} = \frac{r_{12\bullet 4} - r_{13\bullet 4} \times r_{23\bullet 4}}{\sqrt{1 - r_{13\bullet 4}^2}\ \sqrt{1 - r_{23\bullet 4}^2}}$$

De esta forma, al usar la **ecuación 2.2**, la correlación parcial entre $\mathbf{X_1}$ y $\mathbf{X_2}$ después de la parcelación de los efectos de $\mathbf{F_1}$ y $\mathbf{F_2}$ es igual a cero, como se muestra:

$$r_{12\bullet F_1} = \frac{.60 - .6 \times .50}{\sqrt{1 - .6^2}\ \sqrt{1 - .5^2}} = .43$$

$$r_{12\bullet F_2} = \frac{.60 - .6 \times .50}{\sqrt{1 - .6^2}\ \sqrt{1 - .5^2}} = .43$$

$$r_{12\bullet F_1, F_2} = \frac{r_{12\bullet F_2} - r_{1F_1\bullet F_2} \times r_{2F_1\bullet F_2}}{\sqrt{1 - r_{1F_1\bullet F_2}^2}\ \sqrt{1 - r_{2F_1\bullet F_2}^2}}$$

$$r_{1F_1\bullet F_2} = \frac{r_{1F_1} - r_{1F_2} \times r_{F_1F_2}}{\sqrt{1 - r_{1F_2}^2}\ \sqrt{1 - r_{F_1F_2}^2}} = \frac{.6 - .6 \times 0}{\sqrt{1 - .6^2}\ \sqrt{1 - 0}} = .75$$

$$r_{2F_1\bullet F_2} = \frac{r_{2F_1} - r_{2F_2} \times r_{F_1F_2}}{\sqrt{1 - r_{2F_2}^2}\ \sqrt{1 - r_{F_1F_2}^2}} = \frac{.5 - .5 \times 0}{\sqrt{1 - .5^2}\ \sqrt{1 - 0}} = .58$$

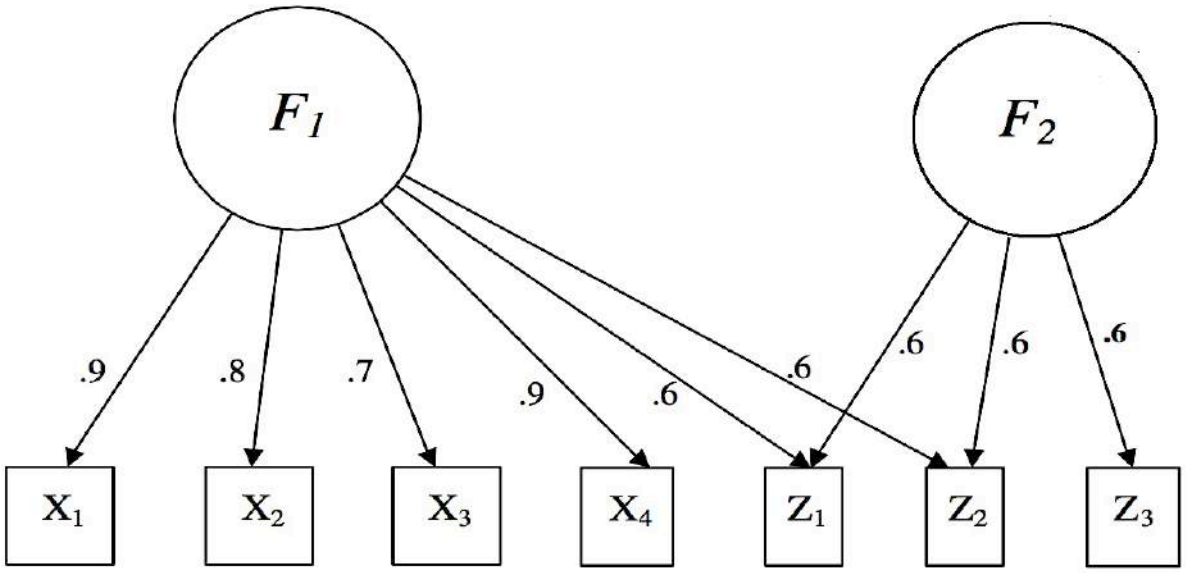
$$r_{12\bullet F_1 F_{21}} = \frac{.43 - .75 \times .58}{\sqrt{1 - .75^2}\ \sqrt{1 - .58^2}} \approx 0$$

Es posible demostrar fácilmente que las correlaciones parciales entre el conjunto $\mathbf{X_1}$ **a** $\mathbf{X_5}$ de variables, después de controlar para $\mathbf{F_1}$ y $\mathbf{F_2}$ son todas iguales a cero. Es decir, son necesarios dos factores o constructos latentes o subyacentes para explicar las correlaciones entre las variables $\mathbf{X_s}$ y, por lo tanto, dos dimensiones representan el conjunto de variables $\mathbf{X_1}$ **a** $\mathbf{X_5}$. La dimensionalidad de un conjunto dado de variables es igual al número de constructos latentes o subyacentes necesarios para dar cuenta de las correlaciones entre las variables. Observe además que cada uno de los cinco indicadores representa a dos constructos. Es decir, la dimensionalidad de cada indicador o constructo es igual a dos, porque cada uno de estos elementos es una medida de dos constructos. Una vez más, si se toman múltiples conjuntos de elementos del dominio de los dos constructos y las correlaciones parciales entre las variables de cada conjunto después de eliminar el efecto de $\mathbf{F_1}$ y $\mathbf{F_2}$ son iguales a cero, entonces *se dice que el constructo es multidimensional (de dos dimensiones en este caso)*. El razonamiento anterior se puede extender a más de dos constructos. Si dos factores no reducen las correlaciones parciales a cero, entonces el constructo tiene más de dos dimensiones. En general, la dimensionalidad de un conjunto de elementos es igual al número de constructos o factores necesarios para reducir las correlaciones parciales a cero. Además, si se extraen múltiples conjuntos de elementos del dominio del constructo y se necesitan factores *n* para tener en cuenta las correlaciones entre los ítems, entonces *se dice que la dimensionalidad del constructo es n.*

# Implicaciones de la unidimensionalidad

Observe la **Figura 2.3**, la cual muestra como la variable $\mathbf{F_1}$ es la causa del efecto de las variables $\mathbf{X_1}$,$\mathbf{X_2}$,$\mathbf{X_3}$ y $\mathbf{X_4}$ así como del efecto de las variables $\mathbf{Z_1}$ y $\mathbf{Z_2}$. Aunado a lo anterior, como la variable $\mathbf{F_2}$ es la causa del efecto de las variables $\mathbf{Z_1}$,$\mathbf{Z_2}$ y $\mathbf{Z_3}$.

**Figura 2.3. Modelo de dos factores con relación parcial de las variables**




Fuente: propia

La **Tabla 2.3** muestra la matriz de correlación para el modelo de dos factores dado en la **Figura 2.3.**

**Tabla 2.3. Matriz de correlación del modelo de la Figura 2.3**

| | $F_1$ | $F_2$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $z_1$ | $z_2$ | $z_3$ |
|---|---|---|---|---|---|---|---|---|---|
| $F_1$ | 1.00 | | | | | | | | |
| $F_2$ | 0.00 | 1.00 | | | | | | | |
| $x_1$ | 0.90 | 0.00 | 1.00 | | | | | | |
| $x_2$ | 0.80 | 0.00 | 0.72 | 1.00 | | | | | |
| $x_3$ | 0.70 | 0.00 | 0.63 | 0.56 | 1.00 | | | | |
| $x_4$ | 0.90 | 0.00 | 0.81 | 0.72 | 0.63 | 1.00 | | | |
| $z_1$ | 0.60 | 0.60 | 0.54 | 0.48 | 0.42 | 0.54 | 1.00 | | |
| $z_2$ | 0.60 | 0.60 | 0.54 | 0.48 | 0.42 | 0.54 | 0.72 | 1.00 | |
| $z_3$ | 0.00 | 0.60 | 0.00 | 0.00 | 0.00 | 0.00 | 0.36 | 0.36 | 1.00 |

Fuente: SPSS IBM 21, con datos propios.

Observe que $\mathbf{X_1}$ a $\mathbf{X_4}$ son indicadores que sólo provienen a causa de $\mathbf{F_1}$; $\mathbf{Z_1}$ y $\mathbf{Z_2}$ son indicadores provenientes a causa de $\mathbf{F_1}$ y $\mathbf{F_2}$ y $\mathbf{Z_3}$ es un indicador proveniente sólo de $\mathbf{F_2}$. La **Tabla 2.4** muestra las correlaciones parciales entre varios conjuntos de variables pertenecientes a la **Figura 2.3**.
A continuación se presentan algunas de las observaciones que pueden extraerse de los resultados presentados en las **Tablas 2.4a-2.4e**.

**Caso 1:** En el caso mostrado en la **Tabla 2.4a**, el conjunto $\mathbf{X_1}$ a $\mathbf{X_4}$ de indicadores es unidimensional, ya que un factor común explica las correlaciones entre los ítems, y cada ítem es unidimensional, ya que mide uno y sólo un constructo

**Tabla 2.4a. Matriz de correlación del modelo de la Figura 2.3.**
**• Caso 1: Parcializando el efecto de $F_1$ desde $X_1$ a $X_4$**

| | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|---|---|---|---|---|
| $x_1$ | 1.00 | | | |
| $x_2$ | 0.00 | 1.00 | | |
| $x_3$ | 0.00 | 0.00 | 1.00 | |
| $x_4$ | 0.00 | 0.00 | 0.00 | 1.00 |

Fuente: SPSS IBM 21, con datos propios.

**Caso 2:** En el caso mostrado en las **Tablas 2.4b y 2.4c**, el conjunto de elementos $\mathbf{X_1}$ a $\mathbf{X_4}$ y $\mathbf{Z_1}$ es unidimensional, ya que las correlaciones entre los ítems pueden ser contabilizadas por un solo factor. El indicador $\mathbf{Z_1}$ no es unidimensional, sin embargo, es una medida de dos factores. Es decir, es posible que un conjunto de indicadores sea unidimensional, sin embargo, cada elemento del conjunto puede o no ser unidimensional. Del mismo modo, el conjunto de elementos $\mathbf{X_1}$ a $\mathbf{X_4}$ y $\mathbf{Z_2}$ es unidimensional; sin embargo, el indicador $\mathbf{Z_2}$ no es unidimensional.

**Tabla 2.4b. Matriz de correlación del modelo de la Figura 2.3.**
**Caso 2:Parcializando el efecto de $F_1$ desde $X_1$ a $X_4$ y $Z_1$**

| | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $z_1$ |
|---|---|---|---|---|---|
| $x_1$ | 1.00 | | | | |
| $x_2$ | 0.00 | 1.00 | | | |
| $x_3$ | 0.00 | 0.00 | 1.00 | | |
| $x_4$ | 0.00 | 0.00 | 0.00 | 1.00 | |
| $z_1$ | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 |

Fuente: SPSS IBM 21, con datos propios.

**Tabla 2.4c. Matriz de correlación del modelo de la Figura 2.3.**
**• Caso 2: Parcializando el efecto de $F_1$ desde $X_1$ a $X_4$ y $Z_2$**

| | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $z_2$ |
|---|---|---|---|---|---|
| $x_1$ | 1.00 | | | | |
| $x_2$ | 0.00 | 1.00 | | | |
| $x_3$ | 0.00 | 0.00 | 1.00 | | |
| $x_4$ | 0.00 | 0.00 | 0.00 | 1.00 | |
| $z_2$ | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 |

Fuente: SPSS IBM 21, con datos propios.

**Caso 3:** En el caso mostrado en las **Tablas 2.4d y 2.4e**, el conjunto de elementos $\mathbf{X_1}$ a $\mathbf{X_4}$, $\mathbf{Z_1}$ y $\mathbf{Z_2}$ no es unidimensional, ya que la correlación entre $\mathbf{Z_1}$ y $\mathbf{Z_2}$ no puede ser explicada por un factor. Las correlaciones entre este conjunto de indicadores se pueden explicar por dos factores. Por lo tanto, este conjunto de elementos es multidimensional (en este caso, de dos dimensiones). Además, los indicadores $\mathbf{Z_1}$ y $\mathbf{Z_2}$ son multidimensionales y no unidimensionales. El lector puede ver fácilmente que los ítems $\mathbf{X_1}$ a $\mathbf{X_4}$ y $\mathbf{Z_1}$-$\mathbf{Z_3}$ no son unidimensionales, ya que las correlaciones parciales entre ellos no pueden ser contabilizadas por un solo factor.

**Tabla 2.4d. Matriz de correlación del modelo de la Figura 2.3.**

**•Caso 3: Parcializando el efecto de $F_1$ desde $X_1$ a $X_4$, $Z_1$ y $Z_2$**

| | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $z_1$ | $z_2$ |
|---|---|---|---|---|---|---|
| $x_1$ | 1.00 | | | | | |
| $x_2$ | 0.00 | 1.00 | | | | |
| $x_3$ | 0.00 | 0.00 | 1.00 | | | |
| $x_4$ | 0.00 | 0.00 | 0.00 | 1.00 | | |
| $z_1$ | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | |
| $z_2$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.225 | 1.00 |

Fuente: SPSS IBM 21, con datos propios.

**Tabla 2.4e. Matriz de correlación del modelo de la Figura 2.3.**
**•Caso 3: Parcializando el efecto de $F_1$ y $F_2$ desde $X_1$ a $X_4$, $Z_1$ y $Z_2$**

| | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $z_1$ | $z_2$ |
|---|---|---|---|---|---|---|
| $x_1$ | 1.00 | | | | | |
| $x_2$ | 0.00 | 1.00 | | | | |
| $x_3$ | 0.00 | 0.00 | 1.00 | | | |
| $x_4$ | 0.00 | 0.00 | 0.00 | 1.00 | | |
| $z_1$ | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | |
| $z_2$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 |

Fuente: SPSS IBM 21, con datos propios

Con base en las observaciones anteriores, se puede concluir que si la definición conceptual del constructo permanece igual, entonces el constructo **$F_1$** no es unidimensional, ya que todos los conjuntos de elementos del dominio del constructo no son unidimensionales. Normalmente, cuando esto sucede, se suele decir que el *constructo depende del contexto*, ya que su dimensionalidad y quizás el significado conceptual podría cambiar dependiendo de qué conjunto de indicadores se utilizan para medir el constructo.

# Importancia de la unidimensionalidad

Una revisión de la literatura existente utilizando el análisis factorial confirmatorio (**CFA.** *Confirmatory Factor Analysis*), el modelo de estructura de covarianza o modelo de ecuaciones estructurales (**SEM.** *Structural Equations Modeling*) sugiere fuertemente la preferencia por el uso de indicadores que sean unidimensionales y,

sólo en el contexto de un estudio de investigación específico, y sólo en un factor (Gerbing y Anderson, 1988, Neuberg, West, Thompson, y Judice, 1997).
A pesar de que no ha habido una investigación sistemática de por qué estos indicadores deben ser preferidos, a primera vista, la fuerte preferencia por estos elementos parece históricamente fundamentada en la práctica de utilizar la suma de una puntuación de los indicadores en escalas específicas en **ANOVA** y análisis de regresión en lugar de indicadores individuales. De hecho, incluso cuando se usan indicadores individuales en modelos de estructura de covarianza (por ejemplo, **SEM**), la tendencia es reportar estadísticas tales como correlaciones corregidas de indicadores totales (*corrected ítem-to-total correlations*) y coeficientes alfa **(α)**, que son características de, o implican, puntuaciones sumarizadas. Si el investigador utiliza puntajes compuestos en un modelo de estructura de covarianza, entonces, desde un punto de vista conceptual, es deseable que los indicadores que forman la puntuación sumarizada sean unidimensionales (Floyd y Widaman, 1995; Neuberg et al., 1997). Este es el argumento más convincente a favor de usar elementos que sean unidimensionales. Si, por otra parte, el conjunto de elementos es unidimensional, pero los elementos individuales no son unidimensionales, entonces el uso de las puntuaciones sumadas no es apropiado, ya que las puntuaciones sumadas también contendrían el efecto de otros factores sobre los cuales se cargan los ítems. En estas situaciones, los investigadores podrían utilizar las puntuaciones de los factores en lugar de las puntuaciones sumarizadas como insumo para un análisis posterior, a pesar de los problemas conocidos con la indeterminación y la estimación de los puntajes de los factores (Sharma, 1996). Las puntuaciones de los factores se forman como combinaciones lineales de los indicadores de cada factor usando los coeficientes de regresión de la puntuación del factor como pesos. Este enfoque tiene la ventaja de que los elementos no necesitan cargas distintas de cero sólo en un factor latente. Cabe señalar, sin embargo, que las relaciones estructurales basadas en las puntuaciones de los factores aún pueden atenuarse, porque las puntuaciones de los factores contienen error. Sin embargo, este enfoque puede ser menos problemático que el uso de compuestos basados en las puntuaciones sumadas.

# Evaluando la dimensionalidad de los constructos

El coeficiente alfa **(α)** (Cronbach, 1951) es la medida más popular que se informa para evaluar la consistencia interna de las escalas y, en muchos casos, un alto coeficiente alfa **(α)** se considera evidencia de la unidimensionalidad del constructo. Desafortunadamente, esto no es verdad. Por ejemplo, la estimación del *coeficiente de alfa* **(α)** para el conjunto de elementos $\mathbf{X_1}$ a $\mathbf{X_4}$ y $\mathbf{Z_1}$ a $\mathbf{Z_3}$ de la **Figura 2.3** es igual a .852, y se podría concluir que el conjunto de elementos es unidimensional. Ahora bien se acepta que el coeficiente alfa **(α)** es significativo sólo para un conjunto unidimensional de elementos (Clark y Watson, 1995; Cortina, 1993). Para un conjunto de elementos que no son *unidimensionales*, existe confusión en cuanto a lo que es exactamente evaluado por el coeficiente alfa **(α)**. Es muy importante garantizar la relación entre consistencia interna y dimensionalidad así como conocer y aplicar procedimientos para evaluar la dimensionalidad de los constructos. Como

se mencionó anteriormente, la dimensionalidad se define como el número de factores comunes o construcciones latentes o subyacentes necesarias para explicar la correlación entre las variables. Por lo tanto, el análisis de factores es un método apropiado y popular para evaluar la dimensionalidad de las construcciones. Podría utilizarse tanto el análisis factorial exploratorio (**EFA**. *Exploratory Factor Analysis*) como el análisis factorial confirmatorio (**CFA**.*Confirmatory Factor Analysis*) o ambos para evaluar la dimensionalidad de los constructos.

# Análisis Factorial Exploratorio (EFA. *Exploratory Factor Analysis*)

El supuesto implícito que subyace en el uso de la **EFA** es que el investigador generalmente tiene una idea limitada con respecto a la dimensionalidad de los constructos y qué elementos pertenecen o cargan sobre qué factor. Además, la **EFA** se lleva a cabo típicamente durante la etapa inicial del desarrollo de la escala. Aún así, el **EFA** se puede usar para obtener información sobre la potencial dimensionalidad de los elementos y escalas (Ver Mejía-Trejo 2017c, Capítulo 12. Análisis Factorial). Por ejemplo, suponga la aplicación de **EFA** a los datos de la **Tabla 2.3**, se obtiene la **Tabla 2.5**

**Tabla 2.5. Salida parcial de SPSS para los datos de la Tabla 2.3**

*Initial Eigenvalues*

| *Factor* | *Total* | *% of Variance* | *Cumulative %* |
|---|---|---|---|
| 1 | 3.961 | 56.580 | 56.580 |
| 2 | 1.337 | 19.094 | 75.674 |
| 3 | .500 | 7.141 | 82.815 |
| 4 | .437 | 6.249 | 89.064 |
| 5 | .296 | 4.222 | 93.286 |
| 6 | .280 | 4.000 | 97.286 |
| 7 | .190 | 2.714 | 100.000 |

Extraction Method: Principal Axis Factoring.

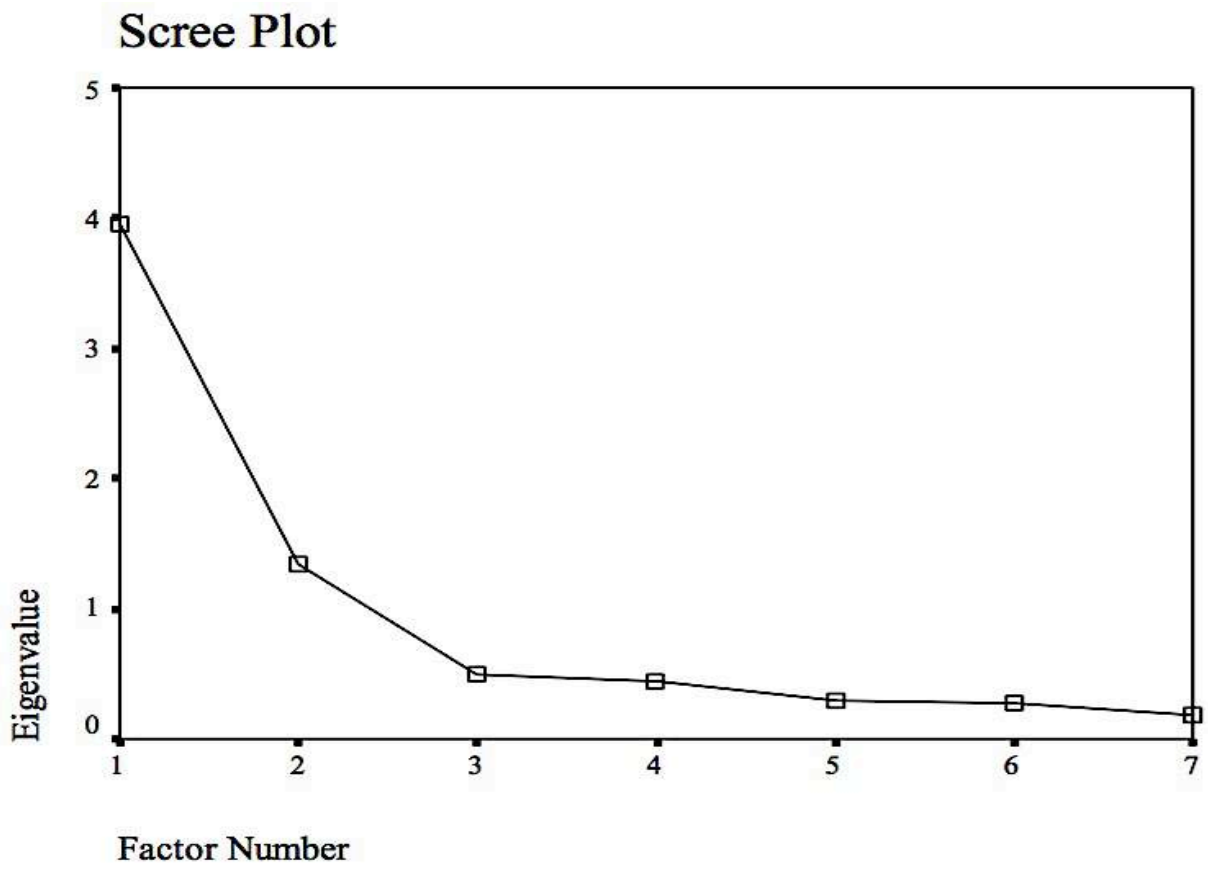
Fuente: SPSS IBM 21, con datos propios.

El número de factores que explican las correlaciones entre las variables representa la dimensionalidad de un conjunto de variables. Para determinar el número de factores se utilizan una serie de reglas empíricas o heurísticas, tales como los criterios:

**a.** El eigenvalor con valor >1

**b.**El gráfico de autovalor para el criterio de contraste de caída (Ver Mejía-Trejo 2017c, p22)
**c.** El gráfico de autovalor para el criterio de contraste de caída con análisis paralelo.
Así, es posible afirmar:
-De acuerdo al criterio **a,** el número de factores es igual al número de autovalores mayores a 1. La razón es que un factor dado debe dar cuenta de por lo menos tanta variación como puede ser explicado por un solo elemento o variable. En el caso que nos ocupa, la regla del inciso **a** sugiere la presencia de dos factores, por lo que se puede concluir que el conjunto de elementos *no es unidimensional*. Cabe destacar que este criterio ha tenido críticas de uso muy interesantes, sobre su utilidad de identificar el número de factores (Cliff,1988).
-Sobre el criterio **b** propuesto por Cattell (1966) es otra técnica popular. Es un diagrama de los valores propios en función del número de factores, y uno busca un *codo* que *significa una fuerte caída en la varianza explicada por un factor dado.* Se supone que los factores en o más allá del codo son factores que ya no aportan más valor y tienden a representar errores o componentes únicos. Ver **Gráfico 2.1**
Por ejemplo, en la **Gráfico 2.1** el codo sugiere la presencia de dos factores.

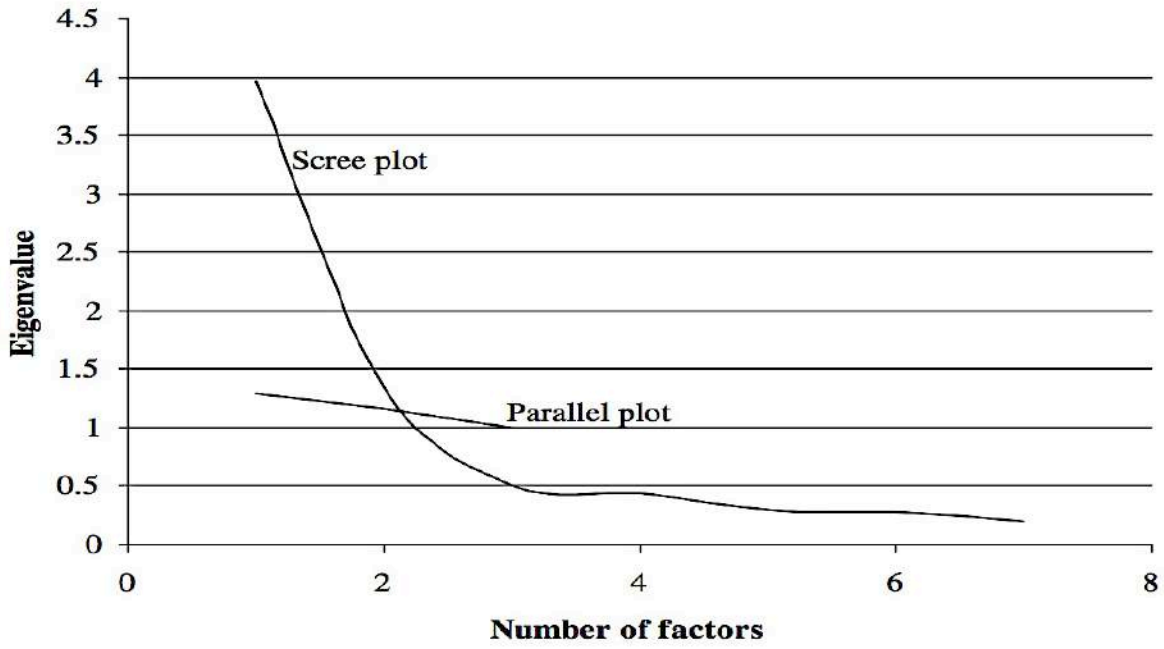
**Gráfico 2.1. El gráfico de autovalor para el criterio de contraste de caída, con codo declarado (*scree test*) o gráfico de sedimentación**




Fuente: SPSS IBM 21, con datos propios.

Es decir, el conjunto de elementos es multidimensional. En muchos casos, sin embargo, no es posible identificar completamente el codo. Ver **Gráfico 2.2**

**Gráfico 2.2. El gráfico de sedimentación (de autovalor para el criterio de contraste de caída, sin codo declarado)**




Fuente: SPSS IBM 21, con datos propios.

En tales casos, se puede utilizar el procedimiento de trazado paralelo sugerido por Horn (1965). Con este procedimiento, el gráfico en paralelo representa los valores propios que resultarían si el conjunto de datos no tuviera factores comunes. Es decir, las correlaciones entre las variables se deben completamente al *error de muestreo*. Se requieren simulaciones extensas para estimar los autovalores para el gráfico paralelo; sin embargo, basadas en ecuaciones empíricamente derivadas, Allen y Hubbard (1986) han desarrollado la siguiente **Ecuación 2.3** para estimar los valores propios para obtener el gráfico paralelo:

**Ecuación 2.3. Estimación de valores para obtener el gráfico paralelo**

$$\ln \lambda_k = a_k + b_k \ln (n-1) + c_k \ln \{(p-k-1)(p-k+2)/2\} + d_k \ln \lambda_{k-1}$$

Fuente: Allen y Hubbard (1986)

Donde:
$\lambda_k$ es la estimación del késimo valor propio,

**p** es el número de variables o ítems,
**n** es el número de observaciones,
$\mathbf{a_k}$, $\mathbf{b_k}$, $\mathbf{c_k}$ y $\mathbf{d_k}$ son coeficientes de regresión, y
**ln** $\boldsymbol{\lambda_0}$ es considerado como 1.

La **Tabla 2.6** muestra los coeficientes empíricamente derivados para la ecuación anterior.

**Tabla 2.6. Coeficientes de regresión por el método de componentes principales**

| *Root (k)* | *Number of Points*[a] | *a* | *b* | *c* | *d* | $R^2$ |
|---|---|---|---|---|---|---|
| 1 | 62 | .9794 | −.2059 | .1226 | 0.0000 | .931 |
| 2 | 62 | −.3781 | .0461 | .0040 | 1.0578 | .998 |
| 3 | 62 | −.3306 | .0424 | .0003 | 1.0805 | .998 |
| 4 | 55 | −.2795 | .0364 | −.0003 | 1.0714 | .998 |
| 5 | 55 | −.2670 | .0360 | −.0024 | 1.0899 | .998 |
| 6 | 55 | −.2632 | .0368 | −.0040 | 1.1039 | .998 |
| 7 | 55 | −.2580 | .0360 | −.0039 | 1.1173 | .998 |
| 8 | 55 | −.2544 | .0373 | −.0064 | 1.1421 | .998 |
| 9 | 48 | −.2111 | .0329 | −.0079 | 1.1229 | .998 |
| 10 | 48 | −.1964 | .0310 | −.0083 | 1.1320 | .998 |
| 11 | 48 | −.1858 | .0288 | −.0073 | 1.1284 | .999 |
| 12 | 48 | −.1701 | .0276 | −.0090 | 1.1534 | .998 |
| 13 | 48 | −.1697 | .0266 | −.0075 | 1.1632 | .998 |
| 14 | 41 | −.1226 | .0229 | −.0113 | 1.1462 | .999 |
| 15 | 41 | −.1005 | .0212 | −.0133 | 1.1668 | .999 |
| 16 | 41 | −.1079 | .0193 | −.0088 | 1.1374 | .999 |
| 17 | 41 | −.0866 | .0177 | −.0110 | 1.1718 | .999 |
| 18 | 41 | −.0743 | .0139 | −.0081 | 1.1571 | .999 |
| 19 | 34 | −.0910 | .0152 | −.0056 | 1.0934 | .999 |
| 20 | 34 | −.0879 | .0145 | −.0051 | 1.1005 | .999 |
| 21 | 34 | −.0666 | .0118 | −.0056 | 1.1111 | .999+ |
| 22 | 34 | −.0865 | .0124 | −.0022 | 1.0990 | .999+ |
| 23 | 34 | −.0919 | .0123 | −.0009 | 1.0831 | .999+ |
| 24 | 29 | −.0838 | .0116 | −.0016 | 1.0835 | .999+ |
| 25 | 28 | −.0392 | .0083 | −.0053 | 1.1109 | .999+ |
| 26 | 28 | −.0338 | .0065 | −.0039 | 1.1091 | .999+ |
| 27 | 28 | .0057 | .0015 | −.0049 | 1.1276 | .999+ |
| 28 | 28 | .0017 | .0011 | −.0034 | 1.1185 | .999+ |
| 29 | 22 | −.0214 | .0048 | −.0041 | 1.0915 | .999+ |
| 30 | 22 | −.0364 | .0063 | −.0030 | 1.0875 | .999+ |
| 31 | 22 | −.0041 | .0022 | −.0033 | 1.0991 | .999+ |
| 32 | 22 | .0598 | −.0067 | −.0032 | 1.1307 | .999+ |
| 33 | 21 | .0534 | −.0062 | −.0023 | 1.1238 | .999+ |
| 34 | 16 | .0301 | −.0032 | −.0027 | 1.0978 | .999+ |
| 35 | 16 | .0071 | .0009 | −.0038 | 1.0895 | .999+ |
| 36 | 16 | .0521 | −.0052 | −.0030 | 1.1095 | .999+ |
| 37 | 16 | .0824 | −.0105 | −.0014 | 1.1209 | .999+ |
| 38 | 16 | .1865 | −.0235 | −.0033 | 1.1567 | .999+ |
| 39 | 10 | .0075 | .0009 | −.0039 | 1.0773 | .999+ |
| 40 | 10 | .0050 | −.0021 | .0025 | 1.0802 | .999+ |
| 41 | 10 | .0695 | −.0087 | −.0016 | 1.0978 | .999+ |

Fuente: De la Tabla 1 de Regression Equations for the Latent Roots of Random Data Correlation Matrices With Unities on the Diagonal, *Multivariate Behavioral Research, 21,* pp. 393-398, en Netemeyer et al. (2003)
a. The number of points used in the regression.

Usando la **Ecuación 2.3,** el primer valor propio es igual a:

$$\ln \lambda_1 = .974 - .2059 \ln (200-1) + .1226 \ln [(7-1-1)(7-1+2)/2] + 0.00x1 = .257$$

$$\lambda_1 = e^{.257} = 1.293$$

y el segundo valor propio es igual a

$$\lambda_2 = -.3781 + .0461 \ln(200-1) + .004 \ln[(7-2-1)(7-2+2)/2] + 1.0578x.257 = .148$$

$$\lambda_2 = e^{.148} = 1.160$$

El procedimiento anterior se repite para calcular los autovalores restantes (de ser negativos, se consideran cero). El **Gráfico 2.2** muestra el *gráfico paralelo*. Como puede verse, la trama sugiere dos factores; es decir, *la dimensionalidad del conjunto de elementos es dos***.** El número de factores identificados por los procedimientos anteriores debe ser confirmado adicionalmente asegurando que el número de factores es suficiente para explicar todas las correlaciones entre las variables, ya que éste es un objetivo clave del análisis de factores. Es decir, el investigador siempre está atento a determinar el *ajuste* del modelo de factor a los datos. Dos medidas de uso común son *la raíz cuadrada media residual* (**RMSR.** *Root-Mean-Square Residual*) y la *raíz cuadrada media de correlaciones parciales entre las variables* (**RMSP.** *Root-Mean-Square of Partial Correlations among the Variables*).
Estos se calculan a partir de la matriz de *correlación residual* (diferencia entre la matriz de correlación de muestras y la matriz de correlación calculada o estimada usando cargas factoriales estimadas) y la *matriz de correlación parcial* que se ha definido anteriormente. **RMSR** y **RMSP** se pueden calcular como lo indicado en las **Ecuaciones 2.4.**

**Ecuaciones 2.4. Cálculo de RMSR y RMSP**

$$\text{RMSR} = \sqrt{\frac{2\sum_{i=1}^{k}\sum_{j=i}^{k} res_{ij}^2}{k(k-1)}}$$

$$\text{RMSP} = \sqrt{\frac{2\sum_{i=1}^{k}\sum_{j=i}^{k} pc_{ij}^2}{k(k-1)}}$$

Donde $res_{ij}$ y $pc_{ij}$ son, respectivamente, la *correlación residual y la correlación parcial entre las variables* ***i*** y ***j***.
Las **Tabla 2.7 y 2.8** muestran la matriz de correlación residual proporcionada por **SPSS** para los modelos de uno y dos factores.

## Tabla 2.7. Salida SPSS parcial para solución de un factor

**Reproduced Correlations**

| | | X1 | X2 | X3 | X4 | Z1 | Z2 | Z3 |
|---|---|---|---|---|---|---|---|---|
| Reproduced Correlation | X1 | .754[b] | .674 | .592 | .754 | .606 | .606 | .123 |
| | X2 | .674 | .602[b] | .529 | .674 | .541 | .541 | .110 |
| | X3 | .592 | .529 | .464[b] | .592 | .475 | .475 | 9.646E-02 |
| | X4 | .754 | .674 | .592 | .754[b] | .606 | .606 | .123 |
| | Z1 | .606 | .541 | .475 | .606 | .487[b] | .487 | 9.880E-02 |
| | Z2 | .606 | .541 | .475 | .606 | .487 | .487[b] | 9.880E-02 |
| | Z3 | .123 | .110 | 9.646E-02 | .123 | 9.880E-02 | 9.880E-02 | 2.005E-02[b] |
| Residual[a] | X1 | | 4.613E-02 | 3.836E-02 | 5.574E-02 | -6.605E-02 | -6.605E-02 | -.123 |
| | X2 | 4.613E-02 | | 3.142E-02 | 4.613E-02 | -6.145E-02 | -6.145E-02 | -.110 |
| | X3 | 3.836E-02 | 3.142E-02 | | 3.836E-02 | -5.538E-02 | -5.538E-02 | -9.646E-02 |
| | X4 | 5.574E-02 | 4.613E-02 | 3.836E-02 | | -6.605E-02 | -6.605E-02 | -.123 |
| | Z1 | -6.605E-02 | -6.145E-02 | -5.538E-02 | -6.605E-02 | | .233 | .261 |
| | Z2 | -6.605E-02 | -6.145E-02 | -5.538E-02 | -6.605E-02 | .233 | | .261 |
| | Z3 | -.123 | -.110 | -9.646E-02 | -.123 | .261 | .261 | |

Extraction Method: Principal Axis Factoring.

a. Residuals are computed between observed and reproduced correlations. There are 16 (76.0%) nonredundant residuals with absolute values greater that 0.05.

b. Reproduced communalities

Fuente: SPSS IBM 21, con datos Netemeyer et al. (2003)

## Tabla 2.8. Salida SPSS parcial para soluciones de dos factores

**Reproduced Correlation**

| | | X1 | X2 | X3 | X4 | Z1 | Z2 | Z3 |
|---|---|---|---|---|---|---|---|---|
| Reproduced Correlation | X1 | .810[b] | .720 | .630 | .810 | .540 | .540 | 1.893E-04 |
| | X2 | .720 | .640[b] | .560 | .720 | .480 | .480 | 3.751E-05 |
| | X3 | .630 | .560 | .490[b] | .630 | .420 | .420 | 5.103E-05 |
| | X4 | .810 | .720 | .630 | .810[b] | .540 | .540 | 1.893E-04 |
| | Z1 | .540 | .480 | .420 | .540 | .720[b] | .720 | .360 |
| | Z2 | .540 | .480 | .420 | .540 | .720 | .720[b] | .360 |
| | Z3 | 1.893E-04 | 3.751E-05 | 5.103E-05 | 1.893E-04 | .360 | .360 | .359[b] |
| Residual[a] | X1 | | -5.284E-05 | -5.296E-06 | 2.717E-04 | 9.458E-05 | 9.458E-05 | -1.893E-04 |
| | X2 | -5.284E-05 | | -2.337E-04 | -5.284E-05 | 1.876E-05 | 1.876E-05 | -3.751E-05 |
| | X3 | -5.296E-06 | -2.337E-04 | | -5.296E-06 | 2.548E-05 | 2.548E-05 | -5.103E-05 |
| | X4 | 2.717E-04 | -5.284E-05 | -5.296E-06 | | 9.458E-05 | 9.458E-05 | -1.893E-04 |
| | Z1 | 9.458E.05 | 1.876E-05 | 2.548E-05 | 9.458E-05 | | 2.049E-04 | 3.388E-04 |
| | Z2 | 9.458E-05 | 1.876E-05 | 2.548E-05 | 9.458E-05 | 2.049E-04 | | 3.388E-04 |
| | Z3 | -1.893E.04 | -3.751E-05 | -5.103E.05 | -1.893E-04 | 3.388E-04 | 3.388E-04 | |

Extraction Method: Principal Axis Factoring.

a. Residuals are computed between observed and reproduced correlations. There are 0 (.0%) nonredundant residuals with absolute values greater than 0.05.

b. Reproduced communalities

Fuente: SPSS IBM 21, con datos Netemeyer et al. (2003)

Usando la **Ecuación 2.4**, el **RMSR** para un *modelo de dos factore*s es .0002**1,** y para el *modelo de un factor* es .215. Aunque no hay pautas sobre cuán bajo es *bajo*, hay una diferencia sustancial entre la **RMSR** para los modelos de uno y dos factores; por lo tanto, *el modelo de dos factores proporciona la mejor explicación de las correlaciones entre las variables.*

(**Nota**: **SPSS** no proporciona el **RMSP**. **SAS**, por otra parte, proporciona ambas matrices y también calcula los dos índices).

# Análisis Factorial Confirmatorio (CFA. *Confirmatory Factor Analysis*)

Aunque **EFA** da una idea de la dimensionalidad, **CFA**, como su nombre implica, se centra esencialmente en si *un hipotético modelo por factores se ajusta o no a los datos.* **CFA** *es al momento, un método comúnmente aceptado para probar/ confirmar la dimensionalidad. En este método, se especifican a priori el número de factores, la estructura del factor (es decir, qué elementos cargan sobre qué factores) y la relación entre factores (o sea, si los factores están correlacionados).*

Por ejemplo, considere el modelo de factores dado en la **Figura 2.3.** Este modelo plantea la hipótesis de que el conjunto de ítems mide dos factores y que estos factores no están correlacionados. Además, se supone que los ítems $X_1$ a $X_4$ son medidas del primer factor $F_1$, los ítems $Z_1$ y $Z_2$ miden los factores $F_1$ y $F_2$, y $Z_3$ miden $F_2$. El objetivo entonces es determinar si los datos apoyan el modelo hipotético o no. Esencialmente, las hipótesis nula y alternativa son:

$H_0$*: El modelo se ajusta a los datos.*

$H_1$*: El modelo no se ajusta a los datos.*

**Obviamente, se desea no rechazar la hipótesis nula**. Las hipótesis anteriores se pueden probar mediante la estimación de los parámetros del modelo usando la estimación de máxima verosimilitud (**MLE***. Maximum Likelihood Estimation*) o una técnica relacionada a través de **CFA**, y empleando la estadística de bondad de ajuste del Chi-Cuadrado resultante. Una serie de paquetes de software, como **SAS, LISREL, EQS, y AMOS**, se puede utilizar para realizar **CFA**. La **Tabla 2.9** muestra la salida de **PROC CALIS** en **SAS**.

**Tabla 2.9. Salida parcial de resultados del SAS sobre los datos de la Tabla**

**Goodness-of-Fit Results**

| | |
|---|---|
| Fit Function | 0.0000 |
| Goodness of Fit Index (GFI) | 1.0000 |
| GFI Adjusted for Degrees of Freedom (AGFI) | 1.0000 |
| Root Mean Square Residual (RMR) | 0.0000 |
| Parsimonious GFI (Mulaik, 1989) | 0.5714 |
| Chi-Square | 0.0000 |
| Chi-Square DF | 12 |
| Pr > Chi-Square | 1.0000 |
| Independence Model Chi-Square | 797.28 |
| Independence Model Chi-Square DF | 21 |
| RMSEA Estimate | 0.0000 |
| RMSEA 90% Lower Confidence Limit | . |
| RMSEA 90% Upper Confidence Limit | . |
| ECVI Estimate | 0.1675 |
| ECVI 90% Lower Confidence Limit | . |
| ECVI 90% Upper Confidence Limit | . |
| Probability of Close Fit | 1.0000 |
| Bentler's Comparative Fit Index | 1.0000 |
| Normal Theory Reweighted LS Chi-Square | 0.0000 |
| Akaike's Information Criterion | -24.0000 |
| Bozdogan's (1987) CAIC | -75.5798 |
| Schwarz's Bayesian Criterion | -63.5798 |
| McDonald's (1989) Centrality | 1.0305 |
| Bentler & Bonett's (1980) Non-normed Index | 1.0271 |
| Bentler & Bonett's (1980) NFI | 1.0000 |
| James, Mulaik, & Brett (1982) Parsimonious NFI | 0.5714 |
| Z-Test of Wilson & Hilferty (1931) | -7.2124 |
| Bollen (1986) Normed Index Rho1 | 1.0000 |
| Bollen (1988) Non-normed Index Delta2 | 1.0153 |
| Hoelter's (1983) Critical N | . |

Fuente : EQS 6 con datos propios.

El estadístico de la *prueba Chi-Cuadrado es cero con 12 grados de libertad y no es significativo en p <0.05, lo que implica que la hipótesis nula no puede ser rechazada, esto es, el modelo se encuentra suficientemente soportado por los datos* (Cabe señalar que el ajuste del modelo es perfecto, ya que se utilizaron datos hipotéticos para el modelo de la **Figura 2.3**. Normalmente, para datos reales y tamaños de

muestra grandes, el ajuste no será perfecto y la prueba estadística de *Chi-Cuadrado* será grande y significativa).
Por el contrario, la estadística de la prueba *Chi-Cuadrada* para el *modelo de un factor* en el que todos los indicadores cargan en el factor único es 126.25 con 14 grados de libertad y es estadísticamente significativa en $p < 0.05$. *Es decir, el modelo de un factor no está soportado por los datos.* Las cargas factoriales estimadas del modelo de dos factores son, dentro del error de redondeo, los mismos que los reportados en la **Figura 2.3.**
Hay una serie de cuestiones relacionadas para el uso de la *prueba Chi-Cuadrado* estadística para la prueba de hipótesis, el más importante es su sensibilidad al tamaño de la muestra (Bearden, Sharma, y Teel, 1982, Hoyle, 1995, McDonald y Marsh, 1990). Por esta razón, los investigadores han propuesto una serie de pruebas estadísticas de bondad de ajuste alternativo que no son sensibles al tamaño de la muestra ya otros parámetros del modelo.

# Conclusión

El concepto de *unidimensionalidad* fue discutido en este capítulo. Los siguientes puntos, concluyen los conceptos presentados:
**1.**Un conjunto de elementos es unidimensional si las correlaciones entre los ítems después del efecto de un solo factor son parcializados son cero. Es decir, un conjunto de elementos es unidimensional si un modelo de factor único se ajusta a los datos.
**2.**La dimensionalidad de un conjunto de elementos es el número de factores comunes necesarios para reducir las correlaciones parciales entre los elementos a cero.
**3.**Un ítem dado es unidimensional si mide una y sólo un constructo. La dimensionalidad de un elemento es igual al número de factores que mide.
**4.**Es posible que el conjunto de elementos de un conjunto no sea unidimensional.
**5.**Un constructo es unidimensional si cada conjunto de **n** ítems tomados aleatoriamente del dominio del constructo es unidimensional.
**6.**Los análisis factoriales exploratorios y confirmatorios son dos métodos para evaluar la dimensionalidad de los constructos.

# CAPÍTULO 3. CONFIABILIDAD

Hasta ahora, se han discutido problemas relacionados con la dimensionalidad de los constructos latentes y sus elementos de medición. Llegamos a la conclusión de que el conocimiento de la dimensionalidad de un constructo es fundamental para el desarrollo de los indicadores que miden dicho constructo. Después de que los indicadores se han desarrollado y acotado, los siguientes pasos importantes a seguir, son examinar la *confiabilidad y validez de la escala*. El propósito de este capítulo es discutir el concepto de *confiabilidad* y proporcionar procedimientos sugeridos para evaluarlo. La *validez* de una escala y los procedimientos para evaluarla se abordarán en el próximo capítulo.

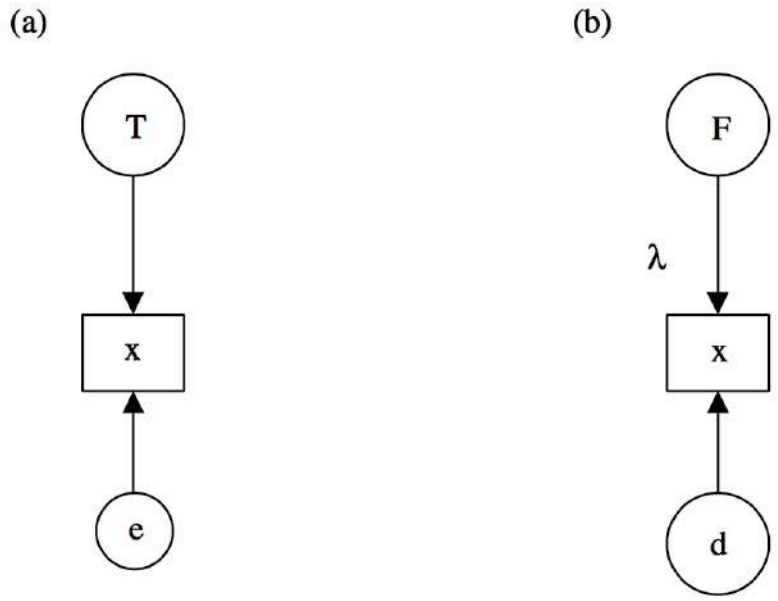
Si un conjunto de indicadores de una escala están midiendo una misma *variable o constructo latente,* cabrá esperar que sus puntuaciones estén fuertemente correlacionadas entre sí, es decir, que sean internamente consistentes. Es muy importante señalar que la *confiabilidad de una escala indica solamente que los distintos ítems que la componen, al estar muy correlacionados entre sí, están midiendo la misma variable o constructo latente***.** Pero que una escala sea confiable no quiere decir que dicha variable o constructo latente que está midiendo, sea la que tiene que medir, es decir, que *sea válida.* Por ejemplo, supongamos que en un examen de *innovación tecnológica* hemos diseñado un test de 50 preguntas y que, por error, el día del examen repartimos otro test que hemos diseñado para el examen de *innovación de procesos.* Este test será probablemente confiable, en cuanto que todas las preguntas obtendrán puntuaciones muy correlacionadas entre sí, dado que están midiendo lo mismo: el conocimiento de la industria por las *prácticas de innovación de procesos.* Lo que ocurre es que no son estos conocimientos que queríamos medir, sino los de *innovación tecnológica. La escala diseñada es fiable, pero no es válida. Sin embargo, una escala no podrá ser válida, si no es confiable, de*

*tal forma que la confiabilidad se convierte en una condición necesaria, aunque no suficiente, de la validez.*

# La importancia de medir la confiabilidad

La confiabilidad de las mediciones está estrechamente asociada con el concepto del modelo de la puntuación verdadera (*true-score model*, Coker y Algina, 1986) de la teoría de la prueba clásica. (**Ver Apéndice**). Considere la **Figura 3.1a.**

**Figura 3.1. Modelos de puntuación verdadera**




Fuente: propia

La cual describe a ***X*** como la medida observada de un constructo latente. Según el modelo de puntuación verdadera, la puntuación observada, ***X***, puede dividirse en verdadera puntuación y *error* (es decir, *error de medición*). Ver **Ecuación 3.1**

**Ecuación 3.1. Puntuación verdadera**

$$X = T + e$$

Donde:
***X*** es la puntuación observada,
***T*** es la puntuación verdadera, y
***e*** es el error de medición aleatorio.
Cuando la puntuación verdadera, ***T***, varía, también lo hace la puntuación observada, ***X***, porque el constructo latente influye en la puntuación observada (es decir, ***X*** es un indicador de *efecto*). Suponiendo que la puntuación verdadera y el error no están correlacionados es decir:

$$\mathrm{Cov}\,(T, e) = 0$$

Puede demostrarse que la *varianza de X es igual a la varianza de la puntuación verdadera más la varianza del error*. Es decir:

**Var(*X*)=Var (*T*)+ Var (*e*)**

Así, la confiabilidad de una medición se define como la relación de la *varianza de la puntuación verdadera relacionada con la varianza de la puntuación observada.* Ver **Ecuación 3.2.**

**Ecuación 3.2. Confiabilidad basada en la varianza**

$$\text{Reliability} = \rho_{xx} = \frac{\text{Var(T)}}{\text{Var(x)}} = \frac{\sigma_T^2}{\sigma_x^2}$$

Como se muestra en la **Figura 3.1b,** en la medición de un modelo de un factor, la relación entre un elemento y el constructo latente está dada por la **Ecuación 3.3.**

**Ecuación 3.3. Relación entre un elemento y el constructo latente**

$$X = \lambda F + \delta$$

Donde:
***X*** es la puntuación observada,
***F*** es el factor latente,
***δ*** es el error y
***λ*** carga factorial, que es la medida en que el factor latente **F** afecta a la puntuación observada ***X***.
Una vez más, como ***F*** varía, también lo hace la puntuación observada; sin embargo, *el grado en que varía está determinado por el valor de λ*. Está claro que el término ***λF*** es equivalente a la puntuación verdadera en el modelo *(true-score model*) y ***δ*** representa el error (en sentido estricto, se podría argumentar *que conceptualmente los modelos de factor y el de verdadera puntuación no son equivalentes.* En el modelo de factor, ***δ*** además del error de medición también incluye el error único o específico. *Empíricamente, sin embargo, es imposible separar los dos.* En consecuencia, a nivel empírico no es posible diferenciar entre la puntuación verdadera y el factor modelo.) Ver **Ecuación 3.4.**

**Ecuación 3.4.Confiabilidad de la medición observada**

$$\text{Reliability} = \frac{\text{Var(T)}}{\text{Var(x)}} = \frac{\lambda^2\text{Var(F)}}{\text{Var(x)}} = \frac{\lambda^2\sigma_F^2}{\sigma_x^2}.$$

Observe que los numeradores en las **Ecuaciones 3.2** y **3.4** son equivalentes en el sentido de que miden la varianza que se debe al efecto del constructo latente en la puntuación observada.

Las técnicas de confiabilidad más utilizadas son: *test-retest, alternativa y de consistencia interna.*

# Confiabilidad *test-retest*

Supongamos que una agencia de mercadotecnia digital, está interesado en medir la propensión de sus clientes potenciales a suscribirse al diseño de una página web, denominada (**CP**) con la siguiente afirmación:
***$X_1$**: Suscribirme a la página web me hace sentir bien***.** También, suponga que pide a 10 sujetos que indiquen su grado de aceptación utilizando una escala de tipo Likert de 7 puntos de acuerdo-desacuerdo y que lo hagan en 2 momentos: hoy y 2 semanas después. Los datos hipotéticos resultantes se presentan en la **Tabla 3.1**

**Tabla 3.1. Datos hipotéticos para la prueba de confiabilidad test-retest**

| Sujeto | Ocasión 1 | Ocasión 2 |
|---|---|---|
| 1 | 4 | 5 |
| 2 | 4 | 5 |
| 3 | 6 | 5 |
| 4 | 5 | 3 |
| 5 | 5 | 4 |
| 6 | 3 | 3 |
| 7 | 2 | 3 |
| 8 | 5 | 4 |
| 9 | 3 | 3 |
| 10 | 7 | 7 |

Fuente: propia

Asumiendo que la *propensión a la suscripción* a la página web, por parte de los encuestados *no cambia* durante el período de 2 semanas y no hay error de medición, entonces la correlación entre las medidas tomadas con 2 semanas de diferencia hipotéticamente debería ser igual a uno. Un valor de menos de uno para el coeficiente de correlación puede atribuirse al *error de medición.* La confiabilidad de la medida, pues, viene dada por la correlación de las medidas tomadas en la *ocasión 1* y la *ocasión 2*, que para los presentes datos hipotéticos es igual a 0.74. Esta medida se conoce como *confiabilidad test-retest*.
Tenga en cuenta que la fiabilidad *test-retest* se refiere a la estabilidad de las respuestas a los indicadores a través del tiempo. Un *test-retest o coeficiente de estabilidad* generalmente se estima por la magnitud de la correlación entre las mismas medidas (de la muestra) *en diferentes ocasiones de evaluación***.** Si el coeficiente de estabilidad es *bajo* en magnitud, sin cambios en el constructo a lo largo del tiempo, la *confiabilidad de la medición está en duda.*
Si el coeficiente de estabilidad *es alto* en magnitud, sin cambio en el constructo a lo largo del tiempo, *se mejora la confiabilidad de la medida.* La razón para la confiabilidad *test-retest* es que si una medición refleja verdaderamente el constructo pretendido, debe ser capaz de evaluar el constructo en diferentes ocasiones. La puntuación verdadera en el constructo latente debe ser reflejada por el constructo en dos ocasiones de manera comparable. Por lo tanto, la confiabilidad *test-retest* es útil

porque ofrece información sobre el grado de confianza que se tiene que la medida refleje el constructo y es generalizable a otras ocasiones de evaluación (Haynes et al., 1999). En conclusión, la correlación *test-retest* teóricamente representa el grado en que la constructo latente determina las puntuaciones observadas en el tiempo (DeVellis, 1991; Nunnally y Bernstein, 1994).

Se tienen algunas desventajas clave asociadas con la confiabilidad *test-retest* que limitan su utilidad como un coeficiente de confiabilidad teórica.

1. *¿Cuánto tiempo debe transcurrir entre las administraciones?* No hay respuestas claras, pero para los constructos basados en la opinión (*actitudinal*), se ha propugnado *un período de un mínimo de dos semanas* (Robinson et al., 1991).
2. en relación con la cantidad de tiempo entre administraciones*, si se obtiene un bajo coeficiente test-retest, ¿significa esto que la escala no es confiable, o implica que lo que la escala ha cambiado con el tiempo?* Si este último es el caso, entonces *se ha violado un principio básico de la teoría de la prueba clásica* y el coeficiente *test-retest* no sería una medición adecuada de confiabilidad (Crocker y Algina, 1986). Es decir, debe considerarse la naturaleza del constructo y la posibilidad de que cambie con el tiempo. Es posible que la posición de un encuestado sobre el constructo de interés (su puntuación verdadera) haya cambiado con el tiempo.

   Por otra parte, si el intervalo de tiempo entre las administraciones es demasiado *corto***,** la puntuación del entrevistado del tiempo 1 al tiempo 2 *puede reflejar simplemente la memoria* con una apariencia en consistencia. Por lo tanto, el período de tiempo debe ser lo suficientemente largo como para permitir que los efectos de un sesgo de respuesta de memoria se desvanezcan (un efecto de prueba de ser permanentemente consistente desde el tiempo 1 al tiempo 2), pero no tanto como para reflejar cambios históricos en la puntuación verdadera. Para otras deficiencias asociadas con el coeficiente *test-retest*, consulte a Kelley y McGrath (1988) y Nunnally y Bernstein (1994).

# Confiabilidad alternativa

Suponga que la siguiente afirmación, preferiblemente desarrollada independientemente de la primera, también está disponible para medir el constructo de la propensión a la suscripción de una página web (**CP**) con la afirmación:

***$X_2$**: Me gusta navegar en distintas opciones de páginas web para informarme.*

Asuma que en la *ocasión 1*, la declaración $X_1$ se utiliza, y en la *ocasión 2,* la declaración $X_2$ es la que se emplea. La **Tabla 3.2** muestra los datos para las dos formas alternativas (es decir, las declaraciones) para medir el constructo **CP**.

**Tabla 3.2. Datos hipotéticos para la prueba de confiabilidad alternativa**

| Sujeto | Ocasión 1 (Declaración X1) | Ocasión 2 (Declaración X2 ) |
|---|---|---|
| 1 | 4 | 4 |
| 2 | 4 | 4 |
| 3 | 6 | 5 |
| 4 | 5 | 3 |
| 5 | 5 | 4 |

| | | |
|---|---|---|
| 6 | 3 | 5 |
| 7 | 2 | 4 |
| 8 | 5 | 4 |
| 9 | 3 | 5 |
| 10 | 7 | 7 |

Fuente: propia

La correlación entre las dos respuestas (*ocasión 1 y ocasión 2*), que en el presente caso es igual a *.41*, se denomina *confiabilidad de forma alternativa*. Se deben anotar las similitudes y diferencias entre el *test-retest y la confiabilidad de la forma alternativa*. En ambos casos, la confiabilidad se estima correlacionando dos conjuntos de respuestas tomadas en dos ocasiones. En el caso *test-retest*, los dos conjuntos de respuestas son con respecto a la misma declaración o medida. Por otro lado, en la *confiabilidad de la forma alternativa,* los dos conjuntos de respuestas se refieren a *dos declaraciones diferentes desarrolladas para medir el constructo*. Los problemas relevantes para la confiabilidad de la forma alternativa son los mismos que para la *confiabilidad prueba-retest*.

# Confiabilidad por consistencia interna

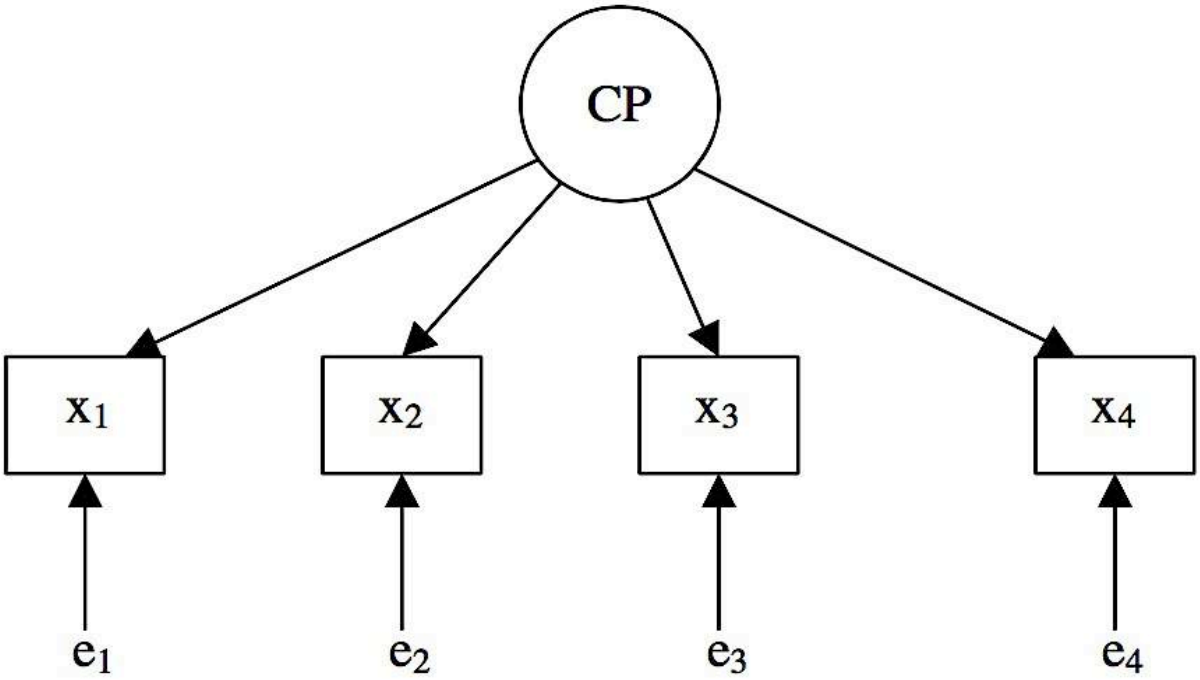
Debido a limitaciones como *el tiempo, el costo y la disponibilidad de los mismos sujetos de estudio en múltiples ocasiones*, no siempre es posible tomar medidas repetidas o usar formas alternativas. En tales casos, el concepto de *consistencia interna* se puede utilizar para estimar la confiabilidad. Este concepto requiere sólo una sola administración de los indicadores a los encuestados; sin embargo, asume la disponibilidad de múltiples medidas o elementos para medir una constructo dado. Antes de discutir la medida más utilizada de *consistencia interna (coeficiente alfa* **.α***)*, se presenta una breve discusión de la confiabilidad de la *división en mitades*, que es una forma de confiabilidad de consistencia interna.

## *Confiabilidad por consistencia interna: división en mitades*

Suponga que existen 4 indicadores que se utilizan para medir el constructo **CP.** Vea **Figura 3.2.**, el cual se basa en 4 afirmaciones:

***$X_1$:****Suscribirme a la página web me hace sentir bien.*
***$X_2$:****Me gusta navegar en distintas opciones de páginas web para informarme.*
***$X_3$:****Cuando logro la suscripción a la página web, siento que he realizado un gran trato*
***$X_4$:****Disfruto cuando me suscribo a una página web, sin importar el tiempo utilizado*

**Figura 3.2. Modelo de 1 factor**




Fuente: propia

Se pide a 10 sujetos que indiquen su grado de acuerdo o desacuerdo sobre las declaraciones anteriores usando una escala de tipo Likert de 7 puntos. Vea **Tabla 3.3.**

**Tabla 3.3. Datos hipotéticos para la prueba de confiabilidad por división en mitades**

| | Declaraciones | | | | Puntuación total | |
|---|---|---|---|---|---|---|
| Sujeto | X1 | X2 | X3 | X4 | X1- X2 | X3- X4 |
| 1 | 4 | 4 | 5 | 5 | 8 | 10 |
| 2 | 4 | 4 | 5 | 5 | 8 | 10 |
| 3 | 6 | 5 | 6 | 5 | 11 | 11 |
| 4 | 5 | 3 | 5 | 3 | 8 | 8 |
| 5 | 5 | 4 | 3 | 4 | 9 | 7 |
| 6 | 3 | 5 | 5 | 3 | 8 | 8 |
| 7 | 2 | 4 | 5 | 3 | 6 | 8 |
| 8 | 5 | 4 | 3 | 4 | 9 | 7 |
| 9 | 3 | 5 | 5 | 3 | 8 | 8 |
| 10 | 7 | 7 | 7 | 7 | 14 | 14 |

Fuente: propia

Ahora suponga, que tomamos las dos primeras y dos últimas declaraciones y calculamos la puntuación total para cada tema, que se muestra en las 2 últimas columnas de la **Tabla 3.3**. La correlación entre las dos puntuaciones totales es igual

a **.77** y se denomina confiabilidad por división de mitades. Es obvio que pueden formarse numerosas mitades divididas y la correlación para cada división será diferente (el número de divisiones que puede formarse es igual a $(2n')!/2(n'!)^2$ donde $n'=n/2$ y $n$ es el número de enunciados).

## Confiabilidad por consistencia interna: coeficiente de alfa (α)

Cronbach (1951) ha demostrado que el promedio de la correlación de todas las divisiones posibles es igual al *coeficiente de alfa* **(α)**.
El coeficiente alfa **(α)** o alfa de Cronbach (**α**) se puede calcular utilizando la **Ecuación 3.4**:

**Ecuación 3.4. Alfa de Cronbach (α)**

$$\alpha = \frac{k}{k-1}\left(\frac{\sum_{i=1}^{k}\sum_{\substack{j=1\\ i\neq j}}^{k} \operatorname{Cov}(x_i x_j)}{\sum_{i=1}^{k}\sum_{\substack{j=1\\ i\neq j}}^{k} \operatorname{Cov}(x_i x_j) + \sum_{i=1}^{k} \operatorname{Var}(x_i)}\right)$$

$$= \frac{k}{k-1}\left(\frac{\sum_{i=1}^{k}\sum_{\substack{j=1\\ i\neq j}}^{k} \sigma_{ij}}{\sum_{i=1}^{k}\sum_{\substack{j=1\\ i\neq j}}^{k} \sigma_{ij} + \sum_{i=1}^{k} \sigma_i^2}\right).$$

Donde:
***α*** es el coeficiente alfa,
$\boldsymbol{X_i}$ es medida para el indicador ***i*** y
***k*** es el número de ítems o indicadores.
Debido a que **α** es la medida de confiabilidad más utilizada, se analiza su relación con la dimensionalidad, la longitud de la escala, las correlaciones entre índices y la redundancia de indicadores.
*El coeficiente alfa* **(α)** *se refiere al grado de interrelación entre un conjunto de indicadores diseñados para medir un solo constructo, con una varianza común entre los ítems.* Para un conjunto de indicadores que componen una escala, la varianza en ese conjunto está compuesta por la varianza *verdadera* (la varianza entre individuos en el constructo que mide la escala) y *error de varianza* (toda otra varianza no contabilizada por la variable *varianza verdadera*, es decir, *varianza no compartida o varianza única*). El cálculo de **α** es un medio para dividir la *varianza de la puntuación total* en sus componentes: *verdadero y de error*. En términos muy simples, *1 - error de varianza = α, y 1 - α = error de varianza*. Por lo tanto, **α** representa la proporción de la *varianza total* de una escala que es atribuible a una fuente común, siendo esa fuente común la puntuación verdadera del constructo latente que se está midiendo.
En la **Figura 3.2,** la varianza en cada ítem o indicador (es decir, $\boldsymbol{X_i}$) que se debe a la variable latente **CP** se considera *varianza común o compartida*. Cuando **CP** varía, también lo hacen las puntuaciones de los indicadores individuales, porque el

constructo latente influye en las puntuaciones de los indicadores. Por lo tanto, las calificaciones de todos los indicadores varían conjuntamente con el constructo latente **CP**, *lo que implica teóricamente que todos los indicadores están correlacionados*. Los términos de error ($e_i$) representados en la **Figura 3.2** se consideran *únicos* para cada indicador. Es decir, *representan una varianza que no es atribuible al constructo latente* **CP**, y de acuerdo con la teoría de la prueba clásica, *los términos de error no están correlacionados*. Tanto la puntuación individual como la puntuación global de toda la escala, varían como funciones de dos fuentes:

**a.** *La fuente de variación común a sí misma (la puntuación global)* y otros indicadores y

**b.** *La variación no compartida o única asociada con ese indicador en particular.* Por lo tanto, la varianza total de la escala y la varianza para cada indicador son las que son atribuibles a fuentes: *comunes y únicas (error).* Así, alfa **(α)** es conceptualmente equivalente a la relación entre la variación de la fuente común y la variación total (Cortina, 1993; DeVellis, 1991; Nunnally y Bernstein, 1994).

Por otro lado y a considerar, tenemos la matriz de covarianza de la escala de CP para 4 ítems, la cual está dada por **C**, de la **Matriz 3.1.**

**Matriz 3.1. Varianzas-Covarianzas**

$$C=\begin{pmatrix} \boxed{\sigma_1^2} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \boxed{\sigma_2^2} & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \boxed{\sigma_3^2} & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \boxed{\sigma_4^2} \end{pmatrix}$$

Donde el arreglo de **C** representa una *matriz de varianzas covarianzas la cual es una generalización de una matriz de correlaciones, sólo que sus elementos no están estandarizados.* Los elementos $\boldsymbol{\sigma_i^2}$ de *la diagonal expresan la varianza de cada ítem i de la escala*. Los elementos $\boldsymbol{\sigma_{ij}}$ de *fuera de la diagonal indican la covarianza entre dos* ***ítems i y j*** *de la escala.* Así, se ha construido la escala para medir el constructo ***X*** (por ejemplo: administración del conocimiento del consumidor) y por ello asumiremos que ***X*** está bien representado por la suma de los ***k*** ítems de la escala que llamaremos ***Y***. Las *matrices de varianzas-covarianzas* tienen un conjunto de propiedades muy útiles. *Si sumamos todos los elementos de la matriz* ***C*** *tenemos la varianza total de la escala Y.*

## Coeficiente alfa (α). Explicación 1.

Supongamos que formamos la puntuación total suma, ***TS***, como:

$$TS= X_1+X_2+X_3+X_4$$

La varianza de ***TS***, es:

$$\begin{aligned}\mathrm{Var}\,(\mathrm{TS}) &= s^2_{\mathrm{TS}} \\ &= \sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2 + \sigma_{12} + \sigma_{13} + \sigma_{14} + \sigma_{21} + \sigma_{23} + \sigma_{24} + \sigma_{31} + \sigma_{32} \\ &\quad + \sigma_{34} + \sigma_{41} + \sigma_{42} + \sigma_{43} \\ &= \sum_{i=1}^{k} \sigma_i^2 + \sum_{i=1}^{k} \sum_{\substack{j=1 \\ i \neq j}}^{k} \sigma_{ij}\,.\end{aligned}$$

Observe que la varianza de la puntuación total suma, ***TS***, es igual a la suma de todos las varianzas y covarianzas en la matriz de covarianza. Para particionar el total de la varianza en, *varianza común* (es decir, *varianza de la puntuación verdadera) y varianza única* (es decir, *variación debida al error de medición*), debe considerarse lo siguiente:

-Los *indicadores en diagonal* representan esencialmente la *covarianza* de un indicador consigo mismo, es decir, la variabilidad en el puntaje de un indicador proveniente de una muestra dada de individuos o casos. Como tales, *los indicadores diagonales son fuentes únicas de varianza y no de la varianza que es común o compartida entre los indicadores.*

-Los *indicadores fuera de la diagonal son covarianzas* que representan la *varianza que es común o compartida por cualquier par de elementos en la escala*. Así, las entradas en la *matriz de covarianza* consisten en: *varianza única (error) y varianza común (compartida / conjunta).*

-Así, lo que es *único*, es representado a lo *largo de la diagonal principal* (**$\Sigma\sigma^2_i$**), lo que es **común** es representado por los indicadores *fuera de la diago***nal,** y la varianza total (***$\sigma^2_{TS}$***) es igual a la suma ***TS*** de todas las entradas de la matriz. Como tal, la proporción de *varianza única (no común)* a la varianza total viene dada por:

$$\frac{\sum_{i=1}^{k} \sigma_i^2}{\sigma_{\mathrm{TS}}^2}\,.$$

Se deduce de la expresión anterior, la **Ecuación 3.5.**

**Ecuación 3.5. Proporción de la variación conjunta/común de alfa (α)**

$$\text{Confiabilidad} = 1 - \frac{\sum_{i=1}^{k} \sigma_i^2}{\sigma_{\text{TS}}^2}$$

$$= \frac{\sigma_{\text{TS}}^2 - \sum_{i=1}^{k} \sigma_i^2}{\sigma_{\text{TS}}^2}$$

$$= \frac{\sum_{i=1}^{k} \sum_{\substack{j=1 \\ i \neq j}}^{k} \sigma_{ij}}{\sum_{i=1}^{k} \sum_{\substack{j=1 \\ i \neq j}}^{k} \sigma_{ij} + \sum_{i=1}^{k} \sigma_i^2}.$$

Que captura el concepto de variación común inherente a la definición del *coeficiente de alfa* **(α)**, *y debe realizarse un ajuste para el número de ítems en una escala*. El número total de indicadores en la matriz de covarianza es igual a $k^2$, donde $k$ es el número de indicadores en la escala. Por lo tanto, el denominador de la **Ecuación 3.5** es la suma de los indicadores $k^2$. Del mismo modo, el numerador está en la suma de $k^2 - k$ indicadores. Para ajustar la confiabilidad dada por la **Ecuación 3.5**, el numerador se divide por $k^2 - k$ y el denominador por $\mathbf{k}^2$, como se muestra en la **Ecuación 3.6.**

**Ecuación 3.6. Ajuste de la proporción de la variación conjunta/común de alfa (α)**

$$\frac{\sum_{i=1}^{k} \sum_{\substack{j=1 \\ i \neq j}}^{k} \sigma_{ij} / (k^2 - k)}{\left( \sum_{i=1}^{k} \sum_{\substack{j=1 \\ i \neq j}}^{k} \sigma_{ij} + \sum_{i=1}^{k} \sigma_i^2 \right) / k^2}$$

$$= \frac{k}{k-1} \left( \frac{\sum_{i=1}^{k} \sum_{\substack{j=1 \\ i \neq j}}^{k} \sigma_{ij}}{\sum_{i=1}^{k} \sum_{\substack{j=1 \\ i \neq j}}^{k} \sigma_{ij} + \sum_{i=1}^{k} \sigma_i^2} \right)$$

El cual, que es el mismo que el *coeficiente alfa* **(α)** dado por la **Ecuación 3.4**. Teóricamente, alfa **(α)** varía de **0 a 1**, por lo que son altamente deseables valores altos de alfa **(α).** También existe una fórmula de indicadores estandarizada análoga a la Ecuación 3.6 que utiliza una matriz de correlación con 1s en la diagonal principal (es decir, la correlación de un indicador consigo misma o desviaciones

estandarizadas) y correlaciones (*r*) entre pares de indicadores como elementos fuera de la diagonal (es decir, *covarianzas estandarizadas*). Así:
**-***El indicador estandarizado* ***α*** *se considera como una evaluación apropiada* de consistencia interna si se suman las puntuaciones estandarizadas de indicadores para formar una escala.
**-***Sin embargo, el indicador estandarizado α no se considera apropiado***,** cuando el puntaje bruto de los indicadores son utilizados para sumar una escala, porque las diferencias en las variaciones de los indicadores pueden afectar la puntuación total (Cortina, 1993). Ver **Ecuación 3.7.**

**Ecuación 3.7. Elemento estandarizado alfa (α)**

$$\frac{kr}{1+(k-1)\,r}$$

Donde:
**k** es el número de elementos en la escala, y
**r** es la correlación promedio entre los indicadores de la escala.
Finalmente, para los indicadores que se califican dicotómicamente, se usa la fórmula de Kuder-Richardson 20 (**KR-20)** para calcular el coeficiente alfa **(α)**. Esta fórmula, es idéntica de la **Ecuación 3.6** con la excepción de que la expresión **Σpq** sustituye a la expresión $\Sigma\sigma^2_i$. La expresión **Σpq** especifica que la varianza de cada indicador se calcula, y luego estas varianzas se suman para todos los indicadores, donde:
**p** representa cada media de indicadores, y
**q** es : 1-media del indicador
Así que la varianza de un indicador se convierte en la expresión **pq**
Para saber más, vea: Crocker y Algina (1986, p. 139-140), en la aplicación de la fórmula **KR-20.** Vea la **Ecuación 3.8**

**Ecuación 3.8. Cálculo de coeficiente de alfa (α) con la fórmula KR20 (indicadores dicotómicos)**

$$\alpha = \frac{k}{k-1}\left(1-\frac{\sum pq}{\sigma_x^2}\right)$$

## *Ejemplo de caso basado en explicación 1*

Para demostrar el cálculo del *coeficiente alfa* **(α)** (**Ecuación 3.6**) y su forma estandarizada de ítem, utilizamos partiremos del ejemplo sobre la agencia de mercadotecnia digital, utilizando la calificación hipotética de 5 indicadores medidos en escala de Likert de 7 puntos siguiente:
***$X_1$****:Suscribirme a la página web me hace sentir bien.*

***$X_2$:****Me gusta navegar en distintas opciones de páginas web para informarme.*
***$X_3$:****Cuando logro la suscripción a la página web, siento que he realizado un gran trato.*
***$X_4$:****Disfruto cuando me suscribo a una página web, sin importar el tiempo utilizado.*
***$X_5$:****Cuando me suscribo a una página web, considero que estoy ganando dinero.*
Las matrices de covarianza y de correlación para estos datos, se reportan en las **Tablas 3.4a y 3.4b**

**Tabla 3.4a. Matriz de covarianza para los 5 indicadores de la escala CP**

| | *CP1* | *CP2* | *CP3* | *CP4* | *CP5* |
|---|---|---|---|---|---|
| CP1 | 3.6457 | | | | |
| CP2 | 2.7831 | 4.3877 | | | |
| CP3 | 2.3934 | 2.2280 | 2.9864 | | |
| CP4 | 2.0706 | 2.3973 | 1.9109 | 3.7786 | |
| CP5 | 1.8338 | 2.0782 | 1.6011 | 1.8813 | 3.3138 |

Fuente: SPSS IBM 21, con datos propios.

**Tabla 3.4b. Matriz de correlación para los 5 indicadores de la escala CP**

| | *CP1* | *CP2* | *CP3* | *CP4* | *CP5* |
|---|---|---|---|---|---|
| CP1 | 1.0000 | | | | |
| CP2 | .6969 | 1.0000 | | | |
| CP3 | .7254 | .6155 | 1.0000 | | |
| CP4 | .5579 | .5888 | .5688 | 1.0000 | |
| CP5 | .5276 | .5450 | .5089 | .5317 | 1.0000 |

Fuente: SPSS IBM 21, con datos propios.

Usando la fórmula de la **Ecuación 3.6**, se tiene:

$$\alpha = \frac{5}{5-1}\left(\frac{42.356}{42.356+18.112}\right) = .876$$

Donde:

$$k = 5,$$

$$\sum \sigma_i^2 = 3.6457 + 4.3877 + 2.9864 + 3.7786 + 3.3138 = \boxed{18.112}$$

$$\sum_{i=1}^{k} \sum_{\substack{j=1 \\ i \neq j}}^{k} \sigma_{ij} = 2(2.7831 + 2.3934 + 2.2280 + 2.0706 + 2.3973 + 1.9109 + 1.8338 + 2.0782 + 1.6011 + 1.8813) = \boxed{42.356.}$$

Observe que la suma de los *indicadores fuera de la diagonal, fue multiplicada por 2* para mostrar ambas covarianzas, tanto abajo como arriba de la diagonal. Al repetir el procedimiento mencionado para la matriz de correlación y con el uso de la **Ecuación 3.7**, obtenemos como resultado .876 para alfa **(α)**. Observe también que no existe diferencia entre ambos valores de alfa **(α)**. Esto será generalmente el caso; sin embargo, si las varianzas de los indicadores son muy diferentes, entonces los 2 valores pierden no ser los mismos. El cálculo (con ***r*=** .585), es como sigue:

$$\frac{kr}{1 + (k-1)\,r} = \frac{5 \times .5865}{1 + (5-1) \times .5865} = .876.$$

## Coeficiente alfa (α). Explicación 2.

Pues bien, el *coeficiente alfa de Cronbach, se define como la proporción de la varianza total de la escala que es atribuible a la variable latente* ***X*.** Cuanto mayor sea este valor, querrá decir que ***X*** está mejor representado por **Y** en la escala porque está *causando (explicando)* la mayor parte de la varianza de esta *(varianza común).* La parte de la *varianza total* que no explica la *variable o constructo latente* es la causada por *los errores de medición de cada indicador* y se denomina *varianza específica*. Hay que hacer notar que, dados los supuestos que asumimos en el apartado anterior, cada término de *error*, provoca sólo *varianza en cada ítem* por separado y *esos errores no están correlacionados unos con otros*. En conclusión, cada indicador (y por suma, el conjunto de la escala) sólo varía como función de:

- **a.** La fuente de variación que supone la variable o constructo latente (varianza común)
- **b.** La fuente de variación que provoca el error (*varianza específica*).

Veamos cómo nos ayuda la *matriz de varianzas covarianzas* a recoger esta información. Como hemos dicho, la varianza total de la escala ($\boldsymbol{\sigma_y^2}$) es la suma de todos los elementos de esa matriz. Asimismo, la suma de los elementos de la diagonal, nos proporcionará la suma de la varianza de los ítems individuales ($\boldsymbol{\Sigma\sigma_i^2}$). Pero como las covarianzas recogen variación entre pares de ítems, y la varianza específica procede de cada ítem por separado, la varianza específica ha de venir recogida necesariamente por $\boldsymbol{\Sigma\sigma_i^2}$.Se había definido el coeficiente alfa **(α)** como la parte de la *varianza total* que era explicada por la variable o constructo latente

(*varianza común*). Como la varianza total es la suma de la varianza común y la varianza específica, la varianza común será la varianza total menos la varianza específica, luego podremos escribir:

$$\alpha = \frac{\sigma_y^2 - \sum \sigma_i^2}{\sigma_y^2} = 1 - \frac{\sum \sigma_i^2}{\sigma_y^2}$$

En la expresión anterior, sin embargo, hace falta introducir una última corrección para tener la expresión del *alfa de Cronbach*. El número total de elementos de la matriz de varianzas covarianzas es $k^2$. El número de elementos de la matriz que son específicos (*la diagonal)* es **k**, mientras que los elementos comunes (fuera de la diagonal) son $\mathbf{k^2 - k}$.
Tenemos, en la expresión anterior, una fracción con un numerador basado en **k** valores y un denominador basado en $\mathbf{k^2}$ valores. Para ajustar los cálculos de tal forma que el ratio exprese las magnitudes relativas, más que el número de términos que hay en numerador y denominador, corregiremos la expresión anterior para contrarrestar el efecto de la diferencia por $\mathbf{k^2/(k^2 - k)}$**,** o lo que es lo mismo, por **k /(k - 1)**, de forma que ahora, á estará acotado entre **0 y 1**. Ver **Ecuación 3.9**

**Ecuación 3.9. Alfa de Cronbach por varianza-covarianza**

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum \sigma_i^2}{\sigma_y^2}\right)$$

La cual, corresponde a una versión reducida de la **Ecuación 3.6.**
Para resumir, la confiabilidad medida por la expresión anterior, nos indica qué proporción de la *varianza total* está provocada por la *variable o constructo latente*, corrigiendo este valor por el número de casos que intervienen en los cálculos. Si se desea trabajar en términos *de correlaciones entre los ítems de la escala, en lugar de en términos de varianzas-covarianzas*, la expresión anterior puede adaptarse de una manera sencilla. La suma de las varianzas individuales de cada ítem $\mathbf{\Sigma\sigma_i^2}$ puede ponerse como el producto entre el número de ítems (**k**) y la media de las varianzas de los mismos **v** (por ejemplo, 10 ítems que suman 50 y 10 veces su media, que es 5, dan el mismo valor). Por lo que respecta al denominador, si llamamos **c** a la media de las covarianzas, como hay **k** varianzas y $\mathbf{k^2 - k}$ covarianzas, podremos poner la expresión anterior del *alfa de Cronbach* como sigue:

$$\alpha = \frac{k}{k-1}\left(1 - \frac{kv}{kv + \left(k^2 - k\right)c}\right)$$

Sustituyendo el 1 por su equivalente ***[kv+(k$^2$-k)c] / [kv+(k$^2$-k)c]*** y operando, se llega a:

$$\frac{k}{k \quad 1} \; \frac{k \quad k \quad 1 \quad c}{k \quad v \quad k \quad 1 \quad c}$$

Y simplificando a:

$$\alpha = \frac{kc}{v + (k-1)c}$$

Dado que se está buscando una expresión que utilice *correlaciones en lugar de varianzas y covarianzas*. Si estandarizamos las variables implicadas en la expresión anterior, la media de las *covarianzas c* se convierte en *la media de las correlaciones ( ρ ) y la media de las varianzas es igual a 1,* por lo que el *alfa de Cronbach* se calcula con la **Ecuación 3.9, resultando:**

**Ecuación 3.10. Alfa de Cronbach por correlación**

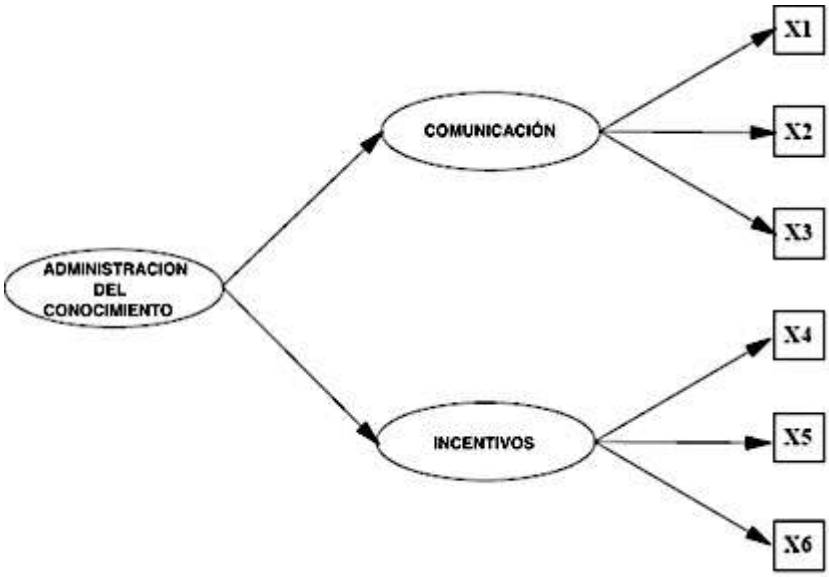
$$\alpha = \frac{k\rho}{1 + (k-1)\rho}$$

Donde:
**ρ** es, como ya hemos indicado, la media de los coeficientes de correlación entre todos los ítems que conforman la escala.
A esta fórmula se la conoce como la fórmula de Spearman-Brown (Crocker y Algina, 1986). *¿Cuáles son los valores del alfa de Cronbach por debajo de los cuales no se puede considerar como confiable a una escala?* Siguiendo a Nunally y Bernstein (1994), es posible afirmar que este nivel depende de para qué vaya a utilizarse la escala. En etapas preliminares de desarrollo de una escala, un nivel de .7 puede ser suficiente y, tras las depuraciones oportunas de la escala*, este valor no debe bajar nunca de .8.* Si en función de los valores de la escala se van a tomar decisiones que afecten a los individuos (asignar empleados a capacitación diversa según los resultados de su test de inteligencia, por ejemplo), *el alfa* **(α)** *no podrá ser inferior a .9.* (Ver más adelante discusión al respecto de las alfas a utilizar).

## *Ejemplo de caso basado en explicación 2*

Expuesto el desarrollo conceptual del análisis de confiabilidad mediante el *alfa de Cronbach,* veamos un ejemplo de aplicación por parte del constructo propuesto de administración del conocimiento (OECD, 2003) que resolveremos mediante **SPSS**. Suponga que se está realizando una investigación que pretende medir el grado de administración del conocimiento que se logra a partir de los incentivos y la comunicación con los empleados de una organización. **Ver Figura 3.3.**

**Figura 3.3. Dlagrama de trayectoria constructo administración del conocimiento**




Fuente: OECD 2003

Así como la descripción el constructo e indicadores propuestos en la **Tabla 3.5**

**Tabla 3.5. Constructo con indicadores de la administración del conocimiento**
**Instrucciones:**
Abajo le señalamos una serie de actividades que se suponen se tienen dentro de su firma y que se realizan o no con cierta regularidad. Para cada una de ellas, le solicitamos que indique con qué frecuencia se han realizado en los últimos seis meses. Tome en cuenta que los números significan: 1= Nunca; 2=Casi Nunca; 3= Rara Vez; 4= Ocasionalmente; 5= Con Alguna frecuencia; 6= Frecuentemente; 7= Con Mucha frecuencia

| | COMUNICACIÓN | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| $X_1$ | 14.-En su firma, los trabajadores comparten el conocimiento con documentación por escrito | | | | | | | |
| $X_2$ | 15.-En su firma, los trabajadores comparten su conocimiento regularmente al actualizarse todas las bases de datos de sus productos | | | | | | | |
| $X_3$ | 16.-En su firma, los trabajadores comparten su conocimiento al realizar trabajo colaborativo en equipos virtuales | | | | | | | |
| | INCENTIVOS | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| $X_4$ | 17.-En su firma, el conocimiento se comparte mediante recompensas con reconocimiento monetario e incentivos | | | | | | | |
| $X_5$ | 18.- En su firma, el conocimiento se comparte mediante reconocimientos con incentivos no monetarios | | | | | | | |
| $X_6$ | 19.-En su firma, se cuenta con un sistema que apoya el flujo del *Know-How* | | | | | | | |

Fuente: OECD (2003)

Supongamos que se ha pasado ese cuestionario a 50 gerentes de operación del sector electrónico de cierto país (empresas medianas y grandes) y que se quiere analizar la confiabilidad de la escala que mide la variable o constructo latente

COMUNICACIÓN. Sin embargo el investigador ha cometido un error y al calcular la confiabilidad ha incluido en la escala de COMUNICACIÓN el ítem $\mathbf{X_4}$, que corresponde al constructo INCENTIVOS. A continuación, recogemos y comentamos la salida de **SPSS** donde se muestran los resultados del análisis de confiabilidad.
En primer lugar se ofrecen estadísticos descriptivos de la aplicación de la escala, con las medias, desviaciones típicas y número de casos obtenidos por cada ítem en la aplicación a los 50 individuos. Ya puede apreciarse que la media del ítem $\mathbf{X_4}$ es claramente distinta a las demás. Ver **Tabla 3 .6.**

**Tabla 3.6. Estadística descriptiva del caso**

| | Media | Desviación típica | N |
|---|---|---|---|
| X1 | 5.18 | 1.395 | 50 |
| X2 | 5.40 | 1.107 | 50 |
| X3 | 5.52 | 1.216 | 50 |
| X4 | 3.64 | 1.793 | 50 |

Fuente: SPSS IBM 21

La siguiente información que se proporciona son las *matrices de varianzas covarianzas y de correlación* que, como se indicó al desarrollar el cálculo del *alfa de Cronbach*, sirven de base indistintamente para su cálculo. Si se observan las correlaciones del ítem $\mathbf{X_4}$ con todos los demás, se observará que son muy bajas, incluso negativas (-.1 con $\mathbf{X_1}$; .05 con $\mathbf{X_2}$ y -.02 con $\mathbf{X_3}$), avanzando ya que la escala, así construida va a tener problemas de confiabilidad. Ver **Tablas 3.7a y 3.7b**

**Tabla 3.7a Matriz de covarianzas interítem del caso**

| | X1 | X2 | X3 | X4 |
|---|---|---|---|---|
| X1 | 1.947 | | | |
| X2 | .763 | 1.224 | | |
| X3 | 1.211 | .522 | 1.479 | |
| X4 | -.260 | .106 | -.054 | 3.215 |

Fuente: SPSS IBM 21

**Tabla 3.7b Matriz de correlaciones interítem del caso**

| | X1 | X2 | X3 | X4 |
|---|---|---|---|---|
| X1 | 1.000 | | | |
| X2 | .494 | 1.000 | | |
| X3 | .713 | .388 | 1.000 | |
| X4 | -.104 | .053 | -.025 | 1.000 |

Fuente: SPSS IBM 21

La siguiente información a considerar, es posible ayude a calcular manualmente el valor del *alfa de Cronbach* tanto, con la expresión tradicional de las *varianzas covarianzas*, como mediante la expresión que implica correlaciones y valores estandarizados (y que en **SPSS** *se denomina standardized ítem alpha*). Ver **Tablas 3.8a y 3.8b.**

**Tabla 3.8a Estadísticos de resumen de los elementos**

| | Media | Mínimo | Máximo | Rango | Máximo/mínimo | Varianza | N de elementos |
|---|---|---|---|---|---|---|---|
| Medias de los elementos | 4.935 | 3.640 | 5.520 | 1.880 | 1.516 | .765 | 4 |
| Varianzas de los elementos | 1.966 | 1.224 | 3.215 | 1.990 | 2.625 | .782 | 4 |
| Covarianzas inter-elementos | .381 | -.260 | 1.211 | 1.471 | -4.649 | .278 | 4 |
| Correlaciones inter-elementos | .253 | -.104 | .713 | .818 | -6.853 | .097 | 4 |

Fuente: SPSS IBM 21

**Tabla 3.8b. Estadísticos de la escala**

| Media | Varianza | Desviación típica | N de elementos |
|---|---|---|---|
| 19.74 | 12.441 | 3.527 | 4 |

Fuente: SPSS IBM 21

Mediante la **Ecuación 3.9,** para calcular el alfa basta con conocer la varianza total de la escala, que aparece en el cuadro anterior bajo el título de *varianza* y que toma el valor 12.441 y la suma de las varianzas de los ítems que se pueden obtener de la diagonal de la matriz de varianzas covarianzas que también hemos ofrecido:

$$\sum \sigma_i^2 = 1,9465 + 1,2245 + 1,4792 + 3,2147 = 7,8649$$

De tal forma que, sustituyendo en la **Ecuación 3.8**, el *alfa de Cronbach*, es:

$$\alpha = \frac{k}{k-1}\left(1-\frac{\sum \sigma_i^2}{\sigma_y^2}\right) \quad \alpha = \frac{4}{3}\left(1-\frac{7,8649}{12,4412}\right) = 0,4904$$

Si aplicamos la fórmula de Spearman-Brown para datos estandarizados que mostramos en la **Ecuación 3.9** la información a requerir proviene de la **Tabla 3.8a**. Así, la media de los *coeficientes de correlación* aparece en la última fila de *correlaciones inter-elementos* bajo la columna *media* con el valor de .2535. Sustituyendo en la **Ecuación 3.10,** se tiene:

$$\alpha = \frac{k\rho}{1+(k-1)\rho} \quad \alpha = \frac{4\times 0,2535}{1+(4-1)\times 0,2535} = 0,5759$$

El ejercicio sólo es de comprobación ya que SPSS, reporta los datos resultado de manera directa. Ver **Tabla 3.9.**

**Tabla 3.9. Estadísticos de confiabilidad**

| Alfa de Cronbach | Alfa de Cronbach basada en los elementos tipificados | N de elementos |
|---|---|---|
| .490 | .576 | 4 |

Fuente: SPSS IBM 21

Hasta aquí, el investigador puede determinar sus propias conclusiones. Si se compara el valor del *alfa de Cronbach* (.4904) con el límite mínimo permitido de acuerdo con lo indicado en el desarrollo teórico *(.7 si la escala está en fase de desarrollo o .8 si ya es una escala contrastada),* se aprecia *que la confiabilidad de la misma no es suficiente.* El paso siguiente es analizar si la *supresión* de algún ítem mejoraría estos resultados o si, por el contrario, no hay forma de lograr un nivel adecuado de confiabilidad. En nuestro caso, dado que hemos cometido el *error* de incorporar el ítem $\mathbf{X_4}$ que medía un constructo diferente, cabe esperar que suprimiéndolo de la escala la confiabilidad mejore. Vea **Tabla 3.10.** como medidas de sensibilidad a la eliminación de ítems o indicadores**.**

**Tabla 3.10. Estadístico total-elemento**

| | Media de la escala si se elimina el elemento | Varianza de la escala si se elimina el elemento | Correlación elemento-total corregida | Correlación múltiple al cuadrado | Alfa de Cronbach si se elimina el elemento |
|---|---|---|---|---|---|
| X1 | 14.56 | 7.068 | .462 | .575 | .244 |
| X2 | 14.34 | 8.433 | .433 | .257 | .319 |
| X3 | 14.22 | 7.604 | .501 | .513 | .240 |
| X4 | 16.10 | 9.643 | -.037 | .029 | .777 |

Fuente: SPSS IBM 21

La columna *correlación elemento-total corregida* nos muestra el *coeficiente de correlación entre cada ítem y la suma de los restantes ítems que constituyen la escala.* Si uno de estos valores es mucho más *bajo* que los demás, esto ya aporta una evidencia de que *probablemente no esté midiendo el mismo constructo*. Como se observa, el valor correspondiente a $\mathbf{X_4}$, *además de pequeño es negativo***,** con lo que *es un firme candidato a ser eliminado de la escala.* Esta conclusión queda reforzada por la información que proporciona la última columna del cuadro anterior *alfa de Cronbach si se elimina el elemento* que nos muestra qué valor tomaría el *alfa de Cronbach* si se rehiciesen los cálculos pero eliminando cada uno de los ítems consecutivamente. Se comprueba que si se eliminasen los ítems $\mathbf{X_1}$, $\mathbf{X_2}$ o $\mathbf{X_3}$, el alfa **(α)** sería mucho menor que el que hemos obtenido (.244; .319 y .240 respectivamente) pero, si eliminásemos $\mathbf{X_4}$, la confiabilidad sería muy superior, entrando en los límites aceptables (**.777**). Por lo tanto, en conveniente, eliminar el ítem $\mathbf{X_4}$ en cuyo caso la escala para medir el constructo COMUNICACIÓN sería confiable. Por último, se muestra la salida correspondiente a **SPSS** que analiza la confiabilidad para los **tres** *primeros ítems.* Obsérvese que el valor del alfa **(α)** coincide con el pronosticado (**.777**).

Se propone, con los datos por último y como ejercicio, repetir el análisis de confiabilidad efectuado pero, ahora para la segunda variable latente (INCENTIVOS) a través de sus correspondientes ítems: $\mathbf{X_4}$**,** $\mathbf{X_5}$ **y** $\mathbf{X_6}$. (*Solución alfa de Cronbach = .7998; alfa* **(α)** *mediante la fórmula de Spearman-Brown = .7992*). Ver **Tabla 3.11.**

**Tabla 3.11 de secuencia de análisis de alfa de Cronbach de los 3 ítems del caso**

**Resumen del procesamiento de los casos**

| | | N | % |
|---|---|---|---|
| Casos | Válidos | 50 | 100.0 |
| | Excluidos[a] | 0 | .0 |
| | Total | 50 | 100.0 |

a. Eliminación por lista basada en todas las variables del procedimiento.

**Estadísticos de fiabilidad**

| Alfa de Cronbach | Alfa de Cronbach basada en los elementos tipificados | N de elementos |
|---|---|---|
| .777 | .773 | 3 |

**Estadísticos de los elementos**

| | Media | Desviación típica | N |
|---|---|---|---|
| X1 | 5.18 | 1.395 | 50 |
| X2 | 5.40 | 1.107 | 50 |
| X3 | 5.52 | 1.216 | 50 |

**Matriz de correlaciones inter-elementos**

| | X1 | X2 | X3 |
|---|---|---|---|
| X1 | 1.000 | .494 | .713 |
| X2 | .494 | 1.000 | .388 |
| X3 | .713 | .388 | 1.000 |

**Matriz de covarianzas inter-elementos**

| | X1 | X2 | X3 |
|---|---|---|---|
| X1 | 1.947 | .763 | 1.211 |
| X2 | .763 | 1.224 | .522 |
| X3 | 1.211 | .522 | 1.479 |

**Estadísticos de resumen de los elementos**

| | Media | Mínimo | Máximo | Rango | Máximo/mínimo | Varianza | N de elementos |
|---|---|---|---|---|---|---|---|
| Medias de los elementos | 5.367 | 5.180 | 5.520 | .340 | 1.066 | .030 | 3 |
| Varianzas de los elementos | 1.550 | 1.224 | 1.947 | .722 | 1.590 | .134 | 3 |
| Covarianzas inter-elementos | .832 | .522 | 1.211 | .688 | 2.317 | .098 | 3 |
| Correlaciones inter-elementos | .532 | .388 | .713 | .325 | 1.838 | .022 | 3 |

**Estadísticos total-elemento**

| | Media de la escala si se elimina el elemento | Varianza de la escala si se elimina el elemento | Correlación elemento-total corregida | Correlación múltiple al cuadrado | Alfa de Cronbach si se elimina el elemento |
|---|---|---|---|---|---|
| X1 | 10.92 | 3.749 | .731 | .565 | .557 |
| X2 | 10.70 | 5.847 | .481 | .247 | .828 |
| X3 | 10.58 | 4.698 | .657 | .511 | .650 |

**Estadísticos de la escala**

| Media | Varianza | Desviación típica | N de elementos |
|---|---|---|---|
| 16.10 | 9.643 | 3.105 | 3 |

Fuente: SPSS IBM 21

# Coeficiente de alfa (α) y dimensionalidad

Aunque el alfa (**α**) está referido al grado de interrelación que existe entre un conjunto de indicadores diseñados para medir un sólo constructo, existen varias descripciones adicionales del coeficiente alfa (**α**), siendo de las más relevantes, las mencionadas por Cortina (1993), tales como:

**a.α** *es la media de todas las confiabilidades por mitades divididas,*
**b.α** *es la confiabilidad del límite inferior de una medición,*
**c.α** *es una medida de la saturación del primer factor,*
**d.α** *es igual a la confiabilidad bajo un supuesto equivalente de tau*, y
**e.α** *es una versión general del coeficiente* **KR-20** *para indicadores dicotómicos*.

Aunque estas descripciones han sido ampliamente utilizadas**,** la validez de algunas de ellas descansa en ciertas suposiciones (por ejemplo, el uso de **α** en su forma estandarizada o no normalizada, como se ha indicado).

Una conclusión que puede extraerse de las diversas descripciones de **α** en su relación con la dimensionalidad es la siguiente (Cortina, 1993, p.100): *Es una función de la medida en que los ítems en una prueba tienen comunidades altas y, por tanto, una singularidad baja. También es una función de la interrelación, aunque hay que recordar que esto no implica una unidimensionalidad u homogeneidad.*

Por lo tanto, debe recordarse que *es posible que un conjunto de indicadores estén interrelacionados pero no sean homogéneos*. Como tal, el ***coeficiente alfa* (α)** no es una medida de la unidimensionalidad y debe usarse para evaluar la consistencia interna sólo después de que se establezca la unidimensionalidad (Clark y Watson, 1995, Cortina, 1993, Hattie, 1985, Schmitt, 1996).

Como se indicó en el capítulo 2, la *unidimensionalidad puede definirse como la existencia de un rasgo o constructo latente o subyacente de un conjunto de indicadores/medidas* (Hattie, 1985). Este conjunto de indicadores también se ha denominado como *medición congenérica*. La importancia de *establecer la dimensionalidad antes de establecer otras propiedades psicométricas (por ejemplo, la consistencia interna y la validez nomológica) no debe subestimarse* (Cortina, 1993; Hattie, 1985; Schmitt, 1996).

Para operacionalizar los constructos latentes, los investigadores usan a menudo *puntuaciones compuestas, sumando o promediando indicadores diseñados para medir el constructo de interés*. El cálculo y el uso de tales puntuaciones son significativos sólo si los *indicadores tienen una unidimensionalidad aceptable***.**

El uso de escalas *multidimensionales como si fueran unidimensionales* (es decir, sumando o promediando compuestos de indicadores) puede resultar en *ambigüedades interpretativas* de las relaciones entre constructos en una prueba de teoría. Es decir, *si un constructo es multidimensional*, pero todas las puntuaciones de los ítems se suman/promedian entre las dimensiones en una sola puntuación compuesta y se correlacionan con una variable de criterio, *tal correlación es ambigua*.

Neuberg et al. (1997) ofrecen una exposición interesante del costo de tratar una escala multidimensional como si fuera unidimensional a través de analogía experimental. Debe recordarse que un objetivo primario de la experimentación es crear una manipulación *inconfundible* de una variable independiente para evaluar con

precisión su efecto sobre la variable dependiente. Si dos constructos están siendo manipulados por una experimentación diseñada para influir en un constructo, **el** *efecto de un constructo sobre la variable dependiente no puede ser medido con precisión,* ya que desentrañar su efecto de la variación no deseada del segundo constructo es problemático. Del mismo modo, los investigadores tienen el objetivo de desarrollar escalas que sí estén evaluando un constructo. La justificación detrás de la unidimensionalidad es que la *interpretación de cualquier medida -ya sea que represente un rasgo, un estado de ánimo, una capacidad o una necesidad- es más clara si sólo una dimensión subyace a la medida* (Neuberg et al., 1997, p.1022). Cuando sólo una dimensión subyace a una medida, la correlación de esa medición con un criterio, es más clara.

Cuando existe más de una dimensión, posiblemente sugiriendo que se está evaluando más de una variable rasgo/diferencia individual, *la correlación de una medición con un criterio puede confundirse.* En suma, *el problema que surge de tratar una escala multidimensional como si fuera unidimensional (por ejemplo, sumando una puntuación total) es que más de una dimensión subyace a la puntuación total.* Esto plantea las posibilidades de que cualquier efecto sobre una variable de criterio pueda atribuirse a la *dimensión errónea o que todas las dimensiones sean necesarias para producir el efecto*. Dado que el objetivo de las investigaciones es construir y probar teorías, *se requiere un significado de constructo válido*. Por lo tanto, *la unidimensionalidad es una condición necesaria para la consistencia interna, la validez del constructo y la prueba teórica.*

Cabe señalar que hay trabajos de investigadores que han ofrecido técnicas para *combinar múltiples dimensiones* de un constructo que intentan *hacer inconfundibles* los efectos de combinar las dimensiones en una sola puntuación compuesta (Carver, 1989; Hull, Lehn, y Tedlie, 1991). En general, estos enfoques intentan estimar el efecto de cada dimensión de un constructo sobre las variables de criterio a través de medios estadísticos (por ejemplo, *regresión y enfoques de modelos de ecuación estructural*). También hay que señalar que *hay casos especiales para los cuales se puede justificar la suma de las puntuaciones a través de las dimensiones y la formación de una puntuación compuesta.* Por ejemplo, si las dimensiones *no tienen efectos diferenciales* sobre las variables de criterio de interés, entonces puede ser necesario crear un compuesto global (sumando las puntuaciones a través de las dimensiones, en Carver, 1989; Richins y Dawson, 1992).

# Coeficiente de alfa (α), longitud de escala, correlación y redundancia entre indicadores

También hay que recordar que alfa **(α)** *es una función de la longitud de la escala, el promedio de la correlación interítem (covarianza) y la redundancia de los indicadores*. Así, consideramos:

1. *La longitud de la escala.* Las fórmulas para **α** expresadas en las **Ecuaciones 3.6** y **3.7** sugieren que *a medida que aumenta el número de indicadores,* **alfa (α)** *tenderá a aumentar.* Debido a que la parsimonia también es una referencia en la medición, una pregunta importante es *¿qué es una longitud de escala apropiada en relación con el coeficiente alfa* **(α)***?.* La respuesta a esta pregunta depende

parcialmente del dominio y las dimensiones del constructo. Naturalmente, un constructo con un dominio de contenido amplio y múltiples dimensiones requerirá más elementos para aprovechar adecuadamente el dominio/dimensiones que una constructo con un dominio estrecho y una dimensión. Dado que la mayoría de las escalas se auto-administran se debe considerar *la fatiga y/o la falta de cooperación de los encuestados, por lo que la brevedad de la escala es a menudo una preocupación* (Clark y Watson, 1995; Cortina, 1993; DeVellis, 1991; Nunnally y Bernstein, 1994). Para resolverla, varios investigadores han sugerido el uso de la *teoría de la generalidad* en el diseño de escalas.

**2.** *La redundancia de los ítems y el promedio de la correlación interítem del coeficiente α.* Se recomienda ampliamente que la formulación de los indicadores sea simple y directa, que exploten el contenido del constructo y que el encuestado obtenga el significado deseado del investigador. También se recomienda que se utilicen varios elementos para aprovechar adecuadamente el dominio del constructo. *Sin embargo, cuando el texto de los indicadores es demasiado similar, el coeficiente α (así como la validez del contenido y la dimensionalidad) puede ser artificialmente mejorado.* Los indicadores que están redactados de forma demasiado *similar aumentarán el promedio de la correlación interítem, que en efecto incrementa el coeficiente α, pero sin añadir sustancialmente a la validez del contenido de la medida* (Boyle, 1991, Clark y Watson, 1995). Aunque es necesaria cierta similitud entre los indicadores de una escala para aprovechar el dominio, varios de estos que son sólo variaciones ligeras de redacción de otros elementos son redundantes y contienen muy poca información nueva sobre el constructo (Clark y Watson, 1995). Los indicadores redundantes pueden contribuir a la *paradoja de la atenuación* (*paradox attenuation*) en la teoría psicométrica, por *lo que el incremento del coeficiente α más allá de un cierto punto no aumenta la consistencia interna* (Boyle, 1991, Loevinger, 1957).

Dada esta posibilidad, *los investigadores deben tener cuidado en su interpretación de **α**,* considerando su relación con el número de indicadores de una escala, el nivel de correlación interítem, redundancia de indicadores y dimensionalidad.

**3.** Como evidencia adicional de la interrelación de la *longitud de la escala, el nivel promedio de la correlación interítem y la redundancia de la redacción de los indicadores (y la dimensionalidad),* considere lo siguiente: un nivel ampliamente preconizado de adecuación para ***α*** ha sido .70, independientemente de los temas que acabamos de discutir. De nuevo, ***α*** nos da información sobre la medida en que cada indicador de un conjunto se correlaciona con otros ítems en ese conjunto, y a medida que aumenta el número de indicadores y el incremento promedio de la correlación interítems (*ceteris paribus*), **α** *aumentará.* Además, el número de indicadores en una escala puede tener un efecto pronunciado en los niveles más bajos de correlación entre indicadores. Por ejemplo, en su meta-análisis, Peterson (1994) encontró que el nivel medio de α para una escala de 3 ítems con un promedio de correlación interítem de .47 fue de .73**.** Si el nivel de correlación entre indicadores de .47 se aplica a una escala de 9 ítems, esta escala mostraría un nivel α de .89. Sin embargo, para las escalas de 9 ítems

revisadas, Peterson (1994) encontró un promedio correlación interítem de .31 y un nivel promedio de α de .80.

Cortina (1993) demostró además las relaciones entre el coeficiente **α,** el promedio de correlación interítem, la longitud de la escala y la dimensionalidad. Sus hallazgos se basan en escalas con 1, 2 o 3 dimensiones; 6, 12 o 18 indicadores; y el promedio de correlaciones interítem de .30, .50 y .70. Tres de sus hallazgos clave son los siguientes:

**a.** *El número de indicadores afecta en gran medida* ***α***. Si una escala tiene un gran número de indicadores, **α** puede estar por encima de .70 incluso con correlaciones interítem bajas. Además, para una medida unidimensional con una promedio de correlación interítem de .50, alfa (***α***) será superior a .70 independientemente del número de indicadores en la medición.

**b.** *Para una escala que tenga más de 14 ítems, un α de .70* o más alto resultará aun si se combinan dos dimensiones ortogonales (*no relacionadas*) con una promedio de correlación interítem de .30. Si las dos dimensiones tienen el promedio de las correlaciones de indicadores que están por encima de **.70,** entonces alfa (**α)** puede ser mayor que .85.

**c.** *Dada una gran cantidad de elementos razonablemente correlacionados, una escala puede tener un* ***α*** *aceptable aunque contenga tres dimensiones ortogonales.* En suma, ***α*** *aumenta a medida que aumenta la correlación interítem, y disminuye como una función de la multidimensionalidad*. Sin embargo**,** alfa (**α**) todavía puede ser alto, a pesar de las correlaciones interítem bajas y la presencia de mutidimensionalidad. Cuando se interpreta alfa (**α**), se debe tener una unidimensionalidad ya establecida, y luego considerar el número de indicadores y el promedio de la correlación interítem.

Sin embargo**,** *maximizar el alfa* es un objetivo encomiable en la construcción de la escala, un objetivo que debe ser mediado considerando la longitud de la escala, la el promedio de correlación interítem, la redundancia del texto del indicador y la dimensionalidad/complejidad de la escala. Aunque para nuestro conocimiento, *No existen criterios estadísticos duros en cuanto a cuál es el número mínimo o máximo de elementos en una escala*, cuál es un aceptable mínimo alfa (**α**), o cuál es un nivel aceptable del promedio de la correlación interítem, por lo que se aplican una o varias heurísticas existentes.

Por ejemplo, Robinson et al. (1991) abogó por un nivel α de .80 o mejor, y el promedio de las correlaciones interítem de .30 o mejor, como ejemplo. Clark y Watson (1995) abogaron por un promedio de correlación interítem de .15 a .50 entre constructos, y para constructos estrechamente definidos, abogaban por un rango de .40 a .50. También abogaron por un nivel de coeficiente **α** de al menos .80 para una *nueva escala.* Una vez que se alcanza el punto de referencia .80, sin embargo, *la adición de elementos es de poca utilidad para la consistencia interna y la* **validez del** *contenido*, particularmente con un constructo estrechamente definido. Para tales constructos, *4 a 7 indicadores podrían bastar*. Cabe decir que estas directrices y *heurísticas* representan sólidos consejos en la construcción de escala.

# Índice de fiabilidad compuesta (IFC)

Como se ha comprobado, el *alfa de Cronbach* permite analizar la confiabilidad de un conjunto de indicadores utilizados para medir un constructo dado. Pero, normalmente, en las investigaciones de ciencias económico-administrativas no aparece implicado un solo constructo, *sino varios***.** *El alfa de Cronbach de cada factor por separado No tiene en cuenta la influencia sobre la confiabilidad del resto de constructos.* Por esta razón, Fornell y Larcker (1981) proponen el cálculo del índice de la fiabilidad compuesta **(IFC)** *para cada factor que, interpretándose exactamente igual que el alfa de Cronbach, sí que tiene en cuenta las interrelaciones*. Su fórmula, para el ***factor i*** *,* es la mostrada en la **Ecuación 3.11.**

**Ecuación 3.11. Índice de la fiabilidad compuesta (IFC)**

$$IFC_i = \frac{\left(\sum_j L_{ij}\right)^2}{\left(\sum_j L_{ij}\right)^2 + \sum_j Var\left(E_{ij}\right)}$$

Fuente: Fornell y Larcker (1981)

Donde:

**L** $_{ij}$ *es la carga factorial* estandarizada de cada uno de los ***j*** indicadores que cargan sobre el ***factor i***, y

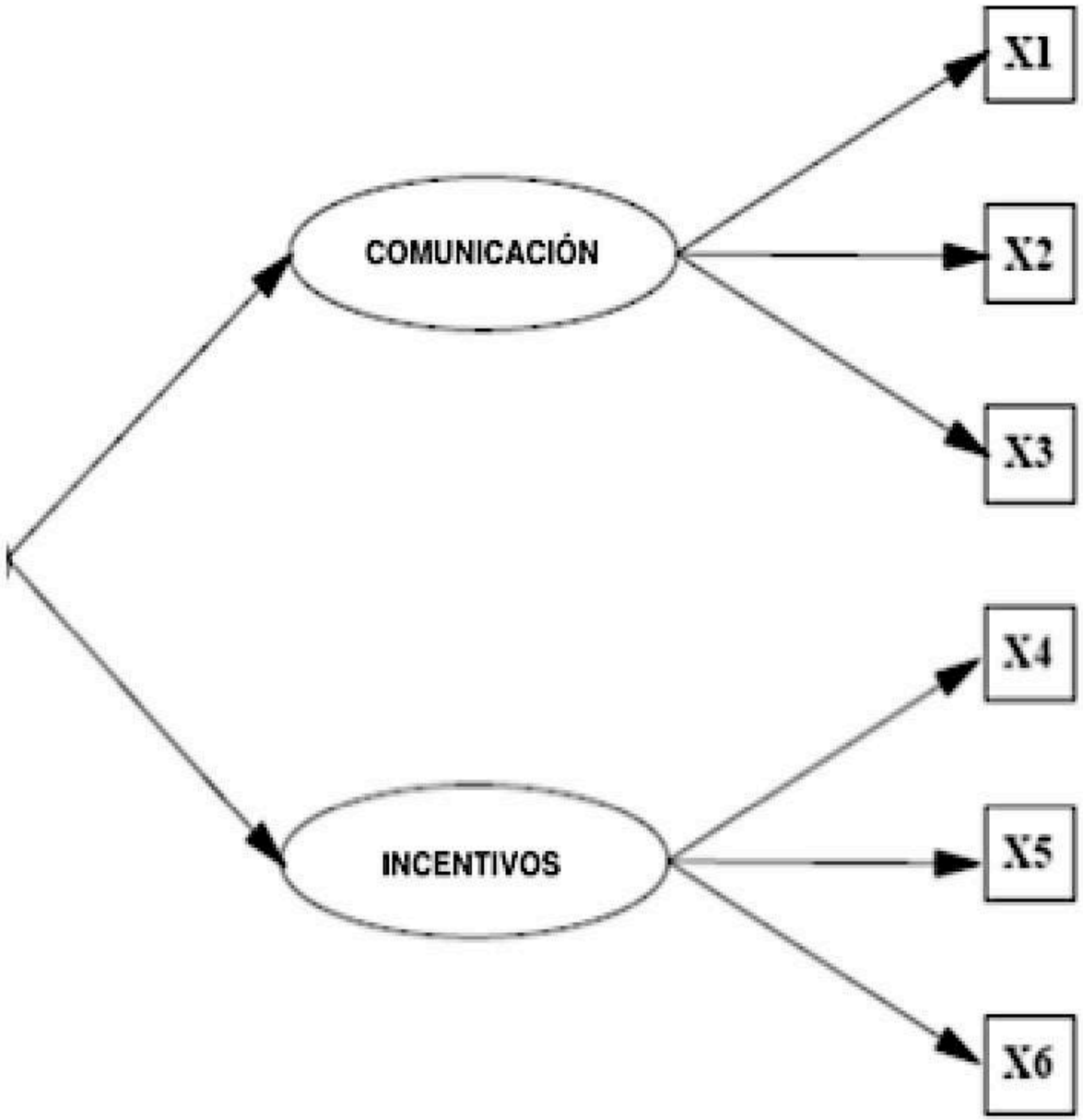
**Var (*E*$_{ij}$)** es la varianza del término de error asociado a cada uno de los **j** indicadores del **factor i** .

Se puede calcular, también como:

$$Var\left(E_{ij}\right) = 1 - L_{ij}^{\ 2}$$

Para obtener toda esta información es necesario realizar el *análisis factorial confirmatorio* entre los instrumentos de medida. Tomando el ejemplo de la **Figura 3.3**, el **modelo de medida** afecta a las variables latentes COMUNICACIÓN e INCENTIVOS, por lo que el análisis factorial confirmatorio a estimar sería el representado en la **Figura 3.4.**

**Figura 3.4. Diagrama de trayectorias para aplicar CFA en el modelo de administración del conocimiento**




Fuente: OECD (2003)

La **Tabla 3.12** muestra la sintaxis de **EQS** que permite estimar este modelo y el la **Tabla 3.13** la parte de la salida necesaria para el cálculo de la confiabilidad compuesta que se obtiene ejecutando dicha sintaxis (*que no es otra que los coeficientes de regresión estandarizados*).

**Tabla 3.12. Sintaxis EQS**

```
/TITLE
ADMINISTRACION DEL CONOCIMIENTO
/SPECIFICATIONS
VARIABLES= 6; CASES= 50; DATAFILE='E:\AFCS.ESS ';
MATRIX=RAW; METHOD=ML; ANALYSIS= COV;

/EQUATIONS
V1=F1+E1;
V2=*F1+E2;
V3=*F1+E3;

V4=F2+E4;
V5=*F2+E5;
V6=*F2+E6;

/VARIANCES
E1=*;
E2=*;
E3=*;
E4=*;
E5=*;
E6=*;

F1 TO F2=*;

/COVARIANCES
F1,F2=*;

/LMTEST

/PRINT
FIT=ALL;

/END
```

Fuente: EQS 6. con adaptación propia

**Tabla 3.13. Información necesaria para calcular el IFC**

```
STANDARDIZED SOLUTION:

X1     =V1   =    .963 F1     + .270 E1
X2     =V2   =    .514*F1     + .858 E2
X3     =V3   =    .741*F1     + .671 E3
X4     =V4   =    .946 F2     + .323 E4
X5     =V5   =    .656*F2     + .755 E5
X6     =V6   =    .672*F2     + .740 E6
```

Fuente: EQS 6. con adaptación propia

Así, los cálculos necesarios para obtener el **IFC** se ilustran en la **Tabla 3.14**:

**Tabla 3.14. Recopilación de información que calcula el IFC**

| Constructo e indicadores | Carga Factorial Estandarizada $L_{ij}$ | Varianza del Término de Error $Var(E_{ij}) = 1 - (L_{ij})^2$ | Índice de Fiabilidad Compuesta IFC> .7 $\sum(L_{ij})^2/[\sum(L_{ij})^2 + \sum Var(E_{ij})]$ |
|---|---|---|---|
| **COMUNICACIÓN** | | | |
| **$X_1$** | .963 | .073 | **.796** |
| **$X_2$** | .514 | .736 | |
| **$X_3$** | .741 | .451 | |
| **Total** | **2.218** | **1.259** | |
| **INCENTIVOS** | | | |
| **$X_4$** | .945 | .107 | **.809** |
| **$X_5$** | .657 | .568 | |
| **$X_6$** | .673 | .547 | |
| **Total** | **2.275** | **1.222** | |

Fuente: EQS 6. Con datos y adaptación propia

Con la información de la **Tabla 3.14** y aplicando la **Ecuación 3.11 para el IFC,** este sería para ambos factores:

$$IFC_i = \frac{\left(\sum_j L_{ij}\right)^2}{\left(\sum_j L_{ij}\right)^2 + \sum_j Var\left(E_{ij}\right)}$$

$$IFC_1 = \frac{\left(\sum_j L_{ij}\right)^2}{\left(\sum_j L_{ij}\right)^2 + \sum_j Var\left(E_{ij}\right)} = \frac{(2,218)^2}{(2,218)^2 + 1,259} = 0,796$$

$$IFC_2 = \frac{(2,275)^2}{(2,275)^2 + 1,222} = 0,809$$

Se puede comprobar cómo ambos índices superan lo que, en investigación, debería ser entendido como un mínimo **el valor de .7**

# Índice de la varianza compuestas extraída (IVE/AVE)

Fornell y Larcker (1981) presentan este índice **IVE (AVE.** *Average Variance Extracted***)** como la relación entre la varianza que es capturada por un factor en relación a la varianza total debida al error de medida de ese factor, esto es:

**Ecuación 3.12. Índice de varianza extraída**

$$IVE_i = \frac{\sum_j L_{ij}^{\ 2}}{\sum_j L_{ij}^{\ 2} + \sum_j Var\left(E_{ij}\right)}$$

Fuente: Fornell y Larcker (1981)

Nótese que la única *diferencia con el* **IFC** *es que cada carga factorial estandarizada es primero elevada al cuadrado antes de ser sumadas*. Forner y Larcker (1981) sugieren que es deseable que el constructo tenga valores de **IVE** iguales a .5 o superiores, es decir, que *sea superior la varianza capturada por el factor que la debida al error de medida.* Este nivel mínimo suele ser muy conservador y es fácil encontrar en la literatura ejemplos de escalas aceptadas con **IVE** inferiores.

La **Tabla 3.13** ya nos ofrecía toda la información necesaria para obtener el **IVE** en conjunto con el recopilado de la información de la **Tabla 3.14**. Los cálculos necesarios se reflejan en la **Tabla 3.15**.

**Tabla 3.15. Recopilación de información que calcula el IVE/AVE**

| Constructo e indicadores | Cuadrado de la Carga Factorial Estandarizada $(L_{ij})^2$ | Varianza del Término de Error $Var(E_{ij}) = 1 - (L_{ij})^2$ | Índice de Varianza Compuesta IVE>= .5 $\sum(L_{ij})^2/[\sum(L_{ij})^2 + \sum Var(E_{ij})]$ |
|---|---|---|---|
| **COMUNICACIÓN** | | | |
| $X_1$ | .927 | .073 | **.580** |
| $X_2$ | .264 | .736 | |
| $X_3$ | .549 | .451 | |
| Total | 1.741 | 1.259 | |
| **INCENTIVOS** | | | |
| $X_4$ | .893 | .107 | **.592** |
| **$X_5$** | .432 | .568 | |
| **$X_6$** | .453 | .547 | |
| **Total** | **1.778** | **1.222** | |

Fuente: EQS 6. Con datos y adaptación propia

Sustituyendo en la **Ecuación 3.12** del **IVE:**

$$IVE_1 = \frac{\sum_j L_{ij}^2}{\sum_j L_{ij}^2 + \sum_j Var(E_{ij})} = \frac{1,741}{1,741 + 1,259} = 0,580$$

$$IVE_2 = \frac{1,778}{1,778 + 1,222} = 0,592$$

Valores que, como se ha señalado, superan los mínimos recomendables.

# Conceptos adicionales de la confiabilidad

Finalmente, en la **Tabla 3.16**, se mencionan algunos para sus consideraciones.

**Tabla 3.16. Conceptos de confiabilidad adicionales**

| Confiabilidad | Tipo de Enfoque | Cuando se aplica | Algunos índices, son: |
|---|---|---|---|
| **Relativa** | Estabilidad | Una misma escala en dos ocasiones sucesivas | Coef. de correlación Coef. W. p. |
| | Equivalencia | En la misma muestra dos formas paralelas de una escala | Coef. de correlación entre dos formas paralelas |
| | Consistencia interna | La escala una sola vez | Spearman - Brown Rulon y Guttman Kuder - Richardson Host α de Cronbach |
| **Absoluta** | Variación interna | La escala una vez | Varianza |
| | Consistencia interna | La escala una vez y se mide el error de la prueba o cuando se aplica en dos pruebas diferentes | **σ** Dif med. |

Fuente: Maldonado-Guzmán (2015)

# Conclusión

Este capítulo ha discutido varias formas de confiabilidad, resaltando las ventajas del coeficiente *alfa de Cronbach*. También discutió la relación entre el coeficiente alfa **(α)**, la unidimensionalidad, la correlación entre los elementos, la duración de la escala y la redundancia de la redacción de los elementos. Destacamos la importancia de establecer la dimensionalidad antes de evaluar la consistencia interna, así como los efectos de la correlación entre las palabras y la redundancia de la redacción en la confiabilidad. Como reconocimos previamente, la dimensionalidad y la confiabilidad son entradas necesarias para establecer la validez de una escala. El siguiente capítulo discute los diversos tipos de validez que se encuentran y los procedimientos para establecer estos tipos de validez en el desarrollo de la escala.

# CAPÍTULO 4. VALIDEZ

¿Cuándo es válida una escala? Cuando lo que está midiendo realmente es la variable latente que se supone que tiene que medir. Esta definición puede sofisticarse mucho más, pero esa es la esencia de la *validez*. Siguiendo la definición de Bohrnstedt (1976) que apunta en esta misma línea: *validez es el grado en que un instrumento mide el concepto bajo estudio***.** La validez, sin embargo, es un concepto multifacético y tiene diversas dimensiones que deben explicarse y analizarse por separado.

Este capítulo aborda cuestiones críticas relacionadas con la *validez de constructo*, tal como se aplican al desarrollo de la medición para *constructos latentes o subyacentes.* La validez de constructo se considera como el grado en el cual, una medición refleja verdaderamente el concepto que se está investigando o el grado en el cual las variables utilizadas para observar la covariación dentro y entre constructos pueden interpretarse en términos de *constructos teóricos* (Calder, Phillips y Tybout, 1982).

La validez de constructo puede y debe considerarse ampliamente como una referencia al grado en que las inferencias se pueden hacer legítimamente, provenientes de las mediciones de estudios de constructos teóricos en los que se basan esas operacionalizaciones. Similar a las opiniones expresadas por Haynes et al. (1999) y Trochim (2002), la validez de constructo representa la calidad global de un estudio de investigación o incluso un programa de estudios. En general, una medición tiene validez de constructo cuando refleja o muestra:

**a.** Un grado de evaluación en magnitud y dirección de una muestra representativa de las características del constructo, y

**b.** Un grado en que la medición *no está contaminada* con elementos del dominio de otras constructos o errores (Peter, 1981).

Es importante señalar *que la validez de constructo de una medición no se evalúa directamente sino que se deduce de la evidencia en que las puntuaciones sustantivas de la medida, se comportan como se esperaba de las pruebas planeadas teóricamente, así como de la calidad de los procedimientos que se emplearon en el desarrollo y la validación de la medición*.
El resto del capítulo describe las diversas maneras de demostrar o proporcionar evidencia de validez de constructo. Estos tipos de validez pueden utilizarse en el desarrollo de la medición para generar y desarrollar ítems o indicadores válidos de constructos no observables y/o para proporcionar evidencia de la calidad de la medición. Usando clasificaciones similares a las propuestas por Haynes et al. (1999) y Trochim (2002), los diversos tipos o pruebas de validez discutidos en el capítulo se agrupan de la siguiente manera:

**1.** Validez de versión (*translation validity*).
- Validez aparente (*face validity*).
- Validez del contenido (*content validity*).

**2.** Validez relacionada con el criterio (*criterion-related validity*).
- Validez predictiva y post-dictiva (*predictive and post-dictive validity*).
- Validez concurrente (*concurrent validity*).
- *Validez convergente (convergent validity)*
- *Validez discriminante (discriminant validity)*
- Validez del grupo conocido (*known-group validity*)

***3.*** • *Validez nomológica (nomological validity)*

A continuación, se presentarán los conceptos de cada tipo de validez arriba anotado, para su conocimiento y exploración. Sin embargo, las que se abordarán a detalle son las marcadas con *, por su utilidad y practicidad. Así, definiremos que la validez del constructo se divide en dos: validez convergente (*convergent validity*) y validez discriminante (*discriminant validity*). Una escala tiene validez de constructo, cuando tiene ambas.

# Validez de versión (*translation validity*)

La validez de contenido (*content validity*) y la validez aparente (*face validity*) reflejan la medida en que un constructo se traduce en la operacionalización del constructo (Trochim, 2002). Los términos a menudo se confunden y/o se utilizan indistintamente. Aunque la distinción entre la validez aparente (*face validity*) y la validez de contenido (*content validity*) frecuentemente no está clara, Rossiter (2001) argumentó que los dos conceptos de validez difieren en formas importantes y no deben ser confundidos.

## *Validez aparente (face validity)*

La evidencia de la validez aparente (*face validity*) se proporciona a partir de una evaluación *post hoc* en la que los ítems o indicadores de una escala miden adecuadamente el constructo (Nunnally y Bernstein, 1994, Rossiter, 2001). Esta validez, puede ser juzgada después de que una medición ha sido desarrollada, a menudo antes de la aplicación en otro estudio, por los usuarios potenciales de dicha medición. Sin embargo, en examen posteriores del contenido de un indicador, para una medición, sólo ofrece pruebas incompletas de validez, porque los evaluadores

sólo ven los indicadores que quedan y tienen que deducir lo que se omitió y por qué. Por lo tanto, la validez aparente (*face validity*) puede considerarse como limitado en validez del contenido, (Nunnally y Bernstein, 1994, p.110). Además, algunos argumentan que la validez aparente (*face validity*) es más similar a lo que un encuestado puede inferir acerca de lo que pretende medir un indicador (Nevo, 1985) y ha sido referido como *la mera suposición de que una medida tiene validez* (Kaplan y Saccuzzo, 1997, p. 1320). Por lo tanto, una alta validez aparente (*face validity*) de un instrumento mejora su uso en situaciones prácticas al inducir la cooperación entre los encuestados a través de la facilidad de uso, el nivel de lectura adecuado, la claridad y los formatos de respuesta apropiados.

## *Validez de contenido (content validity)*

Las garantías de validez de contenido (*content validity*) se basan sobre esfuerzos de producción, a priori, teórica, de la generación de indicadores a medir. Específicamente, la validez del contenido representa el grado en que los elementos de un instrumento de medición son relevantes y representativos del constructo objetivo, para un propósito de evaluación particular (Haynes et al., 1995, 238). La representatividad se refiere al grado en que los elementos son proporcionales a las facetas de la constructo objetivo y al grado en que se ha muestreado todo el dominio de dicho constructo. Es decir, *la medición debe ser coherente con el dominio teórico de la constructo en todos los aspectos, incluyendo la formulación de los ítems, los formatos de respuesta y las instrucciones* (Haynes et al., 1995; Haynes et al., 1999; Netemeyer, et al., 2002; Robinson et al., 1991).

El objetivo básico, entonces, es asegurar que los ítems reflejen las áreas de contenido abarcadas por el constructo objetivo. Como tal, la validez del contenido se manifiesta a partir de procedimientos en las etapas iniciales del desarrollo de la escala que generan ítems o indicadores representativos del dominio del enfoque del constructo. El desarrollo de una medición válida de contenido, se mejora durante las primeras etapas del desarrollo de la escala mediante la creación efectiva de un *pool de indicadores* y la posterior evaluación de los indicadores por parte de expertos.

En primer lugar, los ítems iniciales de dicho *pool*, deben elegirse de modo que muestren todas las áreas de contenido posibles que podrían ser abarcadas por el constructo de acuerdo con el conjunto completo de perspectivas teóricas que subyacen a dicho constructo (Clark y Watson, 1995). El grupo inicial de ítems debe ser amplio en cobertura e incluir un gran número de indicadores o ítems potenciales a través de las dimensiones teóricas realizadas *a priori*. El grupo *pool de indicadores*, inicial debe incluir sin duda, algunos ítems que *posteriormente serán eliminados* en los procedimientos de evaluación de seguimiento y análisis psicométricos. **Un gran** *número de indicadores aumenta la probabilidad de que todas las dimensiones se representen adecuadamente*. Es decir, el enfoque en la amplitud de la agrupación de elementos se aplica a todas las áreas o dimensiones del constructo posibles, de modo que las áreas individuales no estén sub-representadas en la escala final. *La validación de contenido es particularmente importante para constructos ambiguos o complejos.* La validación de contenido se mejora mediante la definición precisa y la conceptualización del constructo, incluyendo la especificación de la dimensionalidad y de las definiciones individuales de las diversas dimensiones que comprende el

constructo. Además, se garantiza la validez del contenido en la medida en que los expertos coinciden en que los ítems que sean de *efecto o reflectivos (reflective)* en toda la construcción global y que estos jueces estén de acuerdo en que los ítems son representativos del dominio y las facetas del constructo.
Los indicadores o ítems que se incluirán en un grupo. se pueden obtener utilizando cualquier número de fuentes. Las fuentes de indicadores más frecuentemente utilizadas en el desarrollo de escalas incluyen declaraciones, previamente empleadas de investigaciones previas que involucran el constructo, casos de muestras ya cerrados, de sujetos representativos y declaraciones generadas por investigadores basadas en el conocimiento y comprensión del dominio del constructo y sus fundamentos teóricos.
Como ejemplo de validez de contenido, se tiene el trabajo de *Los factores determinantes del modelo de negocios abierto* (**OBM**) de Mejía-Trejo (2017), el cual, muestra la creación de un modelo empírico de **OBM**, que consiste en 5 principales factores/22 variables/139 ítems o indicadores distribuídos en administración del negocio (**BMG**, 10 variables/76 indicadores), estrategia (**STR**, 3 variables/14 indicadores), tecnología (**TEC**, 3 variables/24 indicadores), nuevos emprendimientos (**NWE**, 3 variables /7indicadores) y orientación de la innovación abierta (**OIO**, 3 variables/18 indicadores). Conclusión. Aunque el modelo empírico final de **OBM** tiene un efecto positivo significativo entre sus variables, también mostró diferentes niveles de carga de factores, lo que significa oportunidades para mejorar el modelo para el sector de las TIC de la Zona Metropolitana de Guadalajara (**ITSMZG**), México. Se trató de un estudio documental para seleccionar las principales variables entre los especialistas de las **ITSMZG** que practican el proceso **OBM** mediante el proceso de jerarquía analítica (**AHP.** *Analysis Hierarchy Process*) y el **Panel de Delphi** a fin de contrastar los términos académicos con la experiencia de los especialistas. Es un estudio descriptivo, exploratorio, correlacional, transeccional, cualitativo-cuantitativo para obtener un cuestionario final en **escala Likert**, con confiabilidad a través de prueba piloto (*alfa de Cronbach*>0.7), aplicado entre enero 2015-mayo 2016 a una población total de: 600 especialistas en el **ITSMZG** (150 profesores de IT; 150 representantes de consultores de firmas IT como *parte consultora* ; 290 CEO PyME y 10 CEO de empresas grandes como parte de *toma de decisiones*, con 1 año en el mercado, 80% con licenciatura, 20% con postgrado, 20% mujeres y 80% hombres). Se diseñó un modelo de ecuaciones estructural de primer orden (**SEM**) como técnica de análisis factorial confirmatorio (**CFA**), mediante el software **EQS 6** para analizar las variables subyacentes de **OBM**, y determinar un modelo final.

## *Estudio de caso: administración del conocimiento OECD (2003)*

Al final, una escala no es sino un conjunto de indicadores que se supone que reflejan todas las dimensiones del concepto que pretenden medir. Se define la validez de contenido *(content validity)* como el grado en que la escala recoge todas esas dimensiones. Retomando el ejercicio de la administración del conocimiento mostrado en la **Figura 3.3.** Diagrama de trayectoria constructo administración del conocimiento de la vemos que la escala propuesta por la OECD (2003), se compone de dos factores: COMUNICACIÓN e INCENTIVOS. La escala tendrá validez de contenido (*content validity)* si los ítems seleccionados de entre todos los posibles para

construirla, recogen adecuadamente esas dos factores.
Queda claro, de la explicación precedente, que es muy difícil determinar si una escala tiene validez de contenido (*content validity)*, por cuanto que se basa en el juicio del experto que se supone que ha considerado todas las posibles factores y contenidos del concepto, fenómeno o comportamiento analizado. Un siguiente paso ahora es estimar el modelo mediante la utilización del software **EQS**, aunque no se trata de enseñar el manejo del programa, sí que es muy importante que se vea cómo la sintaxis refleja el conjunto de relaciones que se han puesto de manifiesto en la **Figura 3.3**, y la inmediatez de las instrucciones.
Considerando la **Tabla 3.12 Sintaxis EQS,** para hacerla comprensible, expliquemos algunas instrucciones. En la primera línea, se indica únicamente que se ejecute el programa, teniendo en cuenta que los datos son correlaciones, pero que deben transformarse en *varianzas-covarianzas*. Las ecuaciones **(/EQUATIONS**) se corresponde a la formulación de Bentler (Bentler y Wu, 1993). Tomemos las ecuaciones que definen el modelo, por ejemplo:

**V3 = *F1 + E3;**

Si nos fijamos en la **Figura 3.3**, vemos que **V3** se obtiene multiplicando por una cierta carga factorial **L** el factor **F1**, sujeta a un *error* que hemos llamado $\mathbf{E_3}$. Pues bien ***** es como **EQS** indica que la carga factorial **L** debe ser estimada, ese parámetro aparece multiplicado por **F1** y sumándosele **E3** nos daría el indicador **V3**.
Bajo ***VARIANCES*** aparecen las varianzas de los dos factores que tienen que ser estimadas, por eso se observa que la varianza de **F1** y la de **F2** están con un asterisco (*).

**/VARIANCES**

**F1 TO F2 = *;**

Mientras que las varianzas de los términos de **error** $\mathbf{E_1}$ hasta $\mathbf{E_6}$ también deben ser estimadas: **E1 TO E6 = *;**

Finalmente, bajo **/COVARIANCES** se le indica al programa qué parámetros de las covarianzas se fijan a una cantidad y cuáles hay que estimar. Como en nuestro caso todas deben ser estimadas, *no aparecerá* **F1, F2 = 1**, sino **F1, F2 = *,** es decir que el parámetro de *covarianzas entre* **F1 y F2** *debe estimarse*. Una forma de no tener que escribir todas las combinaciones entre los 6 factores es utilizar el término ***TO***:

**/COVARIANCES F1 TO F6 = *;**

Pues bien, con esas sintaxis **EQS** estima el modelo. El primer paso en la interpretación de los resultados pasa por determinar cuál es la *bondad de ajuste* del mismo. Ver **Tabla 4.1.**

**Tabla 4.1. Estadísticos de bondad de ajuste proporcionados por EQS.**

```
GOODNESS OF FIT SUMMARY FOR METHOD = ML

INDEPENDENCE MODEL CHI-SQUARE        =     110.190 ON    15 DEGREES OF FREEDOM

INDEPENDENCE AIC =      80.190   INDEPENDENCE CAIC =      36.509
       MODEL AIC =      -4.301          MODEL CAIC =     -27.597

CHI-SQUARE =       11.699 BASED ON       8 DEGREES OF FREEDOM
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS       .16515

THE NORMAL THEORY RLS CHI-SQUARE FOR THIS ML SOLUTION IS         10.841.

FIT INDICES
-----------
BENTLER-BONETT     NORMED FIT INDEX =      .894
BENTLER-BONETT NON-NORMED FIT INDEX =      .927
COMPARATIVE FIT INDEX (CFI)         =      .961
BOLLEN'S          (IFI) FIT INDEX   =      .964
MCDONALD'S        (MFI) FIT INDEX   =      .964
JORESKOG-SORBOM'S  GFI  FIT INDEX   =      .931
JORESKOG-SORBOM'S AGFI  FIT INDEX   =      .820
ROOT MEAN-SQUARE RESIDUAL (RMR)     =      .165
STANDARDIZED RMR                    =      .081
ROOT MEAN-SQUARE ERROR OF APPROXIMATION (RMSEA)    =      .097
90% CONFIDENCE INTERVAL OF RMSEA  (        .000,          .206)
```

Fuente: EQS 6. con datos y adaptación propia

## *Interpretación.*

De su análisis se desprende lo siguiente.

• *La prueba Chi cuadrado analiza la hipótesis nula de que el modelo es NO significativo.* Esta hipótesis debería *rechazarse (p > 0.05)*. *Sin embargo, con grandes muestras, el estadístico Chi-Cuadrado tiene tendencia a ser significativo* (James, Mulaik y Brett, 1982). Por ello, además de analizar otros estadísticos que luego veremos, algunos autores proponen que el estadístico es aceptable *si el ratio entre el valor de la Chi-Cuadrado y el número de grados de libertad es inferior a 2.*

En nuestro caso, la *Chi-Cuadrado es significativa*, pero el *ratio es inferior a 2 (11.699/8 = 1.46).*

• Parece indicar que el modelo de *administración del conocimiento* tiene un buen ajuste, pero podemos seguir analizando otros estadísticos.

Entre el conjunto de indicadores de ajuste que proporciona **EQS**, se recomienda analizar aquellos *no normalizados, dado que tienen menor tendencia a ofrecer resultados sesgados en casos de pequeñas muestras* (Bentler y Wu, 1993).

• Se observa que, tanto el *Non Normed Fit Index (**NNFI**)* (Bentler y Bonnet, 1980) como el *Comparative Fit Index (**CFI**)* (Bentler y Wu, 1993**)**, que cumplen este requisito, ya que tienen *valores superiores a 0.9*, indicando un ajuste razonable.

Con esto, ya hemos estimado el modelo de medida y hemos comprobado que ofrece un buen ajuste. Pues bien, sobre este modelo de medida, ya podemos analizar la *validez convergente* y la *validez discriminante* de las escalas implicadas. Para

analizar la *validez nomológica,* será necesario estimar el modelo teórico.

# Validez relacionada con el criterio (*criterion-related validity*)

Los esfuerzos para demostrar la validez relacionada con el criterio (*criterion-related validity*) implican procesos en los que se emplean mediciones externas al instrumento de medición propuesto (Nunnally y Bernstein, 1994). Las fuentes de validez específicas que pueden utilizarse bajo la etiqueta ésta etiqueta, incluyen la *validez predictiva y post-dictiva, la validez concurrente, la validez convergente y discriminante y la validez del grupo conocido.*

## *Validez predictiva y post-dictiva (predictive and post-dictive validity)*

El término validez predictiva a menudo se utiliza indistintamente como *validez de criterio.* En aplicaciones tradicionales que implican desarrollo de escala, la validez predictiva *se refiere a la capacidad de una medida para predecir con eficacia algún criterio subsiguiente y temporalmente ordenado.*

Un ejemplo, es el desarrollo de una prueba de conocimiento de la persuasión del consumidor para su uso con adolescentes (véase Friestad y Wright, 2001) puede usarse para ilustrar el problema. Así, la validez de tal prueba podría ser apoyada por el tamaño de las correlaciones entre una prueba de conocimiento de persuasión del consumidor e índices o una observación de los comportamientos efectivos del consumidor que se producen después de que la prueba se ha administrado.

La *validez del criterio post-dictivo ocurre cuando la variable de resultado se mide antes de la variable independiente*.

Por ejemplo, los hogares de consumidores pueden ser observados por comportamientos que demuestren las *necesidades únicas de consumidores* (**CNFU***. Consumers Need for Uniqueness*). Posteriormente, una escala diseñada para medir el **CNFU** es administrada a un adulto de cada hogar. Una correlación significativa entre un índice que comprende los comportamientos observados y la escala del **CNFU** ofrecería alguna evidencia de *validez post-dictiva* para la escala de **CNFU** (Tian, Bearden y Hunter, 2001).

## *Validez concurrente (concurrent validity)*

La evidencia de la validez concurrente de una medición es proporcionada por correlaciones considerables entre la medida de constructo en desarrollo y una medida de criterio recogida simultáneamente o *concurrentemente.* A modo de ejemplo, Bearden et al. (2001) aportaron pruebas de la validez concurrente de sus mediciones de su indicador *auto-confianza del consumidor* utilizando correlaciones entre las escalas de dicha variable y las mediciones de otra variable llamada *conocimiento subjetivo del producto* (Park, Mothersbaugh y Feick, 1994). Además, estos esfuerzos de desarrollo incluían también la demostración de la validez de la escala relativa, en las que las mediciones de confiabilidad que se proponían estaban más fuertemente correlacionadas que de las mediciones fueron realizadas del

indicador *autoestima* (Rosenberg, 1965) y el procesamiento del indicador *autoconfianza* (Wright, 1975) . En estas comparaciones posteriores, se obtuvieron pruebas de *relativa validez concurrente* a partir de diferencias significativas en las pruebas de *correlaciones dependientes,* lo que sugirió que las mediciones evaluadas estaban más fuertemente correlacionadas con el criterio concurrente (es decir, *evaluaciones subjetivas de conocimiento)* que las medidas competitivas .

## *Validez convergente (convergent validity).*

Se dice que una medición posee validez convergente si las mediciones independientes del mismo constructo convergen o están altamente correlacionadas. La evidencia de validez convergente es ofrecida por correlaciones significativas y fuertes entre diferentes medidas del mismo constructo.

## *Estudio de caso: administración del conocimiento OECD (2003)*

La *validez convergente* existe cuando se emplean distintos instrumentos para medir un mismo constructo (*distintos indicadores para una misma variable latente*) y estos instrumentos están fuertemente correlacionados. Como veremos al desarrollar el ejemplo que vendrá a continuación, en este tema la validez convergente se determinará revisando los test *t de las cargas factoriales*.

Si todas las cargas factoriales de las variables manifiestas que miden el mismo constructo son *estadísticamente significativas, será una evidencia que apoya la validez convergente de estos indicadores* (Anderson y Gerbing, 1988). Hair, Anderson, Tatham y Black (1998**)** recomiendan que, además de ser significativos, el promedio de las cargas sobre cada factor sean superiores a *.7*.

También deberíamos analizar si el modelo mejoraría significativamente si se añadiera alguna relación entre una variable manifiesta diseñada para medir un factor dado y otro factor para el que no se ha diseñado. Los *multiplicadores de Lagrange* nos permiten verificar esta posibilidad.

Si analizamos las **Tablas 4.2, 4.3 y 4.4** para el *modelo de medida* observamos:

**a.** Todas las *cargas factoriales son significativas*, como se desprende de sus respectivas **t** (3.073 para la carga entre **V2 y F1**; 3.810 para la carga entre **V3** y **F1**...). Si este valor de **t** es superior a 1.96, el parámetro *es significativo para p < 0.05,* si es superior a *2.576 es significativo para p < 0.01* y si es superior *a 3.291* para *p < 0.001*. Puede comprobarse en la misma **Tabla 4.2** que el parámetro p < 0.01 es significativo para la variable $\mathbf{V_2}$ y de *p < 0.001* para las variables $\mathbf{V_3}$, $\mathbf{V_5}$ y $\boldsymbol{V_6}$. Asimismo, si calculáramos el promedio de los mismos, *veríamos que superan claramente .7*

**b.** El ajuste del modelo mejoraría significativamente *si la variable V4 también cargara sobre el factor* **F1** (primera línea **Figura 4.4.**).

**Tabla 4.2. Estimación de las cargas factoriales**

```
MEASUREMENT EQUATIONS WITH STANDARD ERRORS AND TEST STATISTICS
STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.

 X1     =V1   =  1.000 F1     +1.000 E1


 X2     =V2   =   .423*F1     +1.000 E2
                  .138
                 3.073@

 X3     =V3   =   .671*F1     +1.000 E3
                  .176
                 3.810@

 X4     =V4   =  1.000 F2     +1.000 E4


 X5     =V5   =   .646*F2     +1.000 E5
                  .164
                 3.932@

 X6     =V6   =   .616*F2     +1.000 E6
                  .154
                 3.994@
```

Fuente: EQS 6. con datos y adaptación propia

**Tabla 4.3. Estimación de las cargas factoriales estandarizadas**

```
STANDARDIZED SOLUTION:

X1     =V1   =    .963 F1     + .270 E1
X2     =V2   =    .514*F1     + .858 E2
X3     =V3   =    .741*F1     + .671 E3
X4     =V4   =    .946 F2     + .323 E4
X5     =V5   =    .656*F2     + .755 E5
X6     =V6   =    .672*F2     + .740 E6
```

Fuente: EQS. 6. con datos y adaptación propia

**abla 4.4. Test de los multiplicadores de Lagrange multivariados**

| | CUMULATIVE MULTIVARIATE STATISTICS | | | | UNIVARIATE INCREMENT | | | |
|---|---|---|---|---|---|---|---|---|
| STEP | PARAMETER | CHI-SQUARE | D.F. | PROB. | CHI-SQUARE | PROB. | HANCOCK'S SEQUENTIAL D.F. | HANCOCK'S SEQUENTIAL PROB. |
| 1 | V4,F1 | 4.143 | 1 | .042 | 4.143 | .042 | 8 | .844 |

Fuente: EQS 6. con datos y adaptación propia

## *Interpretación*

Estos últimos resultados cuestionan la validez convergente *(convergent validity) de la escala.* El que la *Chi-Cuadrado mejorara significativamente si se añadiera la relación entre* $\mathbf{V_4}$ y $\mathbf{F_1}$ (se reduciría en un valor de 4.143), es coherente con el patrón de residuos elevados que hemos visto con anterioridad **(RMSEA** superior a 0.08). Ante una situación como ésta pueden hacerse varias cosas:

- Lo más fácil es añadir esa relación, pero esto es tanto *como asumir que el modelo no posee validez convergente.*
- Otra alternativa es asignar **V4 a F1** y eliminarla de **F2**, pero vemos que la carga factorial sobre **F1** es muy elevada, lo que hace esta opción también inviable.

Por tanto la opción más razonable es asumir que $\mathbf{V_4}$ no es un buen ítem de la escala y eliminarlo completamente. *Por lo tanto en la sintaxis eliminaremos la expresión:*

**V4 = * F1 + E4.**

A continuación reestimamos el modelo y volvemos a analizar la validez convergente a partir de las mismas salidas de bondad de ajuste (**Tabla 4.5**), significatividad de las cargas factoriales (**Tabla 4.6**) y valor de las cargas estandarizadas (**Tabla 4.7.**).

**Tabla 4.5. Indicadores de ajuste del modelo corregido**

```
GOODNESS OF FIT SUMMARY FOR METHOD = ML

INDEPENDENCE MODEL CHI-SQUARE        =      66.051 ON    10 DEGREES OF FREEDOM

INDEPENDENCE AIC =       46.051   INDEPENDENCE CAIC =      16.931
       MODEL AIC =       -2.897          MODEL CAIC =     -14.545

CHI-SQUARE =         5.103 BASED ON        4 DEGREES OF FREEDOM
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS        .27691

THE NORMAL THEORY RLS CHI-SQUARE FOR THIS ML SOLUTION IS          4.949.

FIT INDICES
-----------
BENTLER-BONETT      NORMED FIT INDEX =       .923
BENTLER-BONETT NON-NORMED FIT INDEX =       .951
COMPARATIVE FIT INDEX (CFI)          =       .980
BOLLEN'S            (IFI) FIT INDEX  =       .982
MCDONALD'S          (MFI) FIT INDEX  =       .989
JORESKOG-SORBOM'S  GFI  FIT INDEX    =       .961
JORESKOG-SORBOM'S AGFI  FIT INDEX    =       .854
ROOT MEAN-SQUARE RESIDUAL (RMR)      =       .093
STANDARDIZED RMR                     =       .050
ROOT MEAN-SQUARE ERROR OF APPROXIMATION (RMSEA)     =       .075
90% CONFIDENCE INTERVAL OF RMSEA  (          .000,          .237)
```

Fuente: EQS. 6. con adaptación propia

**Tabla 4.6 Estimación de las cargas factoriales del modelo corregido**

```
MEASUREMENT EQUATIONS WITH STANDARD ERRORS AND TEST STATISTICS
STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.

 X1     =V1   =  1.000 F1     +1.000 E1


 X2     =V2   =   .442*F1     +1.000 E2
                  .137
                 3.221@

 X3     =V3   =   .707*F1     +1.000 E3
                  .176
                 4.024@

 X5     =V5   =  1.000 F2     +1.000 E5


 X6     =V6   =   .614*F2     +1.000 E6
                  .955
                  .643
```

Fuente: EQS. 6. con datos adaptación propia

**Tabla 4.7. Estimación de las cargas factoriales estandarizadas modelo corregido**

```
STANDARDIZED SOLUTION:

 X1     =V1   =    .938 F1     + .347 E1
 X2     =V2   =    .523*F1     + .852 E2
 X3     =V3   =    .761*F1     + .648 E3
 X5     =V5   =    .831 F2     + .557 E5
 X6     =V6   =    .548*F2     + .837 E6
```

Fuente: EQS 6. con  datos y adaptación propia

Analizando los indicadores de ajuste vemos que la *Chi-Cuadrado* ha disminuido sustantivamente y que los errores son menores (**RMSEA** = 0.075). Pero en lo que nos interesa para analizar la *validez convergente*, todos los *coeficientes de regresión son significativos, con lo que esta validez puede afirmarse.*

## Validez discriminante (discriminant validity)

La validez discriminante requiere que una medición no se correlacione demasiado con mediciones de las que se supone debe diferir (Churchill y Iacobucci, 2002, p. 413). En términos de desarrollo y validación de mediciones, la evidencia de validez convergente típicamente se proporciona a partir de las correlaciones entre la nueva medición que se está desarrollando y las mediciones existentes. Obviamente, los problemas para generar evidencia de validez convergente a menudo ocurren en el desarrollo de mediciones cuando no hay mediciones alternativas disponibles.

Estas instancias se encuentran frecuentemente en constructos que no han sido estudiados previamente o que han sido investigadas usando métodos de operacionalización *ad hoc* o desarrollados de manera inadecuadaZ.

*Cuando las escalas previamente validadas no están disponibles para proporcionar la variedad necesaria en la medición, a menudo se emplean medidas de otros campos, como la psicología* (Richins, 1983).

Entre los métodos empleados más frecuentemente para investigar la validez convergente, así como la validez discriminante, se encuentra la *matriz multitrait-multimétodo* (**MTMM**. *MultiTrait-MultiMethod Matrix*) propuesta por Campbell y Fiske (1959). Como se resume en Cook y Campbell (1979, p. 61), como centro de la evaluación de la validez del constructo hay dos procesos:

1. *Probar la convergencia* entre diferentes mediciones del mismo concepto y,
2. *Probar la divergencia* entre mediciones relacionadas pero conceptualmente distintas en conceptos.

La matriz **MTMM** típica muestra una serie de coeficientes de correlación de Pearson entre mediciones de constructos no observables. Casos de esto, son:

1. Richins (1983, p. 77) ofreció un ejemplo de *comportamiento del consumidor* en el que el análisis incluía 2 rasgos (es decir, *defensa* y *agresión del consumidor*, como términos mercadotécnicos) y 2 métodos (es decir, las escalas tipo Likert desarrolladas y diferentes escalas que evalúan la *defensa* y la *agresión* como términos de la psicología).
2. Nunnally y Bernstein (1994, p.93) ofrecieron otro ejemplo, con 2 rasgos: *ansiedad* y *depresión* con 2 métodos: los auto-informes de *ansiedad* y *depresión,* con puntuaciones de observación obtenidas de expertos externos. **Ver Gráfico 4.1.**

**Gráfico 4.1. Matriz Multitrait Multi-Método**

| | | Method 1—Likert Scale | | | Method 2—Thermometer Scale | | |
|---|---|---|---|---|---|---|---|
| | | Job Satisfaction | Role Conflict | Role Ambiguity | Job Satisfaction | Role Conflict | Role Ambiguity |
| Method 1—Likert Scale | Job Satisfaction | 1 | | | | | |
| | Role Conflict | .896 | 2 | | | | |
| | Role Ambiguity | -.236 | .670 | .817 | | | |
| Method 2—Thermometer Scale | Job Satisfaction | -.356 | .075 | | 3 | | |
| | Role Conflict | .450 | -.082 | -.054 | 4 | 2 | |
| | Role Ambiguity | -.244 | .395 | .142 | 4 | -.147 | |
| | | -.252 | .141 | .464 | -.170 | .289 | |

Fuente: adaptación de *A Paradigm for Developing Better Measures of Marketing Constructs*, *Journal of Marketing Research*, *16*(February), p. 17, Churchill, copyright © 1979 by the American Marketing Association.

Como se muestra en el **Gráfico 4.1.** y se describe a continuación, Churchill (1979, p.70-71) describió un ejemplo de *gestión de ventas* que incluía tres rasgos (*satisfacción laboral, conflicto de roles y ambigüedad de roles)* y dos métodos (Likert y termómetro escalas).

El *patrón de correlaciones* dentro de la matriz (es decir, la fuerza relativa y la importancia de las correlaciones incluidas en la matriz) se interpretan para evidencia de validez convergente y discriminante. (Kenny y Kashay [1992] y Widaman [1985] describieron los procedimientos de uso del análisis factorial confirmatorio para analizar datos *multitrait-multimétodo*).

De este segundo caso, en el **Gráfico 4.1**, las estimaciones de confiabilidad de consistencia interna se representan en diagonal 1, en la esquina superior izquierda de la tabla.

Las correlaciones incluidas en la *diagonal 3*, en la esquina inferior izquierda, proporcionan evidencia de *validez convergente*. En este ejemplo, *las 3 correlaciones en la diagonal 3* ofrecen soporte para la *validez convergente*, ya que las estimaciones son significativas y grandes en relación con otras correlaciones en la matriz.

La evidencia de la *validez discriminante* es ofrecida por las correlaciones en el *bloque heterométodo 4* en la esquina inferior izquierda (es decir, las correlaciones entre las mediciones de constructo para las cuales se emplean diferentes

enfoques de medición) y los 2 triángulos *heterotrait-monométodo* en las esquinas superior izquierda e inferior derecha (es decir, las correlaciones entre las medidas de constructo para las cuales se emplea un enfoque de medición único). La evidencia de la *validez discriminante implica 3 comparaciones* (Churchill, 1979, p. 71):

**a.** Las entradas en la *diagonal de validez 3 deben ser mayores que las entradas* en el bloque de *heterométodos 4* que comparten la misma fila y columna.

**b.** *Las correlaciones en la validez diagonal deben ser mayores que las relaciones en los triángulos de heterotrait-monométodo.* Este requisito más estricto sugiere que las correlaciones entre diferentes mediciones para un rasgo deben ser más altas que las correlaciones entre rasgos que tienen métodos en común.

**c.** *El patrón de las correlaciones culturales debe ser el mismo en todos los triángulos de contorno* (es decir, 2 y 4).

En el ejemplo, se cumplen las dos primeras condiciones para la *validez discriminante.*

## Estudio de caso: administración del conocimiento OECD (2003)

Por otro lado, es posible afirmar que la validez discriminante *(discriminant validity*) se consigue cuando, si diversos instrumentos de medida están diseñados para medir distintas variables latentes, entonces las correlaciones entre dichos instrumentos son bajas. *Una escala tiene, pues, validez discriminante cuando no mide un constructo para el que no se diseñó.* Así, volviendo al caso de la **Figura 3.3.** Diagrama de trayectoria constructo administración del conocimiento de la OECD (2003**),** por ejemplo, si la *variable latente* **F1** COMUNICACIÓN y la **F2** INCENTIVOS, estuvieran fuertemente correlacionadas entre ellas, podría ocurrir que sirvieran para medir el mismo concepto (un caso extremo de una correlación 1 entre ambas). Vea **Tabla 4.8**

**Tabla 4.8. Estimación de covarinzas del modelo corregido**

```
COVARIANCES AMONG INDEPENDENT VARIABLES
 ---------------------------------------
 STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.

                V                              F
               ---                            ---
                              I F2  -   F2                .293*I
                              I F1  -   F1                .032 I
                              I                           .883 I
                              I                                I
```

Fuente: EQS 6. con datos y adaptación propia

Donde se muestran las estimaciones de los coeficientes de las trayectorias (*paths*) entre los distintos factores, vemos que la relación entre el factor **F1** y **F2** no es muy

alta (coeficiente 0.293). Podemos plantearnos si los indicadores de **$F_1$** servirían o no para medir **F2**, con lo que la validez discriminante de las dos escalas podría cuestionarse. Plantearemos tres procedimientos para analizar este hecho:

- *Test de la diferencia entre las Chi-Cuadrado*
- *Test del intervalo de confianza*
- *Test de la varianza extraída*

*Test de la diferencia entre las Chi-Cuadrado*

Se procede del siguiente modo:

**a.** Estima mediante un análisis factorial confirmatorio el modelo de medida donde todas las variables latentes pueden covariar (ya lo hemos hecho).

**b.** Crea un nuevo modelo de medida idéntico al anterior, pero donde el coeficiente entre los dos factores cuya validez discriminante suscita dudas, es fijado a 1 (es decir, serían el mismo factor, con lo que sus ítems serían perfectamente intercambiables y no habría validez discriminante).

**c.** Se calcula un test para la diferencia entre las Chi cuadrado de los dos modelos. La

**d.** validez discriminante quedará demostrada sólo si la Chi cuadrado es significativamente menor (mejor ajuste del modelo) para el primer modelo en el cual los dos constructos son vistos como factores distintos, aunque correlacionados (Anderson y Gerbing, 1988).

Dado que el primer modelo ya ha sido estimado, y su Chi cuadrado aparece en el **Tabla 4.5**, modificaremos la sintaxis de **EQS** de ese modelo retocando el subcomando **COVARIANCES**, como se indica a continuación, para forzar que la covarianza entre los factores **F1** y **F2** sea **1** y estimamos con ella el modelo. Ver **Tabla 4.9**.

**Tabla 4.9. Bondad de ajuste del modelo alternativo**

```
GOODNESS OF FIT SUMMARY FOR METHOD = ML

INDEPENDENCE MODEL CHI-SQUARE        =      66.051 ON    10 DEGREES OF FREEDOM

INDEPENDENCE AIC =      46.051   INDEPENDENCE CAIC =      16.931
       MODEL AIC =      -1.479          MODEL CAIC =     -16.039

CHI-SQUARE =         8.521 BASED ON        5 DEGREES OF FREEDOM
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS       .12976

THE NORMAL THEORY RLS CHI-SQUARE FOR THIS ML SOLUTION IS          7.718.

FIT INDICES
-----------
BENTLER-BONETT      NORMED FIT INDEX =       .871
BENTLER-BONETT NON-NORMED FIT INDEX =       .874
COMPARATIVE FIT INDEX (CFI)          =       .937
BOLLEN'S           (IFI) FIT INDEX   =       .942
MCDONALD'S         (MFI) FIT INDEX   =       .965
JORESKOG-SORBOM'S  GFI   FIT INDEX   =       .937
JORESKOG-SORBOM'S AGFI   FIT INDEX   =       .811
ROOT MEAN-SQUARE RESIDUAL (RMR)      =       .342
STANDARDIZED RMR                     =       .162
ROOT MEAN-SQUARE ERROR OF APPROXIMATION (RMSEA)     =       .120
90% CONFIDENCE INTERVAL OF RMSEA  (         .000,           .251)
```

Fuente: EQS 6. con adaptación propia

**/COVARIANCES**

**F1, F2=1;**

Pues bien, la Chi cuadrado del primer modelo era 5.103 con 4 grados de libertad, mientras que la del segundo modelo resulta ser 8.521 con 5 grados de libertad. La diferencia entre ambos estadísticos y sus grados de libertad es:

*Chi cuadrado: 3.418 grados libertad:* 1

Si buscamos en una tabla de la distribución *Chi-Cuadrado* vemos que, para un grado de libertad, el valor crítico es 2.706 para p = 0.10, 3.84 para p = 0.05; 6.635 para p = 0.01 y 10.827 para p = 0.001**.** Por lo tanto, dado que la diferencia es de 3.418, la *diferencia entre los dos modelos es claramente significativa*. En otras palabras, el modelo de medida estándar en el que los factores están correlacionados, pero son retenidos como distintos es significativamente mejor que el que los considera un único factor. *Ello afirmaría la validez discriminante de las escalas utilizadas para medir* **$F_1$ y $F_2$.** *El proceso debería repetirse para cada par de posibles covarianzas.*

*Test del intervalo de confianza*

Anderson y Gerbing (1988) proponen también el llamado este test, para analizar la validez discriminante entre dos escalas. Este test *implica calcular un intervalo de confianza de ±2 errores estándar entre la correlación entre los factores y determinar si este intervalo incluye al 1.0. Si no incluye al 1.0, la validez discriminante quedará confirmada.*

De la **Tabla 4.8**, se observa que el valor del coeficiente de correlación entre **$F_1$ y $F_2$** es de 0.293 y su error estándar de 0.032. De esta forma el intervalo de confianza tendría estos extremos superior e inferior:

**-***Intervalo Inferior:* 0.293 – 2 x 0.032 = 0.229

**-***Intervalo Superior:* 0.293 + 2 x 0.032 = 0.357

El intervalo queda, consecuentemente, [0.229 – 0.357], *no entra el 1*, lo que *confirma la validez discriminante* de las dos escalas según el criterio expuesto. En conclusión, los dos criterios expuestos confirman la validez discriminante de las escalas analizadas.

## *Test de la varianza extraída (IVE/AVE)*

Un tercer procedimiento para establecer la validez discriminante es aplicar este test (Fornell y Larcker, 1981). Según este procedimiento, calcularíamos el *índice de varianza compuesta extraída (***IVE/AVE***.Average Variance Extracted)* por cada uno de los factores cuya validez discriminante se está calculando (tal y como explicamos al tratar la fiabilidad). Este **IVE** se compara con el cuadrado de las correlaciones entre los dos factores. Podemos afirmar la validez discriminante si los **IVE** de los dos factores son mayores que el cuadrado de la correlación.

En nuestro ejemplo la correlación entre **$F_1$** y **$F_2$** es de 0.293, cuyo cuadrado es 0.085. Vea la **Tabla 4.10.**

**Tabla 4.10. Recopilación de información que calcula el IVE**

| Constructo e indicadores | Cuadrado de la Carga Factorial Estandarizada $(L_{ij})^2$ | Varianza del Término de Error Var $(E_{ij})= 1 - (L_{ij})^2$ | Índice de Varianza Compuesta Extraída IVE |
|---|---|---|---|
| **COMUNICACIÓN** | | | |
| **$X_1=V_1$** | .880 | .120 | **.578** |
| **$X_2=V_2$** | .274 | .726 | |
| **$X_3=V_3$** | .579 | .421 | |
| **Total** | **1.733** | **1.267** | |
| **INCENTIVOS** | | | |
| **$X_5=V_5$** | .691 | .309 | **.496** |
| **$X_6=V_6$** | .300 | .568 | |
| **Total** | **.991** | **1.009** | |

Fuente: EQS 6. con datos y adaptación propia

De acuerdo con este criterio, como el índice de varianza extraída de $\mathbf{F_2}$ es superior al cuadrado de la correlación (.085), *podríamos afirmar la validez discriminante. El resultado, sin embargo, está tan en el límite que la dirección en que apuntan los otros dos indicadores ofrecidos permitiría afirmar esta validez sin demasiado riesgo de error (0.496).*

## *Validez del grupo conocido (known-group validity)*

Esta validez, se refiere a grado en que una medida difiere según lo previsto entre los grupos, con un nivel de bajo y/o alto puntaje en un rasgo. La evidencia de apoyo de la validez de grupo conocido es típicamente proporcionada por diferencias significativas en las puntuaciones medias a través de muestras independientes. Tian et al. (2001, pp. 56-57) aportaron varios ejemplos interesantes para validar su escala de *necesidades únicas de consumidores* (**CNFU.** *Consumers Need for Uniqueness*). Específicamente, la puntuación media de un indicador para una muestra heterogénea de encuestas por correo para su escala **CNFU** de 31 ítems fue de 2.60 (n = 621).
En la **Tabla 4.11** se presentan tres muestras de comparación de grupos conocidos, el tipo de diferenciación de consumo que representó el grupo y los resultados de cada muestra.

**Tabla 4.11. Resumen caso *validez del grupo conocido (known-group validity)***

| Known-Group Validity Test | Sample | Unique Group | | | Comparison Group | | | Value of *t* | Value of *p* | Validity Support |
|---|---|---|---|---|---|---|---|---|---|---|
| | | *n* | *Mean* | *SD* | *n* | *Mean* | *SD* | | | |
| Tattoo and body piercing artists | 1 | 39 | 3.05 | .70 | 621 | 2.60 | .56 | 3.22 | <.001 | Supported |
| Owners of customized low rider autos | 2 | 22 | 2.99 | .45 | 621 | 2.60 | .56 | 3.22 | <.001 | Supported |
| Members of medievalist reenactment group | 3 | 21 | 2.91 | .44 | 621 | 2.60 | .56 | 2.49 | <.01 | Supported |
| Student art majors | 4 | 22 | 3.06 | .45 | 273 | 2.71 | .50 | 3.15 | <.01 | Supported |
| Student purchasers of unique poster art | 5 | 78 | 2.83 | .43 | 273 | 2.71 | .50 | 1.89 | <.05 | Supported |

SOURCE: Adapted from "Consumers' Need for Uniqueness: Scale Development and Validation," *Journal of Consumer Research*, 28(June), p. 58, Tian, Bearden, and Hunter, copyright 2001 by the University of Chicago Press.

Fuente: Netemeyer et al. (2003)

Las muestras de comparación de grupos conocidos fueron las siguientes: *artistas del tatuaje (tatoo artists)*, que reflejan el dominio de las muestras corporales de bienes de consumo (M = 3.05, n = 39, t = 3.22, p <0,01); *propietarios de mini-camiones* (*owners of customized low rider*), reflejando la posesión de objetos únicos (M = 2.99, n = 22, t = 3.22, p <0,01); y miembros de la *Sociedad de Anacronismo Creativo (Society for*

*Creative Anachronism),* que representan la participación en grupos únicos que se visten de manera diferente y promulgan actuaciones medievales (M = 2,91, n = 21, t = 2.49, p <0.01).
Otro ejemplo interesante de la validez de los grupos conocidos fue el presentado por Saxe y Weitz (1982). Al validar su escala de *orientación del agente de ventas* (**SOCO.** *Salesperson Customer Orientation***),** en el que encontraron diferencias significativas entre las muestras de posiciones de ventas que difirieron ampliamente en su estatus profesional.

# Validez nomológica (*nomological validity*)

Un medio para evaluar la validez de constructo de una medición, viene de determinar hasta qué punto la medida encaja *lícitamente* en una red de relaciones o *una red nomológica* (Cronbach y Meehl, 1955). Es decir, un aspecto de la validez del constructo implica el nivel en que una medición opera dentro de un conjunto de constructos teóricos y sus respectivas medidas (Nunnally y Bernstein, 1994, p. 91).
La validez nomológica se basa en investigaciones de constructos y mediciones en términos de *hipótesis formales derivadas de la teoría.* Como tal, la *validación nomológica es principalmente externa e implica investigar tanto las relaciones teóricas entre constructos diferentes como las relaciones empíricas entre las medidas de esos constructos***.** Así, las mediciones deben demostrar validez nomológica para ser aceptadas como constructos válidas (Peter, 1981, p.135).
La evidencia de *validez nomológica* es proporcionada por la posesión de un constructo con distintos antecedentes de causa, efectos y / o condiciones de modificación, así como diferencias cuantitativas en el grado en que un constructo está relacionado con antecedentes o consecuencias (Iacobucci, Ostrom, Grayson, 1995, Tian et al., 2001, p. 59). Ciertamente, la evidencia proporcionada por los métodos descritos en las secciones anteriores de este capítulo pueden tener implicaciones para *validez nomológica.* Por ejemplo, las diferencias en las puntuaciones para una medición de la fuerza de ventas al consumidor a través de grupos de vendedores ineficaces y efectivos proporcionan algunas implicaciones para las teorías de la venta efectiva (es decir, validez nomológica), así como de validez del grupo conocido. Así, los modelos de ecuación estructural (**SEM.** *Structural Equation Modeling*), métodos basados en regresión y/o experimentales se usan frecuentemente en esfuerzos para proporcionar evidencia corroborativa de validez nomológica (*Bollen, 1989; Hoyle, 1995*). Como ejemplo, Bearden et al. (2001) emplearon la regresión logística para demostrar que un subconjunto de sus mediciones de *confianza del consumidor* moderó (como planteó de hipótesis) la relación entre el esquema de calidad de los precios al consumidor y la elección de opciones más altas o más bajas en un conjunto de opciones. Lastovicka et al. (1999) utilizaron una regresión múltiple para demostrar que su medida de *frugalidad* explicaba un índice de comportamientos *frugales* más allá de explicaciones alternativas para comportamiento frugal.
Como otro ejemplo, Tian (2001) describió un gran número de pruebas de diferencias en las relaciones que implican una serie de antecedentes de rasgos (por ejemplo, *individualismo colectivo*), resultados (por ejemplo, elección de diseños exteriores

únicos vs. comunes) y moderadores situacionales por ejemplo, la popularización de diseños únicos) para medir la necesidad de unicidad de los consumidores.

## *Estudio de caso: administración del conocimiento OECD(2003)*

Finalmente, una escala tiene validez nomológica (*nomological validity*) cuando el constructo que miden es capaz de sacar a la luz relaciones con otros constructos que, conceptual y teóricamente deberían existir. Nótese que para contrastar esta validez *ya no nos basta con tener una escala que mida una variable latente, sino diversas escalas con relaciones teóricas entre ellas***.** Por este motivo el ejemplo que desarrollaremos a continuación envuelve a diversas escalas con estas relaciones teóricas. Si los constructos cuya relación se comprueba se han medido al mismo tiempo que la escala, hablamos de validez concurrente (*concurrent validity*)*,* si se han medido en distintos momentos del tiempo, hablamos de *validez predictiva (predictive validity).*

Hemos de señalar que existen muchos programas estadísticos para abordar el análisis factorial confirmatorio y el análisis de sistemas de ecuaciones estructurales necesarios para el contraste de la validez. Quizás el más conocido sea **LISREL**, desarrollado por Jöreskog y Sörbom (1993). **SPSS**, por su parte, incorpora en sus últimas versiones el programa **AMOS**.

Al final, la elección entre uno y otro depende del gusto particular del investigador. Nosotros, atendiendo a estos gustos, solemos recurrir a **EQS** (Bentler y Wu, 1993) aunque, puestos a encontrar un software óptimo, recomendaríamos la utilización del programa **CALIS** de **SAS** (SAS Institute Inc, 1989; Hatcher, 1994). La razón de esta recomendación es que permite el planteamiento del análisis utilizando la notación más cómoda para el investigador. Admite tanto la notación de Bentler (1980), como la de Jöreskog (1993). En cualquier caso, cualquiera de estos programas tiene la precisión suficiente para efectuar un buen análisis. En nuestro ejemplo hemos utilizado la versión 6 de **EQS**.

El modelo de medida, como hemos visto, es un modelo que satura las posibles relaciones entre las variables latentes estableciendo covarianzas entre todas ellas. El modelo teórico, es mucho más parsimonioso, es decir, espera llegar a los mismos resultados con muchas menos relaciones. El modelo teórico tendrá validez nomológica (*nomological validity*)si, efectivamente, no hay diferencias significativas entre los ajustes del modelo de medida y el teórico, dado que las escalas habrán sido capaces de establecer relaciones predictivas de otras variables tan sustantivas que, siendo menos, igualan la bondad del modelo.

Las T**ablas 4.12, y 4.13,** nos muestran los estadísticos de bondad de ajuste que son necesarios para determinar la validez nomológica. Es posible comprobar que las respectivas *Chi-Cuadrado del modelo teórico y de medida* (que obtuvimos al analizar la validez discriminante) son:

## Tabla 4.12. Bondad de ajuste del modelo teórico inicial

```
GOODNESS OF FIT SUMMARY FOR METHOD = ML

INDEPENDENCE MODEL CHI-SQUARE        =     110.190 ON    15 DEGREES OF FREEDOM

INDEPENDENCE AIC =      80.190   INDEPENDENCE CAIC =      36.509
       MODEL AIC =      -4.301          MODEL CAIC =     -27.597

CHI-SQUARE =        11.699 BASED ON        8 DEGREES OF FREEDOM
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS       .16515

THE NORMAL THEORY RLS CHI-SQUARE FOR THIS ML SOLUTION IS         10.841.

FIT INDICES
-----------
BENTLER-BONETT      NORMED FIT INDEX =       .894
BENTLER-BONETT NON-NORMED FIT INDEX =       .927
COMPARATIVE FIT INDEX (CFI)          =       .961
BOLLEN'S            (IFI) FIT INDEX  =       .964
MCDONALD'S          (MFI) FIT INDEX  =       .964
JORESKOG-SORBOM'S  GFI  FIT INDEX    =       .931
JORESKOG-SORBOM'S AGFI  FIT INDEX    =       .820
ROOT MEAN-SQUARE RESIDUAL (RMR)      =       .165
STANDARDIZED RMR                     =       .081
ROOT MEAN-SQUARE ERROR OF APPROXIMATION (RMSEA)     =       .097
90% CONFIDENCE INTERVAL OF RMSEA  (        .000,          .206)
```

Fuente: EQS 6. con datos y adaptación propia

## Tabla 4.13. Bondad de ajuste del modelo teórico final

```
GOODNESS OF FIT SUMMARY FOR METHOD = ML

INDEPENDENCE MODEL CHI-SQUARE        =      66.051 ON    10 DEGREES OF FREEDOM

INDEPENDENCE AIC =      46.051   INDEPENDENCE CAIC =      16.931
       MODEL AIC =      -2.897          MODEL CAIC =     -14.545

CHI-SQUARE =         5.103 BASED ON        4 DEGREES OF FREEDOM
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS       .27691

THE NORMAL THEORY RLS CHI-SQUARE FOR THIS ML SOLUTION IS          4.949.

FIT INDICES
-----------
BENTLER-BONETT      NORMED FIT INDEX =       .923
BENTLER-BONETT NON-NORMED FIT INDEX =       .951
COMPARATIVE FIT INDEX (CFI)          =       .980
BOLLEN'S            (IFI) FIT INDEX  =       .982
MCDONALD'S          (MFI) FIT INDEX  =       .989
JORESKOG-SORBOM'S  GFI  FIT INDEX    =       .961
JORESKOG-SORBOM'S AGFI  FIT INDEX    =       .854
ROOT MEAN-SQUARE RESIDUAL (RMR)      =       .093
STANDARDIZED RMR                     =       .050
ROOT MEAN-SQUARE ERROR OF APPROXIMATION (RMSEA)     =       .075
90% CONFIDENCE INTERVAL OF RMSEA  (        .000,          .237)
```

Fuente: EQS 6. con datos y adaptación propia

## *Interpretación*

La **Tabla 4.13** nos muestra los estadísticos de bondad de ajuste que son necesarios para determinar la validez nomológica. Se puede comprobar que las respectivas Chi-Cuadrado de los modelos teórico y de medida (que obtuvimos al analizar la validez discriminante *(discriminant validity*) son:

*Teórico: 11.699;*
*Grados de libertad: 8;*
*Medida: 5.103;*
*Grados de libertad: 4*

La validez nomológica (*nomological validity)* se establecería calculando un test de diferencias entre las Chi cuadrado, tal como hicimos para la validez discriminante *(discriminant validity)*:

*Diferencias entre las Chi-Cuadrado: 6.596;*
*Diferencias en grados de libertad: 4;*
*Valor crítico p<0.01: 13.277*

En otras palabras, el modelo teórico ofrece un ajuste significativamente mejor que el de medida (diferencia inferior al valor crítico), lo que permitiría constatar la validez nomológica de las escalas.

# Ejemplo: respuesta socialmente deseable (*socially desirable response bias*)

La respuesta socialmente deseable (**SDR.** *Socially Desirable Responding*) es un tema complejo que los psicólogos han discutido durante años. En este apartado, se plantea la cuestión de reiterar su importancia y recordar a los lectores que la **SDR** justifica la consideración en investigación, en particular cuando el potencial de sesgo de respuesta afecta las relaciones entre constructos. En pocas palabras, el **SDR** puede ser visto como un estilo de respuesta o sesgo que refleja tendencias para proporcionar respuestas favorables con respecto a normas y prácticas (Nederhof, 1985). Mick (1996) definió la respuesta socialmente deseable como la tendencia de los individuos a hacerse ver bien con respecto a las normas culturales al responder a las preguntas de los investigadores**.** Como se discute a continuación, este aspecto del **SDR** es consistente con el concepto de *gestión de impresiones* (*impression management*) de Paulhus (1993), que resalta los intentos de los encuestados de dar forma a sus respuestas para reflejar una imagen más positiva. **SDR** puede afectar la medición de constructos, así como las relaciones entre ellos (Mick, 1996, pp. 109-110). Es decir, el **SDR** puede aumentar las relaciones de tal manera que las correlaciones entre los constructos se deban a la variación compartida de los **SDR**. Este fenómeno se denomina *efecto espuria* . En el *efecto de supresión*, la correlación verdadera entre dos medidas está enmascarada por **SDR**. Un tercer efecto posible de **SDR** ocurre cuando la forma de la relación entre dos variables medidas se afecta.

En estas últimas situaciones, **SDR** modera la relación entre los constructos. Mick (1996) y Ganster, Hennessey y Luthans (1983) describen los procedimientos para investigar estos problemas alternativos asociados con el sesgo de respuesta. Los métodos para hacer frente al *sesgo* **SDR** fueron resumidos recientemente por Tian, Bearden y Manning (2002). Los procedimientos se encuadran en *dos categorías: métodos diseñados para evitar que los participantes en la encuesta respondan de una manera socialmente deseable y métodos para detectar y medir el sesgo de respuesta a la deseabilidad social* (Nederhof, 1985; Paulhus, 1991).

Las técnicas empleadas por los investigadores para prevenir o disminuir la respuesta socialmente deseable en la investigación por encuestas, incluyen el *uso de preguntas neutrales, ítems de elección forzada, la técnica de respuesta aleatoria, cuestionamiento indirecto y autoadministración del cuestionario.*

•Las preguntas neutras, los elementos de elección forzada y la técnica de respuesta aleatoria *no han producido pruebas convincentes* de su efectividad para reducir la respuesta socialmente deseable en la investigación general en ciencias sociales (Nederhof, 1985).

•*Los artículos de elección forzada*, para los cuales se hace que los encuestados escojan entre dos ítems que sean aproximadamente equivalentes en su deseabilidad social pero que estén relacionados con diferentes temas, *son problemáticos* porque algunos individuos pueden percibir todavía una diferencia en la deseabilidad de los ítems (Nederhof, 1985).

•*Los encuestados pueden ser antagonizados* por el uso de técnicas de respuesta al azar, por lo que los encuestados al responder a uno de los dos elementos seleccionados al azar, con el *investigador no saben qué tema se respondió*.

•*El cuestionamiento indirecto***,** una técnica proyectiva en la que se pide a los encuestados que respondan a preguntas estructuradas desde la perspectiva de otra persona o grupo, ha sido demostrado empíricamente como una técnica eficaz para reducir el sesgo de la respuesta de deseabilidad social (Fisher, 1993).

•*La autoadministración en las encuestas por correo y las administraciones de masas anónimas* tienden a disminuir la respuesta socialmente deseable en las investigaciones que involucran la autodescripción de los individuos (Nederhof, 1985; Paulhus, 1984). Sin embargo, la autoadministración y las técnicas relacionadas que proporcionan anonimato reducen sólo la respuesta socialmente deseable que proviene de la *gestión de impresiones*, no reducen el sesgo auto-engañoso (Nederhof, 1985).

La segunda categoría de métodos implica el uso de medidas directas de respuesta socialmente deseable para evaluar las tendencias para que los individuos se presenten favorablemente con respecto a normas y estándares. La medición de la respuesta socialmente deseable permite evaluar hasta qué punto la tendencia a informar excesivamente los comportamientos deseables o subnotificar los comportamientos indeseables, confunde la evaluación precisa de las variables de contenido y suprime, infla o modera las relaciones variables (Ganster et al., 1983; , Zerbe y Paulhus, 1987). Además, las medidas directas permiten investigar la respuesta socialmente deseable como una constructo psicológico por derecho propio (véase Mick, 1996, y Paulhus, 1991). Aunque se han desarrollado una variedad de medidas de respuesta socialmente deseables, éstas varían en su capacidad para

capturar de manera independiente las dimensiones de la *gestión de impresiones* y el *autoengaño* de las respuestas socialmente deseables y en su aplicabilidad al estudio de varias poblaciones de encuestados (Paulhus, 1991).

# Conclusión

Este capítulo trató cuestiones relacionadas con la validez de constructo, aplicada al desarrollo de medidas para conceptos no observables. En general, la validez de constructo *representa la medida en que una medida operacional refleja verdaderamente el concepto que se está investigando o la medida en que las variables operacionales utilizadas para observar la covariación en y entre los constructos se puede interpretar en términos teóricas. La validez de constructo de una medida no se evalúa directamente sino que se deduce* de la evidencia de que las puntuaciones sustantivas de la medición que se comporten como se esperaba, con la evidencia procedente de pruebas teóricamente derivadas y de la calidad de los procedimientos empleados en el desarrollo y validación de la medida. Los diversos tipos o fuentes de evidencia sobre la validez de las medidas destinadas a evaluar la validez de constructo discutidas en el capítulo fueron las siguientes.:

- *Validez de contenido* (*content validity*): Grado en que los elementos de un instrumento de medición son relevantes y representativos del constructo objetivo para un propósito de evaluación particular. Las garantías de validez de contenido se basan en la generación teórica a priori, la generación de artículos y el juicio. Validez de la cara: una evaluación de que los ítems de una escala miden adecuadamente el constructo.
- *Validez aparente* (*face validity*): Se puede juzgar después de que se haya desarrollado una medición, a menudo, antes de la aplicación en otro estudio, por los usuarios potenciales de la medición.
- *Validez predictiva* (*predictive validity*): Capacidad de una medida para predecir con eficacia algún criterio subsiguiente y temporalmente ordenado.
- *Validez concurrente* (*concurrent validity*): Para la cual la evidencia es proporcionada por las correlaciones considerables entre la medida del constructo bajo desarrollo y una medida del criterio recogida simultáneamente o *concomitantemente*.
- *Validez convergente* (*convergent validity*): La medición en que las medidas independientes del mismo constructo convergen o están altamente correlacionadas.
- *Validez discriminante* (*dicriminante validity*): El nivel en que las mediciones divergen de otras operacionalizaciones a partir de las cuales el constructo es conceptualmente distinto.
- *Validez del grupo conocido* (*known-group validity*): la medida en que una medición difiere según lo previsto entre los grupos que deben puntuar bajo y alto en un rasgo.
- *Validez nomológica* (*nomological validity*): El grado en que la medida encaja *legalmente* en una red de relaciones o una *red nomológica;* es decir, el grado en que una medida opera dentro de un conjunto de constructos teóricos y sus respectivas medidas. Además de estas fuentes de evidencia o tipos de validez específicos, el capítulo concluyó con una discusión de los problemas asociados con la respuesta socialmente deseable, así como una descripción de los métodos diseñados para evitar que los participantes en la encuesta respondan de una manera sesgada y

métodos dirigidos a detectar y medir el sesgo de la respuesta de deseabilidad social (Nederhof, 1985; Paulhus, 1991). Ver **Tabla 4.14**

**Tabla 4.14. Resumen de conceptos de tipos de validez**

| Tipos de validez | Concepto |
|---|---|
| **De Contenido** | Se aplica para estimar el grado con el que una escala es representativa del concepto que mide, o por el grado en que los ítems de la escala son representativos de la población de ítems que desarrollan un concepto.<br>Por su naturaleza cualitativa, no se puede expresar mediante un coeficiente; tan sólo aproximaciones procedentes de juicios. |
| **De constructo o de concepto** | Se aplica porque una escala no es un conjunto de cuestiones o ítems inconexos, sino un instrumento diseñado para medir algo concreto. Pretende responder a la pregunta: qué concepto mide la escala.<br>Los procedimientos básicos para probarla son las matrices multiconcepto - multimétodo y el análisis factorial confirmatorio. |
| **Convergente** | Se refiere a que dos o más escalas que tratan de medir el mismo concepto deben converger hacia una misma puntuación.<br>La comparación se ha de realizar en el mismo lapso temporal. |
| **Discriminante** | Se refiere al hecho de que una escala sólo cuantifica las características del concepto que trata de medir y no cualquier otra característica de cualquier otro concepto.<br>Se mide mediante la correlación: si las puntuaciones de dos escalas que miden el mismo concepto han de correlacionarse (validez convergente), las escalas usadas para medir conceptos distintos no deben hacerlo. |
| **Nomológica** | Trata de determinar si el instrumento de medida se comporta según lo esperado con respecto a otras construcciones con las cuales está teóricamente relacionada.<br>Se evalúa mediante la confirmación de relaciones significativas entre construcciones según establezca la teoría. |
| **Factorial** | Se refiere al hecho de que la estructura de relaciones entre ítems y sus correspondientes conceptos o factores se mantiene en sucesivos análisis.<br>Requiere la existencia de validez convergente y discriminante. |
| **Predictiva, pragmática o relacionada con criterios** | Se refiere al hecho de que una escala de medida sea capaz de predecir otras variables (criterios);éstas pueden ser cualquier tipo de variable obtenida con otras escalas y que pueden ser estimadas con la escala bajo estudio. |
| **Predictiva simple** | Se refiere a la capacidad de la escala para pronosticar el valor futuro de otras variables (por ejemplo, intención de compra). Los datos de la escala bajo estudio y de los criterios se recogen en momentos diferentes del tiempo. Para obtenerla, es necesario tenerla también de constructo: si algo no se puede medir con precisión, difícilmente se podrán hacer predicciones válidas. |

| | |
|---|---|
| **Concurrente** | Se refiere a la capacidad de la escala para predecir otras variables o criterios, contrastándose simultáneamente los datos recogidos del concepto bajo análisis y de la variable criterio. Normalmente, la evaluación de la validez concurrente se lleva a cabo incluyendo las medidas de las variables criterio en el mismo cuestionario donde se recogen los datos de la escala bajo estudio, por lo tanto, en el mismo momento del tiempo. |

Fuente: Netemeyer et al. (2003)

# CAPÍTULO 5. DEFINIENDO EL CONSTRUCTO Y SU CONTENIDO (Etapas: 1 y 2)

Los Capítulos 2, 3 y 4 discuten los principios psicométricos de la dimensionalidad, la confiabilidad y la validez principalmente desde un punto de vista conceptual, así como también ofrecen ejemplos de algunas escalas desarrolladas más recientemente. Los Capítulos 5, 6 y 7 ahora ofrecerán aplicaciones de estos principios desde un punto de vista empírico. Estos capítulos cubrirán los pasos y procedimientos recomendados en el desarrollo de la escala y la validación.

# Etapa 1: El Constructo

La primera parte de este capítulo analiza el paso crítico de la definición de constructo y la delimitación del dominio de contenido del mismo. Se resaltan así, los siguientes problemas:

**a.** El papel de la teoría en la definición del constructo y el dominio de su contenido,
**b.** La importancia de una revisión exhaustiva de la literatura y la definición a juzgar por expertos e individuos de poblaciones relevantes.
**c.** El enfoque en los indicadores de efectos (*reflexivos*), y
**d.** La importancia de una dimensionalidad, *a priori,* de constructo

## La importancia de la definición de un constructo claro y preciso

La importancia de un constructo bien definido no puede exagerarse, ya que la validez de lo que se mide se basará en gran parte en su definición y dominio de contenido**.** *Definir claramente el constructo* -sus facetas y dominios- es un primer paso esencial que *muchos consideran el paso más difícil en el proceso de escalamiento* (ver Churchill, 1979; Haynes et al., 1999; Haynes et al., 1995; Nunnally y Bernstein , 1994). Al definir los constructos, se debe tener cuidado en lo que se incluye y excluye en el dominio del constructo ya que *las fuentes de invalidez* pueden tener sus orígenes en el proceso de definición del mismo. Una medición puede ser demasiado estrecha y no incluir facetas importantes. Esto se ha referido como insuficiente representación del constructo (Messick, 1993). Si el dominio está *demasiado definido, se pueden incluir factores extraños o dominios de constructos*

*que no sean el objetivo.* Aunque estos dominios extraños pueden estar altamente correlacionados con los dominios relevantes del constructo, crean lo que se ha denominado como *constructo de varianza irrelevante* (*construct-irrelevant variance*), que a su vez puede crear confusiones entre la constructo objetivo y sus variables de predictor y criterios (Neuberg et al. , 1997). La razón detrás del dominio de una constructo es que la interpretación de cualquier medición es más clara cuando el dominio es claro, o sea , inconfundible. Cuando se incluyen factores extraños, posiblemente se sugiera que más de un rasgo/o variable de diferencia individual está siendo evaluada, por lo que la correlación de ese constructo con un criterio puede ser confundida. En conclusión, cuando se incluyen factores extraños o dominios de otras constructos, más de una constructo subyace al puntaje total y la validez del constructo está amenazada. Por lo tanto, se considera ampliamente que el factor clave para el aspecto del contenido de la validez del constructo es la clara especificación de los límites del dominio del constructo que va a ser evaluado.

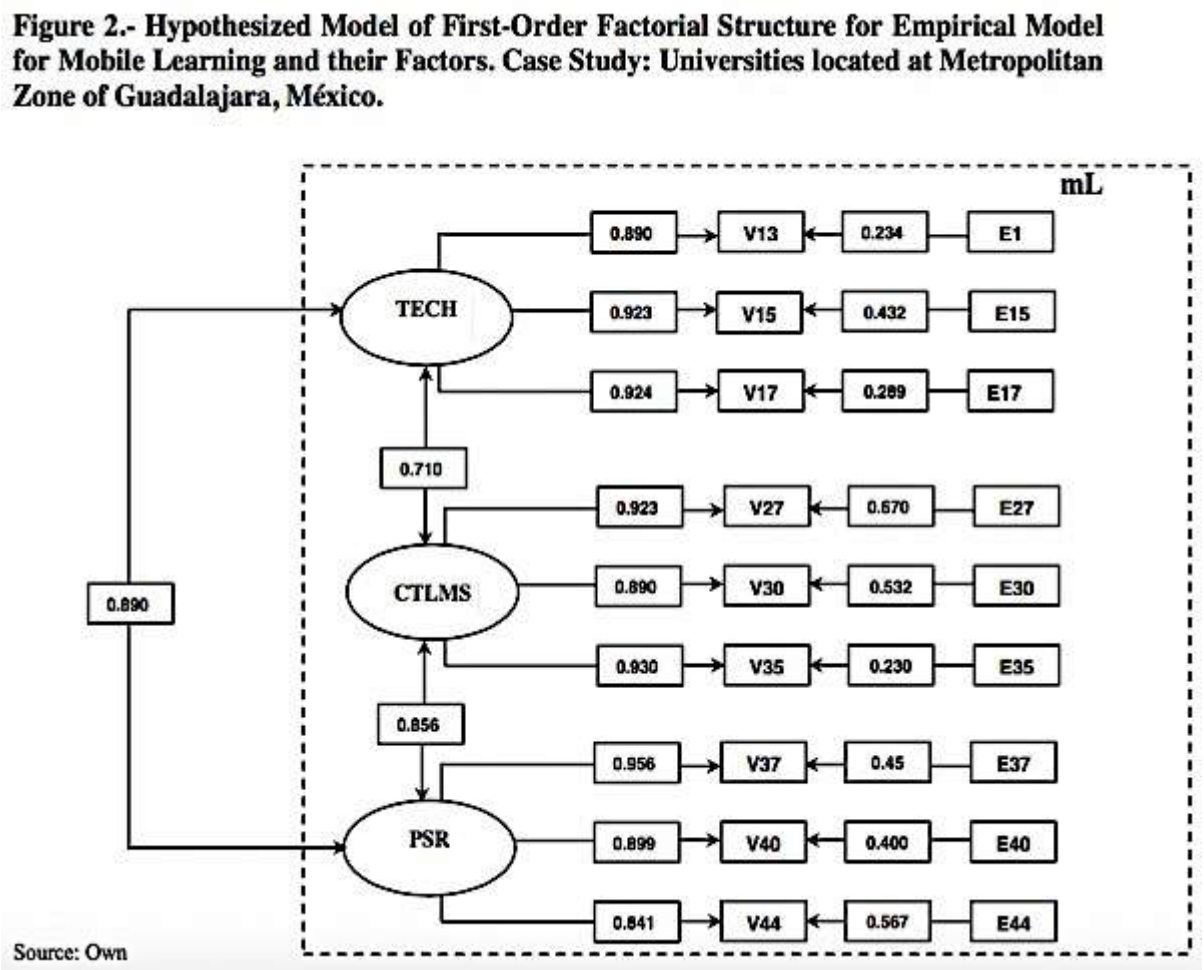
# La importancia de una teoría sólida, revisión de la literatura y su análisis

En sus trabajos clásicos sobre medición y validez, Cronbach y Meehl (1955) y Loevinger (1957) declararon la importancia de la teoría en la medición. Para que las medidas de los constructos latentes tengan relevancia en las ciencias económico-administrativas, los constructos deben basarse en un marco teórico. Incluso los constructos altamente compactos basados en la teoría, son más útiles como antecedentes o consecuencias de otros comportamientos latentes cuando se insertan en la teoría. Como tal, la relevancia de un constructo latente para las ciencias económico-administrativas, depende en parte de las teorías en las que está formulada: *¿qué es lo que hace que el constructo subyacente haga predicciones?; ¿qué predice el constructo subyacente?* Esto se le conoce como *la red nomológica de un constructo latente.*

Un buen ejemplo es el mostrado por Mejìa-Trejo et al. (2015), donde al desarrollar su medida de aprendizaje móvil (*mobile learning*), basado en rasgos (atributos o variables) como la tecnología (**TECH**), la administración de contenidos y estilos de enseñanza aprendizaje (**CTLMS.***Contents Teaching-Learning Management and Styles*); y el rol del profesor y estudiante (**PSR.***Professor and Student Role*).

En éstas de postularon dimensiones e indicadores sumando 60, algunas de ellas como: la *facilidad del uso de la tecnología, las redes sociales, la percepción profesor-estudiante del valor-costo, las políticas de uso, etc.* Lo anterior constituye una red de constructos que conforman al modelo, que es fundamental para guiar el desarrollo de la escala y evaluar la validez. **Ver Figura 5.1.**

**Figura 5.1. Ejemplo de una red nomológica**




Fuente: Mejía-Trejo et al. (2015)

En esencia*, NO puede haber validez de constructo de una medición, sin una teoría bien especificada.* Esta teoría bien pensada comienza con la conceptualización/definición del constructo basada en una revisión exhaustiva de la literatura. Tal revisión de la literatura idealmente sirve para varios propósitos importantes (ver apartado *importancia de la revisión de la literatura del estado del arte en el marco teórico*) ya que:

1. *Una revisión de la literatura debe alertar al investigador sobre las teorías en lasque el constructo puede resultar útil como una variable independiente o dependiente*. Un manejo más preciso de lo que es el constructo, sus límites, dimensiones y dominio de contenido, se puede descubrir a través de una revisión de la literatura. Una revisión de la literatura también puede *revelar intentos previos para medir el constructo y las fortalezas y debilidades de tales intentos.* Los problemas con los intentos previos de medir el constructo se pueden descubrir y, por lo tanto, evitarse en el propio esfuerzo.
2. Una revisión exhaustiva puede *descubrir importantes constructos relacionados y mediciones* que se usarán en la validación convergente y/o discriminante.
3. Finalmente, dado que el desarrollo y la validación de la escala es una tarea que

lleva tiempo y, en ocasiones, es costosa, una revisión exhaustiva de la literatura debe ayudar a responder la siguiente pregunta:

*¿Se necesita una escala?* Si ya existen buenas medidas de un constructo, el valor de una nueva medición puede ser pequeño en relación con los costos involucrados en su desarrollo. *Una nueva medida, para ser útil, debe mostrar alguna ventaja teórica o empírica sobre una medida existente del mismo constructo. Para que una nueva escala tenga validez incremental sobre las mediciones existentes, debe capturar el constructo específico con más precisión o de manera más eficiente* (por ejemplo, es más corto, más barato, más fácil de usar o más fácil de responder) que las medidas existentes (Clark y Watson, 1995; Haynes et al., 1999). Otra forma útil de mejorar la precisión y la exhaustividad de la definición del constructo es la revisión por parte de expertos. Aunque tales revisiones tradicionalmente se han asociado más con la evaluación de los ítems*, pueden ayudar a delimitar las facetas del constructo antes de la generación de indicadores.*

# El enfoque de indicadores reflectivos (*de efecto*) vs. indicadores formativos (*de causa*)

Como se dijo en el Capítulo 1, este texto se centra en las medidas que se consideran indicadores de efecto (*reflexivos o reflectives*) de los constructos más que indicadores formativos (*causales o formatives*). Para los indicadores de efecto, las puntuaciones sobre los propios indicadores, son *hipotéticos* para ser manejados por el constructo latente. Es decir, las diferencias individuales en el constructo cuentan para las diferencias individuales en la aprobación del indicador. Con los indicadores formativos, las diferencias individuales en los indicadores cuentan para las diferencias individuales en la aprobación del constructo. La siguiente discusión, permitirá entenderlo mejor:

El constructo *estrés de la vida en el último año* puede ser interpretado como un constructo de *indicador formativo,* en el cual los indicadores son diferentes fuentes de estrés, incluyendo un cambio de trabajo, la muerte de un ser querido, el nacimiento de un niño y una enfermedad. Es contraintuitivo que los cambios en el estrés de la vida produzcan la muerte de un ser querido o el nacimiento de un niño; Lo opuesto tiene sentido (Smith y McCarthy, 1995). Asimismo, como segundo ejemplo, el estatus socioeconómico (**SES**) se conceptualiza mejor como un constructo de indicador formativo. Las personas tienen SES alto debido a su riqueza o educación; Ellos no se vuelven ricos ni educados como resultado del alto **SES** (Nunnally y Bernstein, 1994). Por lo tanto, los indicadores *forman* el puntaje del constructo. Otras dos diferencias importantes entre el efecto (*reflectivos o reflectives*) y los indicadores formativos (*causales o formatives*) se encuentran tanto en las áreas conceptuales como en las metodológicas, ya que:

-Para los indicadores que se conceptualizan para formar la constructo latente, se cree que la lista de indicadores debe ser completa. Es decir, con indicadores formativos, se necesita un censo de indicadores, no solo una muestra, para medir completamente el constructo. La exclusión de cualquier indicador cambia la composición de la constructo latente y por lo tanto, cada indicador es importante para

la validez del constructo (Bollen y Lennox, 1991; MacCallum y Browne, 1993; Neuberg et al., 1997). Este no es necesariamente el caso con los indicadores de efecto (*reflectivos* o *reflectives*). Con éstos, los indicadores deben representar una muestra razonable de elementos que tocan el dominio del constructo (Nunnally y Bernstein, 1994).

-El hecho de que los indicadores de un constructo formativo (*causal o formatives)* se combinen para producir un índice general no implica necesariamente que todos los puntajes del indicador individual estén intercorrelacionados, y si están o no correlacionados es irrelevante para la confiabilidad de la medida (Bollen y Lennox, 1991; Smith y McCarthy, 1995). Los indicadores formativos (*causal o formatives*) no necesitan ser consistentes internamente, por lo que los métodos de confiabilidad basados en la consistencia interna no se aplican.

-Con los indicadores de efecto (*reflectivos o reflectives*), la interrelación entre ítems y, por lo tanto, la consistencia interna, es motivo de preocupación para la confiabilidad de la medida. No solo existen diferencias conceptuales entre los indicadores formativos (*causal o formatives*) y los de efecto (*reflectivos o reflectives*), también difieren los métodos utilizados para desarrollar tales mediciones. (Para una excelente revisión de estos métodos, vea Diamontopoulus y Winklhofer, 2001.) Aquí nuevamente, la importancia de una definición de constructo bien definida y pensada es muy útil. La teoría y una revisión exhaustiva de la literatura pueden ayudar a determinar si el constructo es una medida de indicador formativa (*causal o formatives*) o de efecto (*reflectivos o reflectives*).

# El papel de la teoría al especificar la dimensionalidad

Como se señaló en los Capítulos 1 y 2, *la dimensionalidad se refiere a la homogeneidad de los ítems o indicadores***.** La unidimensionalidad puede definirse como la existencia de un rasgo, atributo, variable o constructo latente subyacente a un conjunto de indicadores/medidas (Hattie, 1985). Dado que el dominio de contenido de un constructo se puede hipotetizar como unidimensional o multidimensional, su estructura empírica debe reflejar su dimensionalidad hipotética. Para un solo constructo unidimensional, cada elemento/indicador se refleja por su constructo latente, o factor de primer orden. Dicho conjunto de elementos también se ha denominado una *medición congénere (congeneric measure)*, es decir, una por la que una sola dimensión subyace al conjunto de indicadores**.**

Cuando un constructo está compuesto por múltiples facetas que están relacionadas, pero distintas, *el constructo puede clasificarse como multidimensional.*

La escala de confianza del consumidor de seis factores de Bearden et al. (2001) (*the six factor consumer self-confidence scale*) representa una escala multidimensional. Por el contrario, Lastovicka et al. (1999) La escala de la frugalidad del consumidor (*the consumer frugality scale*), las de aprendizaje móvil e innovación de Mejía-Trejo et al. 2015 así como conocimiento del consumidor e innovación de Mejía-Trejo et al., 2016 (*mobile learning and innovation*; *customer knowledge management and innovation*) son unidimensionales.

Con *constructos multidimensionales*, cada dimensión se considera un *factor de primer orden* representado por una variable latente separada, donde cada indicador se usa para operacionalizar su respectiva dimensión hipotética. Dentro de sus respectivas dimensiones, los indicadores deben mostrar evidencia de unidimensionalidad (es decir, cualidades *congéneres o congeneric qualities*).
Se dice que un conjunto de medidas es *congénere* si cada medida en el conjunto pretende evaluar el mismo constructo, excepto los errores de medición. (Jöreskog, 1993). En la medida en que varios factores de primer orden incluyan una variación común, un factor de orden superior (también llamado factor de segundo orden) podría ser conceptualizado para capturar la varianza compartida a través de los factores de primer orden. Los factores de primer orden se consideran indicadores de efecto (*reflectivos o reflectives)* o dimensiones de constructo de orden superior, y el constructo de mayor orden cuenta para la covarianza (correlación) entre los factores de primer orden. Un constructo de orden superior sugiere que las dimensiones (factores de primer orden) miden el mismo concepto jerárquico, a excepción del error aleatorio y la especificidad de medición (Bagozzi y Heatherton, 1994). La importancia de especificar teóricamente y probar empíricamente la dimensionalidad antes de otras propiedades psicométricas (por ejemplo, consistencia interna y validez nomológica) no debería subestimarse (Cortina, 1993; Gerbing y Anderson, 1988; Hattie, 1985). Para operacionalizar los constructos latentes, los investigadores a menudo usan puntajes compuestos, sumando o promediando entre elementos diseñados para medir el constructo de interés. El cálculo y el uso de tales puntuaciones son significativos sólo si los ítems tienen una unidimensionalidad aceptable. Como se señaló en el Capítulo 3, las escalas multidimensionales, cuando se usan como si fueran unidimensional (es decir, compuestos de elementos sumados o promedios), pueden dar lugar a ambigüedades interpretativas de las relaciones entre constructos en una prueba de teoría. Si un constructo es multidimensional pero todos los puntajes de los ítems se suman/promedian entre dimensiones en un solo puntaje compuesto y se correlacionan con una cierta variable criterio, dicha correlación es ambigua (Neuberg et al., 1997). Aquí nuevamente, el papel de la teoría y una revisión exhaustiva de la literatura deberían ayudar a especificar a *priori* la dimensionalidad del constructo de interés.

# Etapa 2: Generando y Analizando los Indicadores de Contenido

La segunda parte de este capítulo trata sobre el segundo paso en el desarrollo de escala: determinar y analizar un conjunto inicial de indicadores que soporten al constructo. Con este paso, cubriremos brevemente los supuestos teóricos del muestreo de dominios. También se discutirán temas en la generación de un grupo de indicadores, opciones de redacción de preguntas/declaraciones, formatos de elección de respuestas y evaluación de los indicadores de contenido y validez de aparente (*face validity*) por parte de expertos y encuestados potenciales de poblaciones relevantes. La última parte de este capítulo usa ejemplos que ilustran y aplican muchos de los procedimientos recomendados para los dos primeros pasos en el desarrollo de la escala.

# La muestra y su contenido

Una vez que el constructo se ha definido y delimitado con precisión, comienza la tarea de *generar indicadores del dominio del constructo. Como modelo de medición de error,* se sugiere que una *medición se componga de una muestra de indicadores de un gran dominio hipotético de elementos***.** Existe una gran agrupación (*pool*) de indicadores o elementos que podrían aprovechar el dominio de la constructo (y/o) para llegar a una medición de escala final, por lo que debe seleccionarse una muestra de elementos de este dominio con *propiedades psicométricas deseables*,que:

**-***En teoría, este grupo se extrae de lo que algunos han denominado un universo o pool de indicadores, pudiendo basarse en expertos*.

-*En la práctica, los indicadores para un grupo inicial se extraen de la literatura existente o generados por los autores de la escala.*

Los *expertos* del tema, también pueden contribuir con indicadores y variables,. Siempre que el grupo de indicadores inicial sea lo suficientemente grande en relación con el dominio de contenido del constructo, los principios de muestreo del mismo todavía son aplicables (Nunnally y Bernstein, 1994).

Por otro lado, *la suficiencia de contenido* existe cuando:

-El contenido de la medición es una muestra aleatoria representativa del dominio a medir,

-Los indicadores finales elegidos (por ejemplo, la muestra de ítems) son considerados representativos del constructo, y cuando

-Tienen las mismas medias y varianzas que los de la totalidad del conjunto del *pool* de indicadores.

Como se señaló en el Capítulo 4, la validez del contenido *(content validity)* es relevante para el muestreo de dominios porque representa parcialmente el grado en que uno puede generalizar de un conjunto determinado de indicadores en el dominio del constructo.

*El error de medición considerado con el muestreo de dominio, entonces, es el error producido al usar una muestra de los indicadores en lugar de todo el dominio de ellos.*

# Generando el *pool* de indicadores

Al generar un conjunto de indicadores, un objetivo importante es muestrear sistemáticamente todas las áreas de contenido del constructo. Se deben tener en cuenta varias cuestiones relevantes para la validez de versión (*translation validity*), ya que:

**1.** Los indicadores de la escala generados deben aprovechar el dominio de contenido del constructo y mostrar la validez del contenido (*content validity*). Como se discutió en el Capítulo 4, este término se ha definido de muchas maneras, y la mayoría de las definiciones enfatizan que los indicadores de una medición son una muestra apropiada del dominio teórico del constructo (Nunnally y Bernstein, 1994; Peter, 1981). La mayoría de las definiciones son consistentes en que la validez del contenido (*content validity*) *refleja el grado en que los elementos de un instrumento de medición son relevantes y representativos del*

*constructo específico para un propósito particular de evaluación* (Haynes et al., 1995, p. 238). Los elementos se refieren al *contenido individual de los indicadores, formatos e instrucciones de respuesta para los encuestados.* La *representatividad* se refiere al grado en que los elementos son proporcionales a las facetas (dominios) del constructo objetivo y al grado en que todo el dominio del constructo ha sido muestreado. Es decir, los elementos deberán aparecer coherentes con el dominio teórico del constructo en todos los aspectos, incluídos los formatos de respuesta y las instrucciones.

**2.** En la generación de indicadores, también se debe considerar la validez aparente (*face validity*). Una escala de alta validez aparente mejora la cooperación de los encuestados debido a su facilidad de uso, nivel de lectura adecuado y claridad, así como a sus instrucciones y formatos de respuesta. Por lo tanto, desde una perspectiva práctica, la validez aparente (*face validity*) puede estar más referida con lo que los entrevistados de poblaciones relevantes infieren con respecto a lo que se está midiendo, y la validez de contenido (*content validity*) es referida con la validez aparente (*face validity*) así como a lo que el investigador cree que está construyendo (Haynes et al., 1995; Nunnally y Bernstein, 1994).

**3.** Incluso, enfocados en ambas validaciones, aparente (*face validity*) y de contenido (*content validity*), se deben considerar otros dos aspectos en la construcción del *pool* de indicadores. Clark y Watson (1995) abogaron por que el desarrollador de la escala vaya más allá de su propia visión al generar una agrupación de indicadores inicial que contenga, en última instancia, los que sólo estarán relacionados tangencialmente. Por lo tanto, es mejor incluir indicadores en exceso al dominio del constructo en lugar de hacerlo en carestía o cortedad. También se debe tener cuidado para garantizar que cada área de contenido del constructo tenga una muestra adecuada de indicadores**.** Aunque es difícil de lograr en la práctica, las áreas de contenido más amplias deben estar representadas por un conjunto de ítems más grande. Con estos problemas en mente, la generación de indicadores puede comenzar con una reflexión cuidadosa sobre:

**a.***Cuál debería ser la fuente de los dichos indicadores,*

**b.***Problemas de redacción de su redacción, y*

**c.***Cuántos artículos deben servir como grupo inicial*.

Para estos problemas, las *respuestas definitivas no existen***,** pero algunas pautas prácticas son evidentes.

# Fuentes de indicadores

En este término, se sugiere que los desarrolladores de escala vayan más allá de su vista del constructo objetivo y consulten varias fuentes. Cabe destacar, que muchas escalas existentes tienen ítems con validez de contenido (*content validity*) y validez aparente (*face validity*) que aprovechan uno de los dominios del constructo.

Observar y analizar cómo los estudios previos han operacionalizado el constructo (o constructos relacionados), puede ser una fuente valiosa en la generación de indicadores. Por ejemplo, para medir el *conflicto trabajo-familia* (**WFC.***Work-Familily Conflict*) y el *conflicto trabajo-familia* (**FWC**.*Family-Work Conflict*) de Netemeyer et al. (1996) muestra la selección de indicadores de las mediciones existentes de **WFC** así

como mediciones de *sobrecarga de roles de trabajo (Work Role Overload)*, y Bearden et al. (2001) convirtieron indicadores de varios estudios de *confianza/experiencia del consumidor* (*Consumer Confidence/Expertise*) para su *pool* inicial de indicadores en evaluar las dimensiones de *autoconfianza del consumidor (Consumer Self-Confidence)*

Otra fuente de generación de ítems es la población de interés (Haynes et al., 1999; Haynes et al., 1995). Aquellos que pueden estar respondiendo a la escala (es decir, miembros de la población) pueden ofrecer hallazgos de lo que podría ser el constructo y cómo medirlo, mejorando así la validez aparente (*face validity*).

Las entrevistas estructuradas y no estructuradas con miembros de la población pueden proporcionar información sobre la redacción de los ítems y los formatos de respuesta. Varias escalas desarrolladas en la literatura de administración de la mercadotecnia, han utilizado este enfoque tanto para ayudar a definir el constructo como para generar ítems (por ejemplo, Bearden et al., 2001; Lastovicka et al., 1999, Mejía-Trejo et al., 2016).

Dos fuentes más de indicadores son: los expertos en el campo el desarrollador de escalas. Los académicos, investigadores, profesionales (consultores) familiarizados con las constructos objetivo, pueden ser excelentes fuentes para indicadores. Los expertos en el campo pueden sugerir, escribir o modificar indicadores coherentes. Obviamente, el desarrollador de la escala también es una fuente de indicadores. Esto será particularmente cierto para constructos nuevos o para las que se dispone de muy poca literatura existente. Incluso con otras fuentes que contribuyen al *pool* inicial de indicadores, la mayoría de los desarrolladores de escala tendrán que escribir varios de sus propios elementos para generar un grupo inicial adecuado.

# Escritura de los indicadores

Una segunda consideración en la generación de indicadores implica la escritura de los mismos. Se deben considerar varias cuestiones:

**a.***La claridad de la redacción,*
**b.***La redacción redundante,*
**c.***La redacción en sentido positivo o negativo, y*
**d.***La elección de formatos de respuesta.*

## *Claridad de la redacción*

La claridad del indicador es sumamente importante. De hecho, Nunnally y Bernstein (1994) sugieren que la regla de oro de la escritura del indicador se resuma en una sola palabra: *claridad.* Un buen indicador debe ser inequívoco hasta el punto de que todos los encuestados comprendan su significado de la misma manera. Se aplican algunas reglas simples, como

**a.***Use un lenguaje común* al nivel de lectura de la población objetivo.
**b.***Mantenga el indicador corto y simple* independientemente del nivel de lectura de la población de interés. Aunque no se quiere sacrificar el significado de un elemento por razones de brevedad, la longitud y la complejidad tenderán a disminuir la claridad de los elementos.

**c.***Evite la jerga o el argot de moda.* El significado de dicho idioma está fechado rápidamente o está sujeto a una interpretación diferente a medida que pasa el tiempo.
**d.***Evite las declaraciones de doble contenido* que, de hecho, abordan dos problemas (o áreas de contenido) en una declaración.
**e.***Escriba indicadores que todos se adhieran ya que de lo contario, serán de poca utilidad.* Tales indicadores producen una pequeña variación de ítems y, por lo tanto, una pequeña variación de escala. Como lo señalaron Clark y Watson (1995), indicadores que todos respaldan de la misma manera de manera positiva (por ejemplo, *a veces, soy más feliz que otras ocasiones)* o de la misma manera, en negativo (por ejemplo, *siempre estoy furioso*) agrega poco a la validez de contenido (*content validity*) de un constructo.

## La redacción redundante

En la fase de generación de indicadores, *se necesita cierta redundancia de redacción entre ellos, acorde al contexto que aborda.* Sin embargo, *el desarrollador de la escala debe ser consciente, sin embargo, de la redundancia que es inútil.*
Crear un nuevo indicador cambiando solo una palabra que no es crítica para el significado del mismo, al no ser importante para el contenido del indicador, refleja una redundancia inútil.
Por otro lado, variando la selección de palabras y la estructura gramatical para crear nuevos indicadores, de modo que el dominio de contenido de la constructo se esté aprovechando de manera diferente, refleja una redundancia útil.
DeVellis (1991, p. 56) ofrece un buen ejemplo para el aprovechamiento del *grado en que un padre ayudará a su hijo*, siendo los indicadores propuestos:
*Haré casi cualquier cosa para asegurar el éxito de mi hijo* y
*Ningún sacrificio es demasiado grande si ayuda a mi hijo a alcanzar el éxito.*

Tienen redundancia útil porque ambas declaraciones expresan el contenido del constructo de diferentes maneras. Sin embargo, al evaluar la actitud hacia los amantes de las mascotas, las declaraciones: *En mi opinión, los amantes de las mascotas son amables* y *Yo estimo, que los amantes de las mascotas son amables* comparten vocabulario y gramática comunes que no aprovechan el constructo, lo que refleja una redundancia inútil. Sin embargo, en general, en la fase de generación de indicadores del desarrollo de la escala, es necesaria cierta redundancia de los mismos, pero no debe reflejar diferencias puramente triviales de redacción/gramática. En la escala final, cuando la redacción de los indicadores es demasiado similar y refleja solo diferencias triviales, algunas propiedades psicométricas en realidad pueden estar comprometidas. Los indicadores que son redactados de manera muy similar aumentarán la media de correlación interítem, lo que de hecho aumenta la consistencia interna sin contribuir sustancialmente a la validez de contenido de la medida. Estos indicadores, también contribuyen a errores de medición correlacionados en el análisis factorial confirmatorio que, de hecho, amenazan la dimensionalidad de una medición (Floyd y Widaman, 1995). Aunque se necesita cierta similitud entre los indicadores de una escala para aprovechar el dominio, los que tengan ligeras variaciones producen escasa información relevante para el

constructo, que cualquier indicador individual (Clark y Watson, 1995). Dado lo anterior, los indicadores con redundancia inútil pueden contribuir a la *atenuación paradójica* (*attenuation paradox*) en la teoría psicométrica, por lo que el aumento de las estimaciones de consistencia interna más allá de un cierto punto, no mejorarán la validez.

## *La redacción en sentido positivo o negativo*

La elección de utilizar todos o algunos de los indicadores redactados positivamente o negativamente, también es de interés. Varios desarrolladores de escala han escrito indicadores que reflejan: niveles bajos de, lo contrario de, o la ausencia del constructo objetivo. El objetivo principal de un procedimiento de este tipo es también *mantener la honestidad del entrevistado* y así evitar el sesgo de respuesta en forma de complacencia, afirmación o tendencia decir sí a todo (*Acquiescence,Affirmation, Yea-saying*). La escala de *Necesidad de Cognición* (*Need for Cognition Scale*) de Cacioppo y Petty (1982) es un ejemplo. Ellos utilizaron indicadores que reflejan el nivel más alto del constructo, como:
*Prefiero los problemas complejos a los simples*, y los indicadores que reflejan un bajo nivel de constructo, tales como:
*Sólo pienso tan duro como tengo que hacerlo*. Sin embargo, y de acuerdo a la experiencia, se tiene que los indicadores redactados negativamente no muestran una confiabilidad tan alta como los elementos redactados positivamente o incluso, pueden ser confusos para los encuestados. Tales indicadores, también pueden contribuir a métodos de modelos analíticos de factor a factor, porque los indicadores redactados positivamente tienden a cargar altamente a un factor y los elementos redactados negativamente tienden a cargarse altamente en otro factor. (Véase Herche y Engellend [1996] para una discusión y ejemplos). Con esa posibilidad en mente, el investigador debe sopesar las ventajas y desventajas potenciales de usar indicadores redactados negativamente en el *pool* de indicadores.

## *La elección de formato de respuesta*

Con respecto a los formatos de respuesta, se deben considerar dos cuestiones básicas:
**a.** *Puntos de escala dicotómicos versus multicótomos, y*
**b.** *Redacción de los puntos de escala de respuesta.*
Aunque existen muchos tipos de formatos de respuesta, incluidas listas de verificación (*checklists*) y análogos visuales (*visual analog*), las dos formas dominantes son los formatos de escala dicotómicos y multicótomica. La mayoría de los formatos dicotómicos utilizan puntuación verdadero-falso o sí-no, y la mayoría de los usos multicotómicos usan tipo Likert o diferenciales semánticos (o alguna variación de) con escalas con tres o más puntos de escala. Dado el dominio de su uso, nuestra discusión se limita a estos dos tipos de escalas (para más información sobre el uso de listas de verificación (*Checklists*), elecciones forzadas (*Forced Choice*) y escalas análogas visuales (*Visual Analog Scales*), vea: DeVellis [1991]; Green, Goldman y Salovey [1993], y Clark y Watson [1995]).

Existen desventajas y ventajas tanto en los formatos dicotómicos como en los multicótomos. Los formatos dicotómicos han sido criticados por:
-Su tendencia a tener *distribuciones de respuesta altamente desequilibradas,* esto es, todos los individuos responden siempre a *verdaderos* o todos los individuos siempre responden a *falso* (Comrey, 1988)**.** Una evaluación cuidadosa de los indicadores durante la etapa de desarrollo de la escala, puede ayudar a eliminar tales ítems, y el formato dicotómico generalmente lleva menos tiempo para que el entrevistado responda, permitiendo que más ítems sean respondidos en un período de tiempo más corto.
-Su relación con la escala de varianza. Una propiedad de medición deseable de una escala es que tenga suficiente varianza en las puntuaciones de la escala global para una muestra dada. Dado que cualquier indicador para una escala dicotómica produce una covarianza limitada con otro indicador, debido al formato binario, la varianza de la escala global será limitada. Para una escala compuesta de indicadores dicotómicos, necesitará un mayor número de ítems para producir una variación de escala similar a la de ítems que usan escalas de tipo Likert y diferenciales semánticas.
Las escalas tipo Likert y los diferenciales semánticos pueden usar una serie de opciones y formatos de respuesta. Las escalas tipo Likert generalmente piden a los encuestados que indiquen su nivel de acuerdo con una declaración declarativa (*diferentes niveles desde totalmente en desacuerdo hasta totalmente de acuerdo en intervalos impares*); el grado en que lo que se expresa en la declaración es el *cierto de una creencia, actitud o característica del entrevistado* o la frecuencia de un comportamiento (es decir: desde nunca hasta siempre). Con las escalas de tipo Likert, también es posible etiquetar cada punto de escala así como los puntos finales de escala. Por ejemplo, para una escala de 5 puntos, cada punto de escala podría etiquetarse de la siguiente manera: *1 = muy en desacuerdo, 2 = algo en desacuerdo, 3 = no está de acuerdo ni en desacuerdo, 4 = está de acuerdo y 5 = muy de acuerdo.* Dicho etiquetado sirve para dar al encuestado una mejor idea del respaldo que está haciendo. Dependiendo de lo que se está evaluando, los puntos de escala pueden tomar muchas palabras diferentes. Por lo tanto, el investigador debe tener cuidado al elegir etiquetar puntos de escala o puntos finales (Clark y Watson, 1995; Comrey, 1988; DeVellis, 1991; Nunnally y Bernstein, 1994). Un indicador correcto elaborado en escala tipo Likert debe expresar la opinión, la actitud, la creencia, el comportamiento o la característica que se evalúa en términos claros y usar la redacción apropiada para los puntos de escala. Una evaluación cuidadosa de los indicadores puede ser muy útil a este respecto.
*Los indicadores diferenciales semánticos* utilizan puntos extremos basados en adjetivos que son de *naturaleza bipolar* (por ejemplo**,** amistoso-hostil, bueno-malo, frío-calor) o de *naturaleza unipolar* (por ejemplo, amigable -no amigable, bueno- no bueno, caliente, no caliente, en Osgood y Tannenbaum, 1955).
Aquí, el encuestado califica el estímulo (por ejemplo, objeto o persona) en una serie de escalas diferenciales semánticas que evalúan algún atributo latente del estímulo. Tanto para los indicadores de escala tipo Likert como para los diferenciales semánticos, los puntajes en los ítems que constituyen la escala se suman para formar un puntaje general.

Una ventaja principal de las *escalas multicotómicas* es que crean más variaciones de escala en relación con una escala dicotómica con un número similar de indicadores.
Sin embargo, con las escalas multicotómicas, se debe tener en cuenta la cantidad de puntos de escala y si se debe usar una cantidad impar o incluso de cuántos intervalos deban ser (3-5-7-9-11).La mayoría de las escalas multicotómicas usan entre 5 y 9 puntos de escala, con algunos tan bajos como 3 y tan altos como 11 puntos de escala. Lo más común es utilizar formatos de 5 o 7 puntos considerados como suficientes, ya que proporcionar más alternativas de respuesta puede no mejorar la confiabilidad o la validez de la escala. Si el investigador quiere proporcionar una *etiqueta* para cada escala, es más fácil y probablemente más significativo, tanto para el desarrollador de la escala como para el encuestado, si se usan formatos de 5 o 7 puntos. Más alternativas pueden requerir más esfuerzo en favor del demandado forzándolo a hacer distinciones más finas. Esto, a su vez, puede producir respuestas aleatorias y más variaciones en el error de escala.
Otra consideración es usar un número impar o incluso de puntos de escala. Con un número impar, al encuestado se le ofrece una escala de punto medio o una respuesta *neutral*. Dicha respuesta, en efecto, expresa una opción de *no opinión, no estoy seguro o ni de acuerdo ni en desacuerdo* dependiendo de lo que se evalúa. El desarrollador de la escala debe ser cuidadoso al elegir una redacción adecuada para dicho punto medio si elige etiquetar cada punto de la escala.
Según lo declarado por Clark y Watson (1995, p. 313), una escala de punto medio de *no puedo decir* puede confundir el punto medio de la escala con una incertidumbre de lo que el artículo significa para el encuestado. Por otro lado, un número par de puntos de escala obliga al encuestado a tener una opinión, o al menos a comprometerse débilmente con lo que se expresa en un ítem (es decir, expresar algún nivel que refleje una preferencia por el endoso o no endosando un indicador), que puede no tener. Aunque algunos investigadores creen que ni un número impar ni un par de puntos de escala son superiores, parecería que para algunos indicadores una respuesta neutral es una respuesta válida, de modo que una cantidad impar de puntos de escala es apropiada.
En general, se ha demostrado que si las escalas dicotómicas y multicotómicas se construyen de manera confiable y válida arrojarán resultados similares. Se recomienda fuertemente que los expertos evalúen cuidadosamente el formato, la redacción de los puntos de escala y el número de puntos de escala, y que *se pruebe el piloto antes de otros pasos de la construcción de escala.*

## *Número de indicadores del pool inicial*

Como se indicó anteriormente, es mejor tener un conjunto de indicadores que incluya en exceso en el dominio del constructo escaso o no inclusivo. Consistente con esta vista, se recomienda un gran conjunto de indicadores. Mucho dependerá del número y la complejidad de las facetas del dominio del constructo. Si éste se define de manera estrecha y la escala resultante es corta (5-10 indicadores), un *pool* tan pequeño como 20 a 30 indicadores puede se suficiente.
De hecho, DeVellis (1991) recomienda que para las constructos estrechos, bastará con una agrupación que sea *el doble del tamaño de la escala final*. Para constructos más amplios y polifacéticos, se necesitarán muchos más indicadores para servir

como grupo (*pool*) inicial. Algunos investigadores abogan por un grupo de *250 indicadores* como ejemplo para la generación de ítems en *constructos multifacéticos* (Robinson et al., 1991, pp. 12-13).
Aún así, uno debe considerar otros asuntos, tales como la *redundancia de indicadores, un nivel deseado de consistencia interna y cooperación de los encuestados.* Es posible que una agrupación (*pool*) sea tan grande que dificulte la cooperación de los encuestados . En conclusión, no hay reglas difíciles y rápidas para el tamaño de un grupo (*pool*) de indicadores inicial. Los constructos de una sola faceta estrechamente definidas requerirán menos ítems que las complejas construcciones multifacéticas. *Sin embargo, en general, se prefiere un número mayor, ya que la inclusión excesiva es más deseable que la falta de inclusión.*

## *Validez de contenido (content validity) y validez aparente (face validity)*

El establecimiento de la validez de contenido (*content validity*) y la validez aparente (*face validity*) *varía según la precisión del constructo y el grado en que los expertos concuerden con el dominio y las facetas del constructo.* Cuanto más amplia sea la discrepancia en cuanto al constructo, más difícil será la validación de contenido (*content validity*). Se dice que dicha *validez* se encuentra amenazada si:

**a.** Los indicadores que reflejan cualquiera de los dominios **(***facetas)* se omitieron de la medición,
**b.** Los indicadores que miden dominios *(facetas)* fuera de la definición del constructo se incluyen en la medición,
**c.** Un puntaje agregado en el constructo refleja de forma desproporcionada un dominio *(faceta)* sobre otros, y
**d.** El instrumento fue difícil de administrar y responder a las poblaciones objetivo.

## *Guía para validez de ambas*

Varios autores ofrecen pautas integrales para juzgar el contenido y la validez de los indicadores en un grupo (*pool*) inicial (Vea Clark y Watson, 1995; Haynes et al., 1999; Haynes et al., 1995; Nunnally y Bernstein, 1994). En el Capítulo 4, se abordaron estas pautas por lo que se agregan más, como:

**1.** El investigador debe tener todos los elementos de los indicadores evaluados con validez aparente (*face validity*) y validez de contenido (*content validity*). *Los ítems en sí, los formatos de respuesta, el número de puntos de escala y las instrucciones* para el encuestado deberían ser evaluados por la representatividad mediante la *intervención de expertos provenientes de la población de estudio* a través de procedimientos *cualitativos y cuantitativos***.** Algunos académicos recomiendan que las valoraciones de *categorización de 3 puntos* deben ser utilizadas por al menos *5 expertos* (para la validez del contenido-*content validity*) y *5 expertos de la población objetivo* (para la validez aparente-*face validity*) para evaluar el grado en que los ítems representan la definición de la constructo y los dominios (es decir, *no representativos***,** *algo representativo, claramente representativo*).

Algunos abogan por escalas de evaluación de 5 o 7 puntos que piden a 5 expertos o más que califiquen cada elemento en términos de *representatividad* (o, *idoneidad), especificidad y claridad*, y luego retienen ítems con altos niveles de acuerdo entre expertos (Haynes et al. , 1995).
*Como regla empírica práctica, se prefieren más expertos (5 o más), ya que la detección de indicadores deficientes o marginales será más confiable si se tiene en cuenta a más evaluadores expertos.* En este sentido, tanto los procedimientos cualitativos como los cuantitativos pueden ayudar a identificar indicadores que deban ser refinados, modificados o eliminados.
Los procedimientos *cualitativos* incluyen expertos que *escriben o verbalizan comentarios* sobre indicadores específicos, así como *grupos de enfoque* (*focus group*) y/o entrevistas personales con miembros de poblaciones relevantes para evaluar la validez aparente (*face validity*).
Los procedimientos cuantitativos incluyen estadísticas de categorización que evalúan el nivel de acuerdo entre los expertos (Perreault y Leigh, 1989).

**2.** Los resultados de las evaluaciones deben notificarse cuando se publiquen nuevos instrumentos. Esto permite al lector juzgar el grado en que los procedimientos fueron adecuados y tomar decisiones sobre la adecuación de la escala para su uso y/o ajustes necesarios al instrumento.

**3.** Finalmente, algunos *psicométricos* abogan por que una vez que una agrupación (*pool*) de ítems ha sido juzgada, modificada y/o recortada por expertos de la población objetivo, se esté probando un indicador piloto en una muestra más grande (n = 100-200) de una población relevante orden (Clark y Watson, 1995; Haynes et al., 1999).

Aquí, los elementos pueden evaluarse inicialmente para la consistencia interna, medias, varianzas, las intercorrelaciones con otros indicadores y la estructura del factor. Los indicadores que no se comporten empíricamente como se esperaba, pueden ajustarse para *redacción, etiquetas de escala, etc*., como recopilación de datos inicial. Algunos indicadores pueden requerir una *eliminación* basada en estos análisis, y si es necesario, se pueden redactar y juzgar más indicadores, de modo que se retenga un conjunto de ítems suficiente para los estudios de desarrollo de escala que siguen.

# Guía para el desarrollo de escalas

El último punto a tratar en el capítulo, es la propuesta de una serie de pasos que todo investigador debería seguir para desarrollar un instrumento de medida (escala) adecuado de una variable latente.

## *Determinar qué es lo que se pretende medir*

Aunque puede parecer qué todo investigador sabe qué es lo que quiere medir, muchas veces se encuentra con que sus ideas son más vagas de lo que creía a la hora de elaborar preguntas. La teoría es la mejor ayuda para esta fase, una escala sólo será correcta si todas las dimensiones del concepto que se quiere medir son conocidas por el investigador, y sólo el conocimiento profundo de la literatura puede

garantizar que esta fase se desarrolla correctamente.

## Generar un listado de ítems

Una vez que el propósito de la escala está claro, el investigador ha de comenzar a construirla. El primer paso es generar un amplio conjunto de ítems de entre los que deberán salir los que conformen definitivamente la escala.

Cada uno de estos ítems tiene que reflejar la variable subyacente que se pretende medir. Asimismo es mejor generar varios ítems para cada una de las dimensiones del constructo que recurrir a una sola pregunta, por cuanto que así puede calcularse su fiabilidad y, además, cuando se vayan sumando las respuestas aflorará el contenido común a ellos y se minimizará el efecto de las particularidades irrelevantes de cada uno por compensación. Así *haría casi cualquier cosa por asegurar la felicidad de mis hijos* y *ningún sacrificio es mucho si ayuda a conseguir la felicidad de mis hijos* podría ser una redundancia útil por cuanto expresan una idea similar de manera distinta.

En cuanto al *número de ítems,* es imposible dar una cifra de cuál debería ser este en el listado inicial. En todo caso debe recordarse que la fiabilidad de la escala es función de la correlación que exista entre los ítems. Como en esta fase no podemos calcular esas correlaciones, es deseable tener ítems suficientes para efectuar sustituciones que mejoren las correlaciones al desarrollar posteriormente la escala. No es extraño comenzar con un listado de ítems tres o cuatro veces superior a la longitud final que se espera dar a la escala.

Referente a *las características de un buen ítem,* es imposible hacer un listado exhaustivo de qué es lo que hace que un ítem sea bueno o malo. En todo caso, debería tenerse en cuenta lo siguiente:

1. *Deben evitarse los ítems excesivamente largos*, por cuanto la longitud incrementa la complejidad y disminuye la claridad. Sin embargo tampoco debe sacrificarse la claridad en aras a la brevedad eliminando preposiciones y conjunciones. En general un ítem del tipo: *Con frecuencia tengo problemas para expresar mis puntos de vista*' es mejor que: *Debo decir que una de las cosas con las que parece que tengo un problema la mayor parte de las veces es la transmisión de mi punto de vista al resto de las personas*.
2. *La complejidad sintáctica y de léxico* de las frases es también muy importante. *Deben evitarse la sucesión de negaciones No estoy a favor de que las empresas no sigan dando fondos a los grupos antinucleares es mucho más confuso que Estoy a favor del apoyo de las empresas a grupos antinucleares.*
3. *Deben evitarse también los ítems con doble argumentación.* Así *estoy en contra de la discriminación racial porque es un crimen contra Dios* es un ejemplo de doble argumentación. Si se está en contra de la discriminación racial por motivo que nada tienen que ver con la religión ¿qué debe contestarse?
4. *Combinación de ítems formulados en positivo y en negativo.* Muchos investigadores abogan por combinar ítems que suponen la presencia del constructo con ítems que suponen su ausencia. Por ejemplo la escala de Rosenberg (1965) de autoestima, incluye frases como: *creo que tengo bastantes buenas cualidades* junto a otras como *ciertamente, muchas veces me siento inútil.* El motivo de esta combinación es evitar el sesgo afirmativo del encuestado, es

decir, la tendencia manifiesta de estar de acuerdo con las afirmaciones independientemente de su contenido. La combinación permite detectar a las personas que han contestado con este sesgo, dado que habrá dado puntuaciones de acuerdo altas tanto a aquellas variables que indican un alto grado de autoestima, como las que la suponen baja. Desgraciadamente, este tipo de formulaciones tienen un aspecto negativo: *provocan confusión en los encuestados.*

## Determinar el formato de medida

En la literatura sobre las ciencias sociales existen muchas formas de hacer preguntas, nosotros presentaremos las más habituales (Aldás-Manzano y Maldonado-Guzmán, 2015).

## Escala de Thurstone

*Las escalas Thurstone* formulan las preguntas generando una serie de ítems que suponen la presencia en distintos grados del constructo que se pretende medir. Es también habitual que estén diseñadas de tal forma que la diferencia de nivel del constructo entre cada par de afirmaciones sea la misma. Una hipotética *escala Thurstone* para medir las aspiraciones de los padres respecto a los logros académicos de sus hijos, sería la mostrada en **Tabla 5.1.**

**Tabla 5.1. Escala de Thurstone. Un ejemplo de caso**

| | | |
|---|---|---|
| 1. Que mi hijo alcance el éxito es lo único que puede compensar mis esfuerzos como padre | Acuerdo<br>Desacuerdo | ☐<br>☐ |
| 2. Ir a una buena universidad y conseguir un buen trabajo es importante, pero no esencial, para la felicidad de mi hijo | Acuerdo<br>Desacuerdo | ☐<br>☐ |
| 3. La felicidad no tiene nada que ver con conseguir objetivos materiales o educacionales | Acuerdo<br>Desacuerdo | ☐<br>☐ |
| 4. Lo que habitualmente se cosidera un éxito es un obstáculo para la verdadera felicidad | Acuerdo<br>Desacuerdo | ☐<br>☐ |

Fuente: Aldás-Manzano y Maldonado-Guzmán (2015)

Como señalan Nunally y Bernstein (1994) es mucho más fácil explicar una *escala Thurstone* que construirla, por la dificultad que entraña generar ítems que aporten grados diferenciales de presencia del atributo.

## Escala de Guttman

*En una escala de Guttman,* los ítems están generados de tal forma que responder afirmativamente a uno de ellos, supone responder afirmativamente a todos los anteriores. Si se pregunta: ¿Fuma?, ¿Fuma usted más de 10 cigarrillos?, ¿Fuma

usted un paquete diario?. Responder afirmativamente a una de estas preguntas supone que también se posee el nivel inferior del atributo medido.
Aunque tanto la escala *Thurstone* con la *Guttman* están formadas por ítems que gradúan la presencia del atributo, *en la primera se busca una afirmación que fije el nivel de atributo poseído, mientras que la segunda se busca el punto de transición entre las respuestas afirmativas y las negativas*. Ver **Tabla 5.2.**

**Tabla 5.2. Escala de Guttman. Un ejemplo de caso**

| | | |
|---|---|---|
| 1. Que mi hijo alcance el éxito es lo único que puede compensar mis esfuerzos como padre | Acuerdo<br>Desacuerdo | ☐<br>☐ |
| 2. Ir a una buena universidad y conseguir un buen trabajo es importante para la felicidad de mi hijo | Acuerdo<br>Desacuerdo | ☐<br>☐ |
| 3. La felicidad es más probable si una persona ha conseguido sus objetivos educacionales y materiales | Acuerdo<br>Desacuerdo | ☐<br>☐ |
| 4. Lo que habitualmente se considera un éxito no es un obstáculo para la verdadera felicidad | Acuerdo<br>Desacuerdo | ☐<br>☐ |

Fuente: Aldás-Manzano y Maldonado-Guzmán (2015)

La dificultad de las *escalas Guttman* está en lograr formular los ítems de tal forma que responder afirmativamente a uno suponga hacerlo también a los anteriores. En el ejemplo anterior se supone que si se contesta afirmativamente a la expresión 2 se debe haber contestado igual a la 3 y 4. Sin embargo, si un entrevistado viera el éxito como un fenómeno complejo que puede ser a la vez una ayuda y un obstáculo para la felicidad, podría dar un patrón de respuestas no esperado.

## *Escalas con ítems del mismo peso (Likert y diferencial semántico)*

En los dos tipos de escalas mostradas, las dificultades de generación suelen superar sus ventajas y no son muy utilizadas. Lo más habitual es generar ítems que sean detectores equivalentes del fenómeno medido, y no de niveles de éste. Una de sus ventajas es que las respuestas pueden recogerse de muchas formas distintas, lo que permite al investigador buscar la más adecuadas para su propósito. Inconvenientes y ventajas, se tienen como:

**a.** *¿Cuántas categorías de respuesta?* La mayor parte de los ítems de una escala, consisten en dos partes: una afirmación y una serie de opciones de respuesta. Estas opciones pueden ir desde un gran número de alternativas (valorar de 1 a 100 un acuerdo, por ejemplo), a unas pocas (Muy de acuerdo, de acuerdo, indiferente, en moderado desacuerdo, muy en desacuerdo). No debemos olvidar que la misión de una escala es detectar la variabilidad en las respuestas, y fracasará en esta tarea si no puede discriminar entre los distintos niveles de presencia del atributo si las alternativas de respuesta son muy limitadas.
Una segunda cuestión, es si el entrevistado tiene capacidad para discriminar de manera significativa entre los distintos niveles de respuesta, lo que depende

claramente de qué se está midiendo. Si se le pregunta sobre su ideología política, puede tener necesidad de matizar su respuesta y requerirá de una escala de bastantes puntos (digamos 10). Pero si se le pregunta acerca de su acuerdo o desacuerdo con la afirmación *Fumar perjudica la salud* esta necesidad de matización no será tan acuciante y es razonable graduar la respuesta en menor número de alternativas. ¿Han de ser estas alternativas un número par o impar? De nuevo depende de lo que se pregunte. El número impar supone la existencia de un punto neutral. Si el investigador considera que, dadas las características de la pregunta, el entrevistado puede buscar conscientemente la indefinición, será recomendable un número par de alternativas para forzar que tome partido.

En otras preguntas, sin embargo, el encontrar el punto de indefinición puede ser muy útil. Veamos el ejemplo que plantea DeVellis (1991). En un estudio para determinar cuál de dos riesgos prefiere la gente asumir (aburrimiento frente a peligro), el investigador puede ir variando la peligrosidad de la actividad hasta dar con el punto de inflexión donde el entrevistado duda entre aburrimiento y peligro. Esa actividad sería un indicador de la propensión al riesgo del entrevistado. Vea la **Tabla 5.3.**

**Tabla 5.3. Escalas con ítems del mismo peso. Un ejemplo de caso**

Señale su preferencia relativa por las actividades A o B entre las alternativas señaladas a continuación.

Actividad A: Leer un libro de filosofía de la ciencia (ningún peligro).

Actividad B: Viajar en un vuelo comercial (poco peligro)

1 = Claramente prefiero A

2 = Prefiero A

3 = No prefiero una ni otra

4 = Prefiero B

5 = Claramente prefiero B

Actividad B: Viajar en una avioneta 1,2,3,4,5

Actividad B: Saltar en paracaídas, teniendo paracaídas de reserva 1,2,3,4,5

Actividad B: Saltar en paracaídas sin paracaídas de reserva 1,2,3,4,5

Actividad B: Saltar de una avión sin paracaídas para intentar llegar a una colchoneta 1,2,3,4,5

Fuente: Aldás-Manzano y Maldonado-Guzmán (2015)

**b.***Escalas de Likert.* Las *escalas Likert* son las más utilizadas en el desarrollo de escalas. El ítem se presenta como una afirmación, seguida por alternativas de respuesta que suponen diversos niveles de acuerdo en ella. A estas escalas se le pueden aplicar las consideraciones anteriores sobre el número par o impar de alternativas y la cantidad de estas. Estas escalas plantean básicamente un

problema. Si las afirmaciones son muy neutras, existe una gran tendencia al acuerdo y, además, pueden estar recogiendo, de hecho, más la ausencia de opinión que la opinión. Por ejemplo, no es lo mismo pedir el acuerdo o desacuerdo con afirmaciones como *los médicos no hacen caso normalmente a lo que les dicen sus pacientes* o *muchas veces los médicos no prestan demasiada atención a sus pacientes* que hacerlo sobre *de vez en cuando, los médicos pueden olvidar algo de lo que les han dicho sus pacientes*. En conclusión, un buen ítem *Likert* debería manifestar la opinión de una manera clara.

**c.** Diferencial semántico. En general, el diferencial semántico va asociado a un estímulo (como un grupo de personas, por ejemplo, los vendedores de coches). Una vez identificado el estímulo se presentan una lista de pares de adjetivos que representan los puntos opuestos de un continuo(*honrado* vs. *no honrado*) como se muestra en la **Tabla 5.4**

**Tabla 5.4. Escalas con ítems en diferencial semántico. Un ejemplo de caso**

| Vendedores de automóviles | | |
|---|---|---|
| Honrados | — — — — — — — | No honrados |
| Apacibles | — — — — — — — | Ruidosos |

Fuente: Aldás-Manzano y Maldonado-Guzmán (2015)

## *El listado de ítems debe ser revisado por expertos*

Esta fase tiene mucho que ver con la difícil tarea de lograr la validez de contenido de la escala que no puede constatarse de ningún otro modo. Tras dar al panel de expertos la definición que el investigador ha hecho del constructo que tiene que medir y del conjunto de ítems que componen la escala, solicitará al panel que realice las siguientes tareas:

**1.** Los expertos confirmarán o invalidarán la definición que el investigador haya hecho del constructo que quiere medir.

**2.** A continuación el panel de expertos debería valorar el nivel de relevancia que creen que cada ítem tiene para medir el fenómeno que se pretende. Estas valoraciones permitirán seleccionar la muestra definitiva de ítems eligiendo entre los más significativos.

**3.** Cuando el constructo tiene distintas dimensiones, los expertos deberán revisar la asignación hecha de afirmaciones a dimensiones.

**4.** expertos también deberán valorar la claridad y precisión de la formulación concreta de cada ítem, apuntando, si es necesario, formas alternativas de redactarlos. Esto hace referencia tanto a problemas de claridad estrictamente l literaria, como a que puedan por esta falta de claridad reflejar factores extraños al constructo que se pretende medir.

**5.** También se pretende que los expertos determinen si se ha dejado fuera de la escala alguna de las dimensiones del fenómeno.

## Considerar la inclusión de ítems de validación

Además de los ítems que pretenden medir la variable subyacente, el investigador debe considerar la incorporación de dos tipos de ítems adicionales (Aldás-Manzano y Maldonado-Guzmán, 2015):

- *Ítems para detectar defectos o problemas de la escala.* A veces los entrevistados pueden no estar respondiendo a las preguntas con la motivación que el investigador presupone. Una de las motivaciones más habituales es presentar una imagen socialmente aceptable, y no su verdadera forma de ser. Suele recomendarse para ciertos tipos de estudios la inclusión de una escala que mida esta tendencia, como la de Strahan y Gerbasi (1972**).**
- *Ítems para evaluar la validez de constructo de la escala***.** Si la teoría apunta que el constructo que se está midiendo está relacionado con otros constructos, la inclusión de escalas de medición de algunos de ellos puede ayudar a analizar la validez de la nueva escala desarrollada, si se confirman esas relaciones.

## Administrar la escala a una muestra de prueba

La muestra tiene que ser lo suficientemente grande como para eliminar la varianza que aporta un individuo como tal, haciendo que no sea representativa de una población determinada, sino de sí misma. Nunally y Bernstein (1994) apuntan la cifra de 300 personas como un número adecuado, aunque DeVellis (1991) señala que se han desarrollado escalas con éxito con un menor número.

## Evaluando los ítems

Con los resultados de administrar la escala a una muestra de prueba, debe analizarse como han funcionado cada uno de los ítems de tal forma que se pueda seleccionar finalmente a los más adecuados (Aldás-Manzano y Maldonado-Guzmán, 2015):

**a.** *Análisis inicial del funcionamiento de cada ítem***.** Cuanto mayores sean las correlaciones entre los ítems, mayor será la fiabilidad de la escala. Por ello serán candidatos a abandonar la escala aquellos ítems que, perteneciendo a una misma dimensión del constructo, guarden poca correlación con los demás integrantes de la misma.

**b.** *Formulación inversa***.** Si algún ítem guarda una elevada correlación negativa con los demás, deberá formularse gramaticalmente en sentido inverso para mantenerlo en la escala. Si esta forma de enunciarlo se ha hecho intencionadamente, deberá cambiarse la codificación de la respuesta.

**c.** *Analizar la varianza de los ítems***.** Si todos los individuos responden igual a un ítem, la varianza será cero. En este caso (o en casos de varianzas muy pequeñas) el ítem no ha sido capaz de discriminar entre los individuos con distintos niveles del constructo que se está midiendo, y debería suprimirse.

**d.** *Analizar la media de los ítems.* Es deseable que la media de los ítems esté entorno al centro de la escala de respuesta. Si están en la zona del acuerdo, indicará que el enunciado de la afirmación era demasiado neutro.

**e.** *Cálculo del alfa de Cronbach.*

## Optimizar la longitud de la escala

Ya se ha indicado que la fiabilidad de una escala viene influenciada por el número de ítems de la misma. Por un lado, escalas cortas reducen los problemas de su administración a los entrevistados, pero las largas son más fiables. Encontrar el punto de equilibrio debe ser el objeto de esta etapa. Si una escala es poco fiable, entonces la brevedad no es una virtud. Dicho de otro modo, se pueden intentar eliminar ítems para facilitar la administración, pero nunca permitiendo que los *alfa de Cronbach* sean inferiores al .7 o .8 recomendados.

Sin embargo, como vimos en el ejemplo, *eliminar malos ítems puede mejorar significativamente la fiabilidad de la escala*.

En ese mismo ejemplo, mostrábamos cómo detectar y eliminar esos ítems. Debe mantenerse, sin embargo un cierto margen de seguridad con el valor de los alfa **(α)**, porque no debe olvidarse que la escala se ha administrado a una muestra de prueba y estos valores pueden caer cuando se administren a la muestra definitiva. Si la muestra de prueba es lo suficientemente grande, debería considerarse su división en dos mitades, de tal forma que una de ellas se utilice para calcular los alfa **(α)**, evaluar los ítems y eliminar los inadecuados, mientras que la segunda mitad se podrá utilizar para validar estos resultados (Aldás-Manzano y Maldonado-Guzmán, 2015).

# Conclusión

En este capítulo, hemos ofrecido varias ilustraciones de los dos primeros pasos importantes en el desarrollo de la escala: el paso 1 de la definición de constructo y el dominio de contenido, y el paso 2 de generar y evaluar los indicadores de medición. Se discutieron varios aspectos importantes relacionados con el primer paso, incluido el papel de la teoría en la definición de constructo y el dominio de contenido, la importancia de una revisión y definición exhaustivas de la literatura y la importancia de una dimensionalidad de constructo *a priori*. También se destacaron los temas relevantes para el segundo paso, incluyendo la generación de un grupo de elementos suficientemente grande para aprovechar el dominio de contenido, los diversos aspectos de la escritura de artículos y el uso de expertos para calificar el la validez aparente (*face validity*) y de contenido (*content validity*). Los Capítulos 6 y 7 se centrarán en estudios empíricos para desarrollar, perfeccionar y finalizar una escala. Por último, se hace una recopilación de recomendaciones de cómo desarrollar una escala.

# CAPÍTULO 6. DISEÑANDO LA ESCALA. ANÁLISIS FACTORIAL EXPLORATORIO (Etapa 3)

El análisis factorial exploratorio (**EFA**) pueden usarse para dos propósitos principales en el desarrollo de la escala:

**a.***Para reducir el número de ítems en una escala,* de modo que los ítems restantes maximicen la varianza explicada en la escala y maximicen la confiabilidad de la escala, y

**b.***Para identificar las posibles dimensiones subyacentes en una escala.*

Estos dos usos están relacionados, y en el caso del desarrollo de escala, pueden usarse de manera complementaria. Típicamente, el análisis de factores por componentes principales (**PCA***.Principal Components Factor Analysis*) se usa para tal fin en lugar del análisis factorial común (*Common Factor Analysis*, Floyd y Widaman, 1995; Hair, Anderson, Tatham y Black, 1998).

Con **PCA**, los componentes (o factores) se estiman de tal manera que representan las varianzas entre los indicadores de la escala de la manera más sencilla (*económica*) posible, con el menor número posible de componentes significativos (*dimensiones*). **PCA** analiza la matriz de correlación entre las variables observadas

(indicadores) con las de la diagonal principal**.** Al hacerlo, **PCA** maximiza todas la varianzas de los indicadores, independientemente de si es común a un factor (*componente*) o a un único un indicador. Aquellos ítems para los que se explica con una baja varianza se consideran candidatos para la *eliminación* (es decir, para *reducir la cantidad de elementos en la escala*).
Por el contrario, el *análisis factorial común* (*Common Factor Analysis* ,en el desarrollo de la escala) se asocia más a menudo con la búsqueda de dimensiones subyacentes para un conjunto de elementos. También utiliza una matriz de correlaciones (o covarianzas) entre los puntajes observados en los ítems para identificar un conjunto de variables (*factores*) latentes que explican las correlaciones entre ellos. Sin embargo, el análisis factorial común usa las estimaciones de comunalidad de los indicadores en la diagonal principal. Además, la varianza en un elemento dado se divide en lo que es común a un factor o variable latente (basado en la varianza compartida con otros elementos en el análisis) y la varianza que es única para un elemento dado, una combinación de factores específicos varianza y varianza de error aleatorio. El *análisis factorial común* puede usarse para identificar el (los) constructo(s) teórico(s) cuyos efectos se reflejen por las respuestas a los ítems en una escala. El *análisis factorial común* también puede ofrecer información sobre qué elementos eliminar o retener, similar a la de **PCA**. Muchos autores informan que las soluciones derivadas de **PCA** y análisis de factores comunes tienden a ser bastante similares. Este es particularmente el caso cuando la cantidad de indicadores supera los 30 o los las comunalidades superan los .60 para la mayoría de los indicadores (Hair et al., 1998). Otros han cuestionado el hallazgo de que **PCA** y *análisis de factores comunes* arrojan resultados similares. Las diferencias pueden ser más pronunciadas con un pequeño número de indicadores y bajas comunalidades. De hecho, en tal caso, **PCA** y *análisis de factores comunes* pueden ofrecer resultados divergentes (Floyd y Widaman, 1995). Por lo tanto, algunos autores sugieren que se prefieren los análisis de factores comunes sobre el **PCA**, ya que la mayoría de las aplicaciones de desarrollo de escala buscan comprender un constructo en términos del número de *factores latentes que lo subyacen, así como también reducir el número de elementos en una escala***.** Además, dado que el análisis factorial confirmatorio (**CFA**) es una herramienta ampliamente utilizada en la finalización de la escala, los *análisis de factores comunes* basados en **EFA** pueden generalizarse mejor a **CFA** que la **PCA** (Floyd y Widaman, 1995). Ver más al respecto en Mejìa-Trejo (2017c, p.18)

# ¿Qué es el análisis factorial?

El Capítulo 5 se centró en la definición de constructos y en la generación, evaluación y refinamiento de un conjunto de indicadores inicial, es decir, los Pasos 1 y 2 en el desarrollo de la escala. Este capítulo se centra en el diseño y la realización de estudios iniciales para su desarrollar y *reducir* aún más el conjunto de indicadores o ítems, es decir, el Paso 3 en el desarrollo de la escala. Específicamente, abordamos los siguientes objetivos: prueba piloto de un grupo de indicadores como reducción de indicadores y procedimiento de prueba de validez inicial, y realización de múltiples estudios para el desarrollo de la escala. Dentro de este segundo objetivo, destacamos:

**a.**La importancia de incluir constructos para evaluar los diferentes tipos de validez,

**b.**El análisis factorial exploratorio (**EFA***. Exploratory Factor Analysis*) sobre múltiples conjuntos de datos para refinar la escala y examinar una estructura teórica a priori de factores iniciales, y

**c.**El análisis de indicadores y su confiabilidad.

Finalmente, ofrecemos ejemplos de escalas recientemente creadas para ilustrar cómo se utiliza **EFA**, el análisis de elementos y la confiabilidad en el desarrollo de una escala.

Esta técnica estadística multivariante ha tenido una creciente utilización desde fines de los 90s en todas las áreas de investigación de las ciencias económico-administrativas de carácter empresarial. A medida que se incrementa el número de variables que intervienen en las técnicas multivariantes, se crea la necesidad también mayor de conocer a profundidad tanto la estructura como las interrelaciones de las variables. El *análisis factorial*, es una técnica especialmente adecuada para analizar las pautas de relaciones complejas y multidimensionales encontradas por los investigadores del campo de las ciencias económico-administrativas. El objetivo en este apartado es el de definir y explicar los aspectos fundamentales de las técnicas de *análisis factorial* en términos conceptuales, lo más amplios posibles. Es posible aplicarlo para examinar las pautas subyacentes o las relaciones para un amplio número de variables y determinar si la información puede ser resumida en una serie de factores o componentes más pequeños, por lo que se necesitará conocer de las directrices básicas para presentar e interpretar los resultados de estas técnicas. Para saber más, ver IBM, 2011a; IBM, 2011b, IBM, 2011c.

El *análisis factorial* es el nombre genérico de aquellos métodos estadísticos multivariantes que se enfocan en *definir la estructura subyacente en una matriz de datos***.** Normalmente, incluye el tratar el problema de *cómo analizar la estructura de las interrelaciones (correlaciones)* entre un gran número de variables (por ejemplo, las puntuaciones de prueba, artículos de prueba, respuestas de cuestionarios) con la definición de una serie de *dimensiones subyacentes comunes, conocidas como factores*. Con el análisis factorial, podrá identificar, en principio, las *dimensiones separadas de la estructura y entonces determinar el grado en que se justifica cada variable por cada dimensión*. Una vez determinadas las dimensiones y la explicación de cada variable, se pueden lograr los dos objetivos principales para el análisis factorial: *el resumen y la reducción de datos*. Al momento de resumir los datos, se obtienen las *dimensiones subyacentes* que, interpretadas y comprendidas, *describen*

*los datos con un número de conceptos mucho más reducido que las variables individuales originales***.**
La *reducción de datos* se obtiene con el *cálculo de la puntuación para cada dimensión subyacente y sustituirlos por las variables originales*. Esta técnica multivariante, es de las primeras en considerar su capacidad para hacer acomodos de las variables múltiples con el objetivo de comprender las relaciones complejas que no son posibles con los métodos univariantes y bivariantes.
Al aumentar el número de las variables también aumenta la posibilidad de que las variables estén no correlacionadas y no sean representativas de unos conceptos distintos. En su lugar, los grupos de variables pueden estar interrelacionados en la medida en que son todos representativos de un concepto más general, tendiendo co9mo causales posibles: el diseño, el intento de medir las muchas facetas de personalidad o la imagen del establecimiento, o puede surgir simplemente de la adición de nuevas variables. Usted tiene que saber cómo se relacionan las variables para interpretar mejor los resultados. Si el número de variables es demasiado grande o deba dar una mejor representación a un número de conceptos más pequeño en vez de las facetas múltiples, ésta técnica ayuda en la selección de un subgrupo de variables representativo o incluso crear nuevas variables como sustitutas para las variables originales mientras mantengan su carácter original.
El análisis factorial es diferente de las técnicas de dependencia (regresión múltiple, el análisis discriminante, el análisis multivariante de la varianza o la correlación canónica), en el que se consideran una o más variables explícitamente como las variables de criterio o dependientes y todas las demás son las variables de predicción o independientes.
El análisis factorial es una técnica de interdependencia en la que se consideran todas las variables simultáneamente, cada una relacionada con todas las demás y empleando todavía el concepto del valor teórico, el compuesto lineal de las variables.
En el análisis factorial, los valores teóricos (los factores) se forman para *maximizar* su explicación de la serie de variables entera*, y no para predecir una(s) variable(s) dependiente(s).* Si hiciéramos una analogía con las técnicas de dependencia, cada una de las variables (originales) observadas sería una variable dependiente, que es una función de una serie de factores (dimensiones) subyacentes y latentes que están compuestas por todas las otras variables. Por tanto, cada variable es predicha por todas las demás. Por el contrario, se puede considerar cada factor (valor teórico) como una variable dependiente que es una función del conjunto entero de las variables observadas.
Cualquiera de estas analogías ilustra las diferencias de propósito entre *las técnicas de dependencia (la predicción) y la interdependencia (identificación de estructura).* Las técnicas analíticas de factores pueden lograr sus propósitos desde una perspectiva *exploratoria o confirmatoria***.** Muchos investigadores lo consideran solamente *exploratorio*, útil para la búsqueda de una estructura entre una serie de variables o como un *método de reducción de datos***.** Desde esta perspectiva, las técnicas de análisis factorial *extraen lo que proporcionan los datos* y no tienen ninguna restricción en la estimación de los componentes o el número de componentes a ser extraído. Para muchas aplicaciones, si no todas, resulta apropiada esta aplicación del análisis factorial. No obstante, en otras situaciones

tendrá sus pensamientos preconcebidos so bre la estructura real de los datos, que se basan en un apoyo teórico o investigaciones previas. Es posible que el investigador quiera probar las hipótesis que implican cuestiones tales como qué variables deberían ser agrupadas en un factor o el número exacto de factores. En estos casos, se re quiere un *análisis factorial que adopte un enfoque confirmatorio, es decir, valorar hasta qué punto los datos se ajustan a la estructura esperada.* Los métodos tratados en esta sección NO proporcionan la estructura necesaria para la prueba de hipótesis formalizada. Abordamos explícitamente la *perspectiva confirmatoria* del análisis factorial está radicada en las *ecuaciones estructurales.* En este capítulo, sin embargo, observamos las técnicas analíticas de factores principalmente desde un punto de vista exploratorio o no confirmatorio.

- **Problema hipotético:**

Supongamos que mediante una investigación cualitativa una empresa de servicios de mercadotecnia digital ha identificado 100 características diferentes de empresas contratantes de sus servicios (consumidores), que como consumidores han mencionado que afectan su elección de en la planeación de la campaña web. La empresa de mercadotecnia digital quiere entender cómo deciden sus consumidores acerca de sus servicios, pero opina que no puede valorar las 100 características individuales o desarrollar planes de acción para tantas variables, porque son demasiado específicos. En su lugar, *a la empresa de mercadotecnia digital le gustaría saber si sus consumidores piensan en una dimensión determinante más general en vez de únicamente en aspectos específicos***.** Para identificar estas *dimensiones*, la empresa de mercadotecnia digital deberá:

- **Realizar una encuesta** solicitando valoraciones de consumidores sobre cada uno de estos aspectos específicos.
- **Emplear el análisis factorial para identificar las dimensiones determinantes** *subyacentes***.** Se considera que los aspectos específicos que se correlacionan en gran medida forman parte de una dimensión más amplia.
- **Convertir éstas dimensiones en compuestos de las variables específicas***,* que a su vez permitan a las dimensiones ser interpretadas y descritas. En nuestro ejemplo, el análisis factorial podría identificar *dimensiones como planeación, estrategia, implementación, tecnología, experiencia de usuario* como las dimensiones determinantes utilizadas por los encuestados.

Cada una de estas dimensiones contiene aspectos específicos que son una faceta de la dimensión determinante más amplia. A raíz de estos resultados la empresa de servicios de mercadotecnia digital puede usar estas dimensiones (factores) para definir áreas generales para la planificación y actuación.

# Prueba piloto

Como se indicó en el Capítulo 5, se debe generar y evaluar un conjunto de indicadores para determinar el contenido y la validez aparente (*face validity*) antes de que se administren a las muestras del sujeto de estudio. Como también se indicó en el Capítulo 5, el conjunto o (*pool*) de indicadores inicial debería ser lo suficientemente grande, con el tamaño apropiado variable dependiendo de la dimensión y la complejidad de la definición de constructo. Con un gran conjunto de indicadores, *será*

*útil reducir el grupo para los estudios de desarrollo que siguen y obtener algunas estimaciones iniciales de confiabilidad y validez. En este sentido, las pruebas piloto pueden ser útiles*. Dado que la escala se administrará a más muestras para refinar, *un estudio piloto puede reducir el número de elementos en un grupo inicial a un número más manejable al eliminar elementos que no cumplen ciertos criterios psicométricos***.**

Las pruebas piloto también pueden servir para el propósito de probar inicialmente la validez. Posteriormente, una *red nomológica* puede evaluarse parcialmente mediante la inclusión de medidas/escalas de constructos relevantes con una muestra piloto. Cuatro cuestiones a tener en cuenta en las pruebas piloto:

- **El tamaño de la muestra.**

Considere el tamaño de la muestra de los encuestados en la prueba piloto. Algunos sugieren que las muestras piloto deberían estar desde n = **300** (DeVellis, 1991), mientras que otras sugieren que n = **100** a **200** serán suficientes (Clark y Watson, 1995). Un problema a tener en cuenta con el tamaño de muestra es cuántos indicadores hay en el grupo inicial. Para las agrupaciones con un pequeño número de elementos (por ejemplo, 20 elementos o menos) asociados con constructos estrechamente definidos, desde n = **100 a 200** parece razonable, ya que los análisis iniciales de elementos y factores pueden realizarse con dichos tamaños de muestra. Sin embargo, para un *pool* más grande (y para constructos complejas y/o multidimensionales), se prefieren tamaños de muestra más grandes desde n= **300.**

- **La composición de la muestra.**

Para las pruebas piloto, las muestras de conveniencia (por ejemplo, estudiantes universitarios) pueden ser suficientes, pero es preferible usar una muestra de una población relevante de interés. Dado que el objetivo de la prueba piloto es reducir el número de indicadores a un número más manejable para los estudios de desarrollo más grandes, la representatividad de la muestra es un plus. Un elemento que se desempeñe bien (o mal) con una muestra de una población relevante será evaluado con más confianza como candidato para la inclusión (o supresión) para las siguientes muestras de desarrollo. Por lo tanto, se prefieren las muestras de poblaciones relevantes.

- **Las estimaciones iniciales de la confiabilidad de los indicadores.**

Los elementos pueden evaluarse inicialmente para la consistencia interna, los medios, las variaciones, la correlación interítem media y la estructura del factor**.**

*Los indicadores que no se comporten empíricamente como se esperaba pueden ajustarse para la reducción, las etiquetas de escala o de otro modo, para la retención en los esfuerzos de recopilación de datos de desarrollo.*

Algunos indicadores pueden requerir una eliminación basada en estos análisis, y si es necesario, se pueden redactar y juzgar más de ellos, de modo que se retenga un conjunto de indicadores suficiente para los estudios de desarrollo de escala que siguen.

- **El número y el tipo de escalas relacionadas con la validez a incluir.**

Se puede hacer una evaluación inicial de algunos aspectos de la validez a través de una muestra piloto. Aquí, algunas mediciones de constructo clave pueden ser muy útiles para obtener estimaciones iniciales de validez (es decir*, parte de la red nomológica del constructo*). Como se dijo anteriormente, una revisión exhaustiva de

la literatura debe alertar al desarrollador de la escala sobre los constructos y las medidas complementarias focales a las que debe relacionarse con (y diferenciarse de). Por lo tanto, deberían incluirse *tres o cuatro mediciones de constructos clave para examinar las estimaciones iniciales de validez***.** Como nota final sobre las pruebas piloto, *es mejor tener un grupo de elementos que incluya en exceso el dominio de contenido del constructo que uno que sea escaso***.**

Dado que el objetivo de las pruebas piloto es reducir indicadores, una pregunta importante es *¿Cuántos artículos se deben retener de la prueba piloto?*. Por lo tanto, si un indicador se desempeña de manera marginal en los análisis de elementos y factores en una prueba piloto, pero se juzga que tiene contenido y validez nominal, *lo mejor es retener ese elemento para otras muestras. Este será particularmente el caso si la muestra piloto fuera de conveniencia y no necesariamente representativa de la población de interés***.**

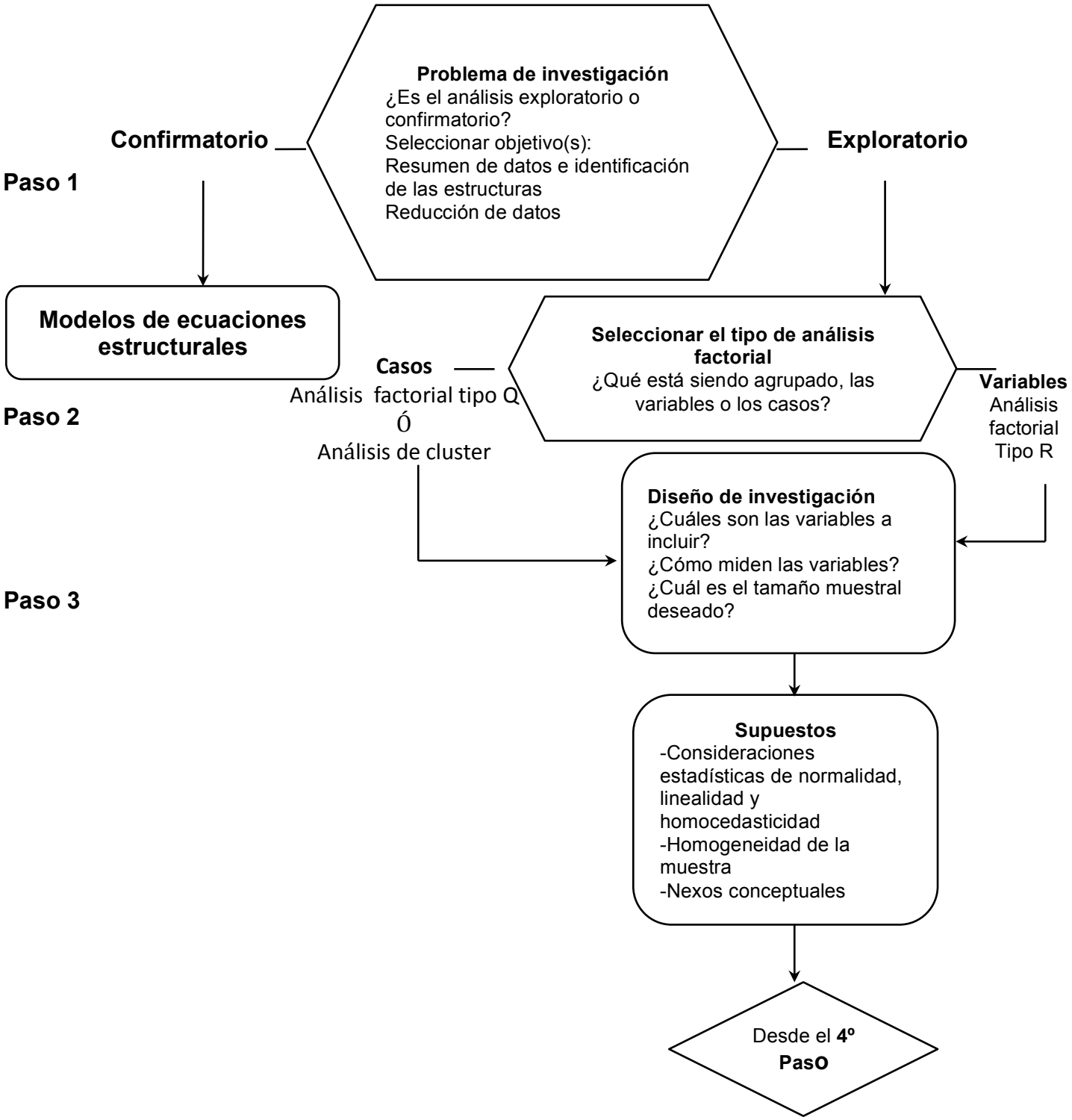
Esto se detallará más adelante

# Análisis factorial y el proceso de decisión

De acuerdo al modelo de seis pasos de Hair et al. (1999) que se introdujo en el Capítulo 2, la **Figura 12.1** muestra tres pasos iniciales de la aproximación estructurada para la construcción de modelos multivariantes.

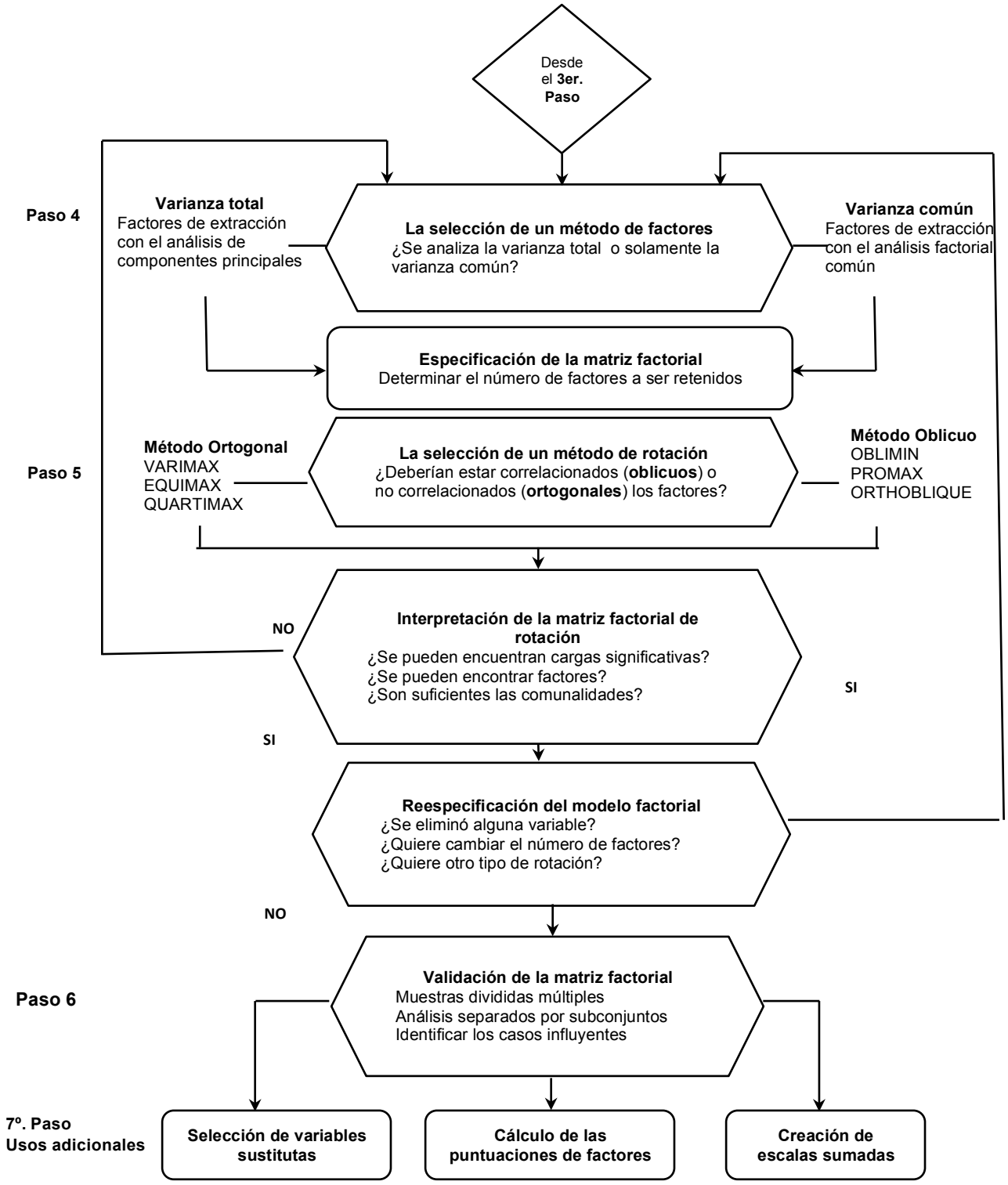
Por otro lado, la **Figura 12.2** muestra en detalle los últimos tres pasos, y uno adicional (el séptimo paso) más allá de la estimación, la interpretación y la validación de los modelos factoriales, que ayuda a la selección de las variables sustitutas, las puntuaciones de factores o la creación de las escalas aditivas para la utilización en otras técnicas multivariantes. A continuación se presenta un análisis de cada paso.

**Figura 3.1 . Diagrama de flujo pasos 1-3 del análisis factorial**




Fuente: Hair (1999)

## Figura 3.2. Diagrama de flujo pasos 4-6 del análisis factorial




Fuente: Hair (1999)

# Objetivos

**Paso 1: establecimiento de objetivos**

Como con otras técnicas estadísticas, inicia con el *problema sujeto de investigación*. El propósito general de las técnicas analíticas de factores es *encontrar una manera de resumir* la información contenida a partir de una serie de variables originales *en una serie más pequeña de dimensiones compuestas o valores teóricos (factores) nuevos* aceptando una mínima pérdida de información (en otras palabras, buscar y definir las construcciones fundamentales o dimensiones que se supone sirven de base para las variables originales **[**Gorsuch, 1983**,** Rurnrnel, 1970**]**.

Las técnicas del análisis factorial deben satisfacer cualquiera de estos dos objetivos:

- **La identificación de estructura mediante el resumen de datos, o bien**
- **La reducción de datos.**

## *La identificación de estructura mediante el resumen de datos*

El análisis factorial puede identificar la estructura de las relaciones entre los encuestados o las variables mediante el *análisis de las correlaciones* entre encuestados y/o variables. Por ejemplo, suponga que se tienen datos sobre 100 encuestados basados en 10 características. Si el objetivo de la investigación fuera:

- **Aplicar el análisis factorial a la matriz de correlación de los encuestados individuales basada en sus característica***s,* se le denomina el *análisis factorial Q*, siendo un método para combinar o condensar grandes grupos de personas en grupos claramente diferentes dentro de una población mayor, no se usa la aproximación del *análisis factorial Q* con mucha frecuencia. En su lugar, la mayoría de los investigadores utilizan algún tipo de *análisis cluster* para agrupar los encuestados individuales. También véase **[**Stewart 1981**]** para otras combinaciones posibles de grupos y tipos de variables
- **El resumen de las características, se aplicaría el análisis factorial a una matriz de correlación de las variables.** *Éste es el tipo de análisis factorial más común, y se denomina el análisis factorial R,* el cual analiza las variables para identificar las *dimensiones que son latentes* (las que no son fácil mente observadas).

## *Reducción de datos*

El análisis factorial también puede:

- **Identificar las variables suplentes** de una serie de variables más grande para su utilización en análisis multivariantes posteriores, o
- **Crear una serie de variables completamente nueva***, mucho más pequeña en número,* para reemplazar parcial o completamente la serie original de variables para su inclusión en técnicas posteriores.

En ambos casos, el propósito es retener la naturaleza y el carácter de las variables originales reduciendo su número para simplificar el análisis multivariante posterior. Aunque las técnicas multivariantes se han desarrollado para utilizar múltiples variables, *Usted debe siempre buscar la serie de variables más reducida para incluirla en el análisis*. Las cuestiones conceptuales y las empíricas deben respaldar la creación de medidas compuestas. El *análisis factorial proporciona la base empírica*

para valorar la estructura de las variables y la potencial para crear estas medidas compuestas o seleccionar una subserie de variables suplentes para el análisis posterior. *El resumen de dato***s** hace que la identificación de las dimensiones subyacentes o los factores sean fines de por sí; las estimaciones de los factores y las contribuciones de cada variable a los factores (*denominadas cargas de los factores*) constituyen todo lo que se necesita para el análisis. La reducción de datos depende también de las cargas de los factores; no obstante, las utiliza como la base para identificar las variables para su análisis subsiguiente con otras técnicas o bien para hacer estimaciones de los factores mismos (puntuaciones de factores o escalas aditivas), que a su vez reemplazan las variables originales en análisis subsiguientes.

## *El uso del análisis factorial con otras técnicas multivariantes*

El análisis factorial *proporciona una visión directa de las interrelaciones entre las variables o los encuestados* y un apoyo empírico para abordar las cuestiones conceptuales que tienen relación con la *estructura subyacente de los datos*. Es un complementario importante con otras técnicas multivariantes mediante *el resumen y la reducción de datos.*

Como *resumen de datos*, el análisis factorial proporciona una comprensión clara de cuáles de las variables podrían actuar juntas y cuántas de las variables realmente se puede esperar que tengan un impacto en el análisis.

• **Por ejemplo,** se esperaría que las variables altamente correlacionadas y miembros del mismo factor tuvieran perfiles similares de diferencia a través de los grupos en el *análisis multivariante de la varianza o en el análisis discriminante***.** Los procedimientos que muestran el impacto de las variables correlacionadas son los basados en etapas (*stepwise*) de la *regresión múltiple o el análisis discriminante. Estas técnicas introducen las variables secuencialmente*, basadas en su *capacidad adicional de predicció***n** sobre las variables en el modelo. Conforme entra la variable de un factor, *es menos probable que variables adicionales del mismo factor sean también incluidas*, porque están altamente correlacionadas y potencialmente tienen *menos capacidad de predicción adicional*, que las variables que no estén en ese factor. Esto no significa que las otras variables del factor sean menos importantes o tengan menos impacto, sino que *su efecto ya ha sido representado por la variable incluida en ese factor.* Esta visión puede ser incorporada directamente a otras técnicas multivariantes mediante cualquiera de las *técnicas de reducción de datos*.

El análisis factorial proporciona la base *para crear una nueva serie de variables* que incorporan el carácter y naturaleza de las variables originales en una *cantidad de nuevas variables más reducida,* sea con la utilización *de variables suplentes*, sea con la *puntuación de factores o las escalas aditivas*.

De esta manera, se pueden reducir los problemas que se asocian con las grandes cantidades de variables o intercorrelaciones altas entre las variables con la sustitución de las nuevas variables. Así la investigación se beneficia de las relaciones y la visión detallada de la base conceptual y la interpretación de los resultados.

## *Selección de variables*

*La reducción y el resumen de datos* pueden ser llevados a cabo tanto con una serie de *variables preexistentes* como con *las variables creadas* por la nueva investigación.

Cuando plantea el uso de una nueva serie, debe realizar una *aproximación conceptual* para determinar qué variables conviene incluir en el análisis. El uso del análisis factorial para la **reducción de datos** es particularmente crítico cuando se requiere la comparabilidad a lo largo de un período de tiempo o en situaciones múltiples. Cuando se usa en una nueva investigación, el análisis factorial puede *determinar también la estructura y/o crear nuevas puntuaciones compuestas* a partir de las variables originales**.**

- **Por ejemplo**, uno de los primeros pasos en la construcción de la escala aditiva es valorar la naturaleza de su dimensión y la conveniencia de las variables seleccionadas mediante el análisis factorial. Por tanto, *aunque no es verdaderamente confirmatorio, el análisis fac torial se puede utilizar para valorar la naturaleza de una dimensión propuesta.*

Una vez que se especifica el propósito del análisis factorial, Usted tiene que definir la serie de variables a examinar**.** Por lo que se refiere tanto al análisis factorial *tipo R o tipo Q,* Usted especifica implícitamente las dimensiones potenciales que se pueden identificar mediante el carácter y la naturaleza de las variables sujetas al análisis factorial.

- **Por ejemplo**, en la valoración de las dimensiones de la *experiencia del usuario*, el análisis factorial no podría identificar esta dimensión si no han sido incluidas preguntas sobre el *acceso, la usabilidad y la visibilidad a la página web*. Debe recordar también que el análisis factorial siempre **producirá factores**. Por tanto, el análisis factorial es siempre un candidato potencial para el fenómeno *basura dentro, basura fuera. Si el investigador incluye indiscriminadamente grandes cantidades de variables y espera que el análisis factorial lo solucione, entonces la posibilidad de obtener malos resultados es alta*. La calidad y el significado de los factores derivados reflejan un acercamiento conceptual a las variables incluidas en el análisis. El uso del análisis factorial como una técnica de resumen de datos **no** *excluye la necesidad de una base conceptual para cualquiera de las variables analizadas*. Incluso si se usa meramente para la reducción de datos, el análisis factorial es más eficiente cuando las dimensiones definidas conceptualmente pueden ser representadas por los factores obtenidos.

# Diseño

**Paso 2: Diseño**

El diseño de un análisis factorial implica tres decisiones básicas:

- **El cálculo de los datos de entrada (matriz de correlación)** para alcanzar los objetivos específicos de la agrupación de variables o encuestados;
- **El diseño del estudio en términos de número de variables, las propiedades de medición** de las variables y los tipos de las *variables permisibles*; y
- **El tamaño de muestra necesario**, tanto en términos absolutos como para la función del número de variables en el análisis.

## *Las correlaciones entre las variables o los encuestados*

La primera decisión en el diseño de un análisis factorial se concentra en la *aproximación* que se usa para calcular la *matriz de correlación* tanto para el análisis factorial de *tipo R* como para el de *tipo Q*. Usted puede:

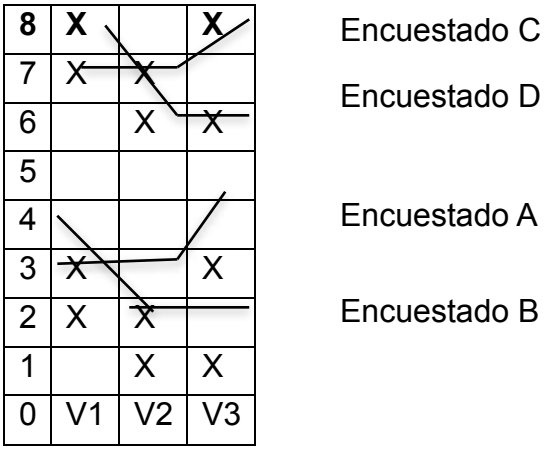
- **Utilizar la matriz de datos de entrada** a partir del cálculo de las correlaciones entre las variables, empleando, por tanto, un *análisis factorial de tipo R*, ó
- **Elegir la matriz de correlación** de las correlaciones entre los encuestados individuales, o *análisis factorial tipo Q*, el resultado será una matriz factorial que identifica individuos similares.
- **Por ejemplo**, si los encuestados individuales se identifican por un número, la pauta de factores de resultado podría indicamos que los individuos 1, 5, 6 y 7 son similares. Del mismo modo, los encuestados 2, 3, 4 y 8 quizá cargarían juntos sobre otro factor, y clasificaríamos estos individuos como similares. A partir de los resultados del *análisis factorial Q,* se pueden *identificar grupos o clusters* de individuos que muestran una pauta similar sobre las variables que se incluyen en el análisis. A estas alturas, una pregunta lógica sería*, ¿cómo se diferencia el análisis factorial tipo Q del análisis cluster, dado que ambas aproximaciones comparan la pauta de respuestas a través de una serie de variables y clasifican a los encuestados en grupos?* La respuesta es que el *análisis factorial tipo Q* se basa en las *intercorrelaciones* entre los encuestados, mientras que el *análisis cluster* forma grupos que se basan en una medida de similitud basada en la *distancia entre las puntuaciones de los encuestados sobre las variables analizadas*.

Para ilustrar esta diferencia, consideremos la **Figura 12.3**, que contiene las puntuaciones de cuatro encuestados sobre tres variables diferentes. Un *análisis factorial tipo Q* de estos cuatro encuestados daría dos grupos con estructuras de *covarianza* similares, agrupando a los encuestados **A y C** frente a **B y D.** Por contraste, la aproximación de **cluster** sería sensible a las distancias reales entre las puntuaciones de los encuestados y llevaría a la agrupación de las parejas más cercanas. Por tanto, con la aproximación del análisis cluster, los encuestados **A y B** estarían situados en un grupo y **C y D** en el otro grupo, por lo que se debe estar consciente de estas diferencias. Con la disponibilidad de otras técnicas de agrupación y el uso general del análisis factorial para la reducción de datos y el resumen, la exposición restante de este capítulo se centra en el *análisis factorial tipo R, la agrupación de variables* en vez de la agrupación de encuestados.

**Figura 3.3. Aproximación cluster**

| | VARIABLES | | |
|---|---|---|---|
| **ENCUESTADOS** | $X_1$ | $X_2$ | $X_3$ |
| **A** | 2 | 2 | 3 |
| **B** | 3 | 1 | 1 |
| **C** | 7 | 7 | 8 |
| **D** | 8 | 6 | 6 |

Fuente: propia




## *La selección de variables y cuestiones de medición*

A estas alturas, deberán resolverse preguntas:

- **¿Cómo se miden las variables?** y
- **¿Cuántas variables deberían ser incluidas?**

Por regla general, las variables a incluir en el análisis factorial tienen que ser de *escala métrica***.** En algunos casos, se pueden utilizar *variables ficticias (codificadas 0-1)*, aunque se consideran como **n***o métricas. Si todas las variables son ficticias*, entonces las formas especializadas del análisis factorial, tales como el *análisis factorial Booleano***,** son más apropiadas **[**BMDP Statistical Software 1992**]**. Además, deberá intentar *minimizar el número de variables que se incluyen*; no obstante, *debe también mantener un número razonable de variables por factor*. Si se está diseñando un estudio para valorar una estructura propuesta, deberá incluir varias variables (**5 o más**) que puedan representar cada *factor propuesto*. El poder del análisis factorial se basa en *encontrar pautas entre grupos de variables y resulta de poca utilidad en la identificación de factores compuestos de una única variable.* Finalmente, cuando se diseña una investigación que utiliza análisis factorial, debería, si cabe, *identificar varias variables claves (denominadas indicadores claves o variables marcadoras)* que reflejan con detalle los factores subyacentes hipotéticos, de forma que se facilite la validación de los factores derivados y la valoración sobre la significación práctica de los resultados.

## *Tamaño muestra*

Se recomienda NO usar el análisis factorial para una muestra inferior a 50 y preferentemente de 100 o más observaciones. Como regla general:

- **El mínimo es tener por lo menos un número de observaciones** 5 veces mayor que el número de variables a ser analizadas
- **El tamaño aceptable deberá estar en un ratio de** diez a uno. Algunos investigadores proponen incluso un mínimo de 20 casos por cada variable. Hay que

recordar, sin embargo, que con 30 variables, por ejemplo, hay 435 correlaciones en el análisis factorial.

- **Con un nivel de** significación de 0.05, es posible que incluso 20 de estas correlaciones sean consideradas significativas y aparecerían en el análisis factorial simplemente por casualidad.
- **Deberá siempre procurar obtener el ratio más alto de casos por variable** para minimizar las posibilidades de *sobreajustar* los datos (es decir, derivar los factores que son específicos a la muestra con poca capacidad de generalización).

De todas formas, se emplea una serie de variables menor al estar guiado por consideraciones conceptuales y prácticas. Aun así, se tienen tamaños muestrales más pequeños y/o ratios más bajos de casos y variables, debiéndose interpretar los resultados con cautela. La cuestión del tamaño muestral será abordada también en una sección posterior sobre la interpretación de las cargas de los factores.

# Supuestos

**Paso 3: supuestos de aplicabilidad**

Los supuestos básicos subyacentes del análisis factorial son *más de tipo conceptual que estadístico*, ya que estadísticamente hablando:

- *Se pueden obviar los supuestos de normalidad, homocedasticidad y linealidad siendo conscientes de que su incumplimiento produce una disminución en las correlaciones observadas.* En realidad, *sólo es necesaria la normalidad* cuando se aplica una prueba estadística a la *significación de los factores; sin embargo, raramente se utilizan estas pruebas. De hecho, es deseable que haya cierto grado de multicolinealidad,* dado que el objetivo es identificar series de variables interrelacionadas.
- **Adicionalmente a las bases estadísticas para las correlaciones de la matriz de los datos**, debe asegurarse también de que la matriz *tiene suficientes correlaciones para justificar la aplicación del análisis factorial*.
- **Si la inspección visual revela que no hay un número sustancial de correlaciones** *> 0.30, entonces el análisis factorial es probablemente inapropiado.*
- **Las correlaciones entre las variables también pueden ser analizadas con el cálculo de las correlaciones** *parciales entre las variables;* esto es, las correlaciones entre variables cuando se tienen en cuenta los efectos de las otras variables.
- **Si los factores verdaderos existen en los datos**, entonces:
  - *Si la correlación parcial es pequeña*, entonces se puede explicar la variable mediante los factores (valores teóricos con cargas para cada variable).
  - *Si la correlación parcial es alta, entonces no existen factores subyacentes verdaderos, y el análisis factorial es inapropiado.* **SPSS** proporciona la *matriz de correlación anti-imagen, que es el valor negativo de la correlación parcial*.
- **En cada caso, las correlaciones parciales o anti-imagen** mayores son indicativas de una matriz de datosque *no es quizá adecuada para el análisis factorial*.

Otra manera de determinar la conveniencia del análisis factorial es:

- **Examinar la matriz de correlación entera**. Utilice el ***contraste de esfericidad de Bartlett***, como prueba estadística para la *presencia de correlaciones entre las variables,* como una de estas medidas. Proporciona la *probabilidad estadística de que la matriz de correlación de las variables sea una matriz identidad.* Debe tener en cuenta, sin embargo, que *el incremento del tamaño muestra* da lugar a que ***la prueba de contraste de Bartlett*** *sea más sensible a la detección de correlaciones entre las variables.*
- **Otra medida para cuantificar el grado de intercorrelaciones** entre las variables y la conveniencia del análisis factorial es la **medida de suficiencia de muestreo (MSA)**. Este índice se extiende de **0 a 1**, con las siguientes directrices:

  **1** cuando cada variable es perfectamente predicha sin error por las otras
  **>=0.80** sobresaliente;
  **>=0.70** regular;
  **>=0.60** mediocre;
  **>=0.50** despreciable y,
  **< 0.50**, inaceptable [Kaiser, 1970 y 1974]
- El **MSA** aumenta conforme:
  - Aumenta el tamaño muestra
  - Aumentan las correlaciones medias
  - Aumenta el número de variables o
  - Desciende el número de factores [Kaiser , 1974].

Las mismas directrices de **MSA** pueden extenderse también a las variables individuales. Usted deberá examinar primero los valores MSA para cada variable y excluir aquellos que caen en la gama de inaceptables. Una vez que las variables individuales logran un nivel aceptable, se puede valorar el MSA general y se puede tomar una decisión sobre la continuación del análisis factorial.

Los *supuestos conceptuales* que subyacen en el análisis factorial se relacionan con la serie de variables seleccionadas y la muestra elegida. Un supuesto básico del análisis factorial es que existe una *estructura subyacente* en la serie de variables seleccionadas. Es su responsabilidad asegurarse de que *las pautas observadas sean válidas y conceptualmente apropiadas* para utilizar el análisis factorial porque *la técnica no tiene medios de determinar la conveniencia, aparte de las correlaciones entre las variables.*

- **Por ejemplo,** la mezcla de *variables dependientes e independientes* en un solo análisis factorial y posteriormente el uso de los factores derivados para apoyar las relaciones de dependencia *es inapropiado.*

Usted debe asegurarse de que *la muestra es homogénea* con respecto a **la** *estructura de factor subyacente.*

- **Por ejemplo***,* la aplicación del análisis factorial resultaría *inapropiada* para una muestra de hombres y mujeres que tienen distintas opiniones sobre una serie de aspectos según el sexo. Cuando se combinan las *dos submuestras (hombres y mujeres),* las correlaciones resultantes y la estructura de factores serán una *representación pobre* de la estructura única de cada grupo. Por tanto, *cuando se esperan grupos diferentes en la muestra, se deben practicar análisis factoriales*

*separados* y los resultados deben ser comparados para identificar las diferencias no reflejadas en los resultados de la muestra combinada.

# Estimación y Ajuste

**Paso 4: estimación y ajuste**

Una vez especificadas las variables y la matriz de correlación preparada, deberá estar listo para *aplicar el análisis factorial que identifique la estructura subyacente de las relaciones* (ver **Figura 12.2**). Para realizarlo, es necesario tomar decisiones con relación a:

- **El método de extracción de los factores (análisis factorial común vs análisis de componentes principales)** La selección del método de extracción depende del objetivo del investigador. Se utiliza el análisis de componentes principales cuando el objetivo es resumir la mayoría de la información original (varianza) en una cantidad mínima de factores con propósitos de pre dicción. Por el contrario, se utiliza el análisis factorial común para identificar los factores subyacentes o las dimensiones que reflejan qué es lo que las variables comparten en común.
- **El número de factores seleccionados para representar la estructura subyacente en los datos.** Para cualquiera de estos métodos, tiene que determinar también el número de factores que representan la serie de variables originales. Tanto las *cuestiones conceptuales como empíricas afectan a esta decisión.*

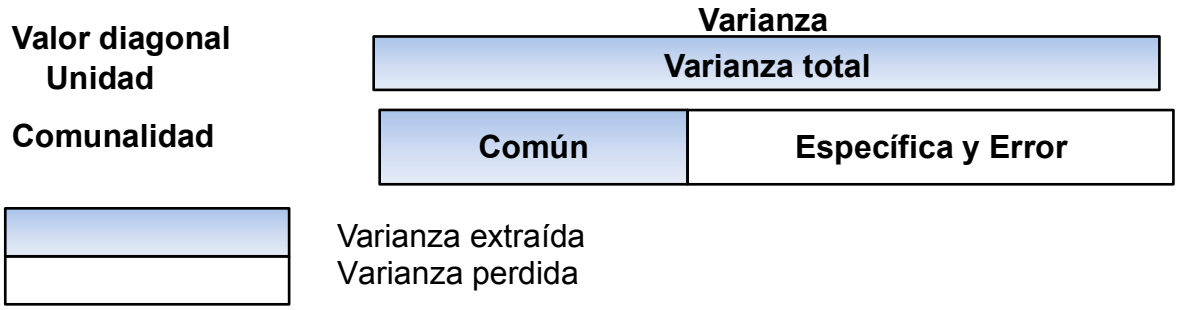
## *Análisis factorial común vs. Análisis de componentes principales*

Usted puede utilizar dos modelos básicos para obtener soluciones factoriales. Éstos se conocen como análisis factorial común y análisis de componentes principales. Con el fin de seleccionar el modelo apropiado, en primer lugar *debe comprender las diferencias entre los tipos de varianza.* Para los propósitos del análisis factorial, existen tres tipos de varianza total:

- **Común**
- **Específica (también conocida como única), y**
- **Error**.

Estos tipos de varianza y su relación con el proceso de selección de modelo factorial se ilustran en **la Figura 3.4.**

**Figura 3.4. Tipos de varianza llevados a la matriz factorial**




Fuente: Hair (1999)

Así, se tienen las siguientes definiciones:

- **Varianza común** en una variable que se comparte conlas otras variables en el análisis.
- **Varianza específica** es aquella asociada solamente con una variable específica.
- **La varianza de error** es aquella que se debe a la poca fiabilidad en el proceso de recolección de datos, al error de medición o un componente aleatorio en el fenómeno medido.

El *análisis de componentes principales* considera la *varianza total* y estima los factores que contienen *proporciones bajas de la varianza única* y, en algunos casos, la *varianza de error***.** No obstante***,*** *los primeros factores no contienen la suficiente varianza única o de error* como para distorsionar la estructura de factores en su conjunto. Específicamente, con el análisis de componentes principales, *se insertan las unidades en la diagonal de la matriz de correlación***,** para que se traiga **la** *varianza completa en la matriz de factores*, tal y como se indica en la **Figura 12.4**.

En el *análisis factorial común, por el contrario***,** se incorporan las varianzas compartidas en la diagonal*. Las comunalidades* son estimaciones de la varianza compartida o común entre las variables. Los factores que resultan del *análisis factorial común* se basan solamente en la **varianza común.**

La selección de un modelo u otro se basa en dos criterios:

- **Los objetivos del análisis factorial**, y
- **En el grado de conocimiento anterior acerca de la varianza** en las variables.

El *análisis de componentes principale***s** es apropiado cuando: -El interés primordial se centra en *la predicción o el mínimo número de factores* necesarios para justificar la *porción máxima de la varianza* representada en la serie de variables original, -Cuando el conocimiento previo sugiere que *la varianza específica y de error* representan una proporción relativamente pequeña de la varianza total. El modelo de *análisis factorial común*, es apropiado cuando:

- **Por el contrario, el objetivo principal es identificar las dimensiones latentes** o las construcciones representadas en las variables originales y *tiene poco conocimiento acerca de la varianza específica y de error y por tanto quiere eliminar esta varianza*
- **Con unos supuestos más restrictivos** y la utilización exclusiva de las dimensiones latentes (varianza compartida), se basa más en *la teoría*.

## Problemas

Aunque teóricamente válido, no obstante, el *análisis factorial común* tiene varios problemas:

- **Adolece de indeterminación de factores.** Esto significa que para cualquier encuestado individual, se pueden calcular varias puntuaciones de factores diferentes a partir de los resultados del modelo factorial **[**Mulaik, y McDonald 1978**].**
- **No existe una sola solución única**, tal y como ocurre con el *análisis de componentes principales*; no obstante, y en la mayor parte de los casos, *las diferencias no son sustanciales.*
- **El cálculo de las varianzas compartidas no siempre se pueden estimar** o pueden no ser válidas (es decir, *valores mayores que **1** o menores que **0***), lo que requiere la *supresión de la variable del análisis.*

Las complicaciones del análisis factorial común han contribuido *al uso generalizado del analisis de componentes principales***.** Aunque todavía los expertos siguen discutiendo acerca de cuál de los modelos factoriales es el más apropiado **[**Borgatta, et al. 1968**,** Gorsuch, 1990**,** Mulaik, 1990**,** Snook y Gorsuch, 1989**],** la investigación empírica ha mostrado resultados similares en muchos casos **[**Velicer y Jackson, 1990**].** En la mayoría de las aplicaciones, ambos análisis llegan a *resultados esencialmente idénticos si el número de variables excede de* **30** **[**Gorsuch, 1983**],o** *las varianzas compartidas exceden de **0.60** para la mayoría de las variables. Si* el investigador está preocupado por los supuestos del análisis de componentes principales, entonces *debe aplicar también el análisis factorial común* para valorar su estructura de representación. Cuando se llega a una decisión acerca del modelo factorial, debe estar preparado para *extraer los factores sin rotación iniciales*. Con el examen de la *matriz de factores sin rotación*, el investigador puede explorar las posibilidades de *reducción de datos* para una serie de variables y obtener una estimación preliminar de los factores a extraer. Así, se debe esperar para determinar el número de *factores final hasta que se haga una rotación* de los resultados y se interpreten los factores .

## Criterios para el cálculo del número de factores a ser extraídos

Existen varios criterios para extraer factores. Hair et al. (1998), Floyd y Widaman (1985) y Sharma (1996) ofrecen exposiciones muy claras y concisas de estos criterios. Solo los revisaremos brevemente aquí. Aunque existen algunas pruebas estadísticas para la extracción de factores, generalmente están restringidas a técnicas de estimación más alineadas con **CFA**. Para **EFA**, los criterios psicométricos y las reglas clave se aplican con mayor frecuencia. Como se señaló en el Capítulo 2, la regla *eigenvalvalor >1* (criterio Kaiser-Guttman o criterio de raíz latente *Latent Root Criterion*) a menudo se utiliza como *criterio psicométrico***.**

Cada componente (factor) tiene un eigenvalor propio que representa la cantidad de varianza contabilizada por el componente, donde la suma de todos los eigenvalores es igual al número de indicadores analizados. Un eigenvalor <1 indica que ese componente tiene menos varianza que cualquier elemento individual. Por lo tanto, con la reducción de datos como objetivo, un componente con eigenvalor<1 no se considera significativo.

Cabe señalar que el criterio de Kaiser-Guttman puede subestimar el número de componentes en algunas circunstancias y puede no ser confiable. Cliff (1988) demostró que la regla la regla del eigenvalor>1 tiene defectos, y probablemente se utilice mejor como una guía y no como un criterio absoluto para extraer **factores**. A menudo se recomiendan otras reglas generales para la extracción de factores, tales como**:**

- **La primer regla clave es el *scree test* o gráfico de sedimentación**. Esta prueba, traza los eigenvalores y muestra la *pendiente* de una línea que los conecta. Los factores son retenidos hasta donde la pendiente de esta línea se *aproxima a cero*, y en ese punto se produce un *codo* agudo. Eliminar un factor muy por debajo de este codo mostrará escasa pérdida de la varianza explicada**.** Tal procedimiento puede ser particularmente el caso de **PCA**, pero con *análisis de factores comunes,* quizás valga la pena examinar los eigenvalores<1 porque el juicio del *codo* es subjetivo. De hecho, Floyd y Widaman (1995) sugirieron que para el análisis factorial común, cuando dos o más factores están cerca del punto de corte del codo, soluciones alternativos con diferentes números de factores deben ser examinados como un esfuerzo para evitar desechar un factor útil. Así, se tiene a Horn (1965) quien propuso el uso del análisis paralelo (*Paralell Analysis*) para identificar y precisar *el codo*. (Consulte el Capítulo 2, apartado *Análisis Factorial Exploratorio (EFA. Exploratory Factor Analysis)* para más detalles).

Dicho análisis calcula los eigenvalores de las variables para un tamaño de muestra dado, suponiendo que las correlaciones entre las variables, son el resultado únicamente del error de muestreo. Es decir, este análisis proporciona una estimación de los eigenvalores para los *indicadores que no tienen factores comunes.* (La **ecuación 2.3** del Capítulo 2 se puede usar para determinar estos valores propios).

- **Una segunda regla clave para la retención de factores** se refiere a la cantidad de indicadores que cargan sustancialmente en un factor. Lo que se considera sustancial es algo abierto para el debate, pero las carga factoriales en desde .40 y superiores se han clasificado como sustanciales (Floyd y Widaman, 1995), y las cargas superiores a .50 se han considerado como *muy significativas* (Hair et al.,1998). Los factores con una sola carga sustancial serán de poca importancia porque solo se tiene en cuenta la varianza del factor específico asociado con ese elemento, y se ha sugerido que *al menos tres indicadores que cargan factorial sean necesarios para identificar un factor* (Comrey, 1988). Sin embargo, muchos autores sugieren que el tamaño de la muestra debe considerarse al juzgar no solo el tamaño de la carga factorial sino también su estabilidad. Las reglas generales para las técnicas de **EFA** varían desde un *tamaño de muestra mínimo de 100 hasta un tamaño de 200 a 300; Otra recomendación es un tamaño de muestra de 5 a 10 encuestados por indicador* (Clark y Watson, 1995; Comrey, 1988; Floyd y Widaman, 1995; Hair et al., 1998).

- **Una tercera regla clave implica que la cantidad de varianza** se explica por un factor extraído en relación con la varianza total explicada por la solución de factor completa. Algunos defienden que el número de factores extraídos debe representar entre *el 50% y el 60%* de la varianza en los ítems y que para que un factor sea significativo, al menos el *5%* de la varianza total explicada debería ser atribuible a ese factor (Hair et. Al., 1998).

•**Una cuarta regla clave que parece apropiada para el desarrollo de la escala implica un criterio a priori.** La mayoría de los desarrolladores de escala tendrán alguna idea sobre la cantidad de factores (*dimensiones*) que subyacen a un conjunto de indicadores. Por lo tanto, restringir la solución del factor a un número de factores pre-especificado (consistente con la teoría) puede ofrecer información valiosa sobre la cantidad de varianza que los factores explican y la fuerza de las cargas de los indicadores sobre los factores respectivos (Floyd y Widaman, 1995; Hair et al ., 1998). Este enfoque también puede proporcionar información sobre el nivel de carga cruzada de un indicador a un factor que no debe cargar (teóricamente) y, por lo tanto, revelar un elemento que puede ser un candidato para su eliminación. En la práctica, *no se debe confiar en una sola regla o criterio psicométrico al decidir la cantidad de factores que se deben extraer*. La mayoría de los criterios de extracción de **EFA** deben usarse junto con otros criterios, y el desarrollo de escala debería usar una teoría *a priori* y el sentido común como guías para decidir la cantidad de factores a extraer. Recomendamos que se realicen análisis factoriales que restrinjan una solución a un número de factores teóricamente *a priori*, con esos resultados comparados con soluciones no restringidas a un número *a priori*. Nuevamente, los desarrolladores de escala deberían tener alguna idea sobre la cantidad de factores que subyacen a un conjunto de indicadores. Restringir la solución a ese número puede ofrecer valiosa información de diagnóstico.

¿Cómo decidimos el número de factores que se deben extraer? Cuando una gran serie de variables se somete a la extracción de factores:

•**En primer lugar el método extrae las combinaciones de las variables que** *explican la cantidad mayor de la varianza* y después continúa con combinaciones que justifican cantidades de varianza cada vez menores.

•**Para decidir cuántos factores se deben extraer,** Usted empieza generalmente con algún criterio predeterminado, *tal como el porcentaje de varianza o el criterio de raíz latente*, para llegar a un número de factores específico

•**Después de estimar la solución inicial, se calculan varias soluciones de prueba adicionales,** normalmente *un factor menos que el número inicial y dos o tres factores más que los que se estimaron inicialmente.*

•**Posteriormente, en función de la información que se obtiene de estos análisis previos, se *examinan las matrices de factores*** *y se escoge el número de factores que represente mejor los datos.*

•**Por analogía, la elección del número de factores es como enfocar un microscopio.** Un ajuste demasiado alto o demasiado bajo hará más oscura una estructura que es obvia cuando el ajuste es acertado.

•**Al examinar un número de estructuras factoriales diferentes que se derivan de varias soluciones de pruebas,** Usted puede comparar y contrastar para llegar a la mejor representación de los datos.

•**Hasta aquí, se puede decir que todavía no se ha desarrollado una base cuantitativa exacta** para decidir el número de factores a extraer.

## *Descripción de los criterios para el cálculo del número de*

## *factores a ser extraídos*

•**Criterio de raíz latente.** Es a más utilizada, por su sencillez de aplicación. La racionalidad que se usa para el criterio de raíz latente es que cualquier factor individual debería justificar la varianza de por lo menos una única variable. Cada variable contribuye con un valor de **1** para el autovalor total. Por tanto, *sólo se consideran los factores que tienen raíces latentes o autovalores mayores que* **1**; explican al menos una variable, se considera que todos los factores con *raíces latentes menores que l (explican menos de una variable) no son significativas* y por tanto, se desestiman a la hora de incorporarlos a la interpretación.

El uso del *autovalo***r** para establecer un corte es más fiable cuando el *número de variables está entre* **20** *y* **50**. Si el número de variables es menor que **20,** hay una tendencia a que este método *extraiga un número de factores conservador* (demasiado poco); por el contrario, *si hay más de* **50** *variables, lo más común es que se extraigan demasiados factores.*

•**Criterio a priori.** Es un criterio simple y a la vez razonable bajo ciertas circunstancias. Con su aplicación, *Usted ya sabe cuántos factores hay que extraer antes de iniciar el análisis factorial*. Usted simplemente instruye al computador para *parar el análisis cuando se haya extraído el número de factores deseado*. Esta aproximación resulta de utilidad cuando:

-Se prueba una teoría o una hipótesis acerca del número de factores para ser extraído. -También puede ser justificado con el objetivo de replicar el trabajo de otro investigador y extraer el mismo número de factores que se encontraron previamente.

•**Criterio de porcentaje de la varianza.** Es una aproximación que se basa en obtener un porcentaje acumulado especificado de la *varianza total extraída.* El propósito es asegurar una *significación* práctica de los factores derivados, asegurando que explican por lo menos una cantidad especificada de la varianza . No existe al momento, un umbral absoluto para todas las aplicaciones. Sin embargo:

•En las **ciencias naturales**, el procedimiento de factores *normalmente no debería ser detenido hasta que los factores extraídos cuenten por lo menos con un 95 % de la varianza* o hasta que el factor justifique solamente una porción pequeña **(***menos del* **5 %)**.

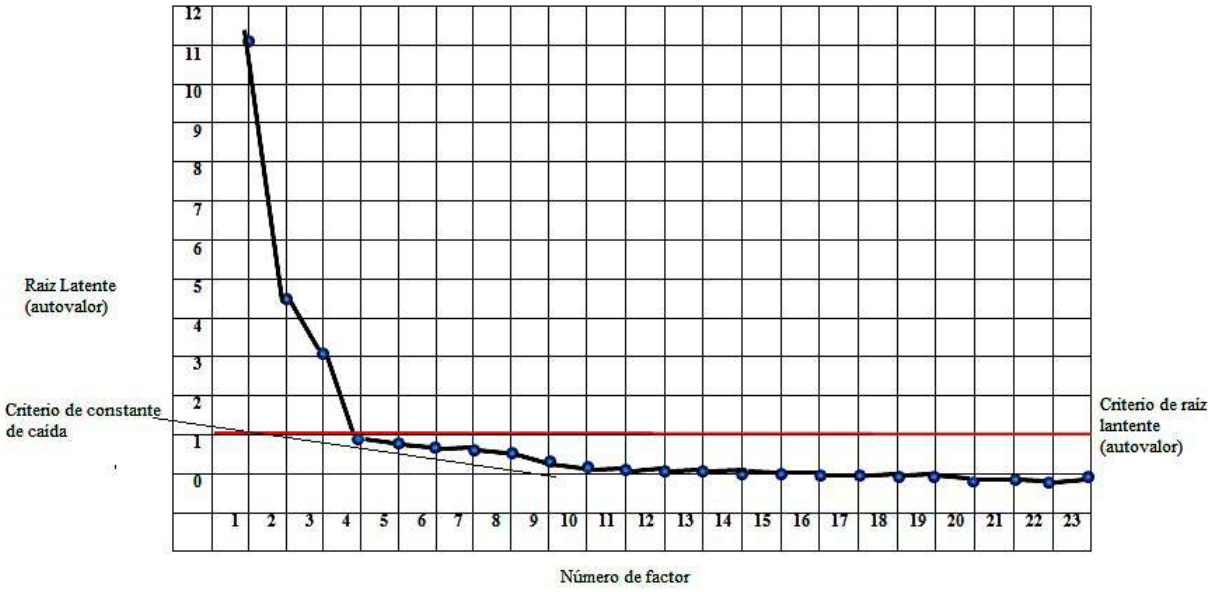
•En las ***ciencias sociales***, donde la información muchas veces es menos precisa, *es normal considerar una solución que represente un 60 % de la varianza total (y en algunos casos incluso menos) como satisfactoria*. Una variante de este criterio implica la selección de suficientes factores para *obtener una comunalidad* para cada una de las variables. Si las razones teóricas o prácticas requieren una cierta varianza compartida para cada variable, entonces la investigación *incluirá tantos factores* como sean necesarios para representar de forma adecuada cada una de las variables originales

•**Criterio de contraste de caída** Recuerde que con el modelo de *análisis de componentes principales*, los factores posteriores que han sido extraídos contienen tanto la *varianza común como la varianza única*. Aunque todos los factores contienen por lo menos alguna *varianza única,* la proporción de la varianza única es sustancialmente *más alta en los factores posteriores que en los primeros*. El contraste de caída se utiliza para *identificar el número óptimo de factores que pueden*

*ser extraídos antes de que la cantidad de la varianza única empiece a dominar la estructura de la varianza común* [Cattell, 1966]. Se estima el contraste de caída con el trazo de *raíces latentes* en función del número de factores en su orden de extracción, y se utiliza la forma de la *curva consiguiente* para evaluar el punto de corte.

La **Figura 3.5** representa los primeros 18 factores extraídos de una investigación realizada.

**Figura 3.5. Gráfico de autovalor para el criterio de contraste de caída.**




Fuente: SPSS 20 IBM

Si empezamos con el primer factor, el trazo tiene inicialmente una inclinación descendente y a continuación se convierte paulatinamente en una línea más o menos horizontal. Se considera que el punto en que *la curva empieza a rectificar se indica el máximo número de factores a extraer.* En el caso que nos ocupa, se incluirán los primeros **10** factores. Por encima de **10**, se incluiría una *proporción de la varianza única demasiado grande, por lo que estos factores no son aceptables***.** Es importante señalar que :

- Con el **uso del criterio de raíz latente**, solamente se tienen en cuenta **ocho**
- Con el **uso del criterio de contraste de caída** nos proporciona **dos factores más.** Por regla general, el contraste de caída normalmente tiene como resultado que se incluyan uno y a veces dos o más factores adicionales que cuando se utiliza el **criterio de raíz latente [**Cattell, 1966**].**

- **Heterogeneidad de la muestra.** La existencia de varianza compartida entre las variables es el núcleo tanto de los modelos *de factores comunes como de los de componentes. Un supuesto subyacente es que la varianza compartida se extiende a lo largo de toda la muestra*. Si la muestra es heterogénea al menos con respecto a un subconjunto de variables, los primeros factores representarán aquellas variables que son más homogéneas a lo largo de toda la muestra. Las variables con mayor capacidad de discriminar entre subconjuntos muestrales cargarán sobre los últimos

factores, en muchas ocasiones aquellos no seleccionados de acuerdo a los criterios comentados más arriba **[**Dillon et al.1989**].** Cuando el objetivo sea identificar factores que discriminen entre subconjuntos muestrales, Usted deberá extraer factores adicionales entre aquellos señalados por los métodos anteriormente expuestos y examinar la capacidad de los factores adicionales para discriminar entre grupos. Si resultan ser peores al discriminar, la solución puede estar en proceder de nuevo y eliminar estos últimos factores.

- **Resumen de los criterios de selección de factores.** En la práctica, rara vez se utiliza un único criterio al determinar cuántos factores extraer. En su lugar, inicialmente se emplea un criterio como el de la *raíz latente como orientación* en un primer intento de interpretación. Después de haber interpretado los factores, como se expone en la siguiente sección*, se valora su carácter práctico***.** También se interpretan los factores identificados mediante otros criterios. Elegir el número de factores *está interrelacionado con la valoración de la estructura***,** lo que se revela en la etapa de *interpretación.* De esta forma*, se examinan varias soluciones factoriales con diferentes números de factores antes de que la estructura esté bien definida.*
- **Nota a la selección del conjunto definitivo de factores:** puede resultar **incoveniente seleccionar tanto muchos como pocos factores** para representar los datos:
  - **Pocos factores**, no se revela la estructura correcta y pueden omitirse importantes dimensiones.
  - **Demasiados factores**, las interpretaciones resultan más difíciles cuando se rotan los resultados. Tal y como ocurre con otros aspectos de los modelos multivariantes, es importante la *parsimonia***.** Una *excepción* a destacar es cuando el análisis de factores se emplea en exclusiva para la reducción de datos y se especifica la extracción de un nivel conjunto de varianza. Usted deberá siempre esforzarse en *conseguir el conjunto de factores lo más representativo y parsimonioso posible.*

# Interpretación de factores

**Paso 5: Interpretación**

Para interpretar los factores y seleccionar la solución factorial definitiva se debe:

- **Calcular la matriz inicial de factores no rotados** para obtener una indicación preliminar acerca del número de factores a extraer. La matriz de factores contiene las cargas factoriales para cada variable sobre cada factor. Al calcularla Usted simplemente *deberá determinar la mejor combinación lineal de variables***,** es decir, *encontrar aquella combinación particular de las variables originales que cuenta con el mayor porcentaje de varianza de los datos*.

En consecuencia, *el primer factor* puede contemplarse como el mejor resumen de las relaciones lineales que los datos manifiestan.

*El segundo facto***r** se define como la segunda mejor combinación lineal de las variables, sujeta a *la restricción de que sea ortogonal al primer factor*. Para ser ortogonal al primer factor, el segundo *factor debe derivarse de la varianza restante tras la extracción del primer factor.* Así, *el segundo factor* puede definirse como la combinación lineal de las variables que da cuenta del mayor porcentaje de varianza

residual una vez se ha eliminado de los datos el efecto del primer factor. Los factores subsiguientes se definen de forma *análoga hasta haber agotado la varianza de los datos. Las soluciones factoriales no rotadas* alcanzan el objetivo de reducción de los datos, pero Usted debe preguntarse si la solución factorial no rotada (que se adecua a los requisitos matemáticos deseables) facilita una información que ofrezca la interpretación más adecuada de las variables examinadas**.** *La mayor parte de las veces no resulta ser así.* Debe recordar que *la carga factorial*: Es el medio para interpretar la función que cada variable desempeña al definir cada factor. Es la correlación o grado de correspondencia entre cada variable y el factor, haciendo a una variable con mayor carga representativa del factor. La *solución factorial no rotada* puede no suministrar un patrón significativo de cargas de las variables. Si se espera que los factores no rotados sean significativos**,** *puede especificar que la rotación no se lleve a cabo. La rotación es deseable* porque simplifica la estructura de los factores, ya que es dificil determinar si los factores no rotados serán significativos.

- **Hacer uso de un método de rotación** para lograr soluciones factoriales más simples y teóricamente más significativas. En muchos casos la rotación de los factores mejora la interpretación disminuyendo alguna de las ambigüedades que a menudo acompañan a las soluciones factoriales inicialmente no rotadas.
- **Valorar** la necesidad de **especificar de nuevo el modelo de factores** debido a:
    - La eliminación de variables en el análisis,
    - El deseo de emplear un método de rotación diferente para la interpretación,
    - La necesidad de extraer un número diferente de fac tores, o
    - El deseo de cambiar de un método de extracción a otro. La especificación nueva del modelo factorial viene acompañada de la vuelta a la etapa de extracción, rotación de factores y de nuevo su interpretación.

## *Métodos de rotación*

Para hacer que los factores sean más interpretables (y que la retención y eliminación de indicadores sea más significativa), los factores se *rotan* después de la extracción. Un objetivo básico para los desarrolladores de escala es buscar una estructura simple después de la rotación. Una estructura simple se produce cuando cada indicador se carga con la mayor cantidad posible de factores o, más preferiblemente, tiene una carga sustancial en un solo factor. La rotación puede especificarse como ortogonal o como oblicua, donde la rotación ortogonal mantiene los factores no correlacionados y la rotación oblicua permite que los factores se correlacionen. **VARIMAX** es la forma más común de *rotación ortogonal para* **EFA** y mostrará una *estructura simple en la mayoría de los casos.* Sin embargo, dado que un objetivo de **EFA** para el desarrollo de la escala es buscar el grado al cual se correlacionan múltiples escalas/dimensiones, se recomiendan métodos de *rotación oblicua* **(como PROMAX)**. La *rotación oblicua* revelará (en la mayoría de los casos) los factores teóricos más significativos.

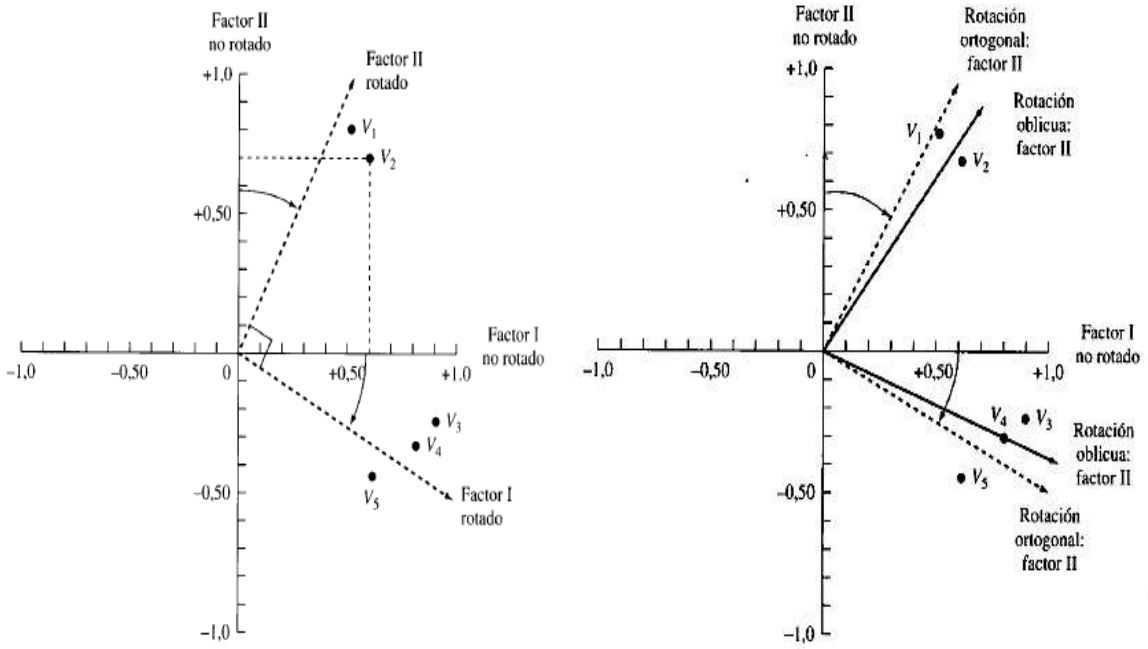
Ambos métodos pueden ayudar a revelar qué elementos conservar y eliminar para futuros estudios, lo que nos lleva a nuestro próximo problema en la aplicación de **EFA** como una herramienta de desarrollo de escala.

## Criterios de rotación de factores

Es una herramienta importante para la interpretación de los factores. Así, *se giran en el origen los ejes de referencia de los factores hasta alcanzar una determinada posición*. Como se indicó previamente, *las soluciones factoriales no rotadas extraen factores según su orden de importancia*:

- **El primer factor** tiende a ser un factor general por el que casi toda variable se ve afectada significativamente dando cuenta del mayor **porcentaje de varianza**. **-El segundo y siguientes factores** se basan en la **varianza residual**. Cada uno explica porcentajes de varianza cada vez menores.
- **El efecto último de rotar la matriz de factores** es redistribuir la varianza de los primeros factores a los últimos para lograr un patrón de factores más simple y teóricamente más significativo.
- **El caso más simple de rotación es la rotación ortogonal,** en la que los ejes se mantienen formando un ángulo de **90** *grados*.
- También es posible rotar los ejes *y no mantener el ángulo de* ***90*** *grados* entre los ejes de referencia. -Cuando no se limita a ser ortogonal*, la rotación se denomina oblicua.* Las rotaciones de factores ortogonal y oblicua están ilustradas en las **Figuras 3.6**, respectivamente.

**Figura 3.6. Rotación factorial ortogonal y Rotación factorial oblicua**




Fuente: propia

De la **Figura 12.6** de rotación factorial ortogonal se han representado cinco variables en un *diagrama de factores bidimensional*., con la siguiente descripción:

- **El eje vertical** representa el **factor no rotado II**, y **el horizontal el factor no rotado I**.
- El **0** indica el origen de coordenadas yendo éstas de **-1.0 a +1.0**.

- **El número sobre los ejes representa las cargas factoriales**.
- Las **5 variables** están denominadas como **VI, V2, V3, V4 y V5.**
- La carga factorial para la **V2** sobre el **factor no rotado II** está indicada horizontalmente mediante una línea discontínua del punto de los datos al eje vertical del **factor II**.
- Análogamente se dibuja una línea vertical de la **V2** al eje horizontal del **factor no rotado I** para determinar la carga de la **V2** sobre el **factor I**.
- Un procedimiento similar para las variables restantes determina las cargas factoriales para las soluciones **no rotadas y rotadas**, como se muestra en la **Tabla 3.1.**

**Tabla 3.1. Comparación entre cargas factoriales rotadas y no rotadas**

| | Cargas factoriales no rotadas | | Cargas factoriales rotadas | |
|---|---|---|---|---|
| | I | II | I | II |
| V1 | 0.5 | 0.8 | 0.03 | 0.94 |
| V2 | 0.6 | 0.7 | 0.16 | 0.90 |
| V3 | 0.9 | -0.25 | 0.95 | 0.24 |
| V4 | 0.8 | -0.30 | 0.84 | 0.15 |
| V5 | 0.6 | -0-50 | 0.76 | -0.13 |

Fuente: elaboración propia

- Sobre el **primer factor no rotado**, todas las variables cargan bastante alto.
- Sobre el **segundo factor no rotado**,**V1** y **V2** cargan muy alto en el lado positivo. La **V5** tiene una carga moderadamente alta en el lado negativo, y las **V3** y **V4** tienen cargas considerablemente inferiores en el lado negativo.
- Como resultado de la inspección de la **Figura 12.7. Rotación factorial ortogonal** se determina que hay dos grupos de variables.**V1 y V2** van juntas, así como las **V3,V4 y V5**.

Sin embargo, este patrón de variables no es tan obvio a partir de las *cargas de factores no rotados.* Rotando los ejes originales en el sentido de las agujas del reloj, como se indica en la **Figura 12.6. Rotación factorial ortogonal**, obtenemos un *patrón de carga factorial completamente distinto.* Nótese que al rotar los factores, los ejes mantienen el ángulo *de **90** grados*. Este procedimiento implica que *los factores son matemáticamente independientes y que la rotación ha sido ortogonal*. Después de rotar el eje de factores, **V3 y V4** cargan muy poco sobre el **factor I**, y **V1** y **V2** cargan mucho sobre el **factor II**. Así, el patrón o agrupamiento de estas variables en *dos grupos* resulta más obvio que antes de la rotación, incluso la posición relativa o con figuración de las variables permanece inalterada.

Los mismos principios generales de las rotaciones ortogonales atañen a las **oblicuas**. El método de la *rotación oblicua es más flexible porque los ejes de factores no necesitan ser ortogonales*. También es más realista porque las dimensiones subyacentes teóricamente más importantes, se su ponen relacionadas entre sí.

En la **Figura 12.6 Rotación factorial oblicua** se comparan los dos métodos de rotación. Observe que la *rotación de factores oblicua* representa el agrupamiento de variables *con más precisión***,** la cual se deriva del hecho de que cada eje de factores

rotado está ahora más cerca del grupo respectivo de variables. Además, la **solución oblicua** provee de información sobre la medida en que los factores realmente están correlacionados uno con otro. La mayor parte de los investigadores están de acuerdo en que *soluciones factoriales no rotadas*, aunque más directas, *no resultan suficientes*; es decir, en muchos casos la rotación mejorará la interpretación paliando alguna de las ambigüedades que a menudo acompañan al análisis preliminar. *Las alternativas disponibles son la rotación ortogonal o la oblicua.*
El objetivo último de toda rotación es obtener algunos *factores teóricamente significativos* y, si es posible*, la estructura de factores más simple. La rotación ortogona***l** se emplea con más frecuencia dada su presencia en todos los programas informáticos de análisis factorial, mientras que *los métodos oblicuos no están tan extendidos.* Además, las *rotaciones ortogonales* se utilizan con más frecuencia porque los *procedimientos analíticos* para llevar a cabo *rotaciones oblicuas no están totalmente desarrollados* y están todavía *sujetos a controversia*. Existen varias aproximaciones distintas *para llevar a cabo rotaciones ortogonales u oblicuas***.** Sin embargo, sólo un número escaso de procedimientos de *cargas factoriales rotación oblicua* está disponible en la mayoría de los programas estadísticos; por eso el investigador tendrá que aceptar probablemente alguno de los provistos.

## *Métodos de rotación ortogonal*

En la práctica, *el objetivo de todos los métodos de rotación es simplificar las filas y columnas de la matriz de factores para facilitar la interpretación*. En una matriz de factores, las columnas representan los factores, con cada fila correspondiendo a las cargas de las variables para cada uno de los factores. Simplificando las filas, queremos *aproximar lo más posible a cero* tantos valores como sea posible *(es decir, maximizar la carga de una variable sobre un único factor*). Simplificando las columnas, queremos *aproximar lo más posible a cero tantos* valores como sea posible (es decir, haciendo que el número de *cargas altas sea el menor posible*). Se han desarrollado principalmente **tres aproximaciones:**

•**QUARTIMAX.** Su objetivo es simplificar las filas de una matriz de factores; esto es, se centra en **rotar los factores iniciales de tal forma que una variable cargue alto sobre un factor y tan bajo como sea posible sobre los otros factores**. En estas rotaciones muchas variables pueden cargar alto o cerca sobre el mismo factor porque la técnica se centra en las filas. El método no ha demostrado gran capacidad para generar estructuras más simples. Su dificultad está en que tiende a producir un factor general, como el primer factor, sobre el que la mayor parte, si no todas las variables, tiene cargas mayores. Con independencia del concepto que cada cual tenga de estructuras ***más simples,*** inevitablemente se ha de tratar con agrupaciones de variables; un método que tiende a producir un factor general grande (por ejemplo, el **QUARTIMAX**) no responde a los objetivos de la rotación.

•**VARIMAX.** En contraste al anterior, se centra en simplificar las columnas de la matriz de factores. Con **VARIMAX** se alcanza la máxima simplificación posible **si sólo hay ceros y unos en una columna**. Esto es, el método maximiza la suma de las varianzas de las cargas requeridas de la matriz de factores. Recuérde que en la aproximación **QUARTIMAX**, muchas variables pueden cargar alto o cerca de lo alto sobre el mismo factor porque la **técnica se centra en simplificar las filas**. Con

**VARIMAX**, tiende a haber **altas cargas factoriales** (esto es, **cercanas a -1 o +1**) y algunas cargas cerca de **0** en cada columna de la matriz. Si la lógica está en que la interpretación es más fácil cuando las correlaciones **variable factor** están:

- **Cercanas a -1 o +1**, indicando así una clara asociación positiva o negativa entre la variable y el factor; o
- **Cercanas a 0** señalando una clara ausencia de asociación. Esta estructura resulta esencialmente sencilla. Aunque la solución **QUARTIMAX** es analíticamente más simple que la solución **VARIMAX**, ésta parece suministrar una separación más clara de factores. En general, el experimento de Kaiser (1974)indica que el patrón factorial obtenido mediante la rotación **VARIMAX** tiende a resultar más robusto que el obtenido por el método **QUARTIMAX** cuando se analizan diferentes subconjuntos de variables. El método **VARIMAX** ha demostrado tener más éxito como aproximación analítica para lograr una rotación ortogonal de factores.
- **EOUIMAX.** Se considera intermedio entre las aproximaciones **QUARTIMAX y VARIMAX**. En lugar de concentrarse bien en la simplificación de las filas, bien de las columnas*, procura cumplir con las dos*. **EQUIMAX** no ha logrado una amplia aceptación y se emplea en muy raras ocasiones.

## *Métodos de rotación oblicua*

Las rotaciones oblicuas *son similares a las rotaciones ortogonales, excepto permiten la existencia de factores correlacionadas en lugar de mantener la independencia entre los factores rotados*. Aunque en la mayor parte de los pro- gramas estadísticos hay varias alternativas de *aproximación ortogonal*, suele haber escasas de rotaciones oblicuas. Por ejemplo, SPSS cuenta con **OBLIMIN**; SAS con **PROMAX** y **ORTOBLIQUE**; y **BMDP** con **DQUART, DOBLIMIN y ORTOBUQUE**.

Los objetivos de simplificación son comparables a los de los métodos ortogonales, con el rasgo añadido de existencia de factores correlacionados. Con esta posibilidad, el investigador ha de tener un cuidado adicional al validar los factores rotados oblicuamente, puesto que cuenta con una forma adicional (*no ortogonalidad*) de proceder, específica a la muestra y no generalizable, especialmente en muestras pequeñas o de bajos ratios *casos/variable***.**

## *Selección del método de rotación*

*No se han desarrollado reglas concretas que guíen al investigador en la selección de una técnica de rotación particular ortogonal u oblicua*. En la mayoría de las ocasiones, el simplemente utiliza la técnica rotacional suministrada por el programa de computador. Muchos programas cuentan por defecto con la rotación **VARIMAX**, pero también resultan fácilmente accesibles los métodos rotacionales más comunes. Sin embargo, *no existe una razón analítica incuestionable a favor de un método de rotación u otro.* La elección de una rotación ortogonal u oblicua debería hacerse según las necesidades concretas de un problema de investigación determinado. Así, de acuerdo al objetivo, se tiene:

- **Si el objetivo del investigador es reducir el número de variables originales,** con independencia de la significación resultante de los factores ó desea reducir un gran número de variables a un conjunto pequeño de variables incorrelacionadas para un

uso posterior en el análisis de regresión u otras técnicas de predicción, la solución ortogonal resulta la más adecuada.

•**Sin embargo, si el objetivo último del análisis factorial es obtener varios factores teóricamente significativos,** resulta apropiada una solución oblicua. Llegamos a esta conclusión dado que, realmente, muy pocos factores están incorrelacionados, como ocurre con la rotación ortogonal.

## *Retención de indicadores*

Dado que **EFA** puede usarse como un método para reducir el número de indicadores en una escala, surgen preguntas sobre cuántos elementos deben eliminarse y qué criterios usar para eliminarlos. Como se indicó anteriormente, la obtención de una estructura simple es el objetivo de **EFA** logrado al buscar cargas que son sustanciales (.**40** y superiores). Este objetivo debe usarse con **EFA** en múltiples conjuntos de datos (es decir, al menos dos). Además, los desarrolladores de escala también deben buscar cargas extremadamente altas, ya que los indicadores con tales cargas pueden ser indicativos de redundancia de redacción que no agrega sustantivamente a la coherencia o validez interna de una escala. Por lo tanto, en general, abogamos por la retención de indicadores a través de múltiples **EFA** *con cargas no inferiores a* ***.40*** *pero no superiores a* ***.90***. En esta etapa, sin embargo, advertimos contra la eliminación de ítems que puedan no cumplir con este criterio, pero se considera que tienen *validez aparente* (*face validity*) y/o contenido (*content validity*). Además, la supresión y retención de ítems en los primeros estudios de desarrollo de la escala, deberían considerar simultáneamente la confiabilidad y las estadísticas tales como las correlaciones de indicadores a total corregidas (*corrected ítem-to-total correlations*), las correlaciones entre elementos promedio (*average interítem correlations)* y las varianzas de ítems (*ítem variances*).

## *Criterios para la significación de la carga factorial*

Al interpretar los factores, debe decidir el criterio en tomo a qué cargas factoriales merece la pena considerar, por lo que se expone a su consideración diversos aspectos relativos a la *significación práctica y estadística*, además de al número de variables, que afectan a la interpretación de las cargas factoriales, como sigue:

•Asegurar la significación práctica. Esta consideración, no está basada en afirmaciones matemáticas, sino que tiene que ver más con la significación práctica. Consiste en un tipo de regla empírica empleado frecuentemente como forma de realizar un examen preliminar de la matriz de factores, donde las cargas factoriales, se consideran:

**> ±:0.30** nivel mínimo; **±:0,40** más importantes; **>=±:0,50** significativas. Así, cuanto mayor sea el tamaño absoluto de la carga factorial, más importante resulta la carga al interpretar la matriz factorial. Dado que la carga fac torial es la correlación entre la variable y el factor, el cuadrado de la carga es la cuantía de la va rianza total de la variable de la que da cuenta el factor. Así, una carga de **0.30** implica aproximadamente una explicación de un 10 %; una carga de **0.50** quiere decir que un 25 % de la varianza es debida al factor. Para que un factor explique un 50% de la varianza ha de contar con una carga que supere el **70 %.** Usted deberá darse cuenta

de que cargas extremadamente elevadas (**>0.80** o más) no son normales y que la significación práctica es un criterio importante. Estas orientaciones son de aplicación cuando el tamaño muestra supera las 100 observaciones.

•**Valoración de la significación estadística.** Como se indicó previamente, la carga factorial representa la correlación entre la variable original y su factor. Al determinar el nivel de significación para la interpretación de las cargas, se debería emplear una aproximación similar a la utilizada para la significación estadística de los coeficientes de correlación. Sin embargo, diversas investigaciones [Cliff y Hamburger,1967] han demostrado que las cargas factoriales cuentan con errores estándar sustancialmente mayores que las correlaciones habituales, por lo que las cargas factoriales deberían evaluarse con niveles considerablemente más estrictos. Usted debe utilizar el concepto de potencia estadística expuesto en el Capítulo 2 para especificar cargas factoriales consideradas significativas según diferentes tamaños muestrales. Con el objetivo establecido en lograr un *nivel de potencia del* **80%,** el uso de un nivel de significación de un **0.05** y la inflación probada de los errores estándar de las cargas factoriales, la Figura 12.8 contiene los ta- maños muestrales necesarios para que cada valor de la carga factorial se considere significativa.

•**Por ejemplo**, en una muestra de **100** observaciones, las cargas factoriales de **0.55** o más son significativas. Sin embargo, en una muestra de **50**, la significación implica una carga factorial de **0.75**.En comparación con la anterior regla empírica que implicaba la significación para cargas del **0.30**, esta aproximación consideraría a una carga de **0.30** significativa si el tamaño muestra fuera de **350** observaciones. Existen varias orientaciones prudentes cuando se comparan con las de la sección previa o incluso con errores estándar asociados a los coeficientes de correlación convencionales. Por ello, estas orientaciones deberían emplearse como punto de partida en la interpretación de las cargas factoriales, considerando significativas cargas factoriales bajas y de forma añadida a la interpretación basada en otras consideraciones. La siguiente sección detalla el proceso de interpretación y la función que pueden desempeñar otras consideraciones. **Ver Tabla 3.2**

**Tabla 3.2. Directrices para la identificación de cargas factoriales significativas basadas en el tamaño muestra**

| Carga factorial | Tamaño muestral necesario para la significación |
|---|---|
| 0.30 | 350 |
| 0.35 | 250 |
| 0.40 | 200 |
| 0.45 | 150 |
| 0.50 | 120 |
| 0.55 | 100 |
| 0.60 | 85 |
| 0.65 | 70 |
| 0.70 | 60 |
| 0.75 | 50 |

**Nota:** La significación se basa en un nivel de significación de 0,05 (ex), un nivel de potencia del 80 por ciento y los errores estándar supuesta mente dos veces mayores que los coeficientes convencionales de correlación.

Fuente: Hair et al. 1999

•**Ajustes basados en el número de variables.** Una *desventaja* de las aproximaciones anteriores es que *no se considera el número de variables y los factores concretos que se analizan.* Se ha comprobado que, a medida que el investigador se *mueve del primer factor a los últimos factores, debe incrementar el grado aceptable para considerar a una carga como significativa***.** El hecho de que la *varianza única y la varianza del error* empiecen a aparecer en los últimos factores significa que debería *incluirse algún ajuste al alza en el nivel de significación* **[**Kaiser, 1970**].** Al decidir qué cargas son significativas *también es importante el número de variables que se analizan*. Según el número de variables analizadas, se incrementa el nivel aceptable para considerar significativa una carga que decrece. *El ajuste por número de variables* crece en importancia según uno se mueve del primer factor extraído a los últimos.

Resumiendo los criterios para la significación de las cargas factoriales, se pueden establecer las siguientes orientaciones:

•**A mayor tamaño muestra, menor puede ser la carga** para ser considerada como **significativa.**

•**A mayor número de variables analizadas, menor ha de ser la carga** para ser considerada como **significativa**;

•**A mayor número de factores, mayor ha de ser el tamaño de la carga** de los últimos factores para considerarse como **significativ**a en la interpretación.

## *Interpretación de la matriz de factores*

La interpretación de las relaciones complejas representadas en la matriz de factores no es una tarea fácil. Sin embargo, siguiendo el procedimiento señalado a continuación, se puede simplificar considerablemente el procedimiento de interpretación, por medio de:

•**El examen de la matriz de cargas factoriales.** Cada columna de números en la matriz de factores representa un factor aislado. Las *columnas* de números son las *cargas factoriales* de cada *variable sobre cada factor*. Con el fin de identificar, el computador normalmente imprime identificando los factores de izquierda a derecha por los números 1, 2, 3, 4, etc. También las *variables* por su número de *arriba a abajo.* Para facilitar aún más la interpretación, *se sugiere escribir el nombre de cada variable en el margen izquierdo* al lado del número de cada variable. Si se ha utilizado una *rotación oblicua*, se presentan *dos matrices de cargas factoriales*:

•La primera es la matriz de *patrones factoriales*, que contiene las cargas que representan la contribución única de cada variable al factor.

•La segunda es la matriz de *estructura factorial*, que contiene las *correlaciones simples* entre variables y factores, pero estas cargas contienen tanto la *varianza única* entre variables y factores *como la correlación entre factores*. Según crece la *correlación entre factores*, es más difícil distinguir qué variables cargan únicamente sobre cada factor en la matriz de estructura factorial. Muchos investigadores suministran los resultados de la *matriz de patrones factoriales.*

•**Identificación de la mayor carga para cada variable.** La interpretación debería comenzar con la *primera variable sobre el primer factor y moverse horizontalmente*

*de izquierda aderech***a**,observandolamayorcargaparacada variable sobre cada factor. *Cuando se identifica la mayor carga (en valor absoluto), debe subrayarse si es significativa.* Entonces la atención debe centrarse en la *segunda variable*, y de nuevo moviéndose de izquierda a derecha *horizontalmente*, comprobar la mayor carga de cada variable sobre cada factor y subrayarla. Este procedimiento debe continuar para toda variable *hasta que todas las variables se hayan subrayado* una vez en la mayor carga sobre un factor. Recuérdese que para *tamaños muestrales menores a* **100**, la menor carga factorial que se considere significativa debería ser en la mayor parte de las ocasiones de **±0,30**. El proceso de subrayar sólo la mayor carga como significativa para cada variable es un ideal que debería perseguirse *pero rara vez se consigue*. Cuando cada variable tiene sólo una carga sobre un factor que es considerado significativo, la interpretación del significado de cada factor se simplifica considerablemente. En la práctica, sin embargo, muchas variables cuentan con *varias cargas de tamaño moderado*, todas las cuales son significativas, y el trabajo de interpretar los factores es mucho más complicado. La dificultad surge porque **una** *variable con varias cargas significativas debe tenerse en cuenta* al interpretar (etiquetar) todos los factores sobre los cuales tiene una carga significativa. Muchas soluciones factoriales no concluyen con una solución de es- tructura simple (una única alta carga para cada variable sólo sobre un factor). Por eso el investigador deberá continuar, tras encontrar la mayor carga para cada variable, evaluando la matriz de factores para encontrar todas las cargas significativas para una variable sobre todos los factores. Por último, *el objetivo es minimizar* el número de cargas significativas sobre cada fila y la matriz de factores (esto es, *hacer que cada variable se asocie sólo con un factor*). Una variable con varias cargas altas es candidata a ser eliminada.

- **Valoración de la comunalidad** Una vez que las variables se han agrupado en sus respectivos factores, Usted debe *examinar la matriz de factores* para *identificar variables que no hayan sido incluidas* en ningún factor. La *comunalidad* representa la *proporción de varianza con la que contribuye cada variable a la solución final*. Debe observar la comunalidad de cada variable para *evaluar si alcanza niveles aceptables de explicación.*
- **Por ejemplo**, puede especificar que al menos *sea explicada la mitad de la varianza de cada variable*. Esto significa *identificar todas las variables con comunalidades menores a 0.50* como carentes de explicación suficiente. Si *hay variables que no cargan* sobre ningún factor o cuyas comunalidades se juzgan **demasiado bajas**, caben **dos alternativas**:
  - **Interpretar la solución tal cual es y simplemente prescindir de esas variables**; o
  - **Evaluar cada una de esas variables para su supresión eventual. Prescindir de variables.** Puede resultar apropiado si el objetivo es únicamente la *reducción de datos,* pero el investigador todavía debe percatarse de que las variables en cuestión están pobremente representadas en la solución factorial. La consideración sobre su eliminación debe depender de la contribución conjunta de las variables para el investigador además del *índice de comunalidad.* Si la variable en cuestión es de escasa importancia para el objetivo del estudio o cuenta con una comunalidad inaceptable, podría ser eliminada y se procedería posteriormente a especificar el

modelo factorial excluyendo esa variable.

• ***Etiquetación de los factores.*** *Cuando se ha obtenido una solución factorial en que todas las variables tienen una carga significativa sobre un factor, el investigador procura atribuir un significado al patrón de cargas factoriales. Las variables con mayores cargas se consideran más importantes y tienen mayor influencia sobre el nombre o etiqueta seleccionada para representar al factor. Por eso, el investigador examinará todas las variables agrupadas en un factor particular y, poniendo mayor énfasis en las variables con mayor carga factorial, intentará asignar un nombre o etiqueta al factor que refleje con precisión las variables cargadas sobre el factor. Los signos se interpretan como otros coeficientes de correlación. Sobre cada factor, signos iguales significan que las variables están positivamente relacionadas, signos opuestos significan que las variables están negativamente relacionadas. En soluciones ortogonales los factores son independientes unos de otros. Por tanto, los signos de las cargas factoriales se relacionan sólo con el factor en el cual apa recen, no con otros factores en la solución. Esta etiqueta no viene asignada por el análisis factorial realizado por el programa de computador; en su lugar, la etiqueta se fabrica intuitivamente de acuerdo a la conveniencia para representar: las dimensiones subyacentes de un factor concreto. El resultado final será el nombre o etiqueta que representa cada uno de los factores obtenidos con tanta precisión como sea posible. En algunas ocasiones, no es posible asignar un nombre a cada uno de los* factores. Cuando sur ge tal situación, el investigador desearía etiquetar un factor o factores derivados de la solución como *indefinidos.* En tales casos el investigador interpreta sólo aquellos factores que son significati- vos y elude aquellos indefinidos o menos significativos. Al describir la solución factorial, el investigador indica que esos factores se obtuvieron pero que eran indefinidos y que sólo se interpretan aquellos factores que representan relaciones significativas.

Como se expuso anteriormente, la selección de un número concreto de factores y el método de rotación están interrelacionados. Se pueden llevar a cabo varias rotaciones adicionales de prueba y comparando la interpretación factorial para las diferentes rotaciones ensayadas, el investigador puede seleccionar el número de factores a extraer. En resumen, la habilidad para asignar algún significado a los factores, o para interpretar la naturaleza de las variables, son consideraciones extremadamente importantes para determinar el número de factores a extraer.

# Validación

**Paso 6: validación** Comprende la evaluación del *grado de generabilidad* de los resultados para la po- blación y la influencia potencial de causas o individuos sobre los resultados globales. Este aspecto es esencial en cada uno de los métodos multivariantes, pero es especialmente relevante en los *métodos de interdependencia* por describir una estructura de datos que debería ser representativa también de la población. El método más directo de validación de los resultados consiste en adoptar una *perspectiva de confirmación*, valorando la *replicabilidad* de los resultados, bien *dividiendo la muestra con los datos originales*, bien con una *muestra adicional. La comparación de los resultados de dos o más modelos factoriales siempre ha sido problemática***.** Sin embargo, existen varias alternativas para realizar una comparación objetiva. El auge del *análisis factorial confirmatorio* (**CFA.** *Confirmatory factor*

*Analysis*) a través de la modelización de *ecuaciones estructurales* supone una alternativa, pero *generalmente es más complicado* y requiere software adicional como **LISREL o EQS [**Bentler, 1992**,** Joreskog y Sorbo 1993**].** Además del **CFA**, se han propuesto otros métodos que van desde un simple índice de adecuación **[**Cattell et al. 1969**]** a programas (**FMATCH**) diseñados especialmente para valorar la correspondencia entre matrices de factores **[**Smith, 1989**].** Estos métodos cuentan con un **uso ocasional**, debido en parte a:

- **La percepción de ausencia de sofisticación y**
- **La no disponibilidad de software** o programas analíticos que automaticen las com- paraciones. Por eso, cuando el **CFA** no es apropiado, estos métodos facilitan una base objetiva para la comparación. Otro aspecto de la *generalización es la estabilidad* de los resultados del modelo factorial, la cual *depende primeramente del tamaño muestra y del número de casos por variable*. Es usual que los investigadores siempre estén obsesionados por contar con el mayor tamaño muestra posible y desarrollar modelos parsimoniosos que *incrementen la ratio-casos-por-variable***.** Si el tamaño de la muestra lo permite, puede *dividir aleatoriamente la muestra en dos subconjuntos y estimar los modelos factoriales de cada uno***.** La comparación de las *dos matrices factoriales* resultantes suministrará una valoración de la robustez de la solución a lo largo de la muestra.

Además de la *generabilidad,* otro aspecto de importancia para la validación del análisis factorial es la *detección de observaciones influyentes*. Las discusiones sobre la *detección de atípicos* y de las *observaciones influyentes en la regresión* se deben aplicar también al análisis factorial. Debe procurar *estimar el modelo con y sin observacione***s** identificadas como *atípicas* para *valorar su influencia* sobre los resultados. Son de aplicación al análisis factorial varias medidas de influencia que reflejan la posición relativa de una observación respecto a las otras (por ejemplo, el *ratio de la covarianza*). Finalmente, se han propuesto métodos específicos de análisis factorial para identificar observaciones influyentes **[**Chatterjee et al.1991**],** pero su complejidad ha restringido su aplicación.

## *Prueba inicial de indicadores y de confiabilidad*

En las etapas más tempranas del desarrollo de la escala, es aconsejable examinar una serie de estadísticas de consistencia interna (*internal consistency*) y las basadas en ítems (*ítem-based)* junto con la **EFA** para usarlas en la toma de decisiones de conservar y/o eliminar indicadores. Estas estimaciones de consistencia interna incluyen: correlaciones interítem promedio de coeficientes alfa (*coefficient alpha average interítem correlations*), correlaciones de indicadores a total corregidas (*corrected ítem-to-total correlations*), varianzas de ítems (*ítem variances*) y redundancia de redacción de ítems (*ítem-wording redundancy*). Se han defendido varias heurísticas o reglas generales para los ítems como candidatos para la retención/eliminación. Robinson et al. (1991) abogó por una *correlación interítem promedio de* ***.30*** *o mejor como ejemplo. Clark y Watson (1995) recomendaron correlaciones interítem promedio de* ***.15 a .50*** *a través de constructos; por constructos estrechamente definidos, abogaron por un rango* ***de .40 a .50****. También propugnaron un nivel de coeficiente alfa* **(α)** *de al menos .80 para una nueva escala y sugirieron que retener ítems con mayores variaciones en relación con otros ítems*

*ayudaría a aumentar la varianza de la escala global.*
Bearden y Netemeyer (1998) *abogaron por correlaciones de elemento a total corregidas (corrected ítem-to-total correlations) de* ***.50*** *y superiores y niveles alfa* **(α)** *de .80 o superiores (sujeto a consideraciones de longitud de escala).* Estas heurísticas tienen sentido, pero también se debe recordar que el coeficiente alfa **(α)** está relacionado con la longitud de la escala, la correlación interítem media (*average interítem correlation*, o covarianza), la redundancia de ítems y la dimensionalidad.
Así que, considere primero, la longitud de la escala. Se observó en el Capítulo 3 que *a medida que aumenta el número de indicadores, alfa* **(α)** *tenderá a aumentar.* Debido a que la parsimonia también es una preocupación en la medición y la mayoría de las escalas son autoadministradas, la brevedad de la escala es a menudo una preocupación.
La redundancia de elementos y la correlación interítem media (*average interítem correlation* o covarianza) también afectan el coeficiente alfa **(α)**. Aunque se defiende que se necesitan varios indicadores para aprovechar adecuadamente el dominio del constructo, cuando la redacción de los elementos es demasiado similar, el coeficiente alfa **(α)** (así como la validez y la dimensionalidad del contenido) puede no mejorarse. Los indicadores que están redactados de manera similar aumentarán la correlación entre elementos promedio y el coeficiente alfa **(α)**; sin embargo, tales elementos pueden solo contribuir a la paradoja de atenuación (*atenuation paradox*) en la teoría psicométrica, por lo que *aumentar el coeficiente de alfa* **(α)** *más allá de cierto punto no aumenta la consistencia interna* (ver el Capítulo 3).
Con estas consideraciones en mente, los investigadores deben ser cuidadosos en su interpretación de alfa **(α)** y considerar su relación con el número de indicadores en una escala, el nivel de correlación interítem y la redundancia de indicadores.
Por último, y como se discutió en los Capítulos 2 y 3, también debe recordarse que *es posible que un conjunto de indicadores esté interrelacionado pero no sea unidimensional. El coeficiente alfa* **(α)** *no es una medida de unidimensionalidad y debe usarse para evaluar la consistencia interna solo después de que se haya establecido la unidimensionalidad* (Clark y Watson, 1995; Cortina, 1993).

# Usos adicionales de los resultados

**Paso 7:**
Dependiendo de los objetivos de partida al aplicar el análisis factorial, puede *detenerse en la interpretación de los factores o proseguir con uno de los métodos de reducción de datos***.** Si el objetivo simplemente consiste en identificar combinaciones lógicas de variables y una mejor comprensión de las relaciones entre variables, *entonces basta con la interpretación de los factores*. Ésta facilita una *base empírica* para enjuiciar la estructura de las variables y la influencia de esta estructura cuando se interpretan los resultados a partir de otras técnicas multivariantes. Si el objetivo, sin embargo, es identificar variables apropiadas para aplicaciones subsiguientes de otras técnicas estadísticas, se empleará alguna forma de reducción de datos, como:

- Examinar la matriz de factores y seleccionar la variable con mayor carga factorial como un representante de una dimensión factorial particular, o
- Reemplazar el conjunto original de variables por uno totalmente nuevo con menos

variables creado a partir de escalas aditivas o de la puntuación de factores.
*Cada alternativa suministrará nuevas variables para ser usadas*, por ejemplo, como variables *independientes* en una *regresión* o en el *análisis discriminante*, o como variables *dependientes* en el **MANOVA**, o incluso como las variables agrupadas en el *análisis cluster.*

## Selección de variables suplentes para el análisis subsiguiente

Si el objetivo del investigador es sencillamente *identificar variables* apropiadas para la aplicación subsiguiente de otras técnicas estadísticas, cuenta con la alternativa de *examinar la matriz factorial y seleccionar la variable con la mayor carga factorial* sobre cada factor para que actúe como variable suplente del factor. Este enfoque es simple y directo sólo cuando una variable tiene una carga factorial que es sustancialmente mayor que las otras cargas factoriales. En muchas ocasiones, sin embargo, el proceso de selección es mucho más difícil *porque dos o más variables tienen cargas que son significativas y bastante cercanas entre sí. Estos casos requieren un examen crítico* de las cargas factoriales de tamaño aproximado y sólo una como representativa de una dimensión concreta. Esta decisión debería basarse en el conocimiento previo de la teoría por parte del investigador que pueda sugerir que una variable con preferencia a otra pueda ser representativa de la dimensión. Además, puede contar con un conocimiento que le sugiera que una variable con una carga ligeramente inferior es de hecho más fiable que la variable con la mayor carga. En tales casos, puede elegir la variable con carga ligeramente inferior como la mejor variable suplente de un factor concreto.

## Problemas

La aproximación de seleccionar una *única variable como representativa del factor* **(**aunque simple y manteniendo la variable original) *cuenta con varios inconvenlentes potenciales.*

- **En primer lugar, no orienta sobre el error de medida** que aparece cuando se emplean medidas únicas (véase la siguiente sección para una discusión más detallada) y *se corre, además, el riesgo de confundir potencialmente los resultados* seleccionando sólo una única variable para representar un resultado posiblemente más complejo.
- **Por ejemplo,** suponga que las variables que representan diseño de página web, analítica web, precio y valor, cargan en varios factores. La selección de cualquiera de estas variables aislada daría lugar a interpretaciones sustancialmente distintas en cualquier análisis subsiguiente, aunque las **4** pueden estar tan próximamente relacionadas como para excluir tal acción.
- **En segundo lugar,** en casos donde varias cargas elevadas complican la selección de una única variable, el investigador puede no tener otra elección que la de emplear el análisis factorial como la base para *calcular una escala aditiva o la puntuación de factores* para usar como variables suplentes. El objetivo, como en el caso de seleccionar una única variable, es representar de la mejor forma posible la naturaleza básica del factor o componente.

## Creación de escalas aditivas

El concepto de escala aditiva se define como aquella escala que está formada por la combinación de varias variables individuales dentro de una única medida compuesta, esto es, *se combinan todas las variables que cargan alto sobre un factor*, y el total (o más normalmente la *puntuación media de las variables*) se emplea como *variable de sustitución.* Una *escala aditiva* cuenta con *tres ventajas concretas*:

- Es una forma de *eludir en alguna forma el error de medida* inherente a todas lasvariables observadas. El error de medida es el grado en el cual los valores observados no son representativos de los valores *verdaderos* debido a cierto número de razones, desde *errores reales* (por ejemplo, errores en la entrada de los datos) *a la incapacidad de los individuos de suministrar información con precisión*. El error de medida *enmascara cualquier relación* (por ejemplo, correlaciones o comparación de medias de grupos) y dificulta la estimación en los modelos multivariantes. La *escala aditiva reduce el error de medida* utilizando *indicadores múltiples (variables)* para *reducir la desconfianza* sobre una única respuesta. Empleando la *media* o la respuesta *típica* a un conjunto de variables relacionadas, *el error de medida que podría tener lugar en una única respuesta se reducirá.*
- **Su capacidad para representar los múltiples aspectos de un concepto** en una única medida. En muchas ocasiones empleamos más variables en nuestros modelos multivariantes en un intento de representar las *muchas facetas* de un concepto que sabemos es bastante complejo. Pero al actuar así, complicamos la interpretación de los resultados debido a la redundancia de la información asociada con el concepto. Por eso, nos gustaría no sólo obtener una descripción mejor de los conceptos utilizando múltiples variables, sino también man tener la parsimonia en el número de variables de nuestros modelos multivariantes. La escala aditiva, cuando se construye apropiadamente, *combina los múltiples indicadores en una medida única representando lo que se mantiene en común a lo largo del conjunto de medidas*.
- El proceso de *construcción de la escala* está fundamentado teórica y empíricamente en una serie de disciplinas que incluyen *la teoría psicométrica, la sociología y el marketing, etc*. Aunque un tratamiento completo de las técnicas y aspectos involucrados están más allá del alcance de este libro, existen fuentes excelentes para un estudio más extenso de estas materias [American Psychological Association 1985, Churchill, 1979, Hattie, 1985, Peter, 1981, Robinson et al.1991]. Adicionalmente hay una serie de compilaciones de escalas existentes que pueden aplicarse en varias situaciones [Bearden et al. 1993, Bruner y Hensel 1993, Robinson y Shaver 1973]. Aquí expondremos, sin embargo, *cuatro aspectos básicos* en la construcción de cualquier escala aditiva: *la definición conceptual, la dimensionalidad, la fiabilidad y la validación.*
- **Definición conceptual.** Es el punto de partida para construir una escala aditiva La definición conceptual especifica las bases teóricas de la escala aditiva definiendo elconcepto que es representado en términos aplicables al contexto de investigación. En la investigación académica, las definiciones teóricas están basadas en investigación previa que define el carácter y naturaleza de un concepto. En un ámbito de gestión empresarial, los conceptos concretos pueden definirse con relación a los objetivos propuestos, tales como la imagen, el valor o la satisfacción. En

cualquier caso, la definición conceptual es la que orienta y concreta el carácter y tipo de ítems que son candidatos a ser incluidos en la escala.

• **La validación del contenido** es la evaluación de la correspondencia de las variables incluidas en la escala aditiva con su definición conceptual. Esta forma de validación, también conocida como la *validación aparente*, sirve para apreciar subjetivamente la correspondencia entre los ítems individuales y el concepto a través de evaluaciones de expertos, contrastes previos con múltiples subpoblaciones, u otros medios. *El objetivo es asegurar que los ítems de la escala abarquen algo más que aspectos empíricos, también consideraciones teóricas y prácticas* **[**Churchill, 1979**,** Robinson et al.1991**].**

• **Dimensionalidad.** Un supuesto subyacente y requisito esencial para construir una escala aditiva es que *los ítems sean unidimensionales*, o sea, *fuertemente asociados unos con otros representando un único concepto* **[**Hattie, 1985**,** McDonald, 1981**]**. El análisis factorial sirve de <1;poyo realizan do una valoración empírica de la dimensionalidad del conjunto de ítems determinando el número de factores y las cargas de cada variable sobre el factor o factores. El contraste de unidimensio nalidad *consiste en que en una escala aditiva los ítems carguen de forma elevada en un único factor* **[**Anderson et al. 1987**,** Hattie,1985**,** McDonald,1981**,** Nunnally, 1979**]**. Si se propone que una escala aditiva cuente con múltiples dimensiones, *cada dimensión reflejará un factor aislado***.** El investigador puede evaluar la *unidimensionalidad* bien con un análisis *factorial exploratorio*, como se discutió en este capítulo, o bien un *análisis factorial confirmatorio*.

• **La fiabilidad** es el *grado de consistencia entre las múltiples medidas de una variable.* Una forma de fiabilidad es el *test-retest* por el cual la consistencia se mide entre las *respuestas de un individuo en dos momentos del tiempo*. El objetivo es asegurar que las respuestas no varían demasiado a lo largo de períodos temporales por lo que una medida tomada en cualquier momento del tiempo es certera.

Una *segunda medida de la fiabilidad* más utilizada es la *consistencia interna* que se aplica a la *consistencia entre las variables en una escala aditiva.* La motivación para la consistencia interna es que los ítems individuales o indicadores de la escala deberían estar midiendo las mismas construcciones y, de esta forma, estar altamente intercorre lacionadas **[**Churchill, 1979**,** Nunnally, 1979**].** Debido a que no hay un único ítem que sea una medida perfecta de un concepto, debemos de legar en un conjunto de medidas de diagnosis para valorar la consistencia interna. En primer lugar, existen varias medidas que se relacionan con cada ítem aislado, incluyendo la correlación ítem-total (la correlación del ítem con la puntuación de la escala aditiva) o la correlación inter-ítem (la correlación entre ítems). La práctica empírica sugiere *que la correlación ítem-total exceda de* **0.50** y que *las correlaciones inter-ítem excedan de* **0.30** **[**Robinson et al.1991**].**

Otro tipo de medida de diagnosis es el *coeficiente de fiabilidad* que valora la consistencia de la escala entera, el *alfa de Cronbach* **[**Nunnally, 1979**,** Peter, 1979**],** que es la medida más extensamente utilizada. El acuerdo general sobre el límite inferior para al alfa de Cronbach es de **0.70** **[**Robinson et al.1991**,** Robinson y Shaver 1973**],** aunque puede bajar a **0.60** en la investigación exploratoria **[**Robinson et al.1991**]**. El **alfa de Cronbach** tiene una relación positiva con *el número de ítems de la escala.*

Debido a que al *aumentar* el número de ítems, incluso con el *mismo grado de intercorrelación, se incrementará el valor de la fiabilidad,* los investigadores deben *imponer requisitos más restrictivos para escalas con un mayor número de ítems.* También están disponibles medidas de fiabilidad derivadas del *análisis factorial confirmatorio.* Dentro de ellas están *la fiabilidad compuesta y la varianza media extraída*, ambas discutidas con mayor detalle la técnica de *ecuaciones estructurales.*
Cada uno de los programas estadísticos más utilizados contiene ahora módulos de *evaluación de la fiabilidad*, de tal forma que el investigador está provisto de un análisis completo de medidas tanto específicas de cada ítem *como medidas de fiabilidad globales. En toda escala aditiva debe analizarse la fiabilidad* para asegurar su adecuación antes de proceder a una evaluación de su validación.

- **Validación.** Habiendo asegurado que una escala: se adecúa a su definición conceptual; es unidimensional, y alcanza los niveles necesarios de fiabilidad.

Usted debe realizar una evaluación final: *la validación de la escala***.** La validación es la medida en que una escala o un conjunto de medidas representa con precisión el concepto de interés. Ya hemos comentado una forma de validación (*de contenido o validación por confrontación*) en relación a las definiciones conceptuales. Otras formas de validación se miden empíricamente por la **correlación** entre los conjuntos de variables definidas teóricamente. Las *3 formas más extensa mente admitidas de validación son: la convergente, la discriminante y la nomológica* **[**Campbell y Fiske 1959**,** Peter, 1981**]**.

- **La validación convergente** valora el grado en el cual dos medidas del mismo concepto están correlacionadas. Aquí debe buscar medidas alternativas de un concepto y correlacionarlas con la escala aditiva. Altas correlaciones indican que la escala está midiendo el concepto que se pretendía.
- **La validación discriminante** es el grado en el cual dos conceptos conceptualmente parecidos difieren. El contraste empírico de nuevo es la correlación entre las medidas, pero esta vez la escala aditiva está correlacionada con una medida parecida, pero conceptualmente distinta. Ahora la correlación debería ser baja, demostrando que la escala aditiva es conceptualmente distinta de otro concepto parecido.
- Finalmente, la **validación nomológica** refleja el **grado** en que la escala aditiva hace *predicciones precisas de otros conceptos en un modelo basado en la teoría*. Así, debeidentificar relaciones apoyadas en investigación previa o en principios aceptados y *evaluar si la escala cuenta con las correspondientes relaciones.*
- En resumen:
  - La **validación convergente** confirma que la escala está correlacionada con otras medidas conocidas del concepto,
  - La va**lidación discriminante** asegura que la escala es suficientemente distinta de otros conceptos parecidos que sean distintos, y
  - La **validación nomológica** determina si la escala demuestra las relaciones cuya existencia se deriva de la teoría y/o de investigación previa***.***

Existen varios método**s** para evaluar la *validació***n**, que van desde *el multirasgo, las matrices multimétodo* **(MTMM)** a las aproximaciones basadas en *ecuaciones*

*estructurales***.** Aunque vayan más allá del alcance de este libro, existen una serie de fuentes que orientan sobre el conjunto de métodos disponibles y de los aspectos implicados por las técnicas concretas **[**Campbell y Fiske 1959**,** Joreskog, y Sorbo 1993**,** Peter, 1981**].**

*Las escalas aditivas*, cuentan con aplicación creciente en investigación aplicada y también en gestión empresarial. La capacidad de la escala aditiva para plasmar conceptos complejos en una única medida *reduciendo el error de medida supone un añadido valioso en cualquier análisis multivariante***.**

*El análisis factorial* ofrece al investigador una *evaluación empírica* de las relaciones entre las variables, esencial en la formación de los fundamentos conceptuales y empíricos de la escala aditiva por medio de la evaluación de la validación del contenido y la dimensionalidad de la escala.

## Cálculo de la puntuación factorial

Otra alternativa para crear un conjunto más pequeño de variables es *reemplazar el conjunto original por medio del cálculo de la puntuación factorial, que constitu- ye también medidas compuestas de cada factor calculadas para cada sujeto. La puntuación factorial* representa el grado en el *cual cada individuo puntúa en el grupo de ítems que cuentan con cargas elevadas sobre un factor. Por tanto, mayores valores en las variables con altas cargas sobre un factor resultarán en una mayor puntuación factorial.*

Una diferencia clave que de la *puntuación factorial vs. la escala aditiva* es que la primera se calcula con base a las *cargas factoriales*, mientras que la segunda lo hace *combinando sólo las variables elegidas*. En consecuencia, aunque el investigador sea capaz de caracterizar un factor por medio de las variables con cargas mayores, *se debe también atender a las cargas de las otras variables*, aunque menores, y su influencia sobre la puntuación factorial.

La mayoría de los programas estadísticos puede calcular *puntuaciones factoriales* fácilmente para cada elemento muestral. Seleccionando la alternativa de *puntuación factorial,* se graban estas puntuaciones para su uso en análisis subsiguientes. *Una desventaja de la puntuación factorial es que no se pueden replicar fácilmente en otros estudios debido a que están basados en la matriz factorial obtenida en cada estudio.* La replicación de la misma matriz factorial en distintos estudios requiere un importante trabajo de programación.

## Selección entre los 3 métodos

Para elegir entre **las 3 alternativas (selección e variables suplentes para el análisis subsiguiente/creción de escalas aditivas/cálculo de la puntuación factorial)** de **reducción de datos** deberá tomar una serie de decisiones:

- Seleccione *una única variable suplente* para cada factor *o calcule una medida compuesta*. La *variable suplente* **única** tiene la *ventajas* de ser der *un sencillo tratamiento e interpretación*, pero tiene las *desventajas de no representar las otras facetas del factor y su propensión al error de medida.*
- Si opta por *calcular una medida compuesta***,** debe elegir entre la *puntuación factorial y las escalas aditivas. Puntuación factorial* tiene la *ventaja* de representar un

*compuesto de las cargas de las variables sobre el factor,* aunque esto supone también una *desventaja potencial* al contar todas las variables con algún *grado de influencia en el cálculo de la puntuación factorial y hacer la interpretación más dificil.* La *escala aditiva* está a medio camino entre *variable suplente y la alternativa de la puntuación factorial.* Es una medida compuesta, como la puntuación factorial, *reduciendo por tanto el error de medida* y representando *múltiples facetas de un concepto***.** Aunque similar a la aproximación de variable suplente incluye sólo las variables que cargan alto sobre el factor y excluye aquellas con escasos efectos**.** Además, su fácil replicación entre muestras es similar al enfoque de *variable suplente.*

- Finalmente, como las **variables suplentes**, las **escalas aditivas** no son necesariamente **ortogonales**, mientras que **los factores pueden ser ortogonales o incorrelados**, si se necesita evitar complicaciones en su uso en otras técnicas multivariantes.
- La regla de decisión, es que:
  - *Si los datos se emplean sólo en la muestra original o se debe mantener la ortogonalidad, la puntuación factorial es la adecuada*.
  - Si se desea la *transferibilidad o la generalización, las escalas aditivas o las variables suplentes son más apropiadas*. -
  - Si la *escala aditiva* está bien construida, validada y es fiable, es probable que sea la mejor alternativa.
  - Si la *escala aditiva* no está contrastada y revisada, con poca o ninguna prueba de fiabilidad o validación, deberían considerarse en su lugar las variables suplentes si no es posible un análisis añadido que *mejore la escala aditiva*.

# Análisis factorial: Resumen para aplicar

**¿Qué es el análisis factorial?**

- El *análisis factorial* tiene como objetivo principal simplificar las múltiples y complejas relaciones que puedan existir en un conjunto de variables observadas $X_1$ , $X_2$ , $X_3$ ,..., $X_N$ . Para lograrlo, trata de encontrar dimensiones comunes o factores que ligan a las variables aparentemente no relacionadas.
- Concretamente, se trata de encontrar un conjunto de **k < p** factores no directamente observables $F_1, F_2, F_3,...,F_k$ que expliquen suficiente a las variables observadas perdiendo el mínimo de información, de modo que sean fácilmente interpretables (*Principio de Interpretabilidad*), y que sean los menos posibles, es decir, k pequeño (*Principio de Parsimonia*).
- Además, los factores han de extraerse de forma que *resulten independientes entre sí*, es decir, que sean *ortogonales*. En consecuencia, el *análisis factorial* es una técnica de reducción de datos que examina la interdependencia de variables, y proporciona conocimiento de la estructura subyacente de los datos.
- Es un conjunto de métodos estadísticos que aborda el problema de *cómo analizar la estructura de las interrelaciones (correlaciones) entre un gran número de variables*, con la definición de *dimensiones subyacentes comunes llamadas factores.*

***¿Para qué se usa?***

- Simplificar un conjunto de datos reduciendo el número de variables,bien por un exceso de variables que dificulta el análisis de la información, o bien por representar la misma información de maneraredundante.
- Encontrar la estructura subyacente en los datos analizados
- En ocasiones se utiliza como un método intermedio para dejar losdatos listos en el empleo de otro método siguiente.
- Se pierde información, pero se gana en la facilidad de interpretación al menor costo posible

**Tipos de Análisis Factorial**

- **Exploratorio:** Se utilizara principalmente para identificar factores,sin restricciones o hipótesis previas, y se caracteriza porque no se conoce a priori el número de factores, y es en la aplicación empírica donde se determina este número.
- **Confirmatorio:** Se utilizan cuando se desea verificar la existencia de una estructura subyacente en los datos, anticipada hipotéticamente, y los factores están fijados a priori, utilizándose contrastaciones empíricas para su corroboración.
- Cuando por *razonamiento teórico, por experiencias u otras* investigaciones similares *se formulan hipótesis sobre la dimensionalidad o estructura subyacente de un fenómeno*, es entonces cuando se utiliza el análisis factorial para confirmar**.**
- El análisis factorial parte de la distinción de la variabilidad de las variables observadas y, por lo tanto de la varianza de las mismas variables, en dos tipos esenciales:
    - **Una parte común:** explicada por un conjunto de factores comunes que afectan a todas las variables. Esto es en el entendido que no captan toda la variabilidad, sino sólo la común.
    - **Una parte específica:** exclusiva para cada variable y sin relación con las demás, explicadas por factores específicos o únicos que informan sobre la especificidad o unicidad de cada variable. Son *factores independientes y ortogonales.*

**Etapas del Análisis Factorial**

- **Primera Etapa:** el propósito general del análisis factorial es encontrar la forma de condensar la información contenida en un amplio número de variables en la menor cantidad posible de factores.
- **Segunda Etapa**: esta etapa involucra *tres decisiones básicas*:
    - Calcular la *matriz de correlación* entre las variables;
    - **Medición** de las variables y;
    - El **tamaño de la muestra** necesaria.
- **Tercera Etapa:** esta etapa se centra en las *características y composición de las variables*, como la *multicolinealidad, la matriz de correlación anti-imagen* y la **prueba de esfericidad de *Bartlett.***
- **Cuarta Etapa:** esta etapa involucra *dos decisiones básicas***:**

- El **método de extracción** de los factores y;
- El **número de factores seleccionados** para mejorar el ajuste de los datos.

- **Quinta Etapa:** la interpretación de los factores es una etapa crucial, y se realiza a través de:
  - La estimación del **factor de la matriz;**
  - La rotación de los factores y;
  - La reespecificación.
- **Sexta Etapa:** esta etapa involucra el grado de generalización de los resultados obtenidos a la totalidad de la población objeto de estudio.
- **Séptima Etapa:** esta etapa proporciona las bases empíricas para el análisis de la estructura de las variables utilizadas, y el impacto en la aplicación de otras técnicas multivariables.

**Análisis Previo de los Datos:**

- Anteriormente ya se ha indicado que es necesario dar una serie de pasos previos antes de aplicar una técnica multivariable determinada. Algunos de ellos tienen que ver con la propia técnica y la comprobación del cumplimiento de sus hipótesis subyacentes: *normalidad, homoscedasticidad y linealidad*; otras comprobaciones son, incluso, previas al uso de la técnica y tienen que ver con la fiabilidad de los datos de partida: existencia de valores perdidos y de observaciones anómalas.
- Asimismo, debe señalarse que algunas de las técnicas de análisis que se expondrán en apartados posteriores, tienen sus propios procedimientos para la comprobación del cumplimiento de sus hipótesis o, por ejemplo, la detección de las observaciones anómalas, y así serán presentadas en su momento (por ejemplo, en la regresión lineal múltiple).

Aunque la **EFA** y las estadísticas tanto de consistencia interna (*internal consistency*) como las basadas en ítem (*ítem-based*) pueden usarse en los primeros estudios de desarrollo de escala como medios para decidir qué ítems retener y eliminar, siempre es mejor retener muchos ítems en esta etapa para estudios posteriores. Es decir, los indicadores que no cumplen ciertos criterios estadísticos o reglas generales, pero que tienen validez aparente (*face validity*) /o contenido (*content validity*) deben mantenerse para la próxima *ronda* de estudios. Si continúan funcionando mal, siempre se pueden eliminar al derivar la forma final de la escala.

# Caso de estudio

## *Problema 1*

**Paso 1: Objetivos**

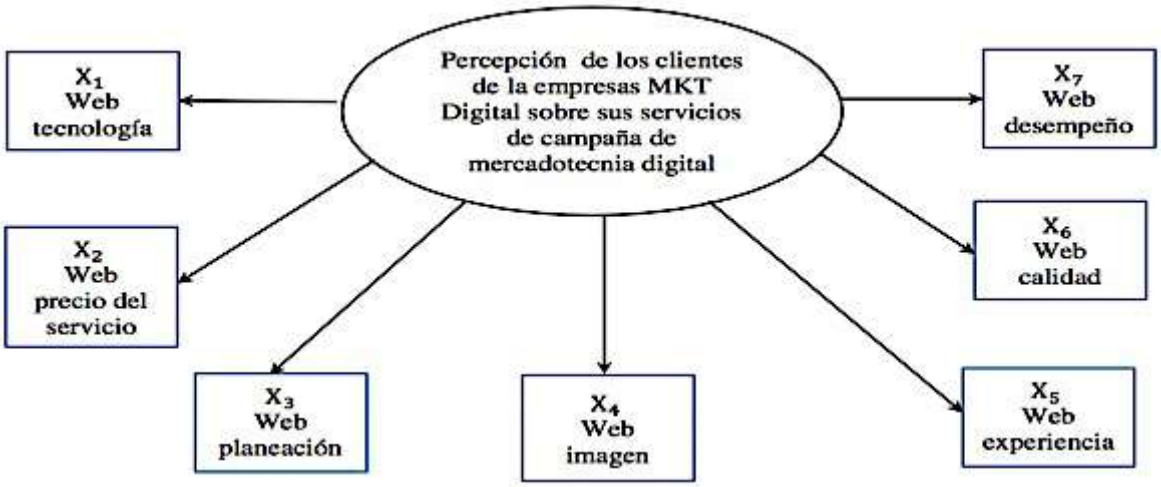
Para explicar de manera más sencilla y clara las diversas técnicas multivariables, se utilizará la base de datos de **WEB Diseño.sav,** de la **compañía MKT Digital**, la cual contiene datos de **100 clientes** a los que ha servido en sus campañas digitales., con la siguiente estructura de base de datos (Ver **Figura 3.3):**

Así, se plantea la siguiente pregunta: ¿Cuál es la percepción que tienen los clientes de la empresa MKT Digital acerca de los servicios ofrecidos en sus campañas de mercadotecnia digital, en 7 atributos($X1$ a$X7$)d eefectividad?.Ver **Figura 3.4.**

**Figura 3.3. Base de datos y diagrama de trayectoria. Caso de estudio inicial .**

| Descripción de variables | Tipo |
|---|---|
| **Clasificación de variables** | |
| V1.- Tipo de consumidor | No métrico |
| V2.- Tipo de Industria | No métrico |
| V3.- Tamaño de la firma | No métrico |
| V4.- Región | No métrico |
| V5.- Sistema de contratación | No métrico |
| **Variables de efectividad** | |
| X1.- Web tecnología | Métrico |
| X2.- Web precio del servicio | Métrico |
| X3.- Web planeación estratégica | Métrico |
| X4.-Web imagen | Métrico |
| X5.-Web experiencia del usuario | Métrico |
| X6.-Web calidad | Métrico |
| X7.-Web desempeño | Métrico |
| X8.-Web analítica | No métrico |
| X9.-Web contrataciones de clientes | Métrico |
| X10.-Web desempeño | Métrico |
| X11.- Web certificación | No métrico |
| X12.-Web plan de implementación | No métrico |
| X13.-Web seguridad | No métrico |
| X14.- Web situación de compra | No métrico |
| **Variables de resultado** | |
| V15.- Probabilidad de recomendación | Métrico |
| V16.- Probabilidad de compra futura | Métrico |
| V17.-Nivel actual de compra | Métrico |
| V18.-Futuras asociaciones/alianzas | No métrico |

**Figura 3.4. Diagrama de trayectoria**




Fuente: elaboración propia

**Paso 2: Diseño**

•La base de datos contiene **100** observaciones con un total de **23** variables, de las cuales se trabajará con **7/14** de las que corresponden al grupo de **Efectividad**. La base de datos **WEB diseño.sav** contiene las percepciones que tienen los **100**

principales clientes

- Asimismo, se recolectó la información de percepciones a través de focus groups y entrevistas a profundidad aplicados a los principales clientes de la empresa MKT Digital
- Los gerentes de mercadotecnia digital, contestaron los **13 atributos** utilizando una escala de **0 – 10**, siendo **10 Excelente y 0 Pésimo**, **con 1 decimal**.
- La información contempla aspectos como: tecnología, precio, planeación estratégica, imagen, experiencia del usuario, calidad, desempeño implicados en una campaña digital
- Con el resultado del análisis factorial, se espera detectar el mejor agrupamiento de éstas variables que le permitan al sector de mercadotecnia digital, ser más competitivo y prestar mejor servicio al cliente
- **Valores Perdidos:** La existencia de valores perdidos en un estudio de las Ciencias Sociales es algo prácticamente inevitable. Las consecuencias para la investigación dependerá del patrón que sigan estos datos ausentes, cuántos son y por qué están perdidos.
- Como señalan Tabachnick y Fidell (1996), el patrón de los valores perdidos es más importante que su cuantía; pues si su distribuciones aleatoria en la matriz de datos no pueden causar mucho daño al análisis; sin embargo, si responden a un patrón determinado sí.
- **Valores perdidos menores o iguales al 10%** del total de los casos por lo general son ignorados y sustituidos por la media o la moda, según sea el caso, excepto cuando los valores perdidos se concentran en una pregunta determinada o un grupo de preguntas del cuestionario
- **Valores Atípicos u Outliers:** Son aquellos casos para los que una, dos o múltiples variables de una investigación determinada toman valores extremos que los hace diferir del comportamiento del resto de la muestra, y permiten al investigador sospechan que han sido alterados o generados por mecanismos distintos al resto de los datos (Hawkins, 1980).
- ¿Por qué es importante detectar los valores atípicos? Por las consecuencias que generan:
    - Distorsionan los resultados al **oscurecer el patrón de comportamiento** del resto de casos y obtenerse conclusiones que, sin ellos, serían completamente distintas y,
    - Pueden afectar gravemente a una de las condiciones de aplicabilidad más habituales de la mayor parte de las técnicas multivariables: la normalidad.

**Paso 3: Condiciones de Aplicabilidad:**

- **Normalidad:** La condición básica que debe asumirse en el análisis multivariable es la normalidad, y se refiere a que todos los datos de las variables métricas deben de seguir una distribución normal. Si la variación de la distribución normal es demasiado amplia, todos los resultados del análisis multivariable serán inválidos, porque la normalidad es un requisito esencial para los **estadísticos F y t.**
- La prueba de normalidad para una sola variable es fácil de realizar y existen

diversos métodos para estimarse, pero la normalidad multivariable es más complicada de realizar y existen relativamente pocos métodos para estimarla.

- **Métodos: Análisis de la Asimetría y Curtosis, Gráficos q-q de residuos y Estadísticos de prueba de Kolmogorov-Smirnov-Lillieforce (KSL).**
- **Solución: Transformación potencial y logarítmica**

## *Problema 2*

**Paso 1: Objetivos**

De la base de datos **WEB Diseño.sav** donde **N>50 muestras**, pruebe qué Hipótesis es aprobada:

$\mathbf{H_0}$.-Lasvariables $\mathbf{X_1}$ **a**$\mathbf{X_7}$ tienen una población con distribución normal

$\mathbf{H_1}$.-La variable $\mathbf{X_1}$**a**$\mathbf{X_7}$ **NO** tienen una población con distribución normal

**Paso 2: Diseño.** Idem dad el uso de la misma base de datos

**Paso 3: Condiciones de aplicabilidad.**

Dado que **N>50** muestras, la normalidad se analizará de acuerdo a **Kolmogorov-Smirnov.**

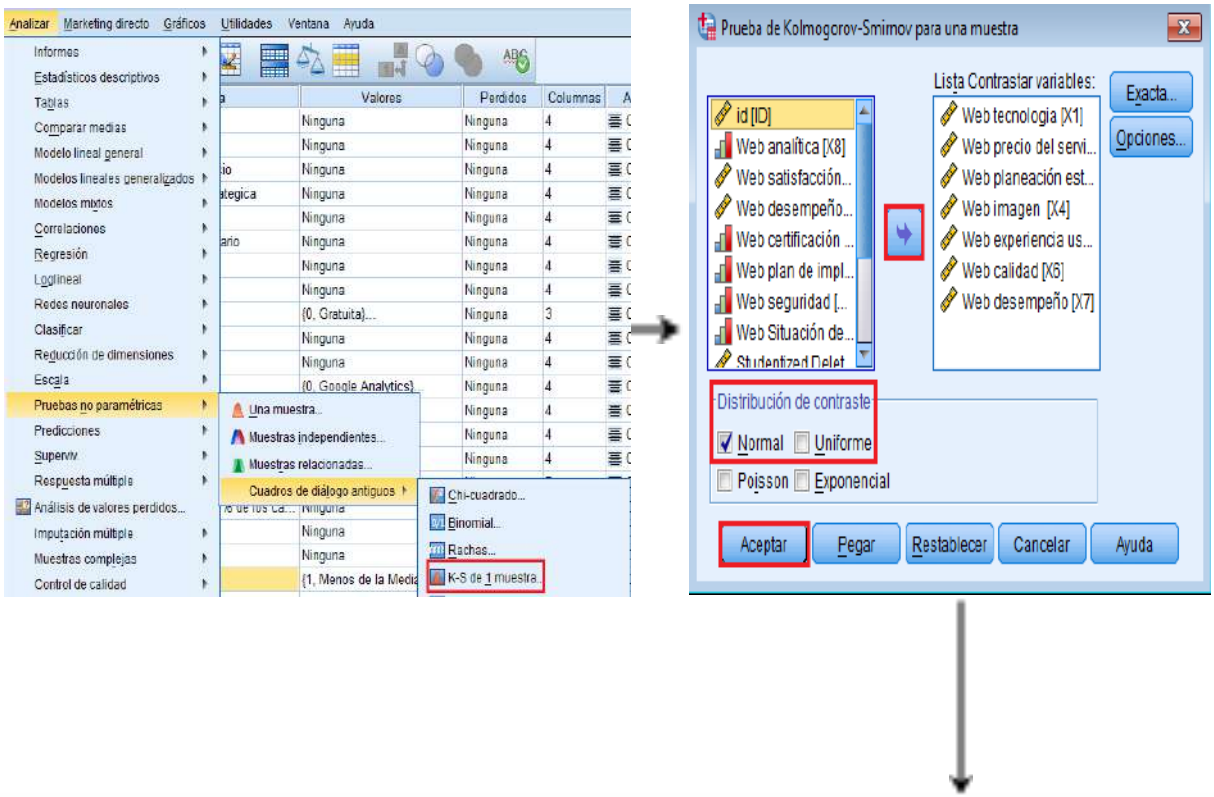
**Paso 4: Ejecución y ajuste**

**Teclear líneas de comando SPSS:**

**Analizar->Pruebas no paramétricas->Cuadro de diálogo antiguos- >K-S de 1 muestra; Seleccionar lista de variables de prueba (**$\mathbf{X_1}$ **a**$\mathbf{X_7}$**);**

**Seleccionar Distribución de prueba (Normal)->Aceptar->Continuar- >Aceptar.** Ver **Figura 3.5.**

**Figura 3.5. Prueba de normalidad**




**Pruebas no paramétricas**

[Conjunto_de_datos1] C:\Users\Juan\Desktop\proy libro mc\WEB diseño copia.sav

**Prueba de Kolmogorov-Smirnov para una muestra**

| | | Web tecnologia | Web precio del servicio | Web planeación estrategica | Web imagen | Web experiencia usuario | Web calidad | Web desempeño |
|---|---|---|---|---|---|---|---|---|
| N | | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| Parámetros normales[a,b] | Media | 3.515 | 2.364 | 7.894 | 5.248 | 2.916 | 2.665 | 6.971 |
| | Desviación típica | 1.3207 | 1.1957 | 1.3865 | 1.1314 | .7513 | .7706 | 1.5852 |
| Diferencias más extremas | Absoluta | .063 | .095 | .095 | .107 | .085 | .122 | .091 |
| | Positiva | .055 | .095 | .085 | .107 | .058 | .122 | .074 |
| | Negativa | -.063 | -.068 | -.095 | -.104 | -.085 | -.062 | -.091 |
| Z de Kolmogorov-Smirnov | | .628 | .945 | .947 | 1.068 | .854 | 1.221 | .909 |
| Sig. asintót. (bilateral) | | .825 | .333 | .331 | .205 | .460 | .101 | .380 |

a. La distribución de contraste es la Normal.

b. Se han calculado a partir de los datos.

Fuente: SPSS 20 elaboración propia

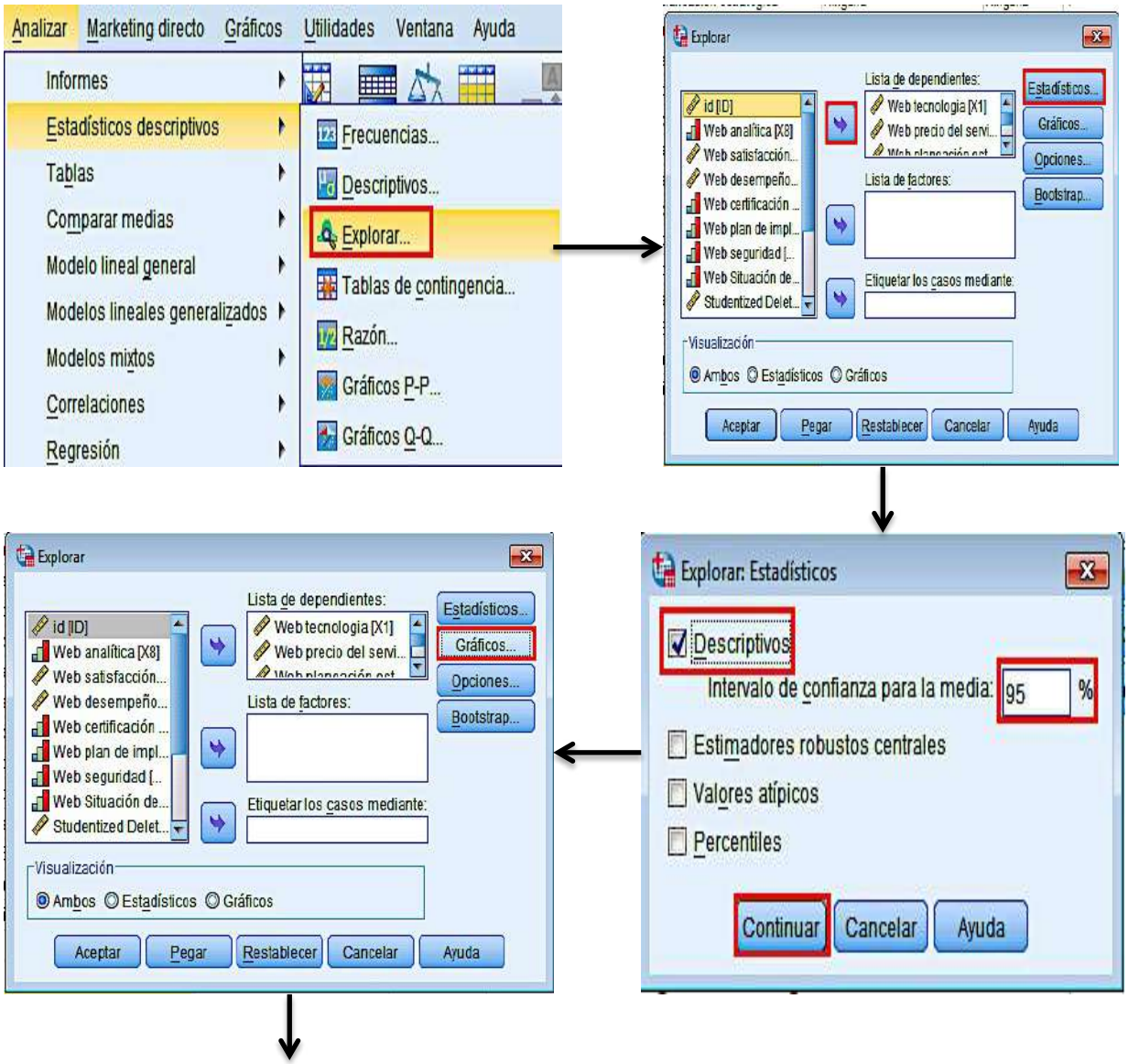
**Paso 5: Interpretación**

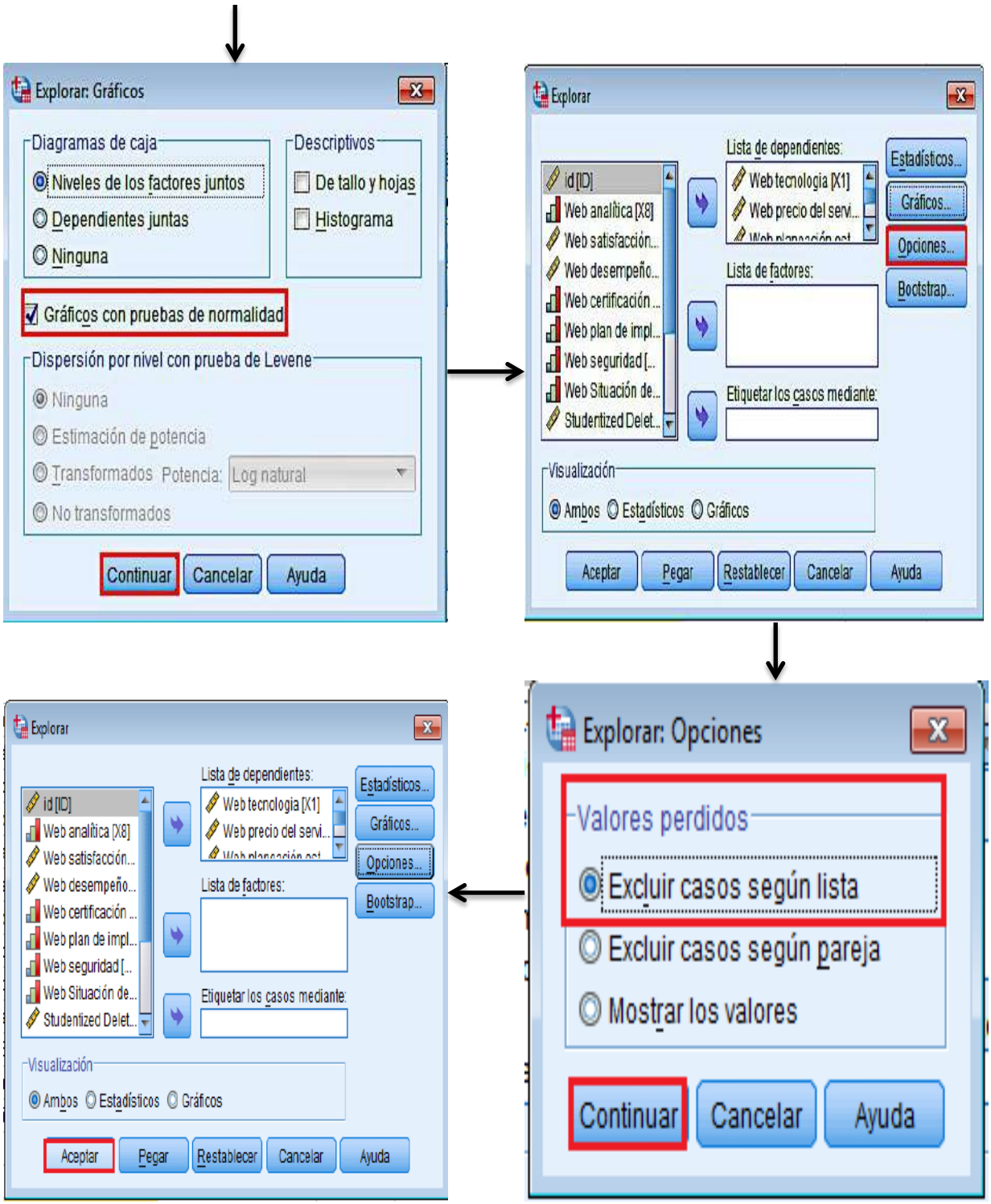
Dado que las **p>0.05** se acepta **Ho.-** Las variables **X1** a **X7 SI** tienen una población con distribución normal

## Problema 3

**Paso 1: Objetivos :** generar análisis de descriptivos y gráficos Q-Q de normalidad
**Paso 2: Diseño**: Idem al anterior
**Paso 3: Condicio0nes de aplicabilidad.** Se consideran sin cambio
**Paso 4: Ejecución y ajuste**
**Teclear líneas de comando SPSS:**
**Analizar->Estadísticos descriptivos->Explorar->Seleccionar lista de dependientes ( $X_1$ a $X_7$ ) -> Estadísticos -> Seleccionar: Descriptivos->Confirmar: 95%-> Continuar->Gráficos->Seleccionar: Gráficos con pruebas de normalidad->Continuar->Opciones->Valores perdidos, seleccionar: Excluir casos según lista- >Continuar->Aceptar**. **Ver Figura 3.6**

**Figura 3.6. Proceso para generar descriptivos y gráficos Q-Q de normalidad**






Fuente: SPSS 20 elaboración propia

•**SPSS** genera la **Figura 3.7,** que contiene las tablas **Resumen de procesamiento de los casos, Descriptivos y Pruebas de normalidad.**

**Figura 3.7. Tabla Resumen del procesamiento de los casos; Descriptivos ; Pruebas de normalidad y Gráficos Q-Q de normalidad**

**Explorar**

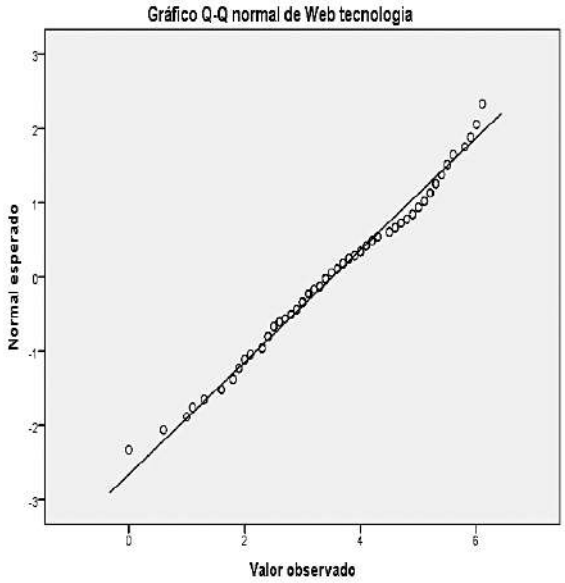
[Conjunto_de_datos1] C:\Users\Juan\Desktop\proy libro mc\WEB diseño copia.sav

**Resumen del procesamiento de los casos**

| | Casos | | | | | |
|---|---|---|---|---|---|---|
| | Válidos | | Perdidos | | Total | |
| | N | Porcentaje | N | Porcentaje | N | Porcentaje |
| Web tecnologia | 100 | 100.0% | 0 | 0.0% | 100 | 100.0% |
| Web precio del servicio | 100 | 100.0% | 0 | 0.0% | 100 | 100.0% |
| Web planeación estrategica | 100 | 100.0% | 0 | 0.0% | 100 | 100.0% |
| Web imagen | 100 | 100.0% | 0 | 0.0% | 100 | 100.0% |
| Web experiencia usuario | 100 | 100.0% | 0 | 0.0% | 100 | 100.0% |
| Web calidad | 100 | 100.0% | 0 | 0.0% | 100 | 100.0% |
| Web desempeño | 100 | 100.0% | 0 | 0.0% | 100 | 100.0% |

**Descriptivos**

| | | | Estadístico | Error típ. |
|---|---|---|---|---|
| Web tecnologia | Media | | 3.515 | .1321 |
| | Intervalo de confianza para la media al 95% | Límite inferior | 3.253 | |
| | | Límite superior | 3.777 | |
| | Media recortada al 5% | | 3.534 | |
| | Mediana | | 3.400 | |




**Pruebas de normalidad**

| | Kolmogorov-Smirnov[a] | | | Shapiro-Wilk | | |
|---|---|---|---|---|---|---|
| | Estadístico | gl | Sig. | Estadístico | gl | Sig. |
| Web tecnologia | .063 | 100 | .200* | .985 | 100 | .341 |
| Web precio del servicio | .095 | 100 | .028 | .969 | 100 | .017 |
| Web planeación estrategica | .095 | 100 | .027 | .950 | 100 | .001 |
| Web imagen | .107 | 100 | .007 | .982 | 100 | .183 |
| Web experiencia usuario | .065 | 100 | .069 | .986 | 100 | .366 |
| Web calidad | .122 | 100 | .001 | .963 | 100 | .007 |
| Web desempeño | .091 | 100 | .041 | .971 | 100 | .028 |

*. Este es un límite inferior de la significación verdadera.

a. Corrección de la significación de Lilliefors

Fuente: SPSS 20 elaboración propia

- **Homoscedasticidad:** Debe definirse de manera distinta según se estén analizando datos no agrupados (caso de una regresión lineal múltiple), o datos agrupados (caso de un análisis de la varianza de un factor).
- En el primer caso la hipótesis de homocedasticidad puede definirse como el supuesto de que cada uno de los valores que puede tomar la distribución se mantiene constante para todos los valores de la otra variable continua.
- En el caso de datos agrupados la homocedasticidad implica que la varianza de la variable continua es más o menos la misma en todos los grupos que conforman la variable no métrica que es la que determina los grupos.
- En resumen, se puede decir que la homocedasticidad es la igualdad de varianza entre las variables independientes.
- **Métodos: Test de Levene**

**•Solución: Transformación logarítmica y potencial**

## *Problema 4*

**Paso 1: Objetivos**

Compruebe la homocedasticidad de la base de datos **WEB Diseño.sav** a partir del tamaño de las empresas vs. cada variable métrica de la misma.

$\mathbf{H_0}$: **No** existen diferencias significativas entre las varianzas de las variables con el tamaño de la empresa

$\mathbf{H_1}$: **Sí** existen diferencias significativas entre las varianzas de las variables con el tamaño de la empresa

**Paso 2: Diseño.** Idem caso anterior.

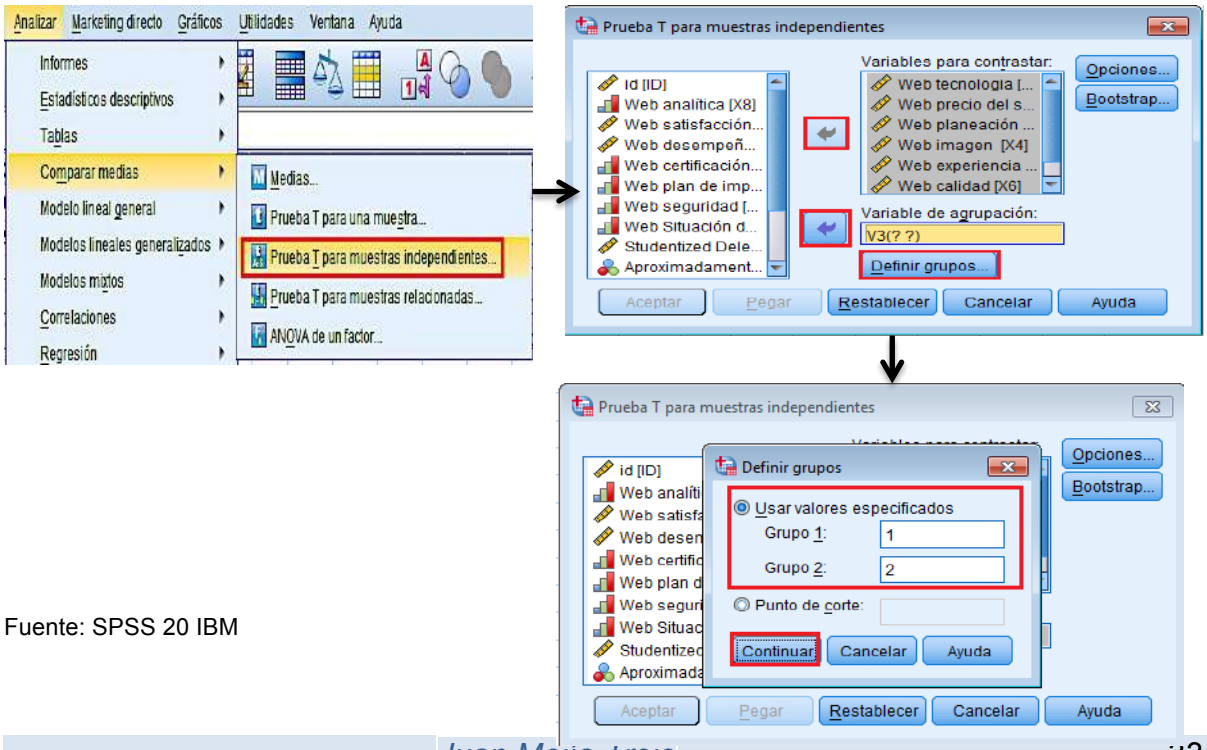
**Paso 3: Condiciones de aplicabilidad.** Se toma en cuenta el ejemplo anterior y que cumple con las condiciones iniciales. Así, se aplicará test de **Levene** para comprobar la **homocedasticidad** (varianzas iguales de las variables dependientes). Se realiza a través de **t de Student** como sigue:

**Paso 4: Ejecución y ajuste.**

**Teclee líneas de comando SPSS:**

**Analizar-> Comparar medias->Prueba T para muestras independientes->Selección Variables para contrastar métricas: X1 ; Selección Variable de agrupación nominal: Tamaño de la empresa (V3)->Definir grupos-> marcar las etiquetas a comparar (1,2,3,4,etc) por pares Continuar->Aceptar. Ver Figura 3.8.**

**Figura 3.8.- Proceso para verificar homocedasticidad de la base de datos**




Fuente: SPSS 20 IBM

**Paso 5: Interpretación**
Dado que **p>0.05** en todos los casos **Se acepta la Ho: No existen diferencias significativas entre las varianzas de las variables con el tamaño de la empresa . Ver Figura 3.9**

**Figura 3.9.- Resultados de Prueba de Levene**

**Prueba de muestras independientes**

| | | Prueba de Levene para la igualdad de varianzas | | Prueba T para la igualdad de medias | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | F | Sig. | t | gl | Sig. (bilateral) | Diferencia de medias | Error típ. de la diferencia | 95% Intervalo de confianza p la diferencia Inferior | Superior |
| Web tecnologia | Se han asumido varianzas iguales | .753 | .389 | 1.244 | 63 | .218 | .4873 | .3916 | -.2952 | 1.26 |
| | No se han asumido varianzas iguales | | | 1.110 | 19.829 | .280 | .4873 | .4391 | -.4291 | 1.40 |
| Web precio del servicio | Se han asumido varianzas iguales | .088 | .768 | -1.554 | 63 | .125 | -.5587 | .3594 | -1.2770 | .15 |
| | No se han asumido varianzas iguales | | | -1.653 | 25.483 | .111 | -.5587 | .3380 | -1.2540 | .13 |
| Web planeación estrategica | Se han asumido varianzas iguales | .918 | .342 | 1.099 | 63 | .276 | .4087 | .3717 | -.3342 | 1.15 |
| | No se han asumido varianzas iguales | | | 1.029 | 21.047 | .315 | .4087 | .3970 | -.4167 | 1.23 |
| Web imagen | Se han asumido varianzas iguales | .553 | .460 | -.061 | 63 | .951 | -.0200 | .3269 | -.6732 | .63 |
| | No se han asumido varianzas iguales | | | -.064 | 24.701 | .950 | -.0200 | .3130 | -.6651 | .62 |
| Web experiencia usuario | Se han asumido varianzas iguales | .170 | .682 | -.152 | 63 | .880 | -.0340 | .2238 | -.4812 | .41 |
| | No se han asumido varianzas iguales | | | -.150 | 22.653 | .882 | -.0340 | .2266 | -.5031 | .43 |
| Web calidad | Se han asumido varianzas iguales | .030 | .862 | .330 | 63 | .742 | .0740 | .2241 | -.3738 | .52 |
| | No se han asumido | | | .313 | 21.426 | .757 | .0740 | .2360 | -.4163 | .56 |

Fuente: SPSS 20 IBM con elaboración propia

- **Linealidad:** el supuesto de linealidad es fundamental para todas aquellas técnicas que se centren en el análisis de *las matrices de correlación o de varianzas – covarianzas, como el análisis factorial, regresión lineal o los modelos de ecuaciones estructurales***.** La razón es sencilla: el coeficiente de **correlación de Pearson** sólo podrá captar una relación si ésta es lineal.
- Si la relación existe y es intensa pero, por ejemplo, *es curvilínea*, el **coeficiente de correlación de Pearson** *tomará un valor relativamente bajo* y el investigador puede interpretarlo como *ausencia de relación cuando, de hecho, ésta existe sólo que no es lineal.*
- Cuando la técnica empleada tiene **una variable dependiente,** como ocurre en el caso de la **regresión lineal múltiple**, existen diversos procedimientos para contrastar la linealidad de las relaciones basadas en el análisis de los residuos o residuales.
- **Métodos: Gráficos: de dispersión entre variables y Estadísticos:coeficientes de correlación** bivariados**.**

## Problema 5

**Paso 1: Objetivos**

Verifique las correlaciones de las variables métricas de la base de datos y analícelas.

**Paso 2: Diseño.Idem** al caso anterior

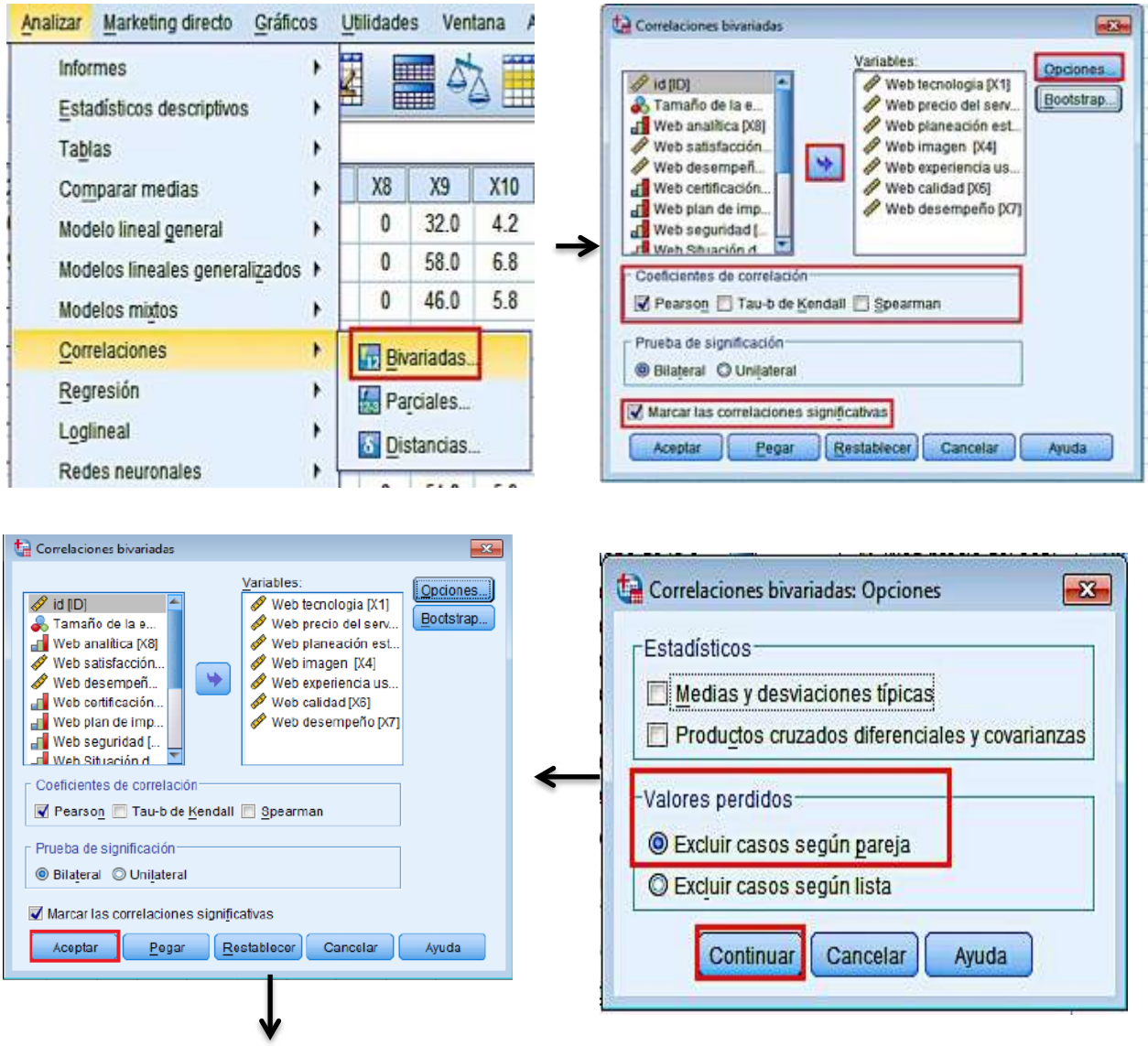
**Paso 3: Condiciones de aplicabilidad.**

Se toma en cuenta el ejemplo anterior y que cumple con las condiciones iniciales.

**Paso 4: Ejecución y ajuste**

**Teclee Analizar->correlaciones->bivariadas->Coeficientes de correlación: Pearson->Marcar las correlaciones significativas->Opciones->Valores perdidos: Excluir casos según pareja->continuar. Ver Figura 3.10.**

**Figura 3.10.- Proceso para determinar la matriz de correlaciones**



↓

**Correlaciones**

| | | Web tecnologia | Web precio del servicio | Web planeación estrategica | Web imagen | Web experiencia usuario | Web calidad | Web desempeño |
|---|---|---|---|---|---|---|---|---|
| Web tecnologia | Correlación de Pearson | 1 | -.349** | .509** | .050 | .612** | .077 | -.483** |
| | Sig. (bilateral) | | .000 | .000 | .618 | .000 | .444 | .000 |
| | N | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| Web precio del servicio | Correlación de Pearson | -.349** | 1 | -.487** | .272** | .513** | .185 | .470** |
| | Sig. (bilateral) | .000 | | .000 | .006 | .000 | .065 | .000 |
| | N | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| Web planeación estrategica | Correlación de Pearson | .509** | -.487** | 1 | -.116 | .067 | -.035 | -.448** |
| | Sig. (bilateral) | .000 | .000 | | .250 | .510 | .731 | .000 |
| | N | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| Web imagen | Correlación de Pearson | .050 | .272** | -.116 | 1 | .299** | .788** | .200* |
| | Sig. (bilateral) | .618 | .006 | .250 | | .003 | .000 | .046 |
| | N | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| Web experiencia usuario | Correlación de Pearson | .612** | .513** | .067 | .299** | 1 | .240* | -.055 |
| | Sig. (bilateral) | .000 | .000 | .510 | .003 | | .016 | .586 |
| | N | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| Web calidad | Correlación de Pearson | .077 | .185 | -.035 | .788** | .240* | 1 | .177 |
| | Sig. (bilateral) | .444 | .065 | .731 | .000 | .016 | | .079 |
| | N | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| Web desempeño | Correlación de Pearson | -.483** | .470** | -.448** | .200* | -.055 | .177 | 1 |
| | Sig. (bilateral) | .000 | .000 | .000 | .046 | .586 | .079 | |
| | N | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

**. La correlación es significativa al nivel 0,01 (bilateral).

*. La correlación es significante al nivel 0,05 (bilateral).

Fuente: SPSS 20 IBM con elaboración propia

## *Problema 6*

**Paso 1: objetivos**

**Objetivo**: Identificar la estructura de un grupo de variables a través de la reducción de datos. **Así, problema 5: determinar la percepción que tienen los gerentes de las empresas del sector de mercadotecnia digital acerca de los servicios ofrecidos a sus consumidores medidos en 7 atributos (X1 a X7**) por las siguientes razones:

- Entender el agrupamiento de estas percepciones sobre los **7 atributos** generan una matriz de **correlaciones** que serán agrupadas en términos de lo que los gerentes de las empresas del sector de mercadotecnia digital tienen sobre los servicios ofrecidos a sus clientes.
- Reducir las **7 variables en el menor número de factores:** Si las **7 variables** se pueden representar en un pequeño número de factores o dimensiones, entonces se eligió a la técnica multivariable correcta y las otras técnicas ya no serían correctas.

**Paso 2: Diseño**

- ¿Cómo deben medirse las variables a utilizar? **Deben ser métricas en general**, pero pueden introducirse no métricas (Variables Dummies), si bien no gozan de las mismas propiedades.
- ¿Cuál es el tamaño de la muestra? **No menos de 50 observaciones,** aunque lo mejor es **contar con más de 100 observaciones. No menos de 5 observaciones por cada variable,** si bien el **ratio óptimo es 10 a1** En el caso de **WEB Diseño.sav**

se tienen **100 observaciones** y **7 variables**, lo que da un ratio adecuado de **14 a 1.** (**100/7**)

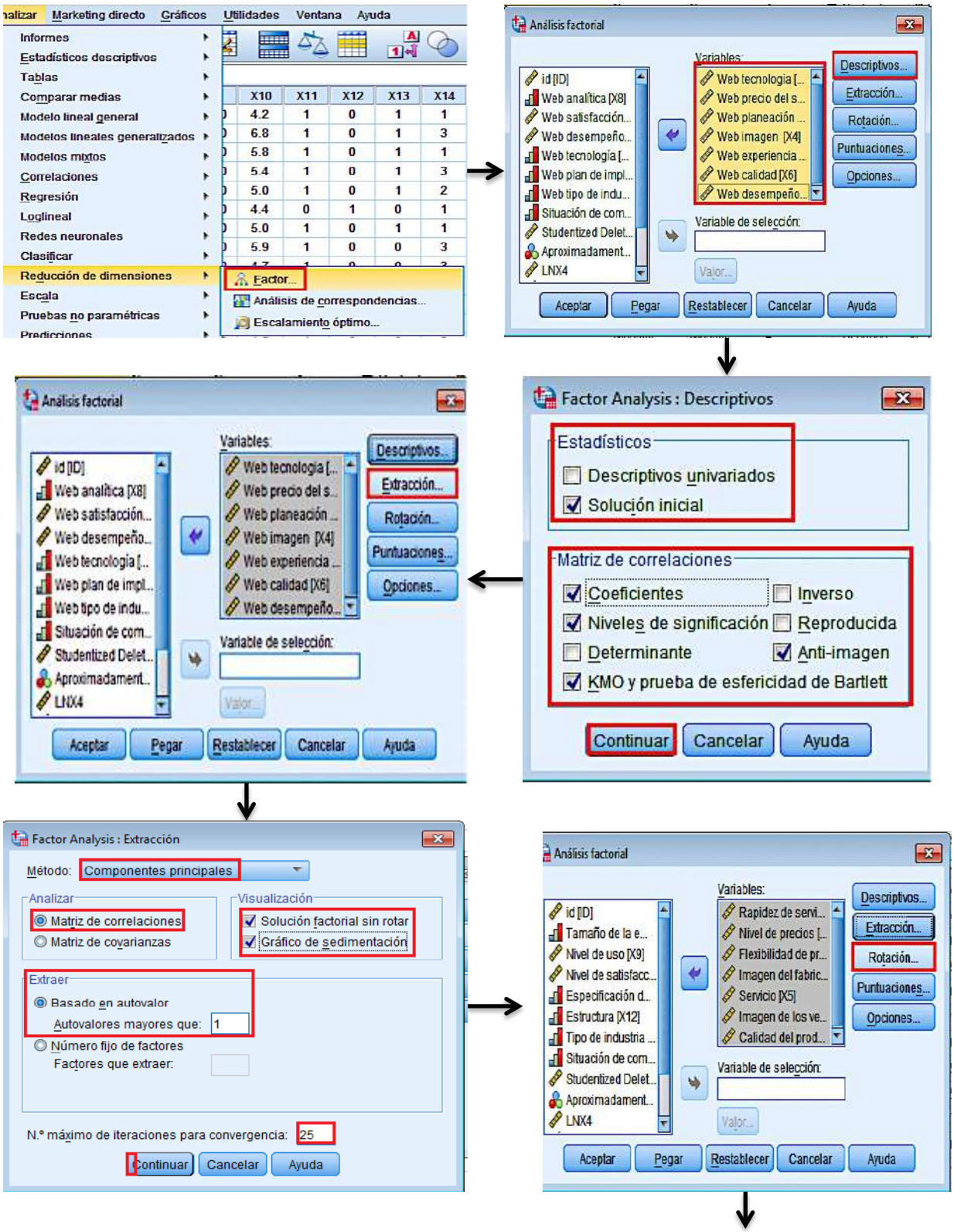
**Paso 3: supuestos de Aplicabilidad**
**No se considera determinante la falta de normalidad, homocedasticidad pero sí la linealidad. Así que se debe cumplir:**

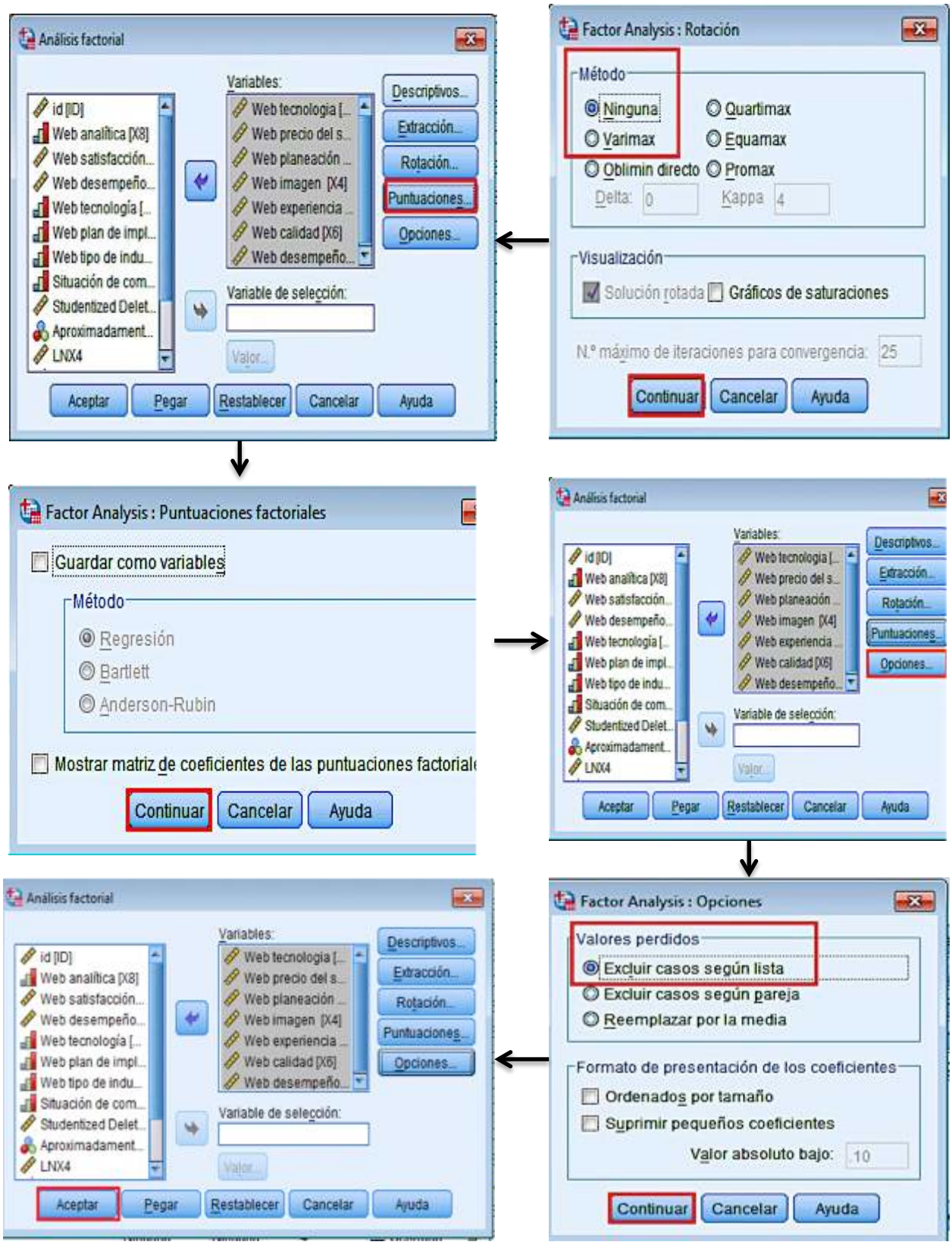
- Hay que asegurarse que existen suficientes correlaciones significativas entre las variables (**incorrelación implica un solo factor por cada variable**). Por linealidad **50%+1**
- Estas correlaciones **deben ser importantes** (**> a 0.3**) **La diagonal** de la **matriz de correlación anti-imagen** debe ser superior a **0.5**
- El test de **esfericidad de *Bartlett*** debe ser significativo (**0.01 ó 0.05 error / 99% o 95% confiabilidad**)
- El ***KMO*** debe ser **>0.5( lo ideales>a0.7)**
- Su **p** debe ser **inferior a los niveles críticos 0.05 o 0.01**. Debe saberse, sin embargo, que es un test muy sensible a incrementos en el tamaño de la muestra. Cuando esta se incrementa es más fácil que encuentra correlaciones significativas..
- Método de extracción con valores a **1**; **Varianza extraída superior al 60%** por gráfico de sedimentación

Deben cumplirse si no es así, **el análisis factorial no será posible.**

**Paso 4: Ejecución y ajuste**
**Teclear las líneas de comando SPSS:**
**Analizar->Reducción de dimensiones->Selección de variables métricas->Descriptivos-> Seleccionar: Estadísticos: Solución inicial; Matriz de correlaciones: Coeficientes; Niveles de significación-> *KMO* y prueba de esfericidad de *Bartlett*; Anti-imagen->Continuar->Extracción->Método: componentes principales->Analizar: Matriz de correlaciones->Visualización: solución factorial sin rotar-> Seleccionar: Gráfico de sedimentación->Extraer: Autovalores mayores que: 1->Numero máximo de interacciones: 25->Continuar->Rotación->Método: Varimax->Visualización: solución rotada->No. máximo de iteraciones para convergencia: 25->Continuar->Puntuaciones->Continuar->Opciones->Valores perdidos (nota: no reemplazar por medias; no pérdida de datos >10%->Excluir casos según lista (nota: evitar falta de datos)->Continuar->Aceptar. Ver Figura 3.11.**

## Figura 3.11. De secuencia de instrucciones del análisis factorial






Fuente: SPSS 20 IBM con elaboración propia

**Paso 5: Interpretación**

**a.SPSS** genera la *Matriz de correlaciones*, de la que debe asegurarse que existen suficientes correlaciones significativas entre las variables (**incorrelación implica un solo factor por cada variable**). Por linealidad **50%+1**

- Se observan **21 datos** por lado de la matriz bajo/sobre la diagonal
- La mayoría deberá ser ***50%+1=21/2+1=11*** *correlaciones deben ser significativas para cumplir.*
- Existen 12 al 0.01 con este cumple también cumple, son significativas
- Existen 15 al 0.05 con este también cumple, son significativas. **Se recomienda usar.**
- **SI CUMPLE CONDICION. Ver Figura 3.12.**

**Figura 3.12. Matriz de correlaciones**

Matriz de correlaciones

| | | Web tecnologia | Web precio del servicio | Web planeación estrategica | Web imagen | Web experiencia usuario | Web calidad | Web desempeño |
|---|---|---|---|---|---|---|---|---|
| Correlación | Web tecnologia | 1.000 | -.349 | .509 | .050 | .612 | .077 | -.483 |
| | Web precio del servicio | -.349 | 1.000 | -.487 | .272 | .513 | .185 | .470 |
| | Web planeación estrategica | .509 | -.487 | 1.000 | -.116 | .067 | -.035 | -.448 |
| | Web imagen | .050 | .272 | -.116 | 1.000 | .299 | .788 | .200 |
| | Web experiencia usuario | .612 | .513 | .067 | .299 | 1.000 | .240 | -.055 |
| | Web calidad | .077 | .185 | -.035 | .788 | .240 | 1.000 | .177 |
| | Web desempeño | -.483 | .470 | -.448 | .200 | -.055 | .177 | 1.000 |
| Sig. (Unilateral) | Web tecnologia | | .000 | .000 | .309 | .000 | .222 | .000 |
| | Web precio del servicio | X .000 | | .000 | .003 | .000 | .032 | .000 |
| | Web planeación estrategica | X .000 | X .000 | | .125 | .255 | .366 | .000 |
| | Web imagen | .309 | X .003 | .125 | | .001 | .000 | .023 |
| | Web experiencia usuario | X .000 | X .000 | .255 | X .001 | | .008 | .293 |
| | Web calidad | .222 | X .032 | .366 | X .000 | X .008 | | .039 |
| | Web desempeño | X .000 | X .000 | X .000 | X .023 | .293 | X .039 | |

Nota : [ ] Datos <=0.01; X Datos <= 0.05

Fuente: SPSS 20 IBM con elaboración propia

**b.**Estas correlaciones **deben ser importantes** (**> a 0.3**) -Se observan **21 datos** por lado de la matriz bajo/sobre la diagonal

- La mayoría deberá ser **50%+1=21/2+1=11** correlaciones deben ser significativas para cumplir.
- **Existen sólo 9 /11 con valor >=0.3 y 2 muy cercanos. CUMPLE MUY CERRADAMENTE LA CONDICION. Ver Figura 3.13.**

**Figura 3.13. Matriz de correlaciones**

**Matriz de correlaciones**

| | | Web tecnologia | Web precio del servicio | Web planeación estrategica | Web imagen | Web experiencia usuario | Web calidad | Web desempeño |
|---|---|---|---|---|---|---|---|---|
| Correlación | Web tecnologia | 1.000 | -.349 | .509 | .050 | .612 | .077 | -.483 |
| | Web precio del servicio | X -.349 | 1.000 | -.487 | .272 | .513 | .185 | .470 |
| | Web planeación estrategica | X .509 | X -.487 | 1.000 | -.116 | .067 | -.035 | -.448 |
| | Web imagen | .050 | .272 | -.116 | 1.000 | .299 | .788 | .200 |
| | Web experiencia usuario | X .612 | X .513 | .067 | .299 | 1.000 | .240 | -.055 |
| | Web calidad | .077 | .185 | -.035 | X .788 | .240 | 1.000 | .177 |
| | Web desempeño | X -.483 | X .470 | X -.448 | .200 | -.055 | .177 | 1.000 |
| Sig. (Unilateral) | Web tecnologia | | .000 | .000 | .309 | .000 | .222 | .000 |
| | Web precio del servicio | .000 | | .000 | .003 | .000 | .032 | .000 |
| | Web planeación estrategica | .000 | .000 | | .125 | .255 | .366 | .000 |
| | Web imagen | .309 | .003 | .125 | | .001 | .000 | .023 |
| | Web experiencia usuario | .000 | .000 | .255 | .001 | | .008 | .293 |
| | Web calidad | .222 | .032 | .366 | .000 | .008 | | .039 |
| | Web desempeño | .000 | .000 | .000 | .023 | .293 | .039 | |

Nota: [ ].-Muy cercano a 0.3; X.- 9 datos >=0.3

Fuente: SPSS 20 IBM con elaboración propia

**c.SPSS genera la tabla Matrices anti-imagen de la que se observa una diagonal** con un valor que debe ser superior a **0.5**

- Sólo $X_1$,$X_2$ y $X_1$ a $X_5$. Con una sola que no cumpla ya no es posible continuar.
- **NO CUMPLE CONDICION. Ver Figura 3.14.**

**Figura 3.14. Tabla Matriz anti-imagen**

**Matrices anti-imagen**

| | | Web tecnologia | Web precio del servicio | Web planeación estrategica | Web imagen | Web experiencia usuario | Web calidad | Web desempeño |
|---|---|---|---|---|---|---|---|---|
| Covarianza anti-imagen | Web tecnologia | .028 | .028 | .002 | .015 | -.025 | -.006 | -.002 |
| | Web precio del servicio | .028 | .032 | .021 | .014 | -.026 | -.005 | -.020 |
| | Web planeación estrategica | .002 | .021 | .608 | .043 | -.011 | -.039 | .086 |
| | Web imagen | .015 | .014 | .043 | .347 | -.015 | -.275 | -.018 |
| | Web experiencia usuario | -.025 | -.026 | -.011 | -.015 | .023 | .005 | .010 |
| | Web calidad | -.006 | -.005 | -.039 | -.275 | .005 | .371 | -.044 |
| | Web desempeño | -.002 | -.020 | .086 | -.018 | .010 | -.044 | .623 |
| Correlación anti-imagen | Web tecnologia | .345[a] | .957 | .018 | .148 | -.978 | -.059 | -.016 |
| | Web precio del servicio | .957 | .330[a] | .155 | .133 | -.975 | -.043 | -.141 |
| | Web planeación estrategica | .018 | .155 | .914[a] | .094 | -.091 | -.083 | .139 |
| | Web imagen | .148 | .133 | .094 | .558[a] | -.172 | -.766 | -.040 |
| | Web experiencia usuario | -.978 | -.975 | -.091 | -.172 | .288[a] | .051 | .088 |
| | Web calidad | -.059 | -.043 | -.083 | -.766 | .051 | .552[a] | -.091 |
| | Web desempeño | -.016 | -.141 | .139 | -.040 | .088 | -.091 | .927[a] |

a. Medida de adecuación muestral

Nota: Valores de a>0.5

Fuente: SPSS 20 IBM con elaboración propia

**d. test de esfericidad de *Bartlett*** debe ser significativo (0.01 ó 0.05 error / 99% o 95% confiabilidad)...**Sí es significativa al 0.01...CUMPLE**

**e.** El ***KMO*** debe ser **> 0.5** (lo ideal es **> a 0.7**)....

- **NO CUMPLE. Ver Figura 3.15.**

**Figura 3.15. De resultados del análisis factorial**

**KMO y prueba de Bartlett**

| | | |
|---|---|---|
| Medida de adecuación muestral de Kaiser-Meyer-Olkin. | | .446 |
| Prueba de esfericidad de Bartlett | Chi-cuadrado aproximado | 567.467 |
| | gl | 21 |
| | Sig. | .000 |

Fuente: SPSS 20 IBM con elaboración propia

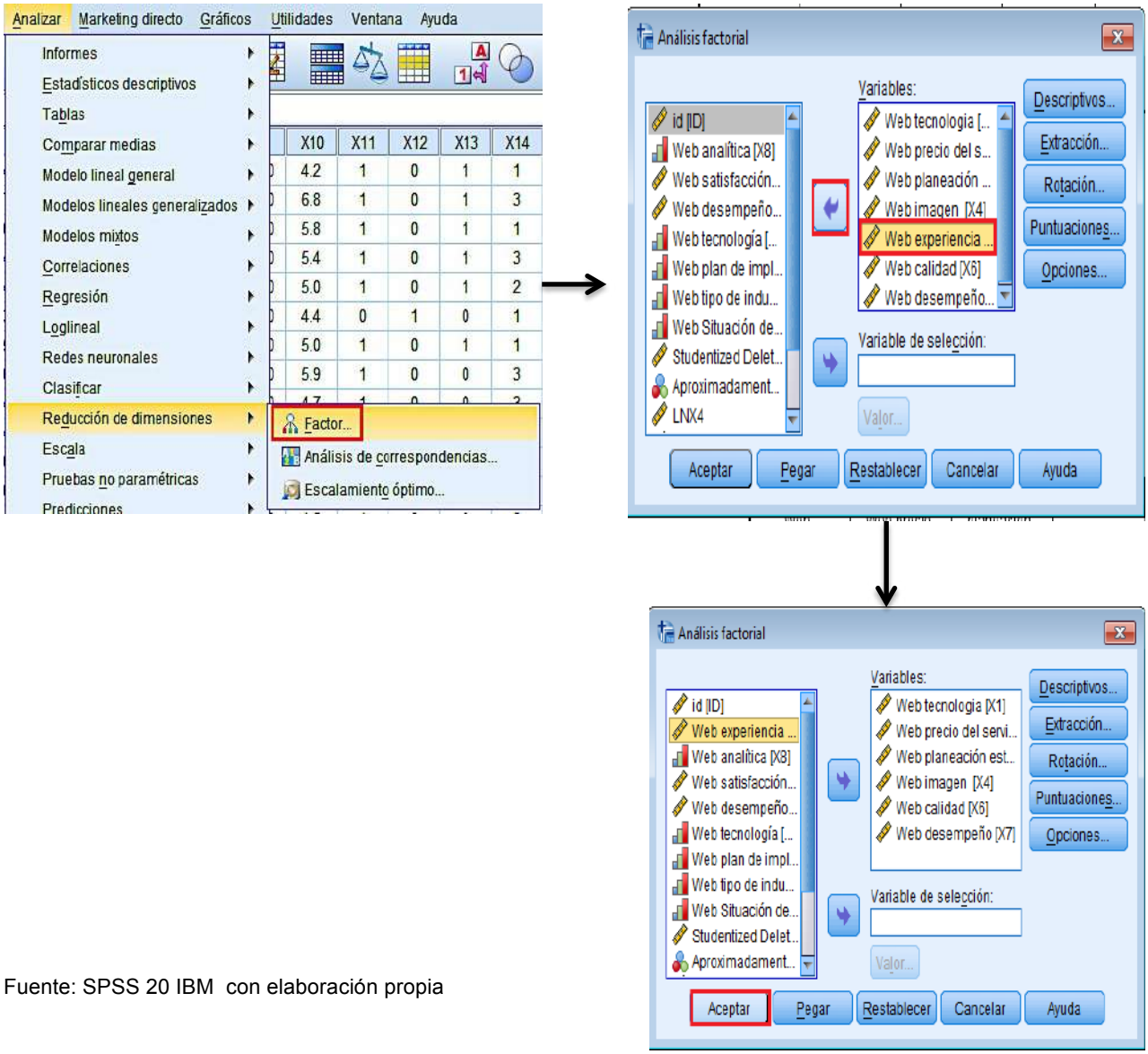
**f.**Se concluye que hay problemas con las variables, por lo que se recomienda revisar

las que se reportan en la matriz anti-imagen, que son: **X1** , **X2**, **X3** La que se recomienda remover es la que tiene un valor bajo en la matriz anti-imagen: **X5** (.288).

- No significa que la variable no sirva.
- De remover **X5** se requiere realizar **TODO el proceso anterior para volver a revisar. Ver Figura 3.16.**

**Figura 3.16. Separación de la variable X5 (Web experiencia) de WEB diseño.sav para reinicio de análisis factorial. Ver Paso 4.**

**Paso 4. Ejecución y ajuste**
**Teclear línea de comando SPSS: Analizar->Reducción de dimensiones->Factor->Seleccionar variable a remover->Flecha->Aceptar->Aceptar**




Fuente: SPSS 20 IBM con elaboración propia

**g.** Analizar la Matriz de correlaciones resultante. Debe asegurarse que existen

suficientes correlaciones significativas entre las variables (*incorrelación* implica un solo factor por cada variable). Por linealidad **50%+1**:

- Se observan 15 datos por lado de la matriz bajo/sobre la diagonal
- La mayoría deberá ser 50%+1=15/2+1=8 correlaciones deben ser significativas para cumplir
- Existen 8 al **0.01** con este cumple también cumple, son significativas
- Existen **11 al 0.05** con este también cumple, son significativas. Se recomienda usar.
- **SI CUMPLE CONDICION**. Ver **Figura 3.17.**

**Figura 3.17. De resultados del análisis factorial**

**Matriz de correlaciones**

| | | Web tecnologia | Web precio del servicio | Web planeación estrategica | Web imagen | Web calidad | Web desempeño |
|---|---|---|---|---|---|---|---|
| Correlación | Web tecnologia | 1.000 | -.349 | .509 | .050 | .077 | -.483 |
| | Web precio del servicio | -.349 | 1.000 | -.487 | .272 | .185 | .470 |
| | Web planeación estrategica | .509 | -.487 | 1.000 | -.116 | -.035 | -.448 |
| | Web imagen | .050 | .272 | -.116 | 1.000 | .788 | .200 |
| | Web calidad | .077 | .185 | -.035 | .788 | 1.000 | .177 |
| | Web desempeño | -.483 | .470 | -.448 | .200 | .177 | 1.000 |
| Sig. (Unilateral) | Web tecnologia | | .000 | .000 | .309 | .222 | .000 |
| | Web precio del servicio | X .000 | | .000 | .003 | .032 | .000 |
| | Web planeación estrategica | X .000 | X .000 | | .125 | .366 | .000 |
| | Web imagen | .309 | X .003 | .125 | | .000 | .023 |
| | Web calidad | .222 | X .032 | .366 | X .000 | | .039 |
| | Web desempeño | X .000 | X .000 | X .000 | X .023 | X .039 | |

Nota: ☐ Datos con valor <=0.01; X.-Datos con valor <=0.05

Fuente: SPSS 20 IBM con elaboración propia

**h.** Estas correlaciones *deben ser importantes* (**> a 0.3**).

- Se observan 21 datos por lado de la matriz bajo/sobre la diagonal
- La mayoría deberá ser 50%+1=15/2+1=8 correlaciones deben ser significativas para cumplir
- **Existen sólo 7/8 con valor >=0.3 y 1 muy cercanos. CUMPLE MUY CERRADAMENTE LA CONDICION. Ver Figura 3.18.**

## Figura 3.18 De resultados del análisis factorial

**Matriz de correlaciones**

| | | Web tecnologia | Web precio del servicio | Web planeación estrategica | Web imagen | Web calidad | Web desempeño |
|---|---|---|---|---|---|---|---|
| Correlación | Web tecnologia | 1.000 | -.349 | .509 | .050 | .077 | -.483 |
| | Web precio del servicio | -.349 | 1.000 | -.487 | .272 | .185 | .470 |
| | Web planeación estrategica | .509 | -.487 | 1.000 | -.116 | -.035 | -.448 |
| | Web imagen | .050 | X 272 | -.116 | 1.000 | .788 | .200 |
| | Web calidad | .077 | .185 | -.035 | .788 | 1.000 | .177 |
| | Web desempeño | -.483 | .470 | -.448 | .200 | .177 | 1.000 |
| Sig. (Unilateral) | Web tecnologia | | .000 | .000 | .309 | .222 | .000 |
| | Web precio del servicio | .000 | | .000 | .003 | .032 | .000 |
| | Web planeación estrategica | .000 | .000 | | .125 | .366 | .000 |
| | Web imagen | .309 | .003 | .125 | | .000 | .023 |
| | Web calidad | .222 | .032 | .366 | .000 | | .039 |
| | Web desempeño | .000 | .000 | .000 | .023 | .039 | |

Nota: Valores >=0.3 ; X.- Valor cercano a 0.3

Fuente: SPSS 20 IBM con elaboración propia

**i. La diagonal** de la **matriz de correlación anti-imagen** debe ser superior a **0.5** **CUMPLE CONDICION. Ver Figura 3.19.**

## Figura 3.19. De resultados del análisis factorial

**Matrices anti-imagen**

| | | Web tecnologia | Web precio del servicio | Web planeación estrategica | Web imagen | Web calidad | Web desempeño |
|---|---|---|---|---|---|---|---|
| Covarianza anti-imagen | Web tecnologia | .629 | .047 | -.210 | -.046 | -.022 | .208 |
| | Web precio del servicio | .047 | .650 | .190 | -.078 | .013 | -.162 |
| | Web planeación estrategica | -.210 | .190 | .613 | .037 | -.038 | .092 |
| | Web imagen | -.046 | -.078 | .037 | .358 | -.281 | -.012 |
| | Web calidad | -.022 | .013 | -.038 | -.281 | .372 | -.046 |
| | Web desempeño | .208 | -.162 | .092 | -.012 | -.046 | .628 |
| Correlación anti-imagen | Web tecnologia | .721[a] | .074 | -.338 | -.098 | -.046 | .331 |
| | Web precio del servicio | .074 | .787[a] | .301 | -.161 | .027 | -.253 |
| | Web planeación estrategica | -.338 | .301 | .749[a] | .079 | -.079 | .149 |
| | Web imagen | -.098 | -.161 | .079 | .542[a] | -.769 | -.025 |
| | Web calidad | -.046 | .027 | -.079 | -.769 | .532[a] | -.096 |
| | Web desempeño | .331 | -.253 | .149 | -.025 | -.096 | .779[a] |

a. Medida de adecuación muestral

Nota: Valores >0.5

Fuente: SPSS 20 IBM con elaboración propia

- **El test de esfericidad de** *Bartlett* debe ser significativo **(0.01 ó 0.05 error / 99% o 95%** *confiabilidad***)...**
- **Sí es significativa al 0.01...CUMPLE**
- **El *KMO* debe ser > 0.5 (lo ideal es > a 0.7)....**
- **SÍ CUMPLE .**

**Ver Figura 3.20.**
**Figura 3.20. De resultados del análisis factorial**

**KMO y prueba de Bartlett**

| | | |
|---|---|---|
| Medida de adecuación muestral de Kaiser-Meyer-Olkin. | | .665 |
| Prueba de esfericidad de Bartlett | Chi-cuadrado aproximado | 205.902 |
| | gl | 15 |
| | Sig. | .000 |

Fuente: SPSS 20 IBM con elaboración propia

I.Se concluye que es correcto y es posible **continuar el análisis**.

**Paso 4: Estimación y ajuste.**
Continuación del análisis.

- Método de obtención de los Factores: Componentes Principales, para resumir la mayor parte de la información en un menor número de factores.
- ¿Cuántos factores van a ser extraídos?:
  - Autovalores mínimo de 1
  - En la **Figura 3.21**, se aprecian los valores a **1**; **la variable que más información perdió es la Web precio del servicio**

**Figura 3.21. Comunalidades**

**Comunalidades**

| | Inicial | Extracción |
|---|---|---|
| Web tecnologia | 1.000 | .658 |
| Web precio del servicio | 1.000 | .580 |
| Web planeación estrategica | 1.000 | .646 |
| Web imagen | 1.000 | .882 |
| Web calidad | 1.000 | .872 |
| Web desempeño | 1.000 | .616 |

Método de extracción: Análisis de Componentes principales.

Fuente: SPSS 20 IBM con elaboración propia

•En la **Figura 3.22**. se observa **la creación de 2 grupos** , cuya varianza conjunta es del **70.879% >60% o 71% de explicación**, (Porcentaje de la varianza explicada **no menor del 60%**). sobre la percepción que tienen los gerentes de las empresas de mercadotecnia digital.

**Figura 3.22. Varianza total explicada**

**Varianza total explicada**

| Componente | Autovalores iniciales | | | Sumas de las saturaciones al cuadrado de la extracción | | | Suma de las saturaciones al cuadrado de la rotación | | |
|---|---|---|---|---|---|---|---|---|---|
| | Total | % de la varianza | % acumulado | Total | % de la varianza | % acumulado | Total | % de la varianza | % acumulado |
| 1 | 2.513 | 41.885 | 41.885 | 2.513 | 41.885 | 41.885 | 2.370 | 39.499 | 39.499 |
| 2 | 1.740 | 28.994 | 70.879 | 1.740 | 28.994 | 70.879 | 1.883 | 31.380 | 70.879 |
| 3 | .598 | 9.959 | 80.838 | | | | | | |
| 4 | .530 | 8.830 | 89.667 | | | | | | |
| 5 | .416 | 6.928 | 96.595 | | | | | | |
| 6 | .204 | 3.405 | 100.000 | | | | | | |

Método de extracción: Análisis de Componentes principales.

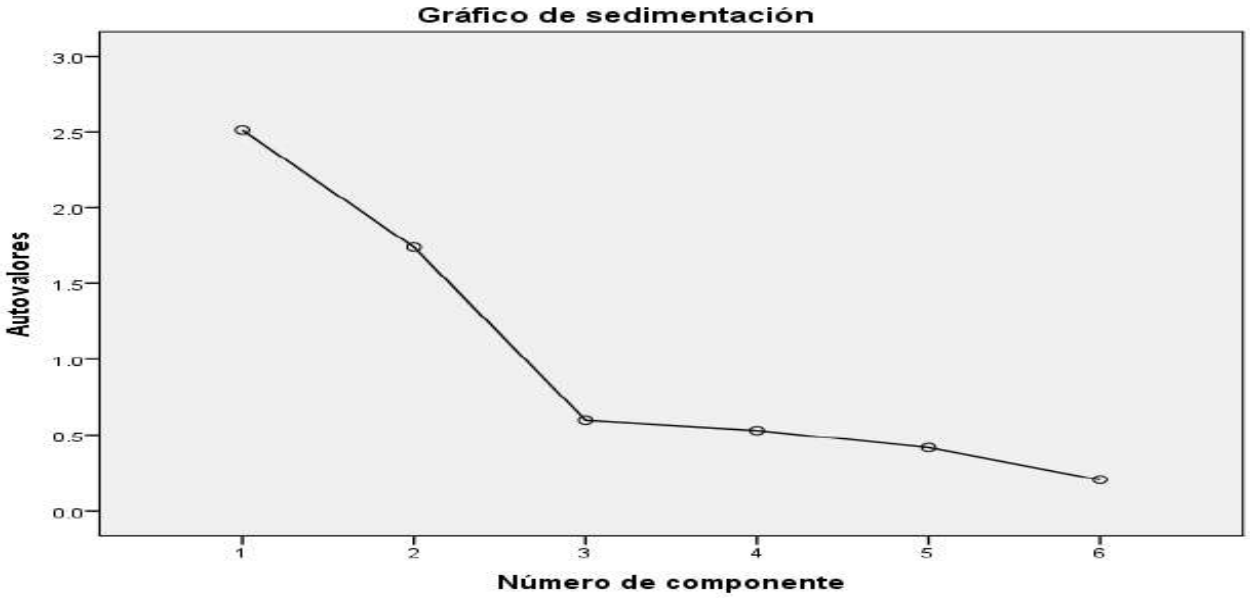
Fuente: SPSS 20 IBM con elaboración propia

**Paso 5: Interpretación**

**•Análisis de gráfico de sedimentación**

Al analizar, se observa que la formación de grupos es determinado por las pendientes que existen entre los nodos siendo representativos los 2 primeros y la baja posibilidad de que se forme un tercer grupo, dado que la pendiente tiende a ser cero. Sin embargo, lo que se pretende es que la menor cantidad de grupos me explique más del **60%** de la varianza extraída.**Ver Figura 3.23.**

**Figura 3.23. Gráfico de sedimentación**



Fuente: SPSS 20 IBM con elaboración propia

- La *Matriz de componentes* reporta los valores que están listos a rotar.
- La *Matriz de componentes rotados* reporta el método de extracción, rotación y las iteraciones de convergencia (en nuestro caso **3/25**). Se deberá cuestionar ¿qué variables son de cada grupo? Como se recordará, la variable **X5** al no tomarse en cuenta, **NO** significa que no sirva, sino que *es tan parecida a los 2 grupos creados que no determina a cual asignarlo***.** Se considera común a los grupos creados.
- Así, ***el resto de las variables se clasificarán en el grupo que les corresponde de acuerdo a aquel que tenga el valor absoluto más alto.***
- **¿Qué variables explican qué factor?**

La rotación ayuda, pues fuerza a que las variables tengan una correlación cercana a **1** con un factor y cercana a **0** con los demás.

- **No hay una rotación menor que otra.** Lo ideal es la que mejor permita interpretar los factores.
- Al analizar la **comunalidad** será capaz de determinar:
    - Qué varianza de cada variable recoge los factores.
    - Las cargas factoriales importantes e interpretar los factores que intervienen .
- Ver **Figuras 3.24** y **3.25**.

**Figura 3.24. Matriz de componentes**

**Matriz de componentes[a]**

| | Componente | |
|---|---|---|
| | 1 | 2 |
| Web tecnologia | -.627 | .514 |
| Web precio del servicio | .759 | -.068 |
| Web planeación estrategica | -.730 | .336 |
| Web imagen | .494 | .799 |
| Web calidad | .424 | .832 |
| Web desempeño | .767 | -.167 |

Método de extracción: Análisis de componentes principales.

a. 2 componentes extraídos

**Matriz de componentes rotados[a]**

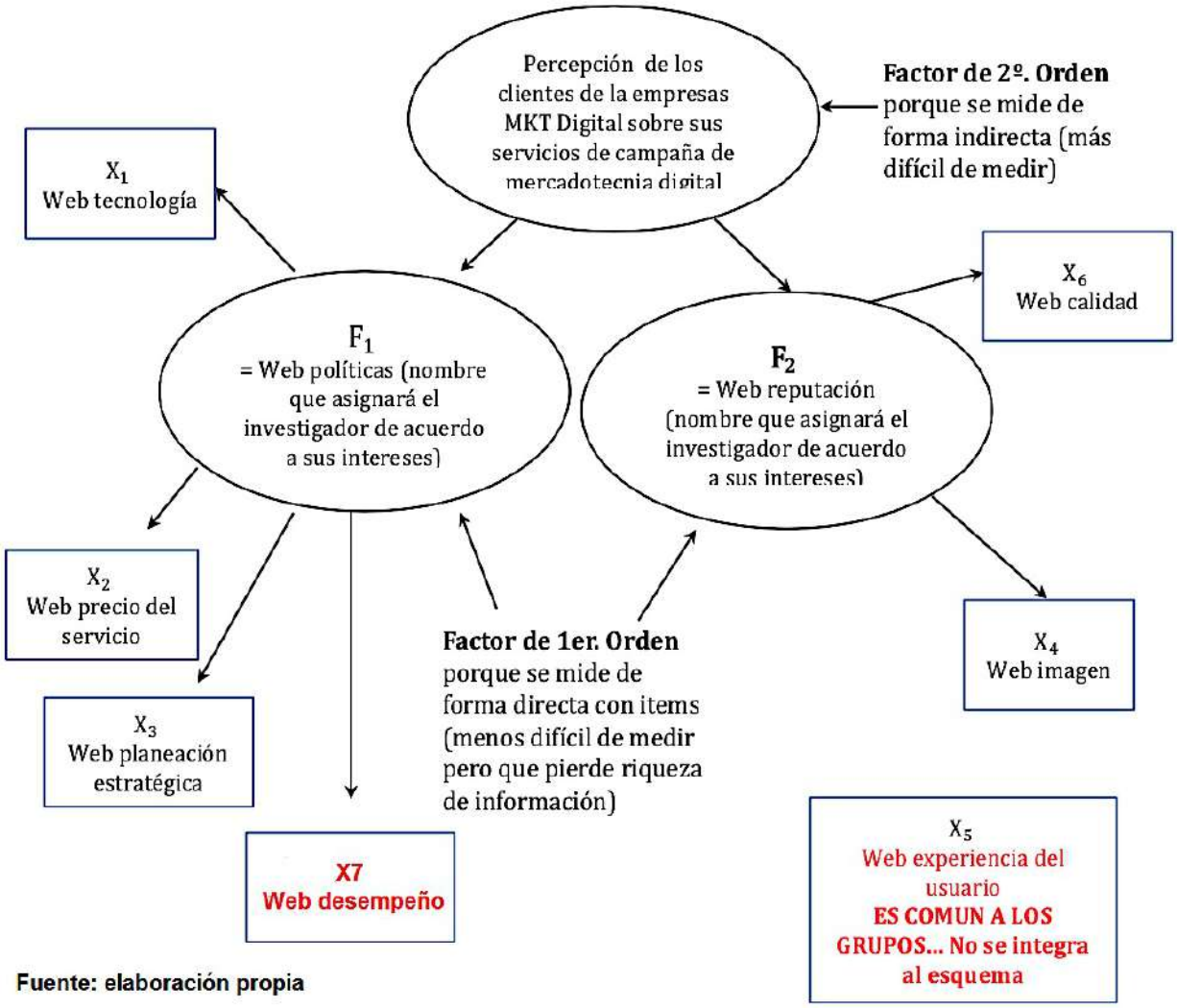
| | Componente | |
|---|---|---|
| | 1 | 2 |
| Web tecnologia | -.787 | .194 |
| Web precio del servicio | .714 | .265 |
| Web planeación estrategica | -.803 | -.011 |
| Web imagen | .102 | .933 |
| Web calidad | .025 | .934 |
| Web desempeño | .764 | .179 |

Método de extracción: Análisis de componentes principales.
Método de rotación: Normalización Varimax con Kaiser.

a. La rotación ha convergido en 3 iteraciones.

Fuente: SPSS 20 IBM con elaboración propia

**Figura 3.25. Modelo final del análisis factorial exploratorio**




**Fuente: elaboración propia**

**Discusión**

- Se genera un nuevo modelo con reagrupación de variables (items) y que en vez de ser explicado con 7 variables se explica con 2 nuevos factores y 6 variables. Se excluye 1 variable que no aporta valor al modelo
- Los 2 nuevos factores generados reciben el nombre que el investigador decida de acuerdo a sus intereses del proyecto
- Los factores finales se dividen en :
  - **Factores de 2o. Orden**, dado que se se mide de forma *indirecta* (la literatura sugiere evitar trabajar con este tipo de factores ya que es más difícil de medir, sin embargo, es capaz de verificar relaciones de estructura, dado que permite ver qué varoable es más importante de los factores de 1er orden y/o cual de éstos aporta más al modelo) y
  - **Factores de 1er. Orden,** porque se mide de forma *directa* con *items* (la

literatura sugiere su uso ya que es menos difícil de medir pero, pierde riqueza de información)

- **Nota:** Si se deseara un factor de 1er. Orden volverlo variable (item) haciendo una *media* de sus variables para incorporarlo al modelo de factor de 1er. Orden a variable*. Esto haría que el factor de 2o. Orden se vuelva incluso de 1er Orden.*
- La variable **X5** se vuelve *común (no es posible agregar)* a los 2 grupos y **No** se incluye en el modelo final, se considera que es básico, vital, primordial que brinde. Además, No puede haber factores con una sola variable (*incorrelacionada*), si se deseara tenerlo como constructo adicional. Lo que se elimina, es muy importante su discusión en el artículo a elaborar. Pudiera **X5** incluirse como factor 2o. Orden midiéndose en términos de % **NO** con escala de Likert. Posible estudiar con ecuaciones estructurales en 2o. Orden (estudios aún escasos al momento)
- Posible publicar en revistas CONACyT
- Por análisis comparativo, se tienen resultados similares de cada una de las mitades respecto al modelo general.
- Lo anterior se refleja en: *matriz de correlaciones*, ***KMO*** y prueba de ***Bartlett***, *matriz anti-imagen, gráfico de sedimentación, comunalidades, varianza total explicada y matriz de componentes rotados* Por lo tanto el modelo se presenta con mínimo error.
- Es posible aplicar ecuaciones estructurales para dar mayor explicación a las variables subyacentes,
- Posible publicar en revistas CONACyT como estudio empírico.

Posibles títulos:

- Construcción de un modelo empírico de percepción de los gerentes...
- Construcción de un modelo teórico es 100% académico

- Siguiente paso: buscar la teoría que justifique las nuevas relaciones los nuevos factores encontrados ya que estadísticamente está bien.

**Paso 6: Validación**

Para validar los resultados de un análisis factorial pueden emplearse dos métodos principales:

- Uno de ellos es llevar a cabo un *análisis factorial confirmatorio* , mediante sistemas de ecuaciones estructurales (**LISREL, AMOS, EQS**) lo que está fuera del objetivo de este curso).
- Otro procedimiento puede ser separar la muestra en dos mitades aleatoriamente y llevar a cabo un análisis factorial con cada una de ellas.
- Si el análisis de las cargas factoriales no difiere sustancialmente, podemos concluir que los resultados son robustos y estables.

## *Problema 7*

**Paso 1: Objetivo**

Validar el modelo resultante con el procedimiento de **separar la muestra en dos mitades aleatoriamente y llevar a cabo un análisis factorial con cada una de ellas**

**Paso 2: Diseño.** Idem al anterior

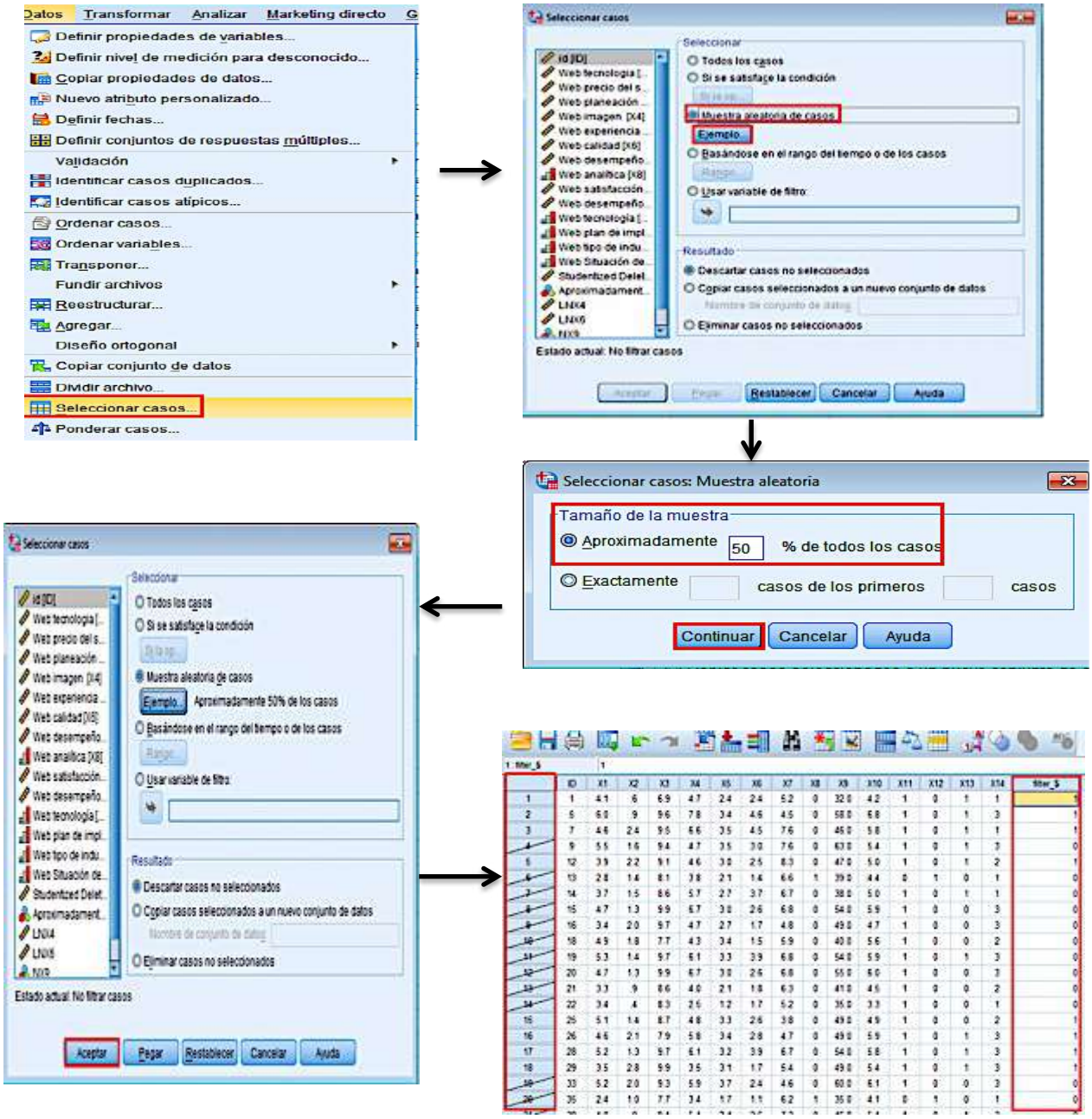
**Paso 3: Condiciones de aplicabilidad.** Se consideran que cumplen al caso.

**Paso 4: Ejecución y ajuste**

**Teclear línea de comandos SPSS:**

**Datos->Seleccionar casos->Seleccionar: muestra aleatoria de casos- >Ejemplo->Tamaño de muestra: 50%->Continuar->Aceptar.** Ir a vista de datos para comprobar (se observaran marcas diagonales en el extremo izquierdo de la pantalla y un nuevo campo: **filter$, con valores: 1/0**). Ver **Figura X. Proceso para validar resultados de análisis factorial por el proceso de las 2 mitades .Ver Figura 3.26**

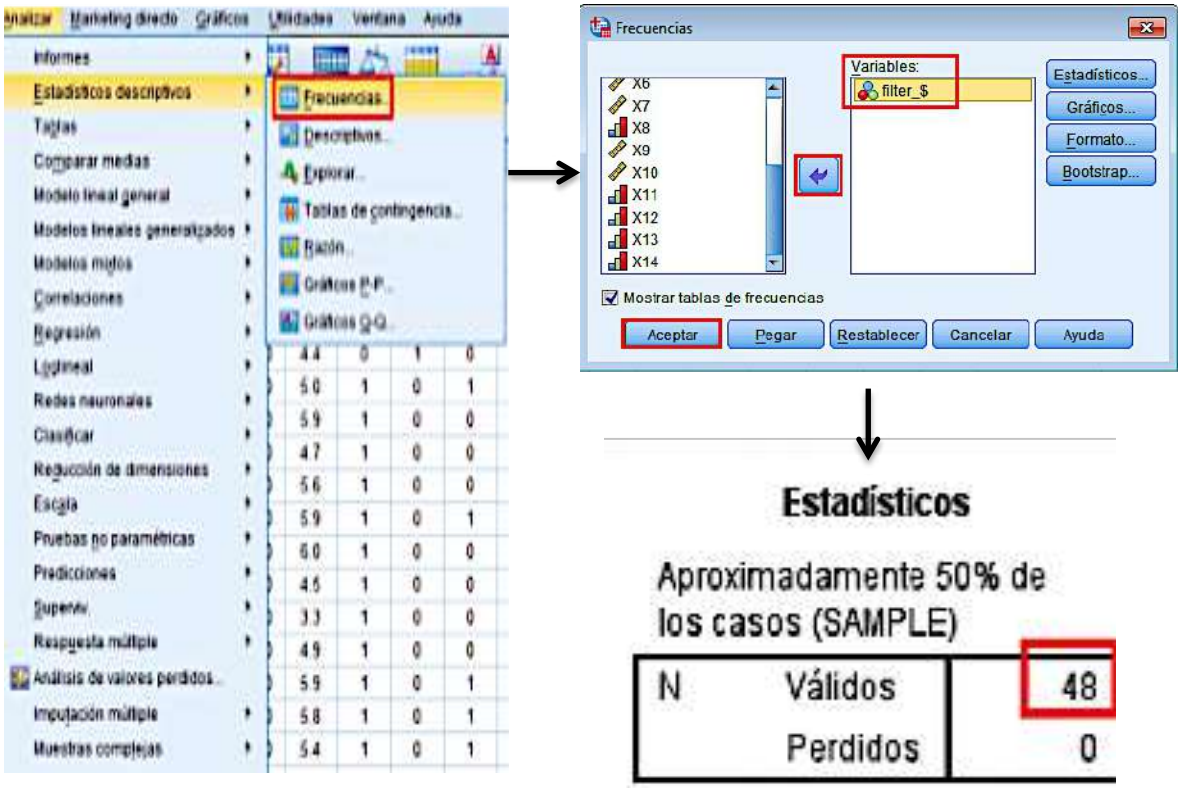
**Figura 3.26. Proceso para validar resultados de análisis factorial por el proceso de las 2 mitades**



Fuente: SPSS 20 IBM con elaboración propia

•Si se requiere saber cuántos datos son del filter$=0 y/o filter$=1,
**Teclear línea de comando SPSS: Analizar->Estadísticos descriptivos->Frecuencias->Seleccionar: filter_$ con flecha->Aceptar.**
•Se revisa la tabla de estadísticos en la que se reporta en nuestro caso, la división aproximada de la muestra como: **48% (1s, es decir, sin rayar ) y 52% (0s, es decir con rayar) por lo tanto como restante. Ver Figura 3.27.**

**Figura 3.27 Proceso de conteo de la división aproximada de la muestra**



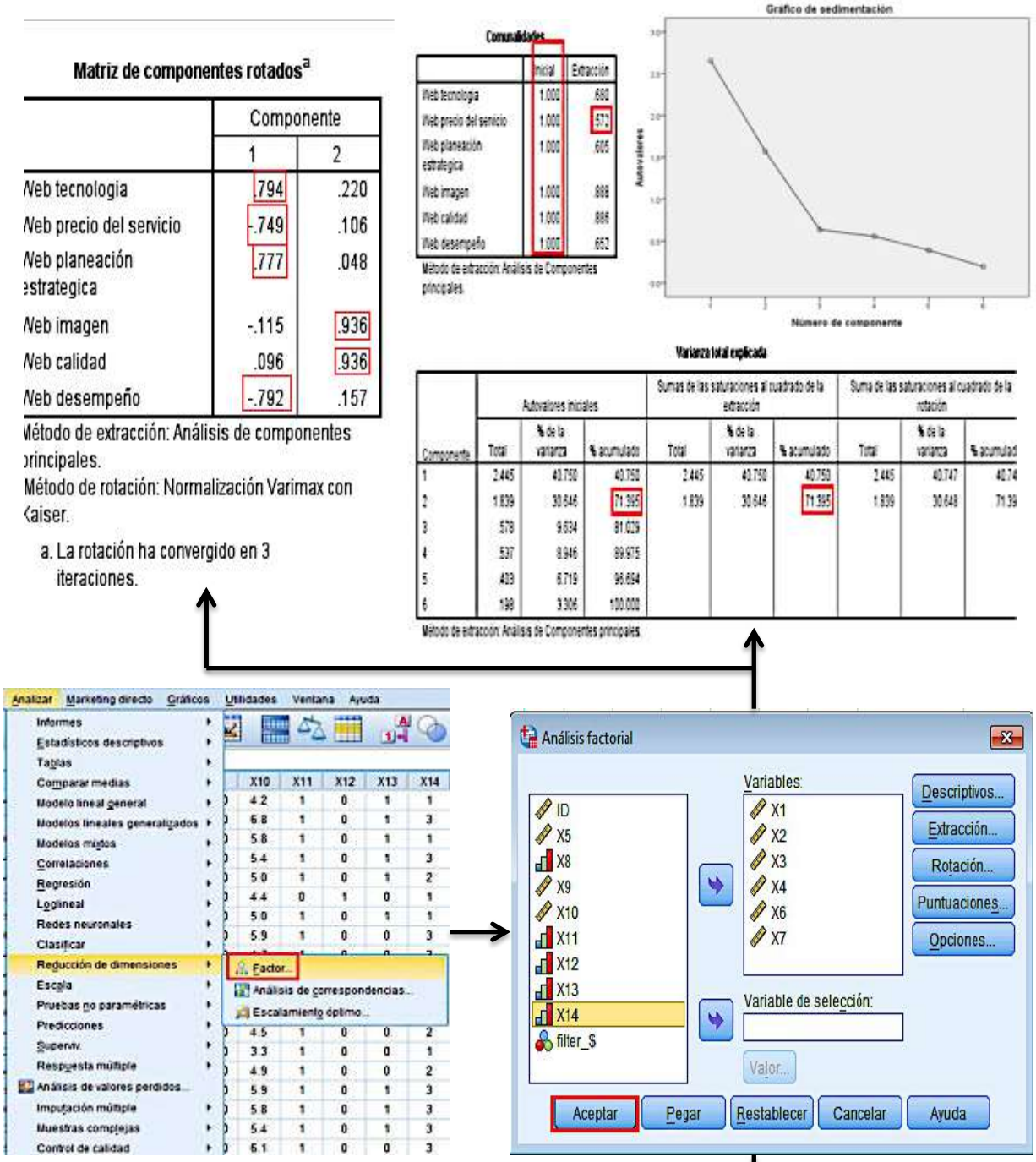

Fuente: SPSS 20 IBM con elaboración propia

•Realizar el proceso de análisis factorial nuevamente pero, con una de las muestras.
**Teclear línea de comando SPSS: Analizar->Reducción de dimensiones->Selección de variables métricas->Descriptivos-> Seleccionar: Estadísticos: Solución inicial; Matriz de correlaciones: Coeficientes; Niveles de significación-> *KMO* y prueba de esfericidad de *Bartlett*; Anti-imagen->Continuar->Extracción->Método: componentes principales->Analizar: Matriz de correlaciones->Visualización: solución factorial sin rotar-> Seleccionar: Gráfico de sedimentación- >Extraer: Autovalores mayores que: 1->Numero máximo de interacciones: 25->Continuar->Rotación->Método: Varimax->Visualización: solución rotada->No. máximo de iteraciones para convergencia: 25->Continuar->Puntuaciones->Continuar->Opciones->Valores perdidos (nota: no reemplazar por medias; no pérdida de datos >10%->Excluir casos según lista (nota: evitar**

**falta de datos)->Continuar->Aceptar**

•**Nota:** Los valores con muestra **filter_$=1/0 deberán ser similares a la muestra total**. No considerar la variable **X5** excluída previamente. Ver **Figura 3.28.**

**Figura 3.28. Proceso para validar resultados de análisis factorial con la primera mitad: filter_$=1.**



**Matriz de componentes rotados**[a]

| | Componente 1 | Componente 2 |
|---|---|---|
| Web tecnologia | .794 | .220 |
| Web precio del servicio | -.749 | .106 |
| Web planeación estrategica | .777 | .048 |
| Web imagen | -.115 | .936 |
| Web calidad | .096 | .936 |
| Web desempeño | -.792 | .157 |

Método de extracción: Análisis de componentes principales.
Método de rotación: Normalización Varimax con Kaiser.

a. La rotación ha convergido en 3 iteraciones.

**Comunalidades**

| | Inicial | Extracción |
|---|---|---|
| Web tecnologia | 1.000 | .680 |
| Web precio del servicio | 1.000 | .572 |
| Web planeación estrategica | 1.000 | .605 |
| Web imagen | 1.000 | .888 |
| Web calidad | 1.000 | .886 |
| Web desempeño | 1.000 | .652 |

Método de extracción: Análisis de Componentes principales.



**Varianza total explicada**

| Componente | Autovalores iniciales: Total | % de la varianza | % acumulado | Sumas de las saturaciones al cuadrado de la extracción: Total | % de la varianza | % acumulado | Suma de las saturaciones al cuadrado de la rotación: Total | % de la varianza | % acumulado |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2.445 | 40.750 | 40.750 | 2.445 | 40.750 | 40.750 | 2.445 | 40.747 | 40.74 |
| 2 | 1.839 | 30.646 | 71.395 | 1.839 | 30.646 | 71.395 | 1.839 | 30.648 | 71.39 |
| 3 | .578 | 9.634 | 81.029 | | | | | | |
| 4 | .537 | 8.946 | 89.975 | | | | | | |
| 5 | .403 | 6.719 | 96.694 | | | | | | |
| 6 | .198 | 3.306 | 100.000 | | | | | | |

Método de extracción: Análisis de Componentes principales.

**Matriz de correlaciones**

| | | Web tecnologia | Web precio del servicio | Web planeación estrategica | Web imagen | Web calidad | Web desempeño |
|---|---|---|---|---|---|---|---|
| rrelación | Web tecnologia | 1.000 | -.419 | .513 | .079 | .221 | -.532 |
| | Web precio del servicio | -.419 | 1.000 | -.452 | .173 | -.028 | .475 |
| | Web planeación estrategica | .513 | -.452 | 1.000 | -.053 | .096 | -.455 |
| | Web imagen | .079 | .173 | -.053 | 1.000 | .774 | .178 |
| | Web calidad | .221 | -.028 | .096 | .774 | 1.000 | .066 |
| | Web desempeño | -.532 | .475 | -.455 | .178 | .066 | 1.000 |
| (Unilateral) | Web tecnologia | | .002 | .000 | .298 | .066 | .000 |
| | Web precio del servicio | .002 | | .001 | .120 | .426 | .000 |
| | Web planeación estrategica | .000 | .001 | | .359 | .259 | .001 |
| | Web imagen | .298 | .120 | .359 | | .000 | .113 |
| | Web calidad | .066 | .426 | .259 | .000 | | .329 |
| | Web desempeño | .000 | .000 | .001 | .113 | .329 | |

**KMO y prueba de Bartlett**

| | | |
|---|---|---|
| Medida de adecuación muestral de Kaiser-Meyer-Olkin. | | .855 |
| Prueba de esfericidad de Bartlett | Chi-cuadrado aproximado | 94.873 |
| | gl | 15 |
| | Sig. | .000 |

**Matrices anti-imagen**

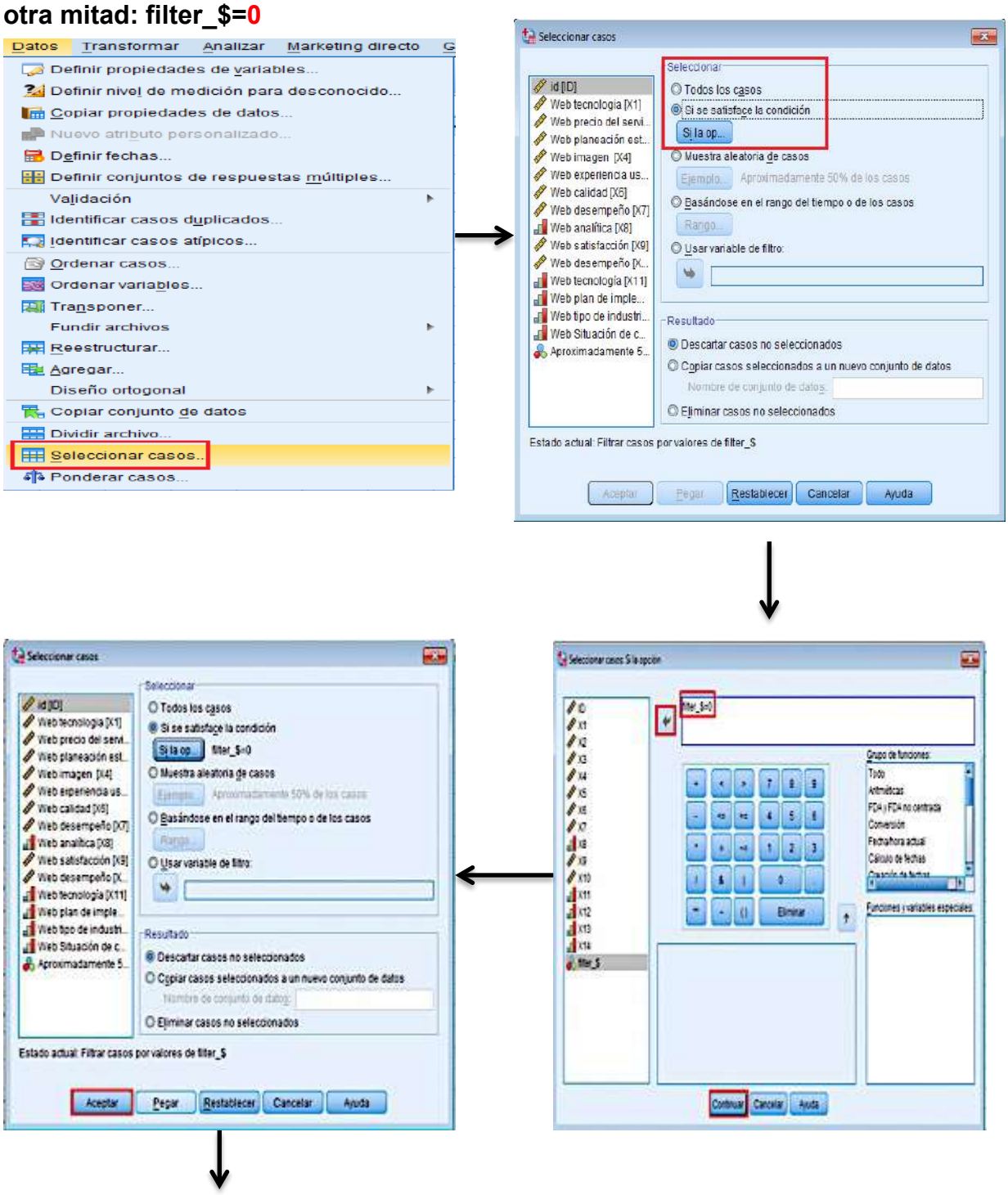
| | | Web tecnologia | Web precio del servicio | Web planeación estrategica | Web imagen | Web calidad | Web desempe |
|---|---|---|---|---|---|---|---|
| Covarianza anti-imagen | Web tecnologia | .568 | .080 | -.177 | -.012 | -.068 | |
| | Web precio del servicio | .080 | .659 | .145 | -.137 | .080 | |
| | Web planeación estrategica | -.177 | .145 | .647 | .030 | -.035 | |
| | Web imagen | -.012 | -.137 | .030 | .360 | -.279 | |
| | Web calidad | -.068 | .080 | -.035 | -.279 | .360 | |
| | Web desempeño | .218 | -.154 | .104 | -.024 | -.044 | |
| Correlación anti-imagen | Web tecnologia | .744[a] | .130 | -.292 | -.026 | -.151 | |
| | Web precio del servicio | .130 | .758[a] | .222 | -.220 | .165 | |
| | Web planeación estrategica | -.292 | .222 | .799[a] | .063 | -.073 | |
| | Web imagen | -.026 | -.220 | .063 | .505[a] | -.776 | |
| | Web calidad | -.151 | .165 | -.073 | -.776 | .499[a] | |
| | Web desempeño | .378 | -.246 | .169 | -.061 | -.095 | |

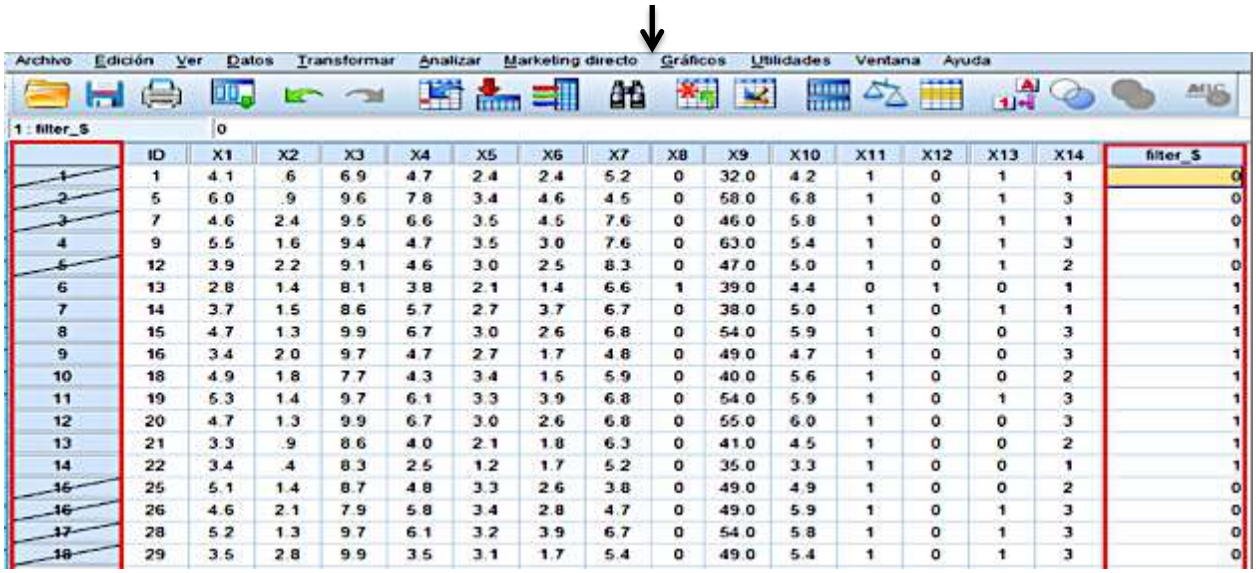
a. Medida de adecuación muestral

Fuente: SPSS 20 IBM con elaboración propia

**Teclear línea de comandos de SPSS:**
**Datos->Seleccionar casos->Seleccionar: si se satisface la siguiente condición->Seleccionar filter_$->flecha->filter_$=0->Continuar->Aceptar. Ir a visor de datos para comprobar rayado de extrema izquierda en los datos y filter_$= 0 rayados/1 sin rayar. Ver Figura 3.29.**

**Figura 3-29.. Proceso para validar resultados de análisis factorial con la otra mitad: filter_$=0**





Archivo Edición Ver Datos Transformar Analizar Marketing directo Gráficos Utilidades Ventana Ayuda

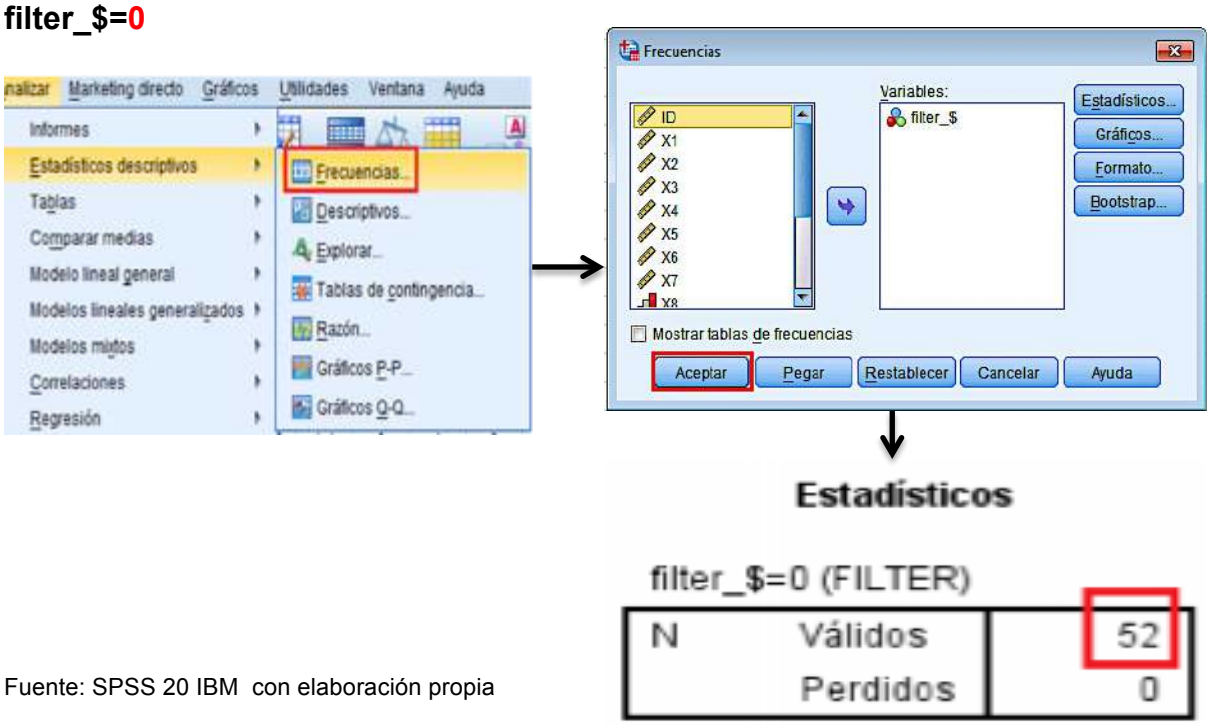
1 : filter_$ 0

| | ID | X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 | X10 | X11 | X12 | X13 | X14 | filter_$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 4.1 | .6 | 6.9 | 4.7 | 2.4 | 2.4 | 5.2 | 0 | 32.0 | 4.2 | 1 | 0 | 1 | 1 | 0 |
| 2 | 5 | 6.0 | .9 | 9.6 | 7.8 | 3.4 | 4.6 | 4.5 | 0 | 58.0 | 6.8 | 1 | 0 | 1 | 3 | 0 |
| 3 | 7 | 4.6 | 2.4 | 9.5 | 6.6 | 3.5 | 4.5 | 7.6 | 0 | 46.0 | 5.8 | 1 | 0 | 1 | 1 | 0 |
| 4 | 9 | 5.5 | 1.6 | 9.4 | 4.7 | 3.5 | 3.0 | 7.6 | 0 | 63.0 | 5.4 | 1 | 0 | 1 | 3 | 1 |
| 5 | 12 | 3.9 | 2.2 | 9.1 | 4.6 | 3.0 | 2.5 | 8.3 | 0 | 47.0 | 5.0 | 1 | 0 | 1 | 2 | 0 |
| 6 | 13 | 2.8 | 1.4 | 8.1 | 3.8 | 2.1 | 1.4 | 6.6 | 1 | 39.0 | 4.4 | 0 | 1 | 0 | 1 | 1 |
| 7 | 14 | 3.7 | 1.5 | 8.6 | 5.7 | 2.7 | 3.7 | 6.7 | 0 | 38.0 | 5.0 | 1 | 0 | 1 | 1 | 1 |
| 8 | 15 | 4.7 | 1.3 | 9.9 | 6.7 | 3.0 | 2.6 | 6.8 | 0 | 54.0 | 5.9 | 1 | 0 | 0 | 3 | 1 |
| 9 | 16 | 3.4 | 2.0 | 9.7 | 4.7 | 2.7 | 1.7 | 4.8 | 0 | 49.0 | 4.7 | 1 | 0 | 0 | 3 | 1 |
| 10 | 18 | 4.9 | 1.8 | 7.7 | 4.3 | 3.4 | 1.5 | 5.9 | 0 | 40.0 | 5.6 | 1 | 0 | 0 | 2 | 1 |
| 11 | 19 | 5.3 | 1.4 | 9.7 | 6.1 | 3.3 | 3.9 | 6.8 | 0 | 54.0 | 5.9 | 1 | 0 | 1 | 3 | 1 |
| 12 | 20 | 4.7 | 1.3 | 9.9 | 6.7 | 3.0 | 2.6 | 6.8 | 0 | 55.0 | 6.0 | 1 | 0 | 0 | 3 | 1 |
| 13 | 21 | 3.3 | .9 | 8.6 | 4.0 | 2.1 | 1.8 | 6.3 | 0 | 41.0 | 4.5 | 1 | 0 | 0 | 2 | 1 |
| 14 | 22 | 3.4 | .4 | 8.3 | 2.5 | 1.2 | 1.7 | 5.2 | 0 | 35.0 | 3.3 | 1 | 0 | 0 | 1 | 1 |
| 15 | 25 | 5.1 | 1.4 | 8.7 | 4.8 | 3.3 | 2.6 | 3.8 | 0 | 49.0 | 4.9 | 1 | 0 | 0 | 2 | 0 |
| 16 | 26 | 4.6 | 2.1 | 7.9 | 5.8 | 3.4 | 2.8 | 4.7 | 0 | 49.0 | 5.9 | 1 | 0 | 1 | 3 | 0 |
| 17 | 28 | 5.2 | 1.3 | 9.7 | 6.1 | 3.2 | 3.9 | 6.7 | 0 | 54.0 | 5.8 | 1 | 0 | 1 | 3 | 0 |
| 18 | 29 | 3.5 | 2.8 | 9.9 | 3.5 | 3.1 | 1.7 | 5.4 | 0 | 49.0 | 5.4 | 1 | 0 | 1 | 3 | 0 |

Fuente: SPSS 20 IBM con elaboración propia

•**Para verificar descriptivos, Teclear: Datos->Seleccionar casos->Seleccionar: si se satisface la siguiente condición->Seleccionar filter_$->flecha- >filter_$=0->Continuar->Aceptar.** Ir a visor de datos para comprobar rayado de extrema izquierda en los datos y filter_$= 0 rayados/1 sin rayar. Resultados: **52 casos, el resto es de 1s. Ver Figura 3.30.**

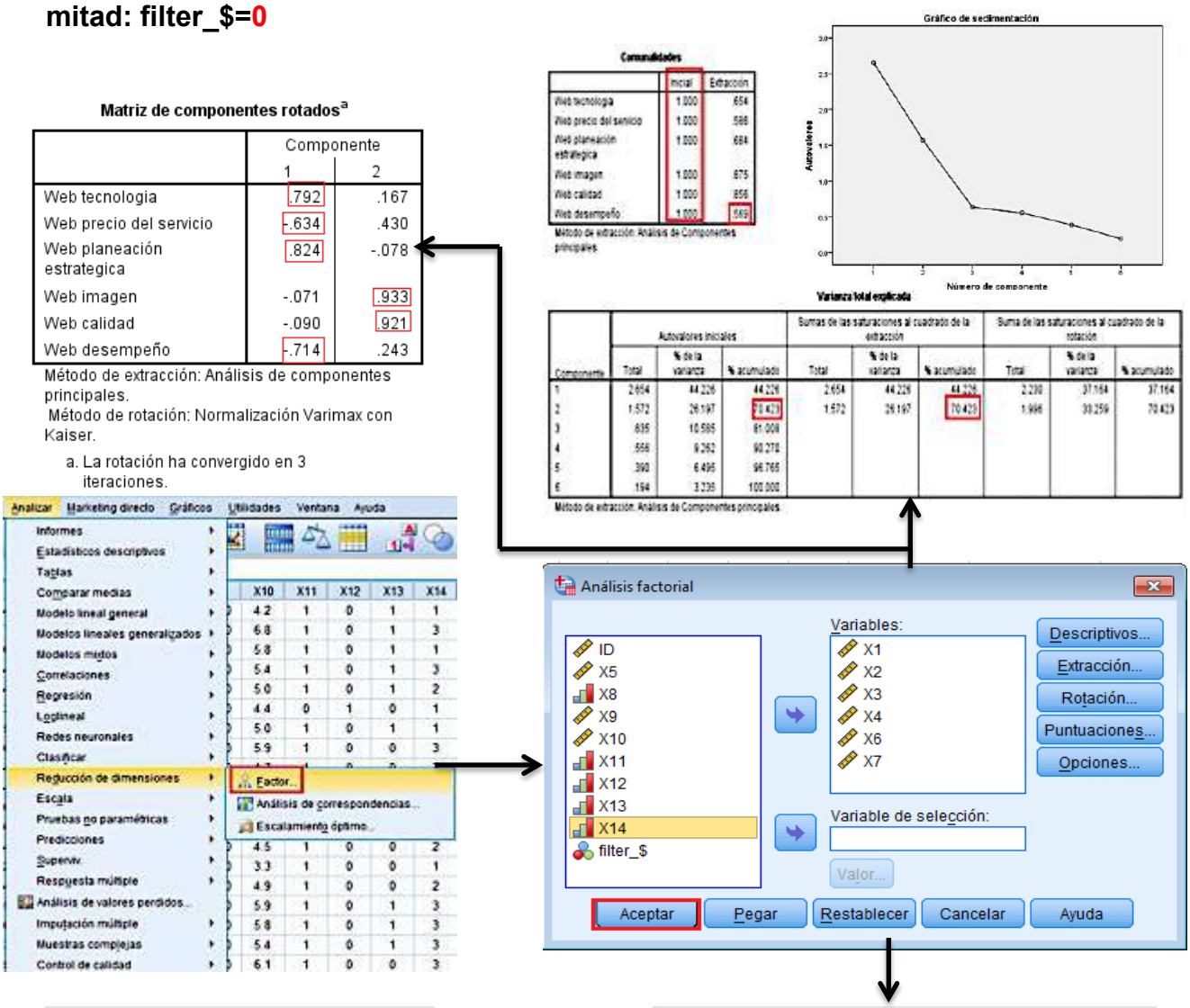
**Figura 3.30. Proceso para verificar las frecuencias de la segunda mitad filter_$=0**




Fuente: SPSS 20 IBM con elaboración propia

•Realizar el proceso de análisis factorial nuevamente pero, con una de las muestras. **Teclear líneas de comando SPSS: Analizar->Reducción de dimensiones->Selección de variables métricas- >Descriptivos-> Seleccionar: Estadísticos: Solución inicial; Matriz de correlaciones: Coeficientes; Niveles de significación-> *KMO* y prueba de esfericidad de *Bartlett*; Anti-imagen->Continuar->Extracción->Método: componentes principales->Analizar: Matriz de correlaciones->Visualización: solución factorial sin rotar-> Seleccionar: Gráfico de sedimentación->Extraer: Autovalores mayores que: 1->Numero máximo de interacciones: 25->Continuar- >Rotación->Método: Varimax->Visualización: solución rotada->No. máximo de iteraciones para convergencia: 25->Continuar->Puntuaciones->Continuar- >Opciones->Valores perdidos (nota: no reemplazar por medias; no pérdida de datos >10%->Excluir casos según lista (nota: evitar falta de datos)->Continuar- >Aceptar. Ver Figura 3.31.**

**Figura 3.31. Proceso para validar resultados de análisis factorial con la primera mitad: filter_$=0**



**Matriz de correlaciones**

| | | Web tecnologia | Web precio del servicio | Web planeación estrategica | Web imagen | Web calidad | Web desempeño |
|---|---|---|---|---|---|---|---|
| Correlación | Web tecnologia | 1.000 | -.268 | .503 | .032 | -.027 | -.422 |
| | Web precio del servicio | -.268 | 1.000 | -.514 | .365 | .348 | .438 |
| | Web planeación estrategica | .503 | -.514 | 1.000 | -.165 | -.128 | -.430 |
| | Web imagen | .032 | .365 | -.165 | 1.000 | .797 | .234 |
| | Web calidad | -.027 | .348 | -.128 | .797 | 1.000 | .271 |
| | Web desempeño | -.422 | .438 | -.430 | .234 | .271 | 1.000 |
| Sig. (Unilateral) | Web tecnologia | | .027 | .000 | .410 | .425 | .001 |
| | Web precio del servicio | .027 | | .000 | .004 | .006 | .001 |
| | Web planeación estrategica | .000 | .000 | | .122 | .184 | .001 |
| | Web imagen | .410 | .004 | .122 | | .000 | .047 |
| | Web calidad | .425 | .006 | .184 | .000 | | .026 |
| | Web desempeño | .001 | .001 | .001 | .047 | .026 | |

**KMO y prueba de Bartlett**

| | | |
|---|---|---|
| Medida de adecuación muestral de Kaiser-Meyer-Olkin | | .665 |
| Prueba de esfericidad de Bartlett | Chi-cuadrado aproximado | 105.729 |
| | gl | 15 |
| | Sig. | .000 |

**Matrices anti imagen**

| | | Web tecnologia | Web precio del servicio | Web planeación estrategica | Web imagen | Web calidad | Web desempeño |
|---|---|---|---|---|---|---|---|
| Covarianza anti-imagen | Web tecnologia | .667 | -.009 | -.239 | -.079 | .030 | .188 |
| | Web precio del servicio | -.009 | .613 | .224 | -.056 | -.040 | -.140 |
| | Web planeación estrategica | -.239 | .224 | .574 | .045 | -.052 | .086 |
| | Web imagen | -.079 | -.056 | .045 | .345 | -.266 | -.004 |
| | Web calidad | .030 | -.040 | -.052 | -.266 | .351 | -.055 |
| | Web desempeño | .188 | -.140 | .086 | -.004 | -.055 | .672 |
| Correlación anti-imagen | Web tecnologia | .660[a] | -.014 | -.387 | -.165 | .063 | .280 |
| | Web precio del servicio | -.014 | .787[a] | .378 | -.121 | -.086 | -.218 |
| | Web planeación estrategica | -.387 | .378 | .690[a] | .102 | -.116 | .138 |
| | Web imagen | -.165 | -.121 | .102 | .572[a] | -.764 | -.008 |
| | Web calidad | .063 | -.086 | -.116 | -.764 | .577[a] | -.113 |
| | Web desempeño | .280 | -.218 | .138 | -.008 | -.113 | .612[a] |

a. Medida de adecuación muestral

Fuente: SPSS 20 IBM con elaboración propia

**Paso 5: Interpretación**

•Por análisis comparativo, se tienen resultados similares de cada una de las mitades respecto al modelo general.

•Lo anterior se refleja en : *matriz de correlaciones, KMO y prueba de Bartlett, matriz anti-imagen, gráfico de sedimentación, comunalidades, varianza total explicada y matriz de componentes rotados*

•Por lo tanto el modelo se presenta con mínimo error. Es posible aplicar ecuaciones estructurales para dar mayor explicación a las variables subyacentes,

•Posible publicar en revistas CONACyT como estudio empírico. Posibles títulos:

- •Construcción de un modelo empírico de percepción de los gerentes...
- •Construcción de un modelo teórico es 100% académico

•Siguiente paso: buscar la teoría que justifique las nuevas relaciones los nuevos factores encontrados ya que estadísticamente está bien.

# Conclusión

Este capítulo abordó los problemas en el Paso 3 del proceso de desarrollo de escala: **prueba piloto** de un conjunto de indicadores como reducción de ítems y validez inicial del procedimiento de prueba y realización de múltiples estudios para el desarrollo de la escala. En el segundo tema, destacamos:

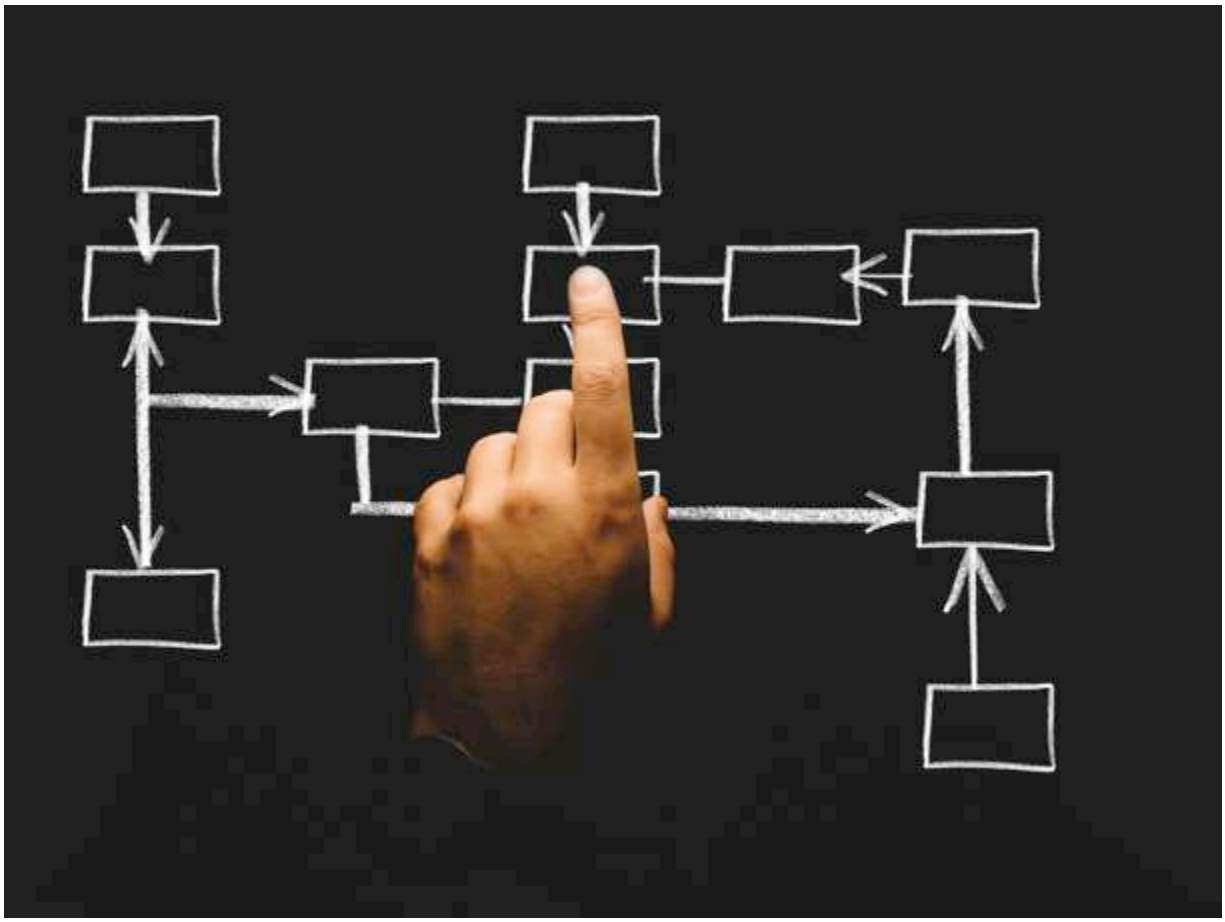
**a.**La importancia de incluir constructos para evaluar los diversos tipos de validez,

**b.** El **EFA** sobre múltiples conjuntos de datos para refinar la escala y examinar una estructura de **factores teóricos *a priori***, así como los

**c. Indicadores y análisis de confiabilidad**.

Finalmente, ofrecimos un ejemplo de una escala desarrollada recientemente con **EFA** y los análisis de ítems y confiabilidad pueden usarse en el desarrollo de la escala. El capítulo 7 se centra en los procedimientos de finalización de la escala.

# CAPÍTULO 7. FINALIZANDO LA ESCALA. ANÁLISIS FACTORIAL CONFIRMATORIO Y ECUACION ESTRUCTURAL (Etapa 4)



El análisis factorial confirmatorio (**CFA.** *Confirmatory Factor Analysis*), es una forma especial de análisis factorial muy utilizado en las ciencias económico-administrativas y sociales. Es utilizado para determinar si las mediciones de un constructo son consistentes con el juicio del investigador, sobre la naturaleza del propio constructo (o factor).

*Su objetivo es probar si los datos se ajustan a un modelo de medida hipotético.* Dicho modelo hipotético se basa en la teoría y/o el análisis previo que la investigación arroja. Hace referencia a las pautas de exploración de las relaciones entre varias variables. Estas pautas se representan por lo que se denominan *componentes principales* o, más comúnmente, *factores.* Como las variables ponderan fuertemente sobre un factor, se hacen descriptores de las dimensiones subyacentes.

En el **CFA**, el investigador primero desarrolla una hipótesis sobre que factores considera se encuentran subyacentes en las mediciones usadas (por ejemplo, **OBM.**Modelo de Negocios Abierto, en Mejía-Trejo, 2017), los cuales, imponen restricciones en las hipótesis previamente propuestas. Ante dicha imposición, el investigador forza el modelo a fin de lograr sea consistente con la teoría. Por ejemplo, si se postula que hay dos factores que explican la covarianza en las medidas, y que estos factores no están relacionados entre sí, el investigador se encuentra en la posibilidad de crear un modelo donde la correlación entre los factores A y B sean limitados a cero. Luego, se podrían obtener medidas de ajuste del modelo para evaluar qué tan bien el modelo propuesto capturó la covarianza entre todos los ítems o medidas en el modelo. *Si las restricciones que el investigador ha impuesto sobre el modelo son inconsistentes con los datos de la muestra, entonces los resultados de las pruebas estadísticas del ajuste del modelo indicarán un ajuste pobre, y el modelo será rechazado*. Existen diversas razones por las que el ajuste del modelo es pobre; una, puede deberse a algunos elementos que miden factores múltiples; o también podría ser que algunos elementos dentro de un factor están más relacionados entre sí que otros.

Para algunas aplicaciones, el requisito de *cargas cero* (para indicadores que no se supone que se cargan en un factor determinado) se ha considerado demasiado estricto. Un método de análisis recientemente desarrollado, *modelado de ecuaciones estructurales exploratorias* (*exploratory structural equation modeling*), especifica las hipótesis sobre la relación entre los indicadores observados y sus supuestos factores primarios latentes, al tiempo también, de que permite la estimación de las cargas con otros factores latentes.

En el **CFA**, los investigadores se encuentran típicamente interesados en estudiar, el grado en el cual, las respuestas de un vector (***px1*** ) de variables observable al azar, tengan la posibilidad de ser usadas para asignar un valor a una o más de las variables subyacentes (**n**). La investigación se logra en gran medida, al estimar y evaluar la carga de cada elemento utilizado para tocar aspectos de la variable latente no observada. Es decir, **y [i]** es el vector de respuestas observadas pronosticadas por la variable latente no observada, que se define como:

$$\mathbf{Y} = \wedge\, \xi + \epsilon$$

Donde:

Y.Representa el vector ***px1*** de variables observadas al azar.

$\xi$.Representa a las variables latentes no observadas.

$\wedge$.Representa a la matriz ***p*x*k***, con **k=** número de variables latentes.

Ya que **Y** es la medición imperfecta de $\xi$, el modelo tiene por lo tanto el error $\epsilon$**.** Al realizar el caso estimado en **ML** (*maximum likelihood*), generado al minimizar iterativamente la función de ajuste, se produce:

Donde:

$$\Lambda\Omega\Lambda' + I - diag(\Lambda\Omega\Lambda')$$

Es la matriz de varianza-covarianza implicada por el modelo de análisis factorial propuesto y ***R*** es la matriz de varianza- covarianza observada.
Es decir, se encuentran valores para los parámetros del modelo liberado que minimiza la diferencia entre la matriz de varianza-covarianza en modo implícito y la matriz de varianza-covarianza observada.
Sólo con el examen de las ponderaciones de las variables sobre los factores se identifica el carácter de la dimensión subyacente. Así, hay una la similitud entre los objetivos del *análisis factorial y el modelo de medida del* **SEM**. Los factores son, en términos de los modelos medidos, las *variables latentes*. Cada variable actúa como indicador de cada factor (dado que cada variable tiene una ponderación para cada factor). Utilizado de esta forma, el análisis factorial es primordialmente una *técnica exploratoria* porque el investigador tiene un *control limitado* sobre las variables consideradas indicadores del constructo latente (es decir, aquellas variables ponderadas sobre cada factor). **SEM**, sin embargo, tiene un papel confirmatorio porque el investigador tiene un *control completo* sobre la especificación de los indicadores de cada constructo. Además, **SEM** permite un *test estadístico de calidad del ajuste* para la solución confirmatoria del factor propuesta, que *no es posible con los componentes principales* o el análisis de factor. El análisis de factor confirmatorio (**CFA**) es particularmente útil en la validación de las escalas para la medida de los constructos específicos.
El modelado de ecuaciones estructurales (**SEM.***Structural Equation Modeling*), tiene una representación, como se muestra en la **Ecuación 7.1.**

**Ecuación 7.1. Representación de la ecuación estructural.**

$$Y_1 = X_{11} + X_{12} + X_{13} + \ldots X_{1n}$$
$$Y_2 = X_{21} + X_{22} + X_{23} + \ldots X_{2n}$$
$$Y_m = X_{m1} + X_{m2} + X_{m3} + \ldots X_{mn}$$

(**Métrica**) (**Métrica, No Métrica**)

El modelo de ecuaciones estructurales tienen una gran oportunidad de aplicación en las ciencias económico-administrativas ya que se tienen trabajos documentados en diversos campos que incluyen la mercadotecnia, la organización, la psicología, la educación, la psicología, la sociología, la gestión, la innovación, la salud, el comportamiento organizacional, la biología e incluso la genética. Las razones de su atractivo para áreas tan diversas son dobles:
**1.**Proporciona un método directo de tratar con múltiples relaciones simultáneamente a la vez que se da eficacia estadística, y
**2.**Su capacidad para evaluar las relaciones exhaustivamente y proporcionar una transición desde el análisis exploratorio al confirmatorio. Esta transición corresponde a los mayores esfuerzos en todos los campos de estudio hacia el desarrollo de perspectivas más sistemáticas y holísticas de los problemas. Tales esfuerzos exigen la capacidad de contrastar una serie de relaciones que constituyen un modelo a gran escala, un conjunto de principios fundamentales o una teoría entera. Estas son tareas para las que el modelo de ecuaciones estructurales se ajusta perfectamente.

Es una metodología estadística que parte de un enfoque confirmatorio (es decir, hipótesis-prueba) para el análisis de una teoría estructural, basado en un fenómeno a explicar algún fenómeno. Típicamente, esta teoría representa procesos *causales* que generan observaciones en múltiples variables (Bentler, 1988). .En nuestro caso, una *causa* es un efecto directo de una variable sobre otra dentro del contexto de un modelo completo. Su magnitud y dirección están dadas por el coeficiente de regresión parcial. Si el modelo completo contiene todas las influencias relevantes en una variable dependiente dada, sus precursores causales son correctamente especificados. En la práctica, sin embargo, los modelos pueden omitir los predictores clave y pueden ser mal especificados, por lo que pueden ser inadecuados como un *modelo causal* en sentido estricto**.**

Por cierto, a lo largo de esta obra, los términos: *variables latentes, no observadas y no medidas se usan como sinónimos para representar un constructo o factor hipotético; los términos: variables observadas, manifestadas y medidas también se usan indistintamente.*

El término **SEM** transmite dos aspectos importantes del procedimiento:

**a.**Que los procesos causales en estudio están representados por una serie de ecuaciones estructurales (*es decir, regresión*), y

**b.**Que éstas relaciones estructurales pueden modelarse gráficamente para permitir una conceptualización más clara de la teoría en estudio.

*El modelo hipotético, puede por tanto, ser probado estadísticamente en un análisis simultáneo de todo el sistema de variables para determinar el grado en que es consistente con los datos***.** Si la *bondad del ajuste* (*goodness-of-fit*) es adecuada, el modelo valida relaciones postuladas entre las variables; si es inadecuado, las relaciones son rechazadas. Varios aspectos del **SEM** lo diferencian de otras generaciones anteriores de procedimientos multivariantes. Ver **Tabla 7.1**

**Tabla 7.1. Aspectos de la SEM**

| Descripción |
|---|
| Como se señaló anteriormente, parte de un enfoque confirmatorio más que exploratorio para el análisis de datos (aunque aspectos de este último pueden abordarse). Además, al exigir que el patrón de relaciones se especifique a priori, **SEM** se presta bien al análisis de datos con fines inferenciales. Por el contrario, la mayoría otros procedimientos multivariantes son esencialmente descriptivos por naturaleza (por ejemplo, **EFA**) por lo que la prueba de hipótesis es difícil, si no imposible. |
| Mientras que los procedimientos multivariantes tradicionales son incapaces de evaluar o corregir errores de medición, **SEM** proporciona estimaciones explícitas de estos parámetros de varianza de error. De hecho, los métodos alternativos (por ejemplo, los que están basados en la regresión o el modelo lineal general-*general lineal model*) suponen que el error en las variables explicativas (es decir, independientes) desaparece. Por lo tanto, aplicando esos métodos cuando hay un error en las variables explicativas equivale a ignorar el error que puede conducir, en última instancia, a inexactitudes serias, especialmente cuando los errores son considerables. Tales errores se evitan cuando se utilizan los análisis SEM correspondientes. |

| |
|---|
| Mientras que los análisis de datos que usan los primeros métodos se basan en mediciones observadas solo aquellos que usan procedimientos **SEM** pueden incorporar variables no observadas (es decir, latentes) y observadas |
| Finalmente, no existen métodos alternativos ampliamente aplicados para el modelado de relaciones multivariantes, aunque sí para el **SEM** |

Fuente: Byrne (2006) con adaptación propia

Dadas estas características altamente deseables, **SEM** se ha convertido en una metodología muy popular para la *investigación no experimental* en las ciencias económico-administrativas, en la que los métodos para probar las teorías no están bien desarrollados y las consideraciones éticas hacen que el diseño experimental sea inviable (Bentler, 1980). **SEM** se puede utilizar de manera efectiva para abordar numerosos problemas de investigación que involucran *investigación no experimental.* Aquí se planteará un problema para su seguimiento en **EQS 6** (Bentler, 2005; Bentler Wu, 2002)**,** aunque existen casos diversos para su uso

Todas las técnicas **SEM** se distinguen por 2 características:

-La estimación de relaciones de dependencia múltiples y cruzadas, y

-La capacidad de representar conceptos no observados en estas relaciones y tener en cuenta el error de medida en el proceso de estimación.

La diferencia más obvia entre **SEM** y otras técnicas de relaciones multivariantes es el uso de relaciones distintas para cada conjunto de variables dependientes. En los términos más sencillos, **SEM** estima una serie de ecuaciones de regresión múltiples distintas pero interrelacionadas mediante la especificación del *modelo estructural* utilizado por el programa estadístico. En primer lugar, el investigador utiliza la teoría, la experiencia previa y los objetivos de investigación para diferenciar qué variables independientes predicen cada variable dependiente. En nuestro ejemplo de los capítulos anteriores, queríamos en primer lugar predecir la imagen del establecimiento. A continuación, queríamos utilizar la imagen del establecimiento para predecir la satisfacción y utilizar a la vez las dos variables para predecir la fidelidad al establecimiento. Por tanto, algunas variables dependientes se convierten en variables independientes en relaciones ulteriores, dando lugar a la naturaleza interdependiente del modelo estructural. Además, muchas de las mismas variables afectan a cada una de las variables dependientes, pero con efectos distintos. El modelo estructural expresa estas relaciones entre variables dependientes e independientes, incluso cuando una variable dependiente se convierte en variable independiente en otras relaciones.

Las relaciones propuestas se trasladan a continuación a series de ecuaciones estructurales (parecidas a las ecuaciones de regresión) para cada variable dependiente. Esta característica sitúa a **SEM** en un lugar aparte de las técnicas discutidas previamente que trataban variables de dependencia múltiples (*análisis multivariante de la varianza y de la correlación canónica*) en la que sólo se permite una *única* relación entre las variables dependientes e independientes.

Sin embargo, es esencial primero revisar algunos conceptos clave asociados con la metodología.

# Variables latentes vs. observadas

La estimación de relaciones múltiples interrelacionadas no es el único elemento de la modelización de ecuaciones estructurales. **SEM** también tiene la habilidad de incorporar *variables latentes* al análisis. Una variable latente es un concepto supuesto y no observado que sólo puede ser aproximado mediante variables medibles u observables. Las variables observadas, que recogemos a través de varios métodos de obtención de datos (es decir, estudios de mercado, experimentos, observación) se conocen como variables manifiestas. Entonces, ¿por qué querríamos utilizar una variable latente no medida en lugar de datos exactos (variables directas, observables o manifiestas) que ofrecen los encuestados? Aunque esto pueda parecer que no tiene sentido o un enfoque de *caja negra* tiene justificaciones tanto prácticas como teóricas, al mejorar la estimación estadística, mejorar la representación de conceptos teóricos y tener en cuenta el error de medida.

En las ciencias económico-administrativas los investigadores a menudo están interesados en estudiar constructos teóricas que no pueden ser observados directamente. Estos fenómenos abstractos se denominan variables latentes o factores. Ejemplos de variables latentes, por ejemplo, en psicología: son el autoconocimiento y la motivación; en educación: capacidad verbal y expectativa de docentes; en economía: capitalismo y clase social. Como las variables latentes no se observan directamente, se deduce que no se pueden medir directamente. Así, el investigador debe definir operacionalizar la variable latente de interés en términos del *comportamiento* que se cree representarlo. Como tal, la variable no observada está vinculada a una que es observable, lo que hace que sea posible que sea medida. La evaluación del *comportamiento*, entonces, constituye la medición directa de una variable observable aunque también la medición indirecta de una variable no observada (es decir, el constructo subyacente). El término *comportamiento* se usa aquí en el sentido más amplio para incluir puntajes en un instrumento de medición particular. Por lo tanto, la observación puede incluir, por ejemplo, respuestas de autoinforme a una escala de *actitud*, puntajes en un prueba de *logro*, puntajes de observación *in vivo* que representan alguna tarea física o actividad, respuestas codificadas a preguntas de la entrevista,etc. Estas puntuaciones medidas (es decir, *mediciones*) se denominan *observadas* o *variables manifiestas*; dentro del contexto de la metodología **SEM**, sirven como indicadores del constructo subyacente que se supone que representan. Dado este necesario proceso de enlace entre *variables observadas y variables latentes no observadas*, ahora debería quedar claro por qué los metodólogos instan a los investigadores a ser *circunspecto* en su selección de medidas de evaluación. Mientras que la elección de los instrumentos *psicométricos* tal selección se vuelve aún más crítico cuando se presume que la *medida observada representa un constructo subyacente.*

# Variables latentes exógenas vs endógenas

Al trabajar con modelos **SEM**, *es útil distinguir entre variables latentes que son exógenas y aquellos que son endógenos. Las variables latentes exógenas* son

sinónimo de variables independientes; ellos *causan* fluctuaciones en los valores de otras variables latentes en el modelo. Cambios en los valores exógenos de las variables no son explicadas por el modelo; más bien, se les considera influenciados por otros factores externo a él, por ejemplo, las variables de clasificación, como el género, la edad y el nivel socioeconómico. Las variables latentes endógenas son sinónimo de variables dependientes y, como tales, son influenciadas por las variables exógenas en el modelo, ya sea directa o indirectamente. La fluctuación en los valores, de las variables endógenas se explican por el modelo porque todas las variables latentes que las influyen se incluyen en la especificación del modelo. En términos de simplicidad y coherencia en todo el resto del capítulo, los términos independientes y dependientes en lugar de exógenos y endógenos Las variables, respectivamente, se usan en los diferentes modelos **SEM** descritos.

# El modelo de análisis factorial

El procedimiento estadístico más antiguo y mejor conocido para investigar las relaciones entre un conjunto de observaciones y las variables latentes son las de *análisis factorial*. Al usar este enfoque para el análisis de datos, el investigador examina el *covariación* entre un conjunto de variables observadas para recopilar información sobre sus *constructos latentes subyacentes* (es decir, *factores*). Existen dos tipos básicos de análisis factorial: *análisis factorial exploratorio* **(EFA)** *y el análisis factorial confirmatorio* **(CFA)**. (Para mayor información ver Byrne 2005a; 2005b**).**

Tanto el **EFA** como **CFA** se emplean para comprender la varianza compartida de las variables medidas que se cree que son atribuibles a un factor o constructo latente. A pesar de esta similitud, sin embargo, **EFA** y **CFA** son análisis conceptuales y estadísticamente distintos.

El objetivo de **EFA** es *identificar factores basados en datos y maximizar la cantidad de varianza explicada*. No se requiere que el investigador tenga ninguna hipótesis específica sobre cuántos factores surgirán y qué elementos o variables comprenderán estos factores. Si estas hipótesis existen, no se incorporan ni afectan los resultados de los análisis estadísticos. Por el contrario, **CFA** *evalúa las hipótesis a priori y en gran medida es impulsado por la teoría*. Los análisis de **CFA** *requieren que el investigador haga una hipótesis, de antemano, del número de factores, si estos factores están correlacionados o no, y qué elementos / medidas se cargan y reflejan qué factores.* Como tal, en contraste con el **EFA**, donde todas las cargas son libres de variar, **CFA** *permite que la restricción explícita de ciertas cargas sea cero.*

La **EFA** está diseñada para la situación en las que las *relaciones entre las variables observadas y latentes son desconocidas o inciertas*. El análisis, por lo tanto, se desarrolla en un *modo exploratorio para determinar cómo y hasta qué punto las variables observadas están vinculadas a sus factores subyacentes*.

La **EFA** en ocasiones se reporta en investigaciones cuando el **CFA** debiera ser una mejor aproxcimación. *Se ha argumentado que* **CFA** *puede ser restrictivo e inapropiado cuando se usa de manera exploratoria*. Sin embargo, la idea de que **CFA** es únicamente un análisis *confirmatorio* a veces puede ser *engañosa*, ya que los modificación de índices utilizados en **CFA** son de naturaleza algo más que de naturaleza exploratoria. *La modificación de índices muestran la mejora en el ajuste*

*del modelo si un coeficiente particular no se restringiera*. Del mismo modo, **EFA** y **CFA** *no tienen que ser análisis mutuamente excluyentes*; Se ha argumentado que **EFA** es un seguimiento razonable de un modelo de **CFA** *pobremente ajustado*.

Típicamente, el investigador desea *identificar un número mínimo de factores que subyacen (o explican) la covariación entre las variables observadas***.** Retomando el caso de estudio mencionado al final del Capítulo 6, luego de la formulación de los ítems del cuestionario diseñados para medir los constructos latentes, el investigador realiza entonces una **EFA** para determinar el grado en que las mediciones de los indicadores (o sea, las variables observadas) se relacionaron con los constructos latentes. En el análisis factorial, estas relaciones son representadas por *cargas factoriales*. El investigador esperaría que los elementos diseñados para medir la percepción de los clientes de una empresa que realiza campañas de mercadotecnia digital, tenga diversas cargas factoriales en el resto de sus factores componentes ($\mathbf{X_1}$.Web Tecnología; $\mathbf{X_2}$.Web precio del servicio; $\mathbf{X_3}$.Web planeación estratégica; $\mathbf{X_4}$.Web imagen; $\mathbf{X_5}$.Web experiencia del usuario; $\mathbf{X_6}$.Web calidad;$\mathbf{X_7}$. Web desempeño) .Este enfoque analítico del factor *se considera exploratorio en el sentido de que el investigador no tiene un conocimiento previo de que los indicadores realmente miden los factores previstos*.

A diferencia de **EFA**, **CFA** *se usa apropiadamente cuando el investigador tiene algún conocimiento de la estructura de la variable latente subyacente*. Basado en el conocimiento de la teoría, la investigación empírica o ambas, el investigador postula las relaciones entre las medidas observadas y los factores subyacentes *a priori* y luego prueba esta estructura hipotética estadísticamente. Por ejemplo, en el caso anterior, el investigador argumentaría basado en las cargas factoriales de los indicadores diseñados para medir la percepción de *los clientes de una empresa que realiza campañas de mercadotecnia digital* que el indicador $\boldsymbol{X_1}$. Web Tecnología tiene más impacto que el resto de los indicadores (si fuera el caso estadístico). En consecuencia, la especificación a priori del modelo **CFA** permitiría que todos los indicadores de $\mathbf{X_1}$**.** Web tecnología sean libres de cargarse en dicho factor, y ser libres así también a tener cargas cero en los factores restantes. El modelo final, debería ser entonces evaluado por medios estadísticos para determinar la adecuación de su b*ondad de ajuste* (*goodness-of-fit*) a los datos de la muestra. (Para discusiones más detalladas de **CFA**, ver, por ejemplo, Bollen, 1989)

En conclusión, el modelo de análisis factorial (**EFA** o **CFA**) se centra únicamente en cómo y en qué medida las variables observadas están vinculadas a sus factores latentes subyacentes. Más específicamente, se enfoca a verificar el grado en que las variables observadas son generadas por las constructos latentes subyacentes y, por lo tanto, en la intensidad de la fuerza de las regresiones desde los factores a las variables observadas (es decir, las cargas factoriales) como Interés primario. Aunque las relaciones entre los factores también son de interés, la estructura de regresión entre ellas no se considera en el modelo de análisis de factorial. Dado que el modelo **CFA** *se enfoca únicamente en el vínculo entre factores y sus variables medidas,* dentro del marco de **SEM**, esto representa lo que se denomina un *modelo de medición o medida*.

# Modelo completo de variable latente

A diferencia del modelo de análisis factorial, el modelo completo de variable latente permite la especificación de estructura por regresión entre las variables latentes. Es decir, el investigador puede hipotetizar el impacto de un constructo latente sobre otro en el modelado de una dirección causal. Este modelo se denomina completo (*full* o *complete*) porque comprende tanto un modelo de medición como un modelo estructural. Así *el modelo de medición representa los enlaces entre las variables latentes y sus mediciones o medidas observadas (es decir, el modelo* ***CFA****) y el modelo estructural representa los enlaces entre las mismas variables latentes.*
Un modelo completo de variable latente que especifica la dirección de la causa desde una única dirección se denomina *recursivo modelo o formativo*; uno que permite efectos recíprocos o de retroalimentación se denomina un *modelo no recursivo*. Sólo aplicaciones de modelos recursivos o formativos se consideran en este capítulo.

# Modelaje estadístico de propósito general y proceso

Los modelos estadísticos proporcionan una manera eficiente y conveniente de describir la estructura latente subyacente de un conjunto de las variables observadas. Expresados ya sea, de forma esquemática o matemática a través de un conjunto de ecuaciones, tales modelos explican cómo las variables observadas y latentes están relacionadas entre sí. Típicamente, los investigadores postulan un modelo estadístico basado en su conocimiento de la teoría relacionada, en investigación empírica en el área de estudio, o en alguna combinación de ambos. Una vez que se especifica el modelo, el investigador prueba su veracidad basado en datos de muestra que comprenden todas las variables observadas en el modelo. La tarea principal en este procedimiento de prueba de modelo, es determinar la *bondad de ajuste* (*goodness-of-fit*) entre el modelo hipotético y los datos de muestra. De esta forma, el investigador impone la estructura del modelo hipotético en los datos de la muestra y luego prueba qué tan bien los datos observados se ajustan a esta estructura restringida. Dado que es altamente improbable que exista un ajuste perfecto entre los datos observados y el modelo hipotético, hay necesariamente un diferencial entre los dos; este diferencial se denomina *residual*. El proceso de ajuste del modelo (*model-fitting process*) puede resumirse de la siguiente manera:

**Datos=Modelo+ Residual**

Donde:
Los *Datos* representan mediciones de puntuación relacionadas tanto con las variables observadas como derivadas de personas que comprenden la muestra.
El *Modelo* representa la estructura hipotética que vincula las variables observadas con las variables latentes y, en algunos modelos, los que vinculan variables latentes particulares entre sí.
El *Residual* representa la discrepancia entre el modelo hipotético y los datos observados. Al resumir el marco estratégico general para probar modelos de ecuaciones estructurales, Jöreskog (1993) distinguió entre 3 escenarios que

denominó *estrictamente confirmatorios (Strictly Confirmatory)*, *modelos alternativos (Alternative Models)* y *generación de modelo (Model-Generating).* Ver Tabla 7.2.

**Tabla 7.2. Escenarios de Jöreskog**

| Escenario | Descripción |
|---|---|
| **Estrictamente Confirmatorio *(Strictly Confirmatory)*** | El investigador postula un modelo único basado en la teoría, recopila los datos apropiados y prueba el ajuste del modelo hipotético a los datos de muestra. De los resultados de esta prueba, el investigador rechaza o no el modelo; no se realizan modificaciones al modelo planteado |
| **Modelos Alternativos *(Alternative Models)*** | El investigador propone varias alternativas (es decir, que compiten) de modelos, todos los cuales están basados en la teoría. Luego del análisis de un conjunto único de datos empíricos, el investigador selecciona un modelo como el más apropiado para representar los datos de muestra. |
| **Generación de Modelo (*Model-Generating*)** | Representa el caso en el que el investigador, teniendo postulado y rechazado un modelo teórico sobre la base de su pobre ajuste a los datos de la muestra, procede de manera exploratoria (más bien que confirmatoria) a modificar y reestimar el modelo. El enfoque principal en esta instancias es localizar la fuente de *desadaptación* en el modelo y determinar uno que describa mejor los datos de la muestra. Jöreskog (1993) señala que aunque la especificación pueda ser teórica o basada en datos, el objetivo final, es encontrar un modelo que sea sustantivamente significativo y estadísticamente apropiado. Aún más, postula que a pesar del hecho de que *se prueba un modelo en cada ronda, el enfoque global está en la generación de modelos, más que las pruebas de modelos* (Jöreskog, 1993, p. 295). Por supuesto, incluso una revisión superficial de la literatura empírica muestra claramente que este escenario es el más común de los tres, por una sencilla razón: los costos asociados con el recopilación de datos, donde es extremadamente raro que un investigador termine su trabajo sobre la base de un modelo hipotético rechazado. En consecuencia, el escenario estrictamente confirmatorio (*Strictly Confirmatory* no es practicado en la realidad. Aunque el enfoque de los modelos alternativos (*Alternative Models*) tiene escasos exponentes. |

Fuente: Jöreskog (1993) con adaptación propia

Aunque se han utilizado numerosos algoritmos para estimar los modelos **CFA**, **ML** (*maximum likelihood*) sigue siendo el procedimiento de estimación básico. Sin embargo, desafortunadamente los modelos **CFA** se aplican frecuentemente a condiciones de los datos, que se apartan de los requisitos teóricos normales para una estimación de **ML** válida. Por ejemplo, los especialistas en ciencias económico-administrativas y sociales a menudo estiman los modelos **CFA** con datos *no-normalizados* e indicadores escalados utilizando categorías ordenadas discretas. En

consecuencia, se han desarrollado algoritmos alternativos que atienden a las diversas condiciones de los datos que enfrentan los investigadores. La estimación alternativa se ha caracterizado en dos tipos generales: (1) *robusto* (*robust*) y (2) de información limitada (*limited information estimator*)

Cuando se implementa **ML** (*maximum likelihood*) con datos que se desvían de las suposiciones de la teoría normal, los modelos **CFA** pueden producir estimaciones parciales de parámetros y conclusiones falsas. La estimación *robusta* normalmente intenta corregir el problema ajustando el modelo teórico normal (**χ2**) y los errores estándar. Por ejemplo, Satorra y Bentler (1994) recomiendan usar la estimación de **ML** (*maximum likelihood*) de la manera habitual y, posteriormente, dividir el modelo (**χ2**) por una medida del grado de curtosis multivariante. Una ventaja adicional de los *estimaciones robustas*, es su disponibilidad en la mayoría del software **SEM** más conocido.

Desafortunadamente, las *estimaciones robustas*, pueden volverse insostenibles bajo condiciones de datos comunes. En particular, cuando los indicadores se escalan utilizando pocas categorías de respuesta (por ejemplo, en desacuerdo, neutral, de acuerdo), las *estimaciones robustas* de **ML** (*maximun likelihood*) tienden a funcionar mal.

Es probable que las *estimaciones de información limitados*, como los mínimos cuadrados ponderados (**WLS.** *Weighted Least Squares*), sean una mejor opción cuando los indicadores manifiestos adquieren una forma ordinal. En términos este tipo de estimación atiende a los indicadores ordinales mediante el uso de *correlaciones policóricas* para ajustarse a los modelos **CFA**. Las *correlaciones policóricas* capturan la covarianza entre dos variables latentes cuando sólo se observa su forma categorizada, lo que se logra en gran parte a través de la estimación de los parámetros del umbral (*estimation of threshold parameters*).

# Notación simbólica

Los modelos de ecuaciones estructurales se retratan esquemáticamente usando configuraciones particulares con 4 símbolos geométricos. Ver **Tabla 7.3**

**Tabla 7.3. Notación utilizada en la representación de ecuaciones estructurales**

| Símbolo | Representan: |
|---|---|
| **Círculo o elipse** | Factores latentes, subyacentes no observados |
| **Cuadrados o rectángulos** | Variables observables |
| **Flechas en un solo sentido** | El impacto de una variable sobre otra |
| **Dobles flechas en dos sentidos** | Covarianzas o correlaciones entre pares de variables |

Fuente: Jöreskog (1993); Byrne 2006, con adaptación propia.

En la construcción de un modelo de una estructura particular en estudio, los investigadores usan estos símbolos dentro de un marco de configuraciones básicas , cada una de las cuales, representa un componente importante en el proceso analítico. Ver **Tabla 7.4**

**Tabla 7.4. Combinación de símbolos en configuración**

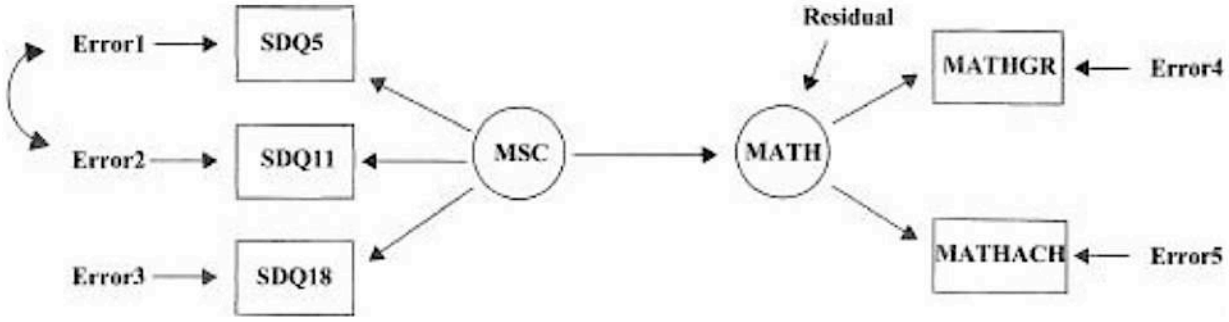
| Configuración | Significa |
|---|---|
| | Coeficiente de trayectoria de regresión de una variable observada vs una variable latente o no observada (o factor) |
| | Coeficiente de trayectoria de regresión de un factor vs otro factor |
| | Medición de error asociado con una variable observada |
| | Error residual en la predicción de un factor no observable |

Fuente: Jöreskog (1993); Byrne (2006) con adaptación propia,

# Diagrama de trayectoria

Las representaciones esquemáticas de los modelos se denominan diagramas de trayectoria porque proporcionan una *descripción visual de las relaciones asumidas entre las variables en estudio*. Esencialmente, como se verá más adelante, un diagrama de trayectoria que representa un modelo de **SEM** en particular es en realidad el *equivalente gráfico de su representación matemática* por lo que un conjunto de ecuaciones relaciona las *variables dependientes con sus variables explicativas***.** A fin de ilustrar cómo las cuatro configuraciones de símbolos anteriores pueden representar un *proceso causal particular,* mostramos la **Figura 7.1**.

**Figura 7.1. Modelo a discusión**




Fuente: Byrne (2006)

Al revisar este modelo, se observa que hay dos factores latentes no observados: el *concepto de matemátcas* (***MSC.****Math Self-Concept*) y el *logro en matemáticas* (***MATH.****Math Achievment.*) con 5 variables observables, donde 3 se consideran para medir el **MSC** (**$SDQ_5$; $SDQ_{11}$; $SDQ_{18}$**) y 2 para medir **MATH** (**MATHGR; MATHACH).** Estos 5 variables observables funcionan como indicadores de sus respectivos factores latentes subyacentes.

Asociado con cada variable observada está un término de error (**$Error_1$-$Error_5$**) y, con el factor de predicción (**MATH**), un término residual (**Residual**). Hay una distinción importante entre estos dos términos de error. El error asociado con las variables observadas representa el *error de medición* el cual se refleja en su adecuación con la medición de los factores subyacentes relacionados (**MSC; MATH**). El error de medición se deriva de dos fuentes: el *error de medición aleatorio* (*random measurment error,* en el sentido psicométrico) y el *error de singularidad* (*error uniqueness*), un término utilizado para describir el error de varianza que surge de alguna característica considerada específica (o única) para una particular indicador variante. Tal error a menudo representa un error de medición no aleatorio (*nonrandom measurement error*).

Los términos residuales representan el error en la predicción de factores dependientes a partir de factores independientes. Por ejemplo, el término residual que se muestra en la **Figura 7.1** representa el error en la predicción de **MATH** (el factor dependiente) de **MSC** (el factor independiente) En esencia, tanto los términos de error de medición como los de error residual representan variables no observadas. Por lo tanto, parece perfectamente razonable que, en consonancia con la representación de los factores, pudieran modelarse como elipses (o círculos). De hecho, este es el enfoque de modelado implementado en al menos una **SEM** programa: **AMOS** (Arbuckle, 2003). Además de los símbolos que representan las variables, otros se utilizan en los diagramas de trayectoria para indicar procesos hipotéticos que involucran a todo el sistema de variables. En particular, las flechas unidireccionales representan coeficientes de regresión estructural, lo que indica el impacto de una variable sobre otra.

En la **Figura 7.1**,por ejemplo, la flecha unidireccional que apunta hacia el factor dependiente, **MATH***,* implica que el factor independiente **MSC** *provoca* el logro matemático (**MATH**). Del mismo modo, las tres flechas unidireccionales desde **MSC** a cada una de las tres variables observadas (**SDQ5, SDQ11, SDQ18**) y las que van desde **MATH** a cada uno de sus indicadores (**MATHGR y MATHACH**) sugieren que estos valores de puntaje son cada uno influenciado por sus respectivos factores subyacentes. Como tal, estos coeficientes de trayectoria representan la magnitud del cambio esperado en las variables observadas para cada cambio en la variable latente relacionada (o factor). Las flechas unidireccionales que apuntan desde los 5 términos de error indican el impacto del error de medición (aleatorio y único) en las variables observadas y, a partir del residuo, el impacto del error en la predicción de **MATH**. Finalmente, como se señaló anteriormente, las flechas curvas de dos vías representan covarianzas o correlaciones entre pares de variables. Así, las flechas bidireccionales enlazan *Error1* y *Error2* , como se muestra en la **Figura 7.1**, implica el error de medida asociado con **SDQ5** en correlación con lo asociado a **SDQ11**.

# La ecuación estructural (SEM) y su historia

El **SEM** es una poderosa técnica estadística multivariable cada vez más utilizada en las ciencias económico-administrativas y en las ciencias sociales en general. Es el resultado de la evolución y conjunción de varias aproximaciones metodológicas desarrolladas por diversas disciplinas: el análisis de variables latentes o no observables de la psicología, los modelos de ecuaciones simultáneas de la economía

o el análisis de trayectorias (*path analysis*) de la sociología, entre otras.
El **SEM** se usa típicamente para realizar análisis de factores confirmatorios**. LISREL, EQS, AMOS, Mplus** y **R** son los programas de software populares. **CFA** también se utiliza con frecuencia como un primer paso para evaluar el modelo de medición propuesto en un modelo de ecuaciones estructurales. *Muchas de las reglas de interpretación con respecto a la evaluación del ajuste del modelo y la modificación del modelo en el modelado de ecuaciones estructurales se aplican por igual al* **CFA**. *Se distingue del modelado de ecuaciones estructurales por el hecho de que en* **CFA***, No hay flechas dirigidas entre factores latentes.* En otras palabras, mientras que en **CFA** no se presume que los factores se causen directamente entre sí, **SEM** a menudo especifica factores y variables particulares que son de naturaleza causal. *En el contexto de* **SEM***, el* **CFA** *a menudo se denomina como el modelo de medición, mientras que las relaciones entre las variables latentes (con flechas dirigidas) se nombra como el modelo estructural.*
El **SEM** abarca una familia entera de modelos conocidos con muchos nombres, entre ellos, el análisis de la estructura de la covarianza, el análisis de variables latentes, el análisis de factor confirmatorio y a menudo simplemente análisis **LISREL** (por cierto, nombre de uno de los programas de *software* más populares). Ver **Cuadro 1.**

**Cuadro 1. Breve Historia SEM**

| Año | Historia |
|---|---|
| **1934** | El primer antecedente de un modelo de ecuaciones estructurales se remonta a 1934, año en el que el biómetra Sewall Wright da a conocer el modelo de trayectoria (*path analysis*) sobre las relaciones de tamaño en mediciones óseas. Esta técnica permitía descomponer la varianza y covarianza de las variables involucradas, en función de los parámetros de un sistema de ecuaciones simultáneas y tenía como fin estudiar el efecto directo e indirecto entre estas variables. |
| **1940-1950** | Anderson y Rubin (1949, 1950) desarrollaron el estimador de máxima verosimilitud de información limitada (*Limited Information Maximum Likelihood Estimator*) para los parámetros de una ecuación estructural única, que incluía indirectamente el estimador de mínimos cuadrados en dos etapas y su distribución asintótica (*Two-stage Least Squares Estimator and its Asymptotic Distribution*), discutidos más tarde por: (Anderson, 2005) y Farebrother (1999). |
| **1951-1960** | Los mínimos cuadrados en dos etapas (*Two-stage Least Squares Estimator*) se propusieron originalmente como un método para estimar los parámetros de una ecuación estructural única en un sistema de ecuaciones simultáneas lineales, siendo introducido por Theil (1953a, 1953b, 1961) y más o menos independientemente por Basmann (1957) y Sargan (1958). La estimación de máxima verosimilitud de información limitada (*Limited Information Maximum Likelihood Estimation*) de Anderson, se implementó finalmente en un algoritmo de búsqueda de computadora, donde compitió con otros algoritmos SEM iterativos. De éstos, los mínimos cuadrados en dos etapas fueron con mucho el método más utilizado en los años sesenta y principios de los setenta. |

| | |
|---|---|
| | El modelado de ecuaciones estructurales, como el término se usa actualmente en sociología, psicología y otras ciencias sociales, evolucionó a partir de los métodos anteriores en el modelado de trayectorias genéticas (*Genetic Path Modeling*) de Sewall Wright.<br>Tanto **LISREL** como **PLS-PA** se concibieron como algoritmos informáticos iterativos (*Iterative Computer Algoritms*), con énfasis desde el principio en la creación de una interfaz gráfica y de entrada de datos accesible y la extensión del análisis de ruta de Wright (1921). El trabajo de la Comisión Cowles (*Cowles Comissio*n) sobre estimación de ecuaciones simultáneas (*simultaneous equations estimation*) se centró en los algoritmos de Koopman y Hood (1953) de la economía del transporte (*Economics of Transportation*) y el enrutamiento óptimo (*Optimal Routing*), con estimación de máxima verosimilitud (**ML.** *Maximum Likelihood*) y cálculos algebraicos de forma cerrada, ya que las técnicas de búsqueda de soluciones iterativas (*Iterative Solution Search Techniques*) eran limitadas. |
| **1961-1970** | No obstante su importancia, no es sino hasta 1960 y principios de los 70 que Blalock, Boundon y Duncan reconocen el potencial del análisis de trayectoria y las técnicas relacionadas de *correlación parcial*, como herramientas para analizar datos no experimentales.<br>Sus formas modernas surgieron con implementaciones intensivas en computación en los años 60 y 70. El **SEM** evolucionó en tres corrientes diferentes: (1) métodos de regresión de sistemas de ecuaciones desarrollados (*Systems of Equation Regression Methods*), desarrollados principalmente por la *Comisión Cowles* (*Cowles Comission*); (2) algoritmos iterativos de máxima verosimilitud para el análisis de trayectorias (*Iterative Maximum Likelihood Algorithms for Path Analysis*) desarrollados principalmente por *Karl Gustav Jöreskog* en la E*ducational Testing Service* y posteriormente en la Universidad de Uppsala; y (3) algoritmos iterativos de ajuste de correlación canónica para el análisis de trayectorias(*Iterative Canonical Correlation Fit Algorithms for Path Analysis*) también desarrollados en la Universidad de Uppsala por *Hermann Wold*. Gran parte de este desarrollo se produjo en un momento en que la computación automatizada ofrecía mejoras sustanciales sobre la calculadora existente y los métodos de computación analógica disponibles, productos ellos mismos de la proliferación de innovaciones en equipos de oficina a fines del siglo XX. |
| **1971-1980** | Un paso decisivo ocurre cuando Jöreskog (1993), Keesling (1972) y Wiley (1973) desarrollan un modelo general de ecuaciones estructurales e incorporan diagramas de trayectoria y otras características del análisis de trayectoria, conocido como modelo **LISREL** o modelo **JKW**.<br>La terminología aún es suelta y confusa, situación que se ha utilizado para oscurecer las debilidades en los métodos. En particular, el **PLS-PA** (el algoritmo de *Lohmoller*) se ha combinado con regresión de |

| | |
|---|---|
| | mínimos cuadrados parciales **PLSR** (*Partial Least Squares Regression*), que es un sustituto de la regresión de mínimos cuadrados ordinarios (*Ordinary Least Squares Regression*) y no tiene nada que ver con el análisis de ruta o trayectoria. **PLS-PA** ha sido promocionado *falsamente* como un método que funciona con pequeños conjuntos de datos cuando fallan otros enfoques de estimación. Westland (2010) demostró decisivamente que esto no era cierto y desarrolló un algoritmo para tamaños de muestra en **SEM**. Desde la década de 1970, se sabe que la afirmación de *tamaño de muestra pequeño es falsa* (Dhrymes, 1972, 1974; Dhrymes y Erlat, 1972; Dhrymes y otros, 1972; Gupta, 1969; Sobel, 1982). |
| **1981- a la actualidad** | Los sistemas de aproximación de ecuación de regresión (*Systems of Regression Equation Approaches*) se desarrollaron en la Comisión Cowles desde la década de 1950 en adelante, extendiendo el modelo de transporte (*Transportation Modeling*) de Tjalling Koopmans. Sewall Wright y otros estadísticos intentaron promover métodos de análisis de ruta en Cowles (entonces en la Universidad de Chicago). Los estadísticos de la Universidad de Chicago identificaron muchas fallas con las aplicaciones de análisis de ruta (*Path Analysis*) a las ciencias sociales; fallas que no plantearon problemas significativos para identificar la transmisión de genes (*Gene Transmission)* en el contexto de Wright, pero que hicieron que los métodos de ruta como **PLS-PA** y **LISREL** fueran problemáticos en las ciencias sociales. Freedman (1987) resumió estas objeciones en el análisis de trayectorias: *la falta de distinción entre suposiciones causales, implicaciones estadísticas y reclamos de políticas ha sido una de las principales razones de la sospecha y confusión en torno a los métodos cuantitativos en las ciencias sociales* (ver también Wold, 1987 su respuesta). El análisis de ruta de Wright nunca obtuvo un gran número de seguidores entre los econometristas de EE. UU., Pero tuvo éxito al influir en Hermann Wold y su alumno Karl Jöreskog. El alumno de Jöreskog, Claes Fornell, promovió **LISREL** en los Estados Unidos.<br>Además de facilitar la difusión del análisis de trayectoria, presenta las ecuaciones que se derivan de las covarianzas entre las variables, a través de operaciones matriciales en lugar de que se *lean* del diagrama de trayectoria, y proporciona una descomposición más clara de los efectos directos, indirectos y totales. El desarrollo de modelos con esta combinación de variables (latentes y medidas) se ha incrementado de forma espectacular. Jöreskog (1993) extendió el análisis factorial exploratorio al factorial confirmatorio, desarrolló el modelo factorial de segundo orden, el análisis factorial multigrupo y el ya citado modelo general de ecuación estructural (**LISREL**).<br>Los avances en las computadoras facilitaron a los principiantes, el acceso y la aplicación de métodos de ecuaciones estructurales en *el análisis intensivo de grandes conjuntos de datos en problemas complejos no estructurados.* Las técnicas de solución más populares |

se dividen en tres clases de algoritmos: (1) algoritmos ordinarios de mínimos cuadrados aplicados independientemente a cada ruta (*Ordinary Least Squares Algorithms Applied Independently to each Path*), como los aplicados en los llamados paquetes de análisis de ruta **PLS** que se estiman con **OLS**; (2) Algoritmos de análisis de covarianza (*Covariance Analysis Algorithms*) que evolucionan a partir del trabajo de Wold y su alumno Karl Jöreskog implementados en **LISREL, AMOS y EQS**; y (3) algoritmos de regresión de ecuaciones simultáneas (*Simultaneous Equations Regression Algorithms*) desarrolladas en la Comisión Cowles por Tjalling Koopmans.

Pearl ha extendido **SEM** de *modelos lineales a no paramétricos*, y ha propuesto interpretaciones causales y contra-actuales de las ecuaciones. Por ejemplo, al excluir una *variable Z* de los argumentos de una ecuación afirma que *la variable dependiente es independiente de las intervenciones en la variable excluída*, una vez que mantenemos constantes los argumentos restantes. Los **SEM** *no paramétrico* permiten la estimación de los *efectos totales, directos e indirectos sin comprometerse con la forma de las ecuaciones ni con la distribución de los términos de error.* Esto amplía el análisis de mediación a sistemas que involucran variables categóricas en presencia de interacciones no lineales. Bollen y Pearl examinan la historia de la interpretación causal de **SEM** *y por qué se ha convertido en una fuente de confusiones y controversias*.

Los métodos de análisis de ruta o trayectoria **SEM,** son populares en las ciencias económico-administrativas y sociales en general, debido a su accesibilidad; Los programas informáticos empaquetados permiten a los investigadores obtener resultados sin la incomodidad de comprender el diseño y control experimental, el efecto y el tamaño de la muestra, y muchos otros factores que forman parte de un buen diseño de investigación. Los partidarios dicen que esto refleja una *interpretación holística, y menos causal*, de muchos fenómenos del mundo real, especialmente en psicología e interacción social, que los que pueden adoptarse en las ciencias naturales*; los detractores sugieren que se han llegado a muchas conclusiones erróneas debido a esta falta de control experimental*.

La dirección en los modelos de redes dirigidas de **SEM** surge de supuestos de *causa-efecto* hechos sobre la realidad. Las interacciones sociales y los artefactos a menudo son *epifenómenos (*fenómenos secundarios que son difíciles de vincular directamente a los factores causales). Un ejemplo de un *epifenómeno* fisiológico es, por ejemplo, el tiempo para completar un *sprint* de 100 metros. Una persona puede mejorar su velocidad de carrera de 12 segundos a 11 segundos, *pero será difícil atribuir esa mejora a cualquier factor causal directo, como la dieta, la actitud, el clima, etc. La mejora de 1 segundo en el tiempo de carrera es una epifenómeno* (*el producto holístico de la interacción de muchos factores individuales*).

| | |
|---|---|
| | La influencia de Jöreskog no sólo se limita a los desarrollos propios. Varios de sus estudiantes de doctorado han realizado importantes contribuciones. Por ejemplo, Sörbom (1974) extiende el modelo multigrupo para incluir medias en las variables latentes; Muthén (1997) introduce métodos para incluir variables observadas categóricas; Hägglund (1985) contribuye con el método de mínimos cuadrados por medio de estimación de dos estados (*two-state last-square methods*); Quiroga (1992) realiza estudios de robustez con *correlaciones policóricas* para desviaciones del supuesto de normalidad, mientras que Yang-Wallentin (1997) desarrolla métodos para estimar relaciones no lineales. Los avances recientes en modelos de ecuaciones estructurales comprenden extensiones para estimaciones en datos que provienen de muestras complejas, modelos lineales generalizados y series de tiempo.<br>Por otro lado Bentler (1980,1990,2006) realiza diversas aportaciones sobre métodos de ajuste creando **EQS,** como software para ecuaciones estructurales. |

Fuente: recopilación y adaptación propia

En varias disciplinas relacionadas con las ciencias sociales es usual intentar medir la inteligencia, motivación, eficiencia, percepción, habilidad verbal, etcétera (fenómenos de una gran complejidad), a partir de percepciones, opiniones, indicadores y variables relativas o aproximadas. Este tipo de variables recibe un nombre genérico: *variables latentes*. Las *variables manifiestas* se pueden medir de manera directa y representan características observables de algún fenómeno subyacente, al contrario de las latentes. Una característica importante de estas variables es que sirven para evidenciar o definir a las variables no observadas o latentes. La forma de modelar un fenómeno que requiere representar relaciones entre variables latentes y variables medidas o manifiestas es a través de los modelos de ecuaciones estructurales.

El modelo de ecuaciones estructurales es una herramienta estadística que busca explicar la relación existente entre diversas variables. Además, examina la relación estructural expresada en una serie de ecuaciones, similar a las series de regresión múltiple. La relación existente entre los constructos (variables dependientes e independientes) se analizan simultáneamente a través de un modelo de ecuaciones estructurales. Los constructos pueden ser factores no observables o variables latentes, quienes comúnmente están representados por múltiples variables o ítems (variables manifiestas). El modelo de ecuaciones estructurales es la única técnica estadística que combina tanto las técnicas de dependencia como las técnicas de interdependencia: *Análisis Factorial y Análisis de Regresión Lineal*. (Ver más en Aldás-Manzano y Maldonado-Guzmán, 2016).

Como se señaló en otros capítulos las *variables exógenas* son aquellas variables latentes equivalentes a las *variables independientes*, las cuales generalmente son medidas a través de una varianza y representan el constructo central de la investigación. Por lo tanto, en un *diagrama de trayectoria de un* **SEM** las variables exógenas son fáciles de reconocer porque de ellas salen las flechas que indicar una relación existente con otro tipo de variables (*endógenas*), y no reciben ninguna flecha, por eso se llaman variables independientes.

Los constructos o *variables endógenas* son aquellas variables latentes multi-ítems equivalentes a las *variables dependientes*, las cuales están teóricamente determinadas por factores o dimensiones.
Este tipo de variables son dependientes de otros constructos y la relación de dependencia que tienen es fácil de visualizarse en un *path model*, porque son los constructos que reciben la flecha de la relación
En **CFA,** se usan varias pruebas estadísticas para determinar qué tan bien el modelo se ajusta a los datos. Tenga en cuenta que un buen ajuste entre el modelo y los datos *no significa que el modelo sea correcto*, *o incluso que explique una gran proporción de la covarianza*. *Un buen ajuste del modelo solo indica que el modelo es plausible*. Al informar los resultados de un análisis factorial confirmatorio, se recomienda informar:
**a.**Los modelos propuestos,
**b.**Las modificaciones realizadas,
**c.**Las medidas que identifican cada variable latente,
**d.**Las correlaciones entre las variables latentes,
e.Cualquier otra información pertinente , como si se usan restricciones.
Con respecto a la selección de las estadísticas de ajuste del modelo para reportar, uno no debe simplemente informar las estadísticas que estiman la mejor opción, aunque esto puede ser tentador. Aunque existen varias opiniones diferentes, que veremos más adelante, por ejemplo, se recomienda aplicar la prueba *Chi-Cuadrado, el error cuadrático medio de aproximación* (**RMSEA**), *el índice de ajuste comparativo* (**CFI***) y la raíz cuadrátrica media residual estandarizada* (**SRMR**)*, etc.* (Ver **Cuadro 7.2**)

**Cuadro 7.2. Descripción de los índices y/o pruebas de ajuste más comunes que aparecen en la SEM**

| Índices y/o Pruebas | Descripción |
|---|---|
| Índice de ajuste absoluto (*Absolute fit indices*) | Los índices de ajuste absoluto determinan qué tan bien encaja el modelo a priori, o reproduce los datos, siendo los más comunes: la prueba **Chi- Cuadrado, RMSEA, GFI, AGFI, RMR y SRMR** |
| Prueba de Chi-Cuadrada (Chi-squared test) | La prueba de *Chi-Cuadrada* indica la diferencia entre las matrices de covarianza observadas y esperadas. Los valores más cercanos a cero indican un mejor ajuste; la diferencia más pequeña entre las matrices de covarianza esperadas y observadas. Las estadísticas de *Chi-Cuadrado* también se pueden usar para comparar directamente el ajuste de modelos anidados a los datos. *Sin embargo, una dificultad con la prueba* ***de Chi cuadrado*** *del ajuste del modelo es que los investigadores pueden no rechazar un modelo inapropiado en tamaños de muestra pequeños y rechazar un modelo apropiado en tamaños de muestra grandes.* Como resultado, se han desarrollado otras medidas de ajuste. |
| Raíz cuadrática media de error por | El error cuadrático medio de aproximación (**RMSEA**) evita problemas de tamaño de muestra al analizar la discrepancia |

| | |
|---|---|
| aproximación. (RMSEA.*Root Mean Square Error*) | entre el modelo hipotético, con estimaciones de parámetros elegidas óptimamente y la matriz de covarianza de la población. El **RMSEA** varía de 0 a 1, con valores más pequeños que indican un mejor ajuste del modelo. Un valor de **0.05** a **.08** o menos es indicativo de ajuste aceptable del modelo. |
| Raíz cuadrática media residual. (RMR.*The Root Mean Square Residual*) | La raíz cuadrática media residual (**RMR**) y la raíz cuadrática media residual estandarizado (**SRMR**) son la raíz cuadrada de la discrepancia entre la matriz de covarianza de la muestra y la matriz de covarianza del modelo. Sin embargo, la **RMR** puede ser algo difícil de interpretar, ya que su rango se basa en las escalas de los indicadores en el modelo (esto se vuelve complicado cuando se tienen múltiples indicadores con escalas variables, por ejemplo, dos cuestionarios, uno en una escala de 0-10) y el otro, está diseñado con respuestas en una escala de 1-3). El residuo cuadrático medio de raíz estandarizado elimina esta dificultad en la interpretación, y varía de 0 a 1, con un valor de **.08** o menos que es indicativo de un modelo aceptable. |
| Índice de Bondad de Ajuste e Índice de Bondad Ajustado (*Goodness of Fit Index and Adjusted Goodness of Fit Index*) | El índice de bondad de ajuste (**GFI**) es una medida de ajuste entre el modelo hipotético y la matriz de covarianza observada. El índice de bondad de ajuste *ajustado* (**AGFI**) corrige el que se vea afectado por el número de indicadores de cada variable latente. El rango de **GFI** y **AGFI** está entre 0 y 1, con un valor de más de **.9** (**nunca 1**) que generalmente indica un ajuste aceptable del modelo |
| Índices de Ajuste Relativo (*Relative fit indices*) | Los índices de ajuste relativo (también llamados *índices de ajuste incremental* e *índices de ajuste comparativo*) comparan el Chi-Cuadrado para el modelo hipotético con uno de un modelo *nulo* o *de referencia*. Este modelo nulo casi siempre contiene un modelo en el que todas las variables no están correlacionadas, y como resultado, tiene un Chi-Cuadrado muy grande (lo que indica mal ajuste). os índices de ajuste relativo incluyen *el índice de ajuste normalizado* (**NFI.***Normed Fit Index*) y el *índice de ajuste comparativo* (**CFI.** *Comparative Fit Index*) |
| Índice de Ajuste Normalizado y No Normalizado (NFI.*Normed fit index and* NNFI.*Non-normed fit index*) | El índice de ajuste normalizado (**NFI**) analiza la discrepancia entre el valor de *Chi-Cuadrado* del modelo hipotético y el valor de Chi-Cuadrado del modelo nulo. Sin embargo, **NFI** tiende a ser negativamente sesgado. El índice de ajuste no normalizado (**NNFI**, también conocido como el *índice Tucker-Lewis*, ya que fue construido sobre un índice formado por Tucker y Lewis, en 1973 resuelve algunos de los problemas de sesgo negativo, aunque los valores **NNFI** pueden a veces caen más allá del rango de 0 a 1.Los valores tanto para **NFI** como para **NNFI** deben oscilar entre 0 y 1, con un valor de |

| | |
|---|---|
| | corte de **.95** o mayor que indica un buen ajuste del modelo |
| Índice de Ajuste Comparativo (CFI.*Comparative Fit Index*) | El índice de ajuste comparativo (**CFI**) analiza el ajuste del modelo al examinar la discrepancia entre los datos y el modelo hipotético, mientras se ajusta a los problemas de tamaño de muestra inherentes en la prueba de Chi-Cuadrado del ajuste del modelo y el índice de ajuste normativo. Los valores de **CFI** varían de 0 a 1, con valores más grandes que indican un mejor ajuste. Anteriormente, se consideraba que un valor de **CF**I de .90 o mayor indicaba un ajuste aceptable del modelo. Sin embargo, estudios recientes han indicado que se **necesita un valor superior a 0.90** para garantizar que los modelos mal especificados no se consideren aceptables (Hu y Bentler, 1999). Por lo tanto, actualmente se acepta un valor de **CFI de .95 o superior** como indicador de buen ajuste (Hu y Bentler, 1999). |

Fuente: recopilación propia

*Para estimar los parámetros de un modelo, el modelo debe identificarse adecuadamente*. Es decir, el número de parámetros estimados (desconocidos) (**q**) debe ser menor o igual que el número de varianzas únicas y covarianzas entre las variables medidas; p (p + 1) / 2. Esta ecuación se conoce como la *regla t.* Si hay muy poca información disponible sobre la cual basar las estimaciones de los parámetros, entonces se dice que el modelo está i*nfraidentificado*, y los parámetros del modelo no se pueden estimar de manera apropiada.

Como se señaló al comienzo de este capítulo, además de prestarse a ser una presentación esquemática de los procesos causales bajo estudio, los modelos de ecuaciones estructurales también pueden ser representados por una serie de ecuaciones de regresión (*es decir, estructurales*) ya que:

**a.** Las ecuaciones de regresión representan la influencia de una o más variables sobre otra, y

**b.** Esta influencia, convencionalmente conocida como **SEM**, es simbolizado por una flecha de una sola cabeza que apunta desde la variable de influencia a la variable de interés, por lo que se puede pensar en cada ecuación como un resumen del impacto de todas las variables relevantes en el modelo (observado y no observado) en una variable específica (observada o no).

Por lo tanto, un enfoque relativamente simple para la formulación de estas ecuaciones, consiste en observar cada variable que tiene una o más flechas apuntando hacia ella y luego registrar la suma de todas esas influencias para cada una de estas variables dependientes. Para ilustrar esta traducción de los procesos de regresión en ecuaciones estructurales, volvamos nuevamente a la **Figura 7.1.** Podemos ver que hay seis variables con flechas apuntando hacia ellas; cinco representan las variables observadas (**SDQ5**, **SDQ11, SDQ18; MATHGR, MATHACH**) y uno representa una *variable no observada* (o factor; **MATH**). Por lo tanto, sabemos que las funciones de regresión simbolizadas en el modelo que se muestra en la **Figura 7.1** pueden ser resumido en términos de *6 representaciones separadas de tipo ecuación de dependencias lineales,* de la siguiente manera:

**MATH = MSC + Residual**
**SDQ5 = MSC + Error1**
**SDQ11 = MSC + Error2**
**SDQ18 = MSC + Error3**
**MATHGR = MATH + Error4**
**MATHACH = MATH + Error5**

La teoría estadística define un *coeficiente de regresión* estar compuesto en realidad de dos elementos: el *verdadero* o coeficiente estructural entre la variable dependiente e independiente y la fiabilidad de la variable predictor. La fiabilidad es el grado en el que la variable independiente está *libre de error*.

En todas las técnicas multivariantes expuestas hasta ahora, hemos asumido que no existía error en nuestras variables. Pero sabemos tanto desde la perspectiva práctica como teórica que no podemos medir perfectamente un concepto y que siempre hay algún grado de *error de medida***.** Por ejemplo, cuando preguntamos acerca de algo tan inmediato como la renta de las economías domésticas, sabemos que algunas personas contestarán incorrectamente, tanto por exceso como por defecto o no lo sabrán con precisión. Las respuestas ofrecidas tienen cierto error de medida y por tanto afectan al verdadero coeficiente estructural.

El impacto del error de medida (y la correspondiente disminución de la fiabilidad) puede observarse de una expresión del coeficiente de regresión como lo mostrado en la **ecuación 7.2:**

**Ecuación 7.2. Coeficiente de regresión**

$$\beta_{V.X} = \beta_s \times \rho_x$$

Donde:
**$\beta_{V.X}$,** es el coeficiente de regresión observado,
**βs,** es el verdadero» coeficiente estructural y
**ρx,** es la fiabilidad de la variable predictor. A menos que la fiabilidad sea del 100 por cien, la correlación observada siempre infravalora la relación *correcta*. Dado que todas las relaciones de dependencia se basan en la correlación observada (y resultante del coeficiente de regresión) entre las variables, esperaremos fortalecen las correlaciones utilizando un modelo de dependencia parahacer estimaciones más precisas de los coeficientes estructurales teniendo en cuenta en primer lugar, la correlación atribuible a cualquiera de los problemas de medida.

El *error de medición* no sólo está provocado por respuestas inadecuadas sino que se produce también cuando utilizamos conceptos teóricos o más abstractos, tales como la actitud hacia un producto o las motivaciones del comportamiento. Con conceptos tales como estos, el investigador intenta diseñar las mejores cuestiones para medir el concepto. Los encuestados también pueden no estar seguros de cómo responder o de cómo interpretar las cuestiones de forma diferente a la que pretende el investigador. Ambas situaciones pueden dar lugar al error de medida. Pero si sabemos la magnitud del problema, podemos incorporar la fiabilidad en la estimación estadística y mejorar nuestro modelo de dependencia.

*¿Cómo tenemos en cuenta el error de medida?* **SEM** proporciona el *modelo de medida***,** que especifica las reglas de correspondencia entre las variables latentes y

manifiestas. El modelo de medida permite al investigador utilizar una o más variables para un único concepto dependiente o independiente y a continuación estimar (o especificar) la fiabilidad. Por ejemplo, la variable de- pendiente puede ser un concepto representado por un conjunto de cuestiones. En el modelo de medida el investigador puede evaluar la contribución de cada ítem de la escala, así como incorporar el grado en que la escala mide el concepto (su fiabilidad) en la estimación de las relaciones entre las variables dependientes e independientes. Este procedimiento es similar al desarrollo por el análisis factorial de los ítems de la escala y utilizar las puntuaciones del factor en la regresión.

# Componentes no visibles de un modelo

Aunque en principio, *SI hay una correspondencia uno a uno entre la presentación esquemática de un modelo y su traducción a un conjunto de ecuaciones estructurales, ninguna de estas representaciones nos informa de la historia completa del modelo*. Algunos parámetros críticos para la estimación del modelo no se muestran explícitamente y por lo tanto pueden no ser obvios para el modelador de ecuaciones estructurales principiante. Por ejemplo, tanto en el diagrama de trayectoria como en las ecuaciones anteriores, *No hay indicación de que las varianzas de las variables independientes sean parámetros en el modelo*; de hecho, tales parámetros son esenciales para todos los modelos de ecuaciones estructurales. Del mismo modo, es igualmente importante notar la inexistencia especificada de ciertos parámetros en el modelo. Por ejemplo, en la **Figura 7.1**, *No hay una flecha curva entre **Error4** y **Error5**, lo que sugiere la falta de covarianza entre los términos de error asociados con las variables observadas **MATHGR** y **MATHACH**. Del mismo modo, No existe una covarianza hipotética entre MSC y el residual.*

La ausencia de esta trayectoria provoca la suposición común de que la *variable predictora (o independiente) no está de ninguna manera asociada con algún error que surja de la predicción de la variable de criterio (o dependiente).*

# Modelo SEM. Composición básica

El modelado de ecuaciones estructurales (**SEM**) incluye un conjunto diverso de modelos matemáticos, algoritmos de computadora y métodos estadísticos que se ajustan a redes de constructos a datos. **SEM** incluye Análisis Factorial Confirmatorio (**CFA.** *Confirmatory Factor Analysis*); Análisis de Ruta (**PA.** *Path Analysis*); Modelado de Ruta de Mínimos Cuadrados Parciales (**PLS**. *Partial Least Squares Path Modeling*) y Modelado de Crecimiento Latente (*Latent Growth Modeling*). El concepto no debe confundirse con el concepto relacionado a modelos estructurales en econometría (*structural models in econometrics*), ni con los modelos estructurales en economía (*structural models in economy*).

Los modelos de ecuaciones estructurales a menudo se usan para evaluar constructos *latente* no observables. *A menudo invocan un modelo de medición que define variables latentes usando una o más variables observadas, y un modelo estructural que imputa relaciones entre las variables latentes.* Los enlaces entre los constructos de un modelo de ecuaciones estructurales pueden estimarse ya sea con

ecuaciones de regresión independientes o mediante enfoques más involucrados, como los empleados en **LISREL**.

El uso de **SEM** se justifica comúnmente en las ciencias económico-administrativas y las sociales, debido a su capacidad para *imputar relaciones entre construcciones no observadas (variables latentes) a partir de variables observables*.

Para proporcionar un ejemplo simple, el concepto de *inteligencia humana* no se puede medir directamente ya que *uno puede medir la altura o el peso.* En cambio, los psicólogos desarrollan una hipótesis de inteligencia y escriben instrumentos de medición con elementos (preguntas) diseñados para medir la inteligencia de acuerdo con su hipótesis. Entonces usarían **SEM** para probar su hipótesis usando datos recopilados de personas que tomaron su prueba de inteligencia. Con **SEM**, *inteligencia sería la variable latente y los elementos de prueba serían las variables observadas.*

La crítica de los métodos **SEM** a menudo aborda las *trampas en la formulación matemática, la débil validez externa de algunos modelos aceptados y el sesgo filosófico inherente a los procedimientos estándar*

A lo largo de la exposición del **SEM**, nos referiremos a la necesidad de una justificación teórica para la especificación de las relaciones de dependencia, modificaciones a las relaciones propuestas y muchos otros aspectos de la estimación de un modelo. La *teoría* ofrece la causa de casi todos los aspectos del **SEM.** La *teoría* puede definirse como *un conjunto sistemático de relaciones que ofrecen una explicación exhaustiva y consistente de un fenómeno*.

A partir de esta definición, podemos ver que *teoría no es el dominio exclusivo de los académicos sino que puede estar basada en la experiencia y en la práctica obtenida por la observación del com portamiento del mundo real***.** La *teoría* es a menudo un objetivo prioritario de la investigación académica, pero la investigación práctica puede desarrollar o proponer un conjunto de relaciones que son tan complejas e interrelacionadas como cualquier teoría de base académica. Por tanto, tanto los investigadores académicos como los de la empresa pueden beneficiarse de unas herramientas analíticas extraordinarias proporcionadas por **SEM.** Desde una perspectiva práctica, un enfoque teórico de **SEM** es una necesidad porque la técnica debe ser casi completamente especificada por el investigador. Mientras que con otras técnicas multivariantes el investigador puede haber sido capaz de especificar un modelo básico y permitir que valores por defecto de los programas estadísticos completen los resultados de estimación restantes, **SEM** no tiene ninguna de estas características, por lo que debe definirse explícitamente cada componente de los modelos y medidas estructurales. Además, cualesquiera modificaciones del modelo deben hacerse a través de acciones específicas dictadas por el investigador. La necesidad de un *modelo teórico* para guiar el proceso de estimación se hace especialmente crítica cuando se hacen modificaciones del modelo.

Dada la flexibilidad de **SEM**, *las oportunidades de sobreajustar el modelo o desarrollar un modelo con escasa generalización son muy elevadas*. Por tanto, cuando resaltamos la necesidad de la justificación teórica, nuestro objetivo es que el investigador reconozca que **SEM** es un método confirmatorio, guiado más por la teoría que por los resultados empíricos.

Uno de los conceptos más importantes que un investigador en ciencias económico-administrativas debe aprender en relación con las técnicas multivariantes es que no existe un único modo *correcto* de aplicarlas. En lugar de eso, el investigador debe formular los objetivos de la investigación y aplicar la técnica de la forma más apropiada para conseguir los objetivos deseados. En algunos casos, las relaciones están especificadas estrictamente y el objetivo es una confirmación de la relación. Otras veces, las relaciones están vagamente reconocidas y el objetivo es el descubrimiento de las relaciones. En cada caso límite y en los casos intermedios, el investigador debe formular el uso de la técnica de acuerdo con los objetivos de investigación. La aplicación de **SEM** sigue este mismo principio. Su flexibilidad ofrece al investigador una poderosa herramienta de análisis apropiada para muchos objetivos de investigación. Pero el investigador, debe definir estos objetivos como líneas maestras de la estrategia de modelización. El uso del término *estrategia* se refiere a un plan de acción orientado a la consecución de un objetivo específico. En el caso de **SEM**, el resultado último es siempre la evaluación de una serie de relaciones. Sin embargo, esto se puede conseguir de muchas maneras. Para nuestros propósitos, definimos 3 estrategias distintas en la aplicación de **SEM**: modelización confirmatoria, modelos rivales y desarrollo del modelo, como se aprecia en el **Cuadro 7.3.**

**Cuadro 7.3. Estrategias de aplicación SEM**

| Estrategia | Descripción |
|---|---|
| **Modelización Confirmatoria** | La aplicación más directa de los modelos de ecuaciones estructurales es una *estrategia de modelización confirmatoria***,** donde el investigador especifica un modelo aislado y **SEM** se utiliza para evaluar su significación estadística. Aquí el investigador está diciendo, *tanto si funciona como si no*. Aunque pueda parecer que es la aplicación más rigurosa, en realidad no es el test más rigu- roso del modelo propuesto. La investigación ha mostrado incluso que las técnicas desarrolladas para evaluar los modelos de ecuaciones estructurales tienen un *sesgo confirmatorio,* que tiende a confirmar que el modelo se ajusta a los datos. Por tanto, si el modelo propuesto tiene un ajuste aceptable por cualquiera de los criterios aplicados, el investigador no ha *probado* el modelo propuesto, sino que sólo ha confirmado que es uno de los varios modelos posibles aceptables. Varios modelos pueden tener ajustes igualmente aceptables. Por tanto, el test más riguroso se consigue comparando modelos rivales |
| **Modelos Rivales** | Obtener un nivel de ajuste aceptable tanto para el modelo conjunto como para los modelos de medida o estructural no asegura al investigador que se haya encontrado el *mejor* modelo. Varios modelos alternativos pueden ofrecer iguales o incluso mejores ajustes. Como medio de evaluar el modelo estimado con modelos alternativos, las comparaciones de modelos conjuntos pueden lle- varse a cabo en una *estrategia de modelos rivales***.** El test más fuerte de un modelo propuesto es identificar y |

contrastar los modelos rivales que representan las verdaderas y distintas relaciones hipotéticas estructurales. Cuando se comparan estos modelos, el investigador se acerca mucho al con- traste de teorías alternativas, que es un contraste mucho más fuerte que una ligera modificación de una única teoría.

¿Cómo genera el investigador este conjunto de modelos alternativos? Un posible origen de modelo rivales son las formulaciones alternativas de la teoría subyacente. Por ejemplo, en una formulación la confianza puede preceder a la lealtad mientras que en la otra la lealtad precede a la confianza. Esto podría constituir la base de dos modelos rivales. Los modelos equivalentes proporcionan una segunda perspectiva sobre el desarrollo de un conjunto de modelos rivales. Se ha mostrado que para cualquier modelo de ecuaciones estructurales, existe al menos otro modelo con el mismo número de parámetros y el mismo nivel de ajuste del modelo que varía en las relaciones representadas. Esto implica que ningún modelo es único en el nivel de ajuste conseguido, y que para cualquier modelo con un ajuste aceptable existe un número indetenninado de modelos alter- nativos con el mismo nivel de ajuste. Se han definido una serie de nonnas para identificar mode- los equivalentes para cualquier modelo estructural . Como nonna general, cuanto más complejo sea el modelo, más modelos equivalentes existen. Una tercera aproximación, el programa **TETRAD**, es un método empírico que está tomando auge y que examina sistemáticamente un modelo estructural e identifica relaciones adicionales que están apoyadas por los datos. Con la matriz de datos de entrada y la especificación del modelo hecha por el investigador, el programa exami- na las pautas de las relaciones (tetrads) y aisla aquellas relaciones que pudieran estar empíricamente fundamentadas. El programa **TETRAD** no estima los parámetros, sino que identifica las relacio- nes que se van a incluir en el modelo original para formar los modelos rivales. El programa funciona mejor empezando con un modelo simple y añadiendo relaciones. Se suele alegar que este modelo es ateórico y muy mecanicista o de caja negra, pero ofrece a los investigadores intuiciones para sus modelos que no se podrían haber obtenido de otra forma. Un ejemplo habitual de la estrategia de modelos rivales es el proceso de evaluación factorial de invarianza, la igualdad de modelos de factor entre grupos. Existe un procedimiento establecido para evaluar el grado de invarianza, comenzando por los modelos más vagamente restringidos y a partir de esos modelos añadir paulatinamente restricciones adicionales hasta que se contraste el modelo más restrictivo. Las restricciones se añaden para representar invarianzas a lo largo de grupos, ponderaciones e incluso intercorrelaciones de

| | |
|---|---|
| | factores. Esto también es un ejemplo de un enfoque de modelos anidados, en el que el número de constructos e indicadores permanece cons-tante, pero el número de relaciones estimadas cambia. Aunque los modelos rivales son normalmente modelos anidados, también pueden no estar anidados (difieren en el número de constructos o indicadores), con lo que se exigen medidas especializadas de ajuste para comparar entre los modelos. |
| **Desarrollo del Modelo** | La estrategia de desarrollo del modelo difiere de las dos anteriores estrategias en que aunque se propone un modelo, el propósito del esfuerzo de modelización es mejorarlo a través de modifica- ciones de los modelos de medida y/o estructurales. En muchas aplicaciones, la teoría sólo puede ofrecer un punto de partida para el desarrollo de un modelo con justificación teórica que pueda ser apoyado empíricamente. Por tanto, el investigador ha de emplear **SEM** no sólo para contrastar el modelo empíricamente sino también para obtener perspectivas acerca de su reespecificación. Debe tomarse, sin embargo, alguna precaución. El investigador tiene que ser cuidadoso no em pleando esta estrategia en la medida en que el modelo final tenga un ajuste aceptable pero que no pueda ser generalizada a otras muestras o poblaciones. Además, la reespecificación del modelo debe hacerse siempre con apoyo teórico en lugar de justificación empírica. |

Fuente: Jöreskog (1993); Hair (et al.,1999); Byrne 2006 con adaptación propia.

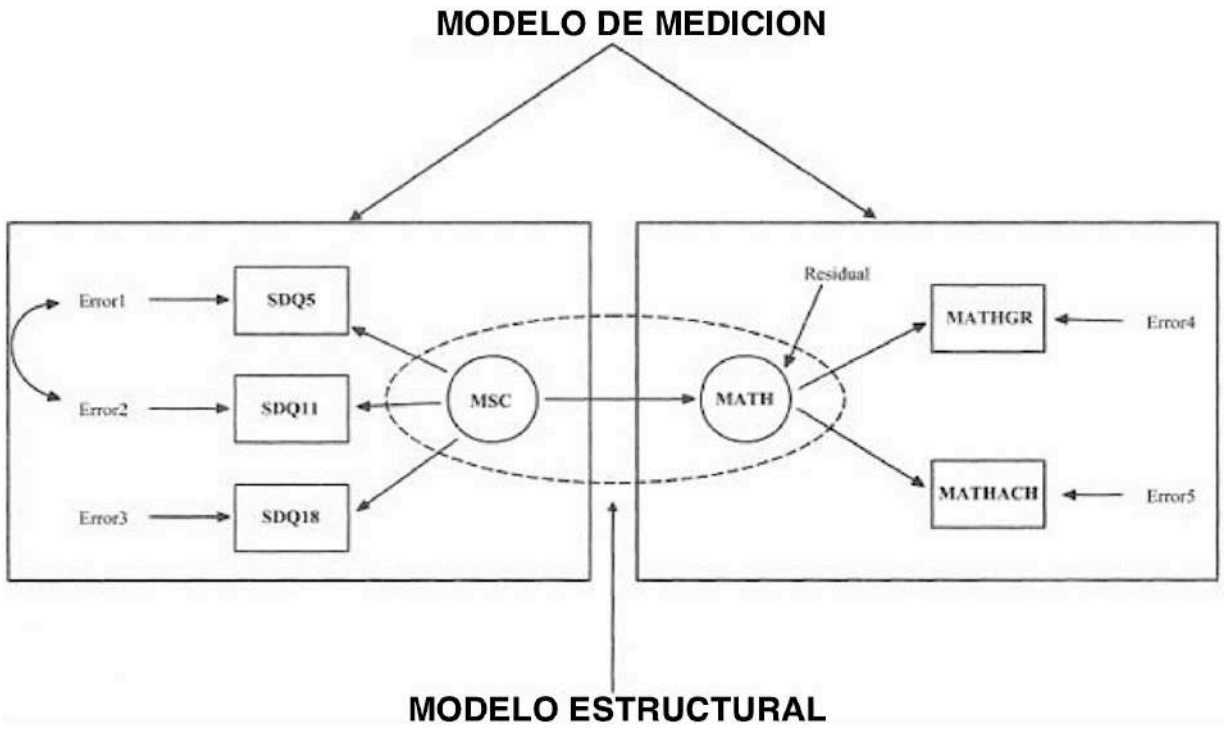
El modelo **SEM** general se puede descomponer en *dos submodelos*:

**a**.*El modelo de medición ,que define las relaciones entre las variables observadas y las no observadas.* En otra palabras, proporciona el vínculo entre las puntuaciones en un instrumento de medición (es decir, las variables indicadoras observadas) y los constructos subyacentes que están diseñados para medir (es decir, las variables latentes no observadas). El modelo de medición, entonces, representa el modelo **CFA** descrito anteriormente en que especifica el patrón por el que cada medida (o medición) carga a un factor particular.

**b.***El modelo estructural***,** por el contrario, define las relaciones entre las *variables* no observadas. En consecuencia, especifica la manera particular en que las variables latentes, directa o indirectamente influyen (es decir, *causan*) cambios en los valores de ciertas otras variables latentes en el modelo.

Con fines didácticos para aclarar este importante aspecto de la composición **SEM**, examinemos ahora la **Figura 7.2**,

**Figura 7.2. Ecuación general estuctural demarcada en sus componentes: medición y estructural**




Fuente: Jöreskog (1993); Hair (et al.,1999); Byrne 2006 con adaptación propia.

el cual, es el mismo modelo presentado en la **Figura 1.1** pero que ha sido demarcado en medición y componentes estructural. Considerados por separado, Los elementos modelados dentro de cada rectángulo en la **Figura 1.2** representan dos modelos **CFA**. Los dos factores dentro de la elipse representa un modelo de variable latente completa y por lo tanto no sería de interés en la investigación **CFA**. El modelo **CFA** a la izquierda del diagrama representa un modelo de un solo factor (**MSC**) medido por tres variables observadas (**SDQ5 - SDQ18**), mientras que el modelo **CFA** a la derecha representa un modelo de un factor (**MATH**) medido por dos variables observables (**MATHGR-MATHACH**). En ambos casos, la regresión de las variables observadas en cada factor y las varianzas tanto del factor como los errores de medición son de interés primario; la covarianza de error sería de interés solo en los análisis del modelo **CFA** relacionado con **MSC**. El verdadero valor del **SEM** proviene de los beneficios de utilizar los modelos estructurales y de medida simultáneamente, jugando cada uno de ellos distintos papeles en el análisis conjunto. Para asegurar que ambos modelos están correctamente especificados y que los resultados son válidos, discutiremos ahora un proceso de modelización de ecuaciones estructurales son:

1. *Desarrollar un modelo fundamentado teóricamente,*
2. *Construir un diagrama de secuencias de relaciones causales,*
3. *Convertir el diagrama de secuencias en un conjunto de modelos y relaciones estructurales,*
4. *Elegir el tipo de matriz de entrada y estimar el modelo propuesto,*
5. *Evaluar la identificación del modelo estructural,*
6. *Evaluar los criterios de calidad del ajuste, y*

7.*Interpretar y modificar el modelo si está teóricamente justificado* (Hair et al., 1999). Ver **Cuadro 7.4.**

**Cuadro 7.4. Etapas sugeridas para modelar CFA y SEM**

| Paso 1<br>Desarrollo de un modelo basado en la teoría |
|---|
| La modelización de ecuaciones estructurales se basa en relaciones causales**,** en las que el cambio en una variable se supone que produce un cambio en otra variable. Encontramos este tipo de enunciado cuando definimos una relación de dependencia, tal y como se encuentra en el análisis de regresión. Las relaciones causales pueden tomar muchas formas y significados, desde la estricta causación encontrada en los procesos fisicos, tales como una reacción química, a las relaciones menos definidas encontradas en la investigación del comportamiento, tales como las «causas» del éxito educativo o las razones por las cuales compramos un producto y no otro. La fuerza y convicción con que el investigador puede asumir la **causal** entre dos variables no descansa en los métodos analíticos escogidos sino en la justificación teórica ofrecida por los análisis. Los «requisitos» para hacer valer la causación tienen profundas raíces en varios enfoques de la filosofia de la ciencia.Hay un acuerdo general con al menos cuatro criterios establecidos para realizar afirmaciones causales: (1) asociaciones suficientes entre dos variables, (2) antecedentes temporales de la causa frente al efecto, (3) falta de alternativas a las variables causales, y ( 4) una base teórica para la relación. Aunque en muchos casos todos los criterios establecidos para la realización de afirmaciones causales no se cumplen estrictamente, probablemente puedan realizarse afinnaciones causales si las relaciones están basadas en una teoría racional. Pero avisamos a cualquier investigador contra la su- posición de que las técnicas discutidas en este capítulo ofrezcan por sí mismas un medio de probar la causal sin tener una perspectiva teórica como guía. Utilizar estas técnicas de forma exploratoria es inútil y equivoca al investigador más que ofrecerle resultados apropiados.El error más crítico en el desarrollo de modelos de fundamentación teórica es la omisión de una o más variables predictivas claves, un problema conocido como error de especificación**.** La consecuencia de omitir una variable significativa consiste en sesgar la evaluación de la importancia de otras variables. Por ejemplo, si suponemos que dos variables (**a y b**) fueran predictores de **c** e incluyéramos tanto a como ben nuestro análisis, haríamos la evaluación correcta de su im tancia relativa como se muestra por sus coeficientes estimados. Pero si dejamos a la variable **b fuera** de nuestro análisis, el coeficiente para **a** sería diferente. Esta diferencia, o sesgo, es el resultado del coeficiente para la variable ague refleja no sólo su efecto sobre *e* sino el efecto que comparte también con **b**. Este efecto compartido, sin embargo, está controlado cuando ambas variables están incluidas en el análisis. (Ver más sobr emanejo de errores de datos en Mejía-Trejo , 2017d).El deseo de incluir todas las variables debe compensarse con las limitaciones prácticas de **SEM**.Aunque no existe un límite teórico sobre el número de variables del modelo, los problemas prác- ticos aparecen incluso antes de llegar a los límites de la mayoría de los problemas informáticos. Muy a menudo, las interpretaciones de los resultados, |

particularmente la significación estadística, se hace bastante difícil a medida que el número de conceptos aumenta (más de 20 conceptos). El investigador nunca debería omitir un concepto solamente porque el número de variables se haga grande, pero también debería de reconocer los beneficios de los modelos teóricos concisos y parsimoniosos.

## Paso 2
## Construcción de un diagrama de trayectorias de relaciones causales

Hasta el momento, hemos expresado las relaciones causales sólo en términos de ecuaciones. Pero existe otro método de representar estas relaciones denominadas diagramas de trayectorias, que son especialmente útiles en el diseño de una serie de relaciones causales. Un diagrama de trayectorias es más que una simple representación visual de las relaciones porque permite al investigador presentar no sólo las relaciones predictivas entre constructos (es decir, las relaciones variable dependiente-independiente), sino también relaciones asociativas (correlaciones) entre los constructos e incluso entre los indicadores.

### Elementos de un diagrama de trayectorias

Recuerde los 2 elementos básicos utilizados en su construcción. Primero, es el concepto de un constructo, que es un concepto teórico que actúa como una piedra angular utilizada para definir relaciones. Un constructo puede representar un concepto tan simple como edad, renta o género o tan complejo como estatus socioeconómico, conocimiento, preferencia o actitud. El investigador en económico-administrativas define diagramas de trayectoria en términos de constructos y a continuación encuentra variables para medir cada constructo. Por ejemplo, podemos preguntar la edad a alguien y utilizar esto como medida del constructo edad. De la misma manera, podemos preguntar una serie de cuestiones acerca de las opiniones de una persona y utilizar esto como una medida de actitud. Ambas series de cuestiones ofrecen valores numéricos para los constructos. Podemos evaluar las cuestiones para la cantidad de error medido que poseen e incluir esto en el proceso de estimación. A partir de este momento, utilizamos el término constructo para representar un concepto particular, no importa cómo se mida. Un constructo se representa normalmente en un diagrama de secuencias mediante un óvalo. El segundo elemento básico es la flecha, utilizada para representar relaciones específicas entre constructos. Una flecha directa indica una relación causal directa de un constructo a otro. Una flecha curvada (o una línea sin cabezas de flechas) entre constructos indica simplemente una co-+rrelación entre constructos. Finalmente, una flecha directa con dos cabezas indica una relación recíproca o no recursiva entre constructos. Sólo con estos tres tipos de flechas, se pueden representar todas las relaciones de los modelos estructurales. Todos los constructos de un diagrama de trayectorias pueden clasificarse en 2 clases: exógenos y endógenos. Los constructos exógenos, también conocidos como variables origen o variables independientes, no están causados o son explicados por ninguna de las variables del modelo; esto es, no existen flechas apuntando a estos constructos. Los constructos endógenos se predicen mediante uno o más constructos pueden predecir otros constructos endógenos (aquí es donde vemos las interrelaciones

que apuntan a la necesidad de los modelos estructurales), pero un constructo exógeno puede estar causalmente relacionado sólo con los constructos endógenos. Por tanto, la distinción entre exógenos y endógenos los determina sólo el investigador, al igual que se hizo al decidir sobre cuáles son variables dependientes e independientes en la regresión.

Dos hipótesis subyacen en los diagramas de trayectorias. En primer lugar, todas las relaciones causales están indicadas. Por tanto, la teoría es la base para la inclusión o la omisión de cualquier relación. Es tan importante justificar por qué no existe una relación causal entre dos constructos como justificar la existencia de otra relación. Pero es importante recordar que el objetivo es modelizar las relaciones entre constructos con el número más reducido de secuencias causales o correlaciones entre constructos que pueden estar teóricamente justificados. El segundo supuesto se refiere la naturaleza de las relaciones causales que se supone que son lineales. Al igual que se encontró en otras técnicas multivariantes, las relaciones no lineales no pueden ser estimadas directamente en la modelización de ecuaciones estructurales, pero modelos estructurales modificados pueden aproximar relaciones no lineales

## Paso 3
## Conversión de un diagrama de secuencias en un conjunto de ecuaciones estructurales y especificación del modelo de medida

Después de desarrollar el modelo teórico y de representarlo en un diagrama de trayectorias, el investigador está preparado para especificar el modelo en términos más formales. Esto se hace a través de una serie de ecuaciones que definen: (1) las ecuaciones estructurales que vinculan los constructos, (2) el modelo de medida que especifica qué variables miden qué constructos y (3) una serie de matrices que indican cualquier correlación supuesta entre constructos o variables. El objetivo es vincular definiciones operacionales de los constructos con la teoría para llegar al contraste empírico apropiado. En esta sección se discute cada tipo de ecuación o matriz. En las discusiones, todas las ecuaciones se especifican en terminología general.

### Modelo estructural

Trasladar un diagrama de trayectorias a una serie de ecuaciones estructurales es un procedimiento directo. En primer lugar, cada constructo endógeno (cualquier constructo con una o más flechas directas hacia él) es la variable independiente en una ecuación separada. Entonces las variables predictor son todos constructos en los principios o colas, de las flechas que apuntan a la variable endógena. Es así de simple.

### Modelo de medida

Hasta ahora nos hemos referido al término de medida en términos generales, pero ahora debemos definirlo en términos específicos. Discutiremos no sólo los procedimientos básicos de especificación de un modelo de medida, sino también los supuestos referentes al número de indicadores por constructo y el proceso de especificación de la fiabilidad del constructo en lugar de en su estimación. Pero antes de hacerlo, revisemos los fundamentos del análisis factorial, que son muy parecidos a los del modelo de medida.

### Correspondencia con el análisis factorial

En el análisis factorial, cada variable individual se *explicaba* por su ponderación en cada factor. El objetivo era representar lo mejor posible todas las variables en un número reducido de factores, es decir, los factores referidos dimensiones subyacentes de los datos, que después tendremos que interpretar y clasificar. El análisis factorial (ver capítulo 12 Mejía-Trejo 2017c), a menudo se clasifica como una técnica exploratoria porque no existen restricciones sobre las cargas de las variables. Cada variable tiene una carga sobre cada factor. El valor de cada factor (puntuación del factor) se calcula mediante las cargas sobre cada variable. También el valor predictor para cada variable se calcula mediante las cargas de la variable para cada factor. Sin embargo, cada variable tiene una carga factorial; por tanto, cada factor es siempre una composición de todas las variables, aunque sus cargas varíen en mag-nitud. Por consiguiente, un factor es en realidad un constructo latente, definido por las cargas de todas las variables.

## Especificación del modelo de medida

Para especificar el modelo de medida, hacemos la transición desde el análisis factorial, en el que el investigador no tiene el control sobre qué variables describen cada factor, a un modo confirmatorio, en el que el investigador especifica qué variables definen cada constructo (factor). Las variables observadas que obtenemos de los encuestados se denominan indicadores en el modelo de medida, porque los utilizamos para medir o indicar, los constructos latentes (factores).

¿Cómo y por qué difiere esta configuración de las cargas del análisis factorial examinadas previamente? La diferencia más evidente es el más reducido número de ponderaciones. En el mode- lo explicativo del análisis factorial, el investigador no puede controlar las ponderaciones. En el modelo de medida sin embargo, el investigador tiene un control completo sobre las variables descritas por cada constructo. En el ejemplo, cada variable era un indicador de un solo constructo; por tanto, existe un número más reducido de ponderaciones. Aunque una variable puede ser indicador de más de un constructo, este método no es recomendable excepto en situaciones específicas con fuerte carga teórica. El investigador especifica un modelo de medida tanto para los constructos exógenos como para los constructos endógenos exactamente de esta forma.

## Determinación del número de indicadores

Ya se ha discutido la lógica y la justificación del uso de indicadores múltiples (variables) para representar un constructo, pero queda una cuestión fundamental: ¿cuántos indicadores deberían utilizarse por constructo? El número mínimo de indicadores por constructo es uno, pero el uso de sólo un único indicador exige al investigador dar estimaciones de fiabilidad. Un constructo puede ser representado por 2 indicadores, pero 3 es el número mínimo preferido de indicadores, porque utilizando sólo dos indicadores aumentan las oportunidades de alcanzar una solución no factible]. Además de los fundamentos teóricos que deberían utilizarse para seleccionar las variables como indicadores de un constructo, no existe límite superior en términos de número de indicadores. A efectos prácticos, sin embargo, la mayoría de los constructos deberían ser representados por un número de entre 5 a 7 y indicadores. La excepción notable es el uso de escalas preexistentes, que pueden contener muchos ítems, cada uno actuando como indicador del constructo. En estos

casos, el investigador debería evaluar la unidimensionalidad del constructo y la posibilidad de múltiples subdimensiones que puedan ser representadas en un modelo factorial de segundo orden .

**Justificar la fiabilidad del constructo**

Una vez que el modelo de medida ha sido especificado, el investigador debe probar después la fiabilidad de los indicadores. En este paso, el investigador debe determinar el método básico en que se establece la fiabilidad de cada constructo. Existen dos métodos principales para establecer la fiabilidad: (1) estimación empírica o (2) especificación por el investigador.

**Fiabilidades estimadas empíricamente**

La estimación empírica de la fiabilidad sólo es posible si el constructo tiene dos o más indicadores. Para un constructo con un sólo indicador, el inVestigador debe especificar la fiabilidad. Para la estimación empírica, el investigador especifica la matriz de ponderaciones tal y como se ha descrito, junto con un término de error para cada variable indicador). Cuando los modelos de medida y estructural están estimados, los coeficientes de ponderación ofrecen estimaciones de las fiabilidades de los indicadores y del constructo conjunto. En este enfoque, la intervención del investigador no tiene impacto sobre el valor de la fiabilidad utilizado en la estimación del modelo excepto en los conjuntos de indicadores incluídos.

**Especificar las confiabilidades**

En algunos casos resulta apropiado especificar o fijar las confabilidades. La especificación de las confiabilidades para el indicador(es) de cualquier constructo latente puede parecer que va en contra de los objetivos de la modelización de ecuaciones estructurales; sin embargo, en al menos 3 situaciones está justificado y es altamente recomendable. En un caso, la estimación empírica de la confiabilidad no es posible, aunque el investigador puede saber que el error de medida todavía existe. En otros, los indicadores pueden haber sido utilizados previamente; por tanto, se conocen las confiabilidades antes de utilizarse. Y finalmente, tenemos un enfoque en 2 pasos en el que las confiabilidades se evalúan en primer lugar y a continuación se especifican en el proceso de estimación. Este enfoque de dos pasos separa explícitamente los 2 procesos empíricos y ofrece perspectivas de los dos por separado.

**Medidas de ítem unico**

Con medidas de ítem único, no es posible estimar empíricamente la confiabilidad; por tanto. el investigador se enfrenta con dos posibilidades. La primera, establecer (fijar) la confiabilidad a 1**,0**, indicando que no existe error de medida en el indicador. Pero como ya hemos discutido, sabemos que esto es erróneo en casi todos los casos, si por ninguna otra razón la confabilidad está afectada por la calidad de la recogida de datos. Por ejemplo. el género puede ser perfecto o muy cercano (99 por ciento), con errores debidos a errores de codificación. Sin embargo, los ingresos pueden tener un nivel elevado de error (es decir, un 10 por ciento) debido a sesgos de las contestaciones y al nivel de medida. Muy a menudo por tanto, el investigador debería realizar alguna estimación de la confiabilidad y especificar el valor de los indicadores de ítem único.

| Uso de medidas o escalas validadas |
|---|
| Muchas veces el investigador emplea una escala o medida que ha sido extensamente contrastada en la investigación previa. Si al usarla el objetivo es obtener una réplica de los efectos encontrado en estudios anteriores, entonces la confiabilidad de la medida o escala debería fijarse en niveles previamente especificados. Este es un ejemplo de las confiabilidades especificadas por el investigador para mantener el control sobre el significado de los constructos. Al fijar la confiabilidad, el investigador fuerza a un indicador para tener la cantidad de varianza apropiada para el constructo y mantener un significado específico para el mismo. |
| **Análisis de dos pasos** |
| Muchos investigadores proponen un proceso de dos pasos de modelización de ecuaciones estructurales en el que el modelo de medida se fija en el segundo paso cuando el modelo estructural está estimado. La razón de este enfoque es que la representación precisa de la fiabilidad de los indicadores se realiza mejor en dos pasos para evitar la interacción de los modelos de medida y estructurales. Aunque no podemos evaluar verdaderamente los modelos de medida y estructural aisladamente, debemos considerar el potencial para den- tro del constructo frente a los efectos entre constructos en estimación, que pueden ser sustanciales y resultar en lo que se denomina *confusión interpretacional.* Un análisis de dos pasos con estimación simultánea tanto de modelos de medida como estructurales es la mejor aproximación cuando el modelo posee tanto fuerte lógica teórica como medidas altamente confiables, resultando en relaciones más precisas y disminuyendo la posibilidad de interacción. Sin embargo, cuando nos enfrentamos con medidas que son menos fiables, o una teoría que es sólo tentativa, el investigador debería considerar un enfoque por pasos para maximizar la interpretabilidad tanto de las medidas como de los modelos estructurales. Se ha producido un debate considerable sobre la conveniencia de este enfoque y sobre aquellos casos en los que está justificado, tanto conceptual como empíricamente. |
| **Métodos para especificar la confiabilidad** |
| Para fijar la confiabilidad de un indicador en una matriz de correlación. el investigador especifica el valor de ponderación como la raíz cuadrada de la fiabilidad deseada o estimada, o especifica el término de error de esa variable como 1.0 menos el valor de fiabilidad deseado. Si se utiliza la matriz de covarianza, entonces el término de error o valor de ponderación se multiplica por el valor de la varianza de la variable. Al especificar las confiabilidades, el investigador puede determinar los valores de carga, el término de error o ambos. Dado que especificar bien la ponderación o bien los términos de error determina automáticamente el otro valor. recomendamos que ambos se fijen para el modelo de mayor parsimonia y que un coeficiente no se utilice para estimar un valor que podría ser especificado. Este procedimiento puede ser realizado simplemente en programas estadísticos a través de un tratamiento único o para cada variable. Una variante de la fijación de la fiabilidad de la escala completa es un enfoque |

propuesto por Hayduk [54, 55] que implica un proceso por pasos. En primer lugar, el investigador selecciona el indicador aislado que se cree es la mejor representación del constructo. Para este indicador, la fia- bilidad (ponderación y término de error) es fija. A continuación, pueden añadirse indicadores adi- cionales y sus ponderaciones y términos de error estimados, permitiendo el cálculo de la fiabilidad del constructo completo. La lógica que hay detrás de este enfoque es especificar el significado de- seado del constructo a través de este indicador y a continuación permitir a otros indicadores añadir significados al concepto ya definido. El inconveniente es que cuando todas las ponderaciones y los términos de error se estiman empíricamente de forma simultánea, el investigador no está se- guro exactamente de qué constructo representa excepto algún concepto subyacente común al resto de los indicadores.

### Correlaciones entre constructos e indicadores

Además de los modelos de medida y estructurales, el investigador también especifica cualquier correlación entre los constructos exógenos o entre los constructos endógenos. Muchas veces los constructos exógenos están correlacionados, representando una influencia compartida sobre las variables endógenas. Las correlaciones entre los constructos endógenos, sin embargo, tienen pocas aplicaciones apropiadas y no están recomendadas para un uso normal debido a que representan correlaciones entre las ecuaciones estructurales que pueden confundir sus interpretaciones. Finalmente, los indicadores en el modelo de medida también pueden estar correlacionadas separadamente de las correlaciones de los constructos. Este método se suele evitar excepto en situaciones específicas, tales como un estudio en que existen efectos conocidos del proceso de recogida de datos o de medida sobre dos o más indicadores, o un estudio longitudinal en el que el mismo indicador se recoge en dos períodos de tiempo

## Paso 4
## Selección del tipo de matriz de entrada y estimación del modelo propuesto

Como se muestra en los pasos precedentes, se exige mucho más al investigador de ciencias económico-administrativas que utiliza **SEM** en términos de especificación del modelo a estimar que con otra técnica multivariante, con la posible excepción del **análisis conjunto**. Ahora el investigador debe llevar el proceso efectivo de estimación del modelo especificado, incluyendo los supuestos de entrada de datos en la forma apropiada y selección del procedimiento de estimación. Las decisiones tomadas en estas áreas tienen un impacto profundo en los resultados conseguidos.

### Introducción de datos

**SEM** difiere de otras técnicas multivariantes en que utiliza sólo la matriz de varianza-covarianza o de correlación como sus datos de entrada. Las observaciones individuales pueden ser introducidas en los programas, pero se convierten en uno de estos dos tipos de matrices antes de la estimación. El interés en SEM no está en las observaciones individuales sino el patrón de relaciones entre los encuestados. La entrada del programa es una matriz de correlación o varianza-covarianza de todos los indicadores utilizados en el modelo. El modelo de medida especifica entonces qué indicadores

corresponden a cada constructo, y las puntuaciones del constructo latente son las empleadas en el modelo estructural.

## Supuestos

**SEM** comparte 3 supuestos con los otros métodos multivariantes que hemos estudiado: observaciones independientes, muestra aleatoria de los encuestados y la linealidad de todas las relaciones. Además, **SEM** es más sensible a las características distribucionales de los datos, particularmente el incumplimiento de la normalidad multivariante (crítica en el uso de **LISREL**) o una fuerte curtosis (apuntamiento) de los datos. Algunos programas informáticos, como EQS, son menos sensibles a los datos no normales, pero los datos deberían ser evaluados sin importar qué programa se está utilizando. Los mínimos cuadrados generalizados (*generalized least squares*) , es un método de estimación alternativo, pueden ajustarse a estas violaciones, pero este método se hace rápidamente impracticable a medida que el modelo aumenta en tamaño y complejidad; por tanto su uso es limitado. La ausencia de normalidad multivariante es particularmente problemática porque infla sustancialmente el estadístico de la Chi-Cuadrado y crea un sesgo alcista en valores críticos para determinar la significación de los coeficientes. Dado que los programas sólo aceptan las matrices de varianza-covarianza o correlación, el investigador debe realizar todos los test de diagnóstico sobre los datos antes de utilizarlos en el procedimiento de estimación. Aunque los programas de ecuaciones estructurales no tienen procedimientos de diagnóstico para contrastar estos tres supuestos, pueden ser contrastados con métodos convencionales como **PRELIS** . El investigador también debería identificar cualquier atípico en los datos antes de convertirlos a forma matricial.

## Datos ausentes

Los datos ausentes pueden tener un profundo efecto sobre el cálculo de la matriz de entrada de datos y su capacidad para ser utilizados en el proceso de estimación. Existen 2 formas en las que se pueden incorporar los valores perdidos en **SEM.** La primera es el método directo, en el que los parámetros del modelo se estiman tanto con datos completos como incompletos. Esta aproximación se utiliza raramente, sin embargo, dada la complejidad del modelo resultante. Más habitual es el método indirecto, para el cual se estima una matriz de datos de entrada utilizando toda o parte de la información disponible. Existen muchos métodos disponibles para resolver el problema de los datos ausentes, que van desde la supresión según el orden de lista a los métodos de imputación. La investigación reciente ha mostrado que el método **SEM** introduce menos sesgo en los modelos estimados. pero que las opciones pareada y en forma de lista se desarrollan adecuadamente si la proporción de datos ausentes no es muy elevada. Una conclusión del método de listas es que puede reducir seriamente el tamaño muestral. La aproximación pareada puede introducir irregularidades en la matriz de datos de entrada que provocarán después serios problemas en el proceso de estimación. Por tanto, no existe un método único que produzca siempre los mejores resultados, y si es posible que el investigador debiera emplear varias aproximaciones para evaluar la estabilidad de los

resultados.

## Matriz varianza-covarianza vs Matriz de correlación

Un tema importante en la interpretación de los resultados es el uso de la matriz de varianza-covarianza frente a la matriz de correlación. **SEM** se formuló inicialmente para ser utilizada con la matriz de varianza-covarianza (de ahí su denominación habitual como análisis estructural de la covarianza). La matriz de covarianza tiene la ventaja de proporcionar comparaciones válidas entre diferente poblaciones o muestras, característica que no es posible cuando los modelos se estiman con una matriz de correlación. La interpretación de los resultados, sin embargo, es de alguna forma más difícil cuando se utilizan covarianzas dado que los coeficientes deben interpretarse en términos de las unidades de medida de los constructos. La matriz de correlación tiene un uso muy extendido en muchas aplicaciones. Las matrices de correlación tienen un rango común que hace posible las comparaciones directas de los coeficientes dentro de un modelo, dado que simplemente es una matriz de varianza-covarianza estandarizada en la que la escala de medida de cada variable se modifica dividiendo las varianzas o covarianzas por el producto de las desviaciones estándar. Utilizar las correlaciones es apropiado cuando el objetivo del investigador es sólo entender el patrón de las relaciones entre constructos, pero no explicar la varianza total del constructo. Otro uso apropiado es hacer comparaciones entre las diferentes variables, dado que las escalas de medida afectan a las covarianzas. Los coeficientes obtenidos de la matriz de correlación están siempre en unidades estandarizadas, similares a las ponderaciones beta de la regresión, y con un rango entre -1,0 y +1 ,0. Además, se ha demostrado que la matriz de correlación ofrece estimaciones más conservadoras de la significación de los coeficientes y no están sesgadas al alza, como se pensaba anteriormente . En conclusión, el investigador debería emplear una matriz de varianza-covarianza siempre que se desarrolle un contraste de la teoría real, en el que las varianzas y las covarianzas satisfacen los supuestos y la metodología y son la forma apropiada de los datos para validar las relaciones chau-sales. Sin embargo, a menudo el investigador se centra sólo en las pautas de las relaciones, no con la explicación total tal y como se necesita en el contraste de la teoría, siendo aceptable la matriz de correlación. Si se utiliza la matriz de correlación, el investigador debería interpretar con cautela los resultados y su generalización a diferentes situaciones. *utilizadas* Uno de los modos más extendidos de cálculo de las correlaciones o covarianzas entre las variables observadas es la correlación de producto-momento de Pearson. También es la forma más habitual de correlación utilizada en el análisis multivariante, simplificando al investigador el cálculo de las matrices de varianzas y covarianzas. El supuesto de la correlación del momento-producto es que ambas variables **se** miden métricamente. Esto hace la correlación momento-producto inadecuada para el uso con medidas no métricas (ordinales o binarias). Para permitir la incorporación de las medidas no métricas en modelos de ecuaciones estructurales, el investigador debe emplear diferentes tipos de correlación. Si ambas variables son ordinales con tres o más categorías (*policotómicas*), entonces es apropiada la correlación *policórica*. Si las medidas no métricas son binarias, entonces se utiliza la correlación *tetracórica*. Para casos en los que la medida

métrica está relacionada con una medida ordinal *policotómica*, la correlación *poliserial* es la que representa la relación. Si una medida binaria está relacionada con una medida métrica, se utiliza una correlación biserial.

**Tamaño muestral**

Incluso aunque las observaciones individuales no sean necesarias, como sí ocurría con el resto de los métodos multivariantes, el tamaño muestral juega un papel importante en la estimación e interpretación de los resultados **SEM**. El tamaño muestral, como con cualquier otro método estadístico, ofrece bases para la estimación del error muestral. La cuestión crítica en **SEM** es el tamaño muestral necesario. Aunque no existe ningún criterio que dicte el tamaño de muestra necesario, existen al menos **cuatro factores** que afectan a los requisitos del tamaño muestral :

- **Mala especificación del modelo,**
- **Tamaño del modelo,**
- **No cumplimiento de la normalidad, y**
- **Procedimiento de estimación.**

**Mala especificación del modelo**

La mala especificación del modelo se refiere a la medida en que el modelo sufre por los errores de especificación. El error de especificación es la omisión de variables relevantes del modelo especificado. Todos los modelos de ecuaciones estructurales sufren de errores de especificación en la medida en que no puede ser incluido cada indicador o constructo potencial. El impacto de los constructos e indicadores omitidos, sin embargo, debería de ser insignificante si el investigador ha incluido todos aquellos relevantes según la teoría. El tamaño muestral afecta a la capacidad del modelo para ser estimado correctamente e identificar el error de especificación si se desea. Por tanto, si el investigador está preocupado por el impacto del error de especificación, las exigencias sobre el tamaño muestral deberían aumentarse por encima de aquellos que se hubiesen requerido en otro modelo.

**Tamaño del modelo**

El tamaño de muestra mínimo absoluto debe ser al menos tan grande como el número de covarianzas y correlaciones de la matriz de datos de entrada. Sin embargo, más ha- bitual es un mínimo de al menos cinco encuestados para cada parámetro estimado, con un ratio de 10 encuestados por parámetro considerado más apropiado. Por tanto, a medida que la complejidad del modelo aumenta, también aumentan los requisitos. Nótese que estos requisitos difieren del concepto de grados de libertady hacen referencia al número de encuestados originales utilizado para calcular la matriz de correlación o covarianza.

**Incumplimientos de la normalidad**

En la medida en que los datos no cumplan los supuestos de normalidad multivariante, el ratio de los encuestados respecto de los parámetros necesita aumentar a una tasa generalmente aceptada de 15 encuestados para cada parámetro. Aunque algunos procedimientos de estimación se diseñan específicamente para tratar con datos no normales, siempre se anima al

investigador a ofrecer un tamaño de muestra suficiente para que el impacto de error muestra) sea mínimo, especialmente para datos no normales

**Procedimiento de estimación**

Se ha demostrado que el método de estimación de máxima verosimilitud (**MLE**), es el procedimiento de estimación más habitual, ofrece resultados válidos con muestras de tan sólo 50 observaciones, pero una muestra de este tamaño *no es recomendable.* Se acepta generalmente que el tamaño de muestra mínimo que asegura el uso apropiado de **MLE** es de 100 a 150. A medida que aumentamos el tamaño muestra por encima de ese valor**,** el método **MLE** aumenta en su sensibilidad al tamaño muestra por encima de ese valor. A medida que el tamaño muestra) aumenta (pasando de 400 a 500), el método se hace más sensible y casi no se detecta ninguna diferencia, haciendo que todas las medidas de calidad del ajuste indiquen un ajuste muy pobre . Aunque no existe un tamaño muestra) correcto, se recomiendan tamaños que vayan entre 100 y 200. Un enfoque es contrastar siempre un modelo con un tamaño de muestra de 200, sin importar cuál fuera el tamaño muestra original, debido a que algunos investigadores han propuesto 200 como un tamaño de muestra crítico. Si se selecciona un procedimiento de estimación asintóticamente libre de distribución, los requisitos del tamaño muestra aumenta sustancialmente, ya que exige mayores tamaños de muestra para compensar la dependencia sobre los supuestos de distribución mediante otros métodos.

Existen muchos factores que afectan al tamaño muestra exigido. Normalmente, recomendamos un tamaño muestral de 200, aumentando si se sospecha que existe mala especificación, el modelo es muy grande o complejo, los datos exhiben características no normales o se utiliza un procedimiento de estimación alternativo. Una evaluación aislada es el diagnóstico de la *N* crítica, que es el tamaño de muestra que haria significativo el nivel de ajuste del modelo (medido por Chi-Cuadrada**)**, llegando al nivel especificado de significación. La medida Chi-Cuadrado está disponible en la mayoría de los programas y es una base para la comparación del tamaño muestral.

**Estimación del modelo**

Una vez que están especificados los modelos estructurales y de medida y que se ha seleccionado el tipo de datos de entrada, el investigador debería elegir cómo se estimará el modelo. En el caso de **SEM**, tiene varias opciones tanto para el procedimiento de estimación como respecto al programa informático a utilizar.

**Técnicas de estimación**

Los intentos iniciales de estimación del modelo de ecuaciones estructurales se realizaron con la regresión de los mínimos cuadrados ordinarios (***OLS****.Ordinary Least Squares*). Pero estos esfuerzos fueron superados rápidamente por la estimación máxima verosímilitud (**ML.***Maximum Likelihood*), que es eficiente y no sesgada cuando se cumplen los supuestos de normalidad multivariante. Como tal y se ha convertirdo en una técnica ampliamente empleada en la mayoría de los programas informátios como **EQS**. La sensibilidad de **MLE** a la no normalidad, sin embargo, creó una necesidad de técnicas de estimación alternativas y pronto aparecieron métodos tales como los mínimos cuadrados ponderados (***WLS.****Weighted Least Squares.*),

mínimos cuadrados generalizados (***GLS.****Generalized Leased Squares*) y asintóticamente libre de distribución (**AGL**) . La técnica **AGL** ha recibido recientemente atención particular debido a su insensibilidad a la no normalidad de los datos. Su principal conclusión es la exigencia de un aumento del tamaño de muestra. Todas las técnicas de estimación alternativas han empezado a difundirse ampliamente a medida que ha aumentado el uso de computadoras personales,·haciéndoles factibles resolver cualquier problema habitual.

## Procesos de estimación

Además de la técnica de estimación empleada, el investigador puede también escoger entre varios procesos de estimación. Estos procesos van desde la estimación directa del modelo, que es similar a lo que hemos visto en el resto de las técnicas multivariantes, a métodos que generan miles de estimaciones del modelo para las cuales se obtienen los resultados finales del modelo. Existen 4 procesos: *estimación directa, bootstrapping, simulación y análisis (jackknife).*

**Estimación directa**

El proceso de estimación más habitual es el de estimación directa, en el cual se estima directamente un modelo con un procedimiento de estimación elegido. En este proceso, se estima primero, el parámetro; a continuación el intervalo de confianza (y error estándar) de cada parámetro estimado que se basa en el error muestrlal. Tanto los parámetros estimados como su intervalo de confianza provienen del modelo estimado de una muestra aislada.

**Bootstrapping**

Sin embargo, existen varias alternativas que no descansan en un único modelo de estimación sino en estimaciones del parámetro calculado y sus intervalos de confianza basados en estimaciones múltiples .La primera opción es *bootstrapping*, que se realiza en 4 pasos:. En el primero, se diseña la muestra original para que actúe como la población a efectos muestrales. En el segundo paso, se vuelve a muestrear la muestra original un número especificado de veces (quizá hasta varios miles) para generar un gran número de nuevas muestras, siendo cada una un subconjunto aleatorio de la muestra original. En el tercer paso, se estima el modelo y para cada muestra nueva se guardan los parámetros estimados. En el último paso, las estimaciones de los parámetros finales se calculan como la media de las estimaciones de los parámetros de todas las muestras. El intervalo de confianza no se estima por un error muestra!, sino que se observa directamente examinando la distribución efectiva de los parámetros estimados alrededor de la media. De esta forma, las estimaciones del parámetro final y sus intervalos de confianza se obtienen directamente de las estimaciones de modelos múltiples para varias muestras aisladas y no descansan en supuestos como la distribución estadística de los parámetros.

**Simulación**

El investigador puede emplear técnicas de simulación, que también descansan en muestras múltiples y modelos estimados. Los procesos de simulación difieren de *bootstrapping* en que durante el proceso de generación de nuevas muestras, el

programa de simulación puede cam-biar cietias características de la muestra para cumplir con lo que se propone el investigador. Por ejemplo, el grado de correlación entre variables puede ser variable a lo largo de las muestras en cierta forma sistemática. De esta forma, el investigador no sólo tiene una variación asintótica de la muestra entre las muestras sino también una pauta sistemática especificada en el procedimiento de simulación. Con estas muestras, se estiman otra vez los modelos para cada muestra y los resultados compilados, como en el proceso de *bootstrapping.*

**Análisis *Jackknife*.**

En este proceso de estimación**,** se crean de nuevo muestras repetidas a partir de la muestra original. Este método difiere de los procedimientos de simulación y *bootstrapping**,*** sin embargo, en el método de creación de nuevas muestras. En lugar de crear un gran número de nuevas muestras aleatorias, el proceso *Jackknife* crea N muestras nuevas, donde N es el tamaño muestral original. Cada vez que se crea una nueva muestra, se omite una observación diferente. Por tanto, cada nueva muestra tiene un tamaño de muestra de N-1 con una observación diferente omitida en cada muestra. La ventaja de este proceso es la facilidad de identificar observaciones influyentes mediante el examen de los cambios en los parámetros estimados. Si se desea, la estimación del último parámetro puede calcularse como el parámetro medio, pero en los casos de tamaños de muestras reducidas no existen nuevas muestras para calcular adecuadamente el intervalo de confianza.

**Programas informáticos**

Una vez que se ha elegido el procedimiento de estimación, el si- guiente paso es escoger el programa informático utilizado para estimar de forma efectiva el modelo. El programa más utilizado es **LISREL** *(Linear Structural RELations)* un modelo verdaderamente flexible para varias situaciones de investigación (sección-cruzada, experimental, cuasi-experimental y estudios longitudinales). **LISREL** ha encontrado aplicaciones a lo largo de todos los campos de estudio y se ha convertido casi en sinónimo de la modelización de ecuaciones estructurales**. EQS** tiene supuestos menos exigentes sobre la normalidad multivariante de los datos y **LVPLS** se ajusta mejor para la predicción aunque es limitado a efectos de interpretación de los resultados. **AMOS** se ha ganado una creciente popularidad en los últimos años debido a la sencillez del interfaz para el usuario y se ha comparado recientemente con **LISREL** y **EQS**. Sin embargo, existen varios programas informáticos alternativos, entre ellos, **PROC CALIS** de **SAS**, **COSAN**. Todos estos programas se encuentran disponibles en versiones que pueden ser utilizadas por computadoras personales.

## Resolución del problema de *definida no positiva*

Un problema habitual para todos los usuarios de **SEM** es el mensaje de error del computador: *la matriz... es definida no positiva*. Lo que ha ocurrido es que tanto la matriz de datos de entrada como la matriz de datos estimados es

singular, lo que significa que existe una dependencia lineal o inconsistencia entre algún grupo de variables. Existen muchas causas de este problema, pero se han encontrado algunas generales. Si el error se produce en la matriz de entrada de datos, las causas más probables son:

- El enfoque para tratar con los datos ausentes utilizado, especialmente la eliminación pareada; o
- Una dependencia lineal entre las variables, incluyendo todos los ítems de la escala y el total de la escala en la matriz de entrada. En estos dos casos, el investigador debería generar una nueva matriz de datos, empleando un proceso de datos ausentes alternativo o eliminando las variables infractoras. Si el problema se produce en la matriz de datos estimados, entonces el investigador debe corregir cualquier varianza de error negativa (conocida como casos Heywood) y que se describen en el paso 6 o intentar valores de entrada distintos.

## Paso 5
## Valoración de la identificación del modelo estructural

Durante el proceso de estimación. la causa más probable de la interrupción súbita del programa de la computadora o de producir resultados sin sentido o ilógicos es un problema de identificación del modelo estructural. Un problema de identificación, en términos sencillos es la incapacidad del modelo propuesto para generar estimaciones aisladas. Se basa en el principio de que debemos tener una ecuación aislada y diferenciada para estimar cada coeficiente, basado en la premisa de que *se deben tener más ecuaciones que incógnitas* como principio básico del álgebra . Sin embargo, a medida que el modelo estructural se hace más complejo, no existe un enfoque garantizado para asegurar que el modelo está identificado. (**Ver** más en apartado *el concepto de identificación de modelo*)

### Grados de libertad

A efectos de identificación, el investigador se centra en el tamaño de las matrices de correlación o covarianzas relativas respecto al número de coeficientes estimados. La diferencia entre el número de correlaciones o covarianzas y el número efectivo de coeficientes en el modelo propuesto se denomina grados de libertad. Al igual que los grados de libertad que encontramos en la regresión múltiple o **MANOVA** (Ver Mejía-Trejo, 2017d), un grado de libertad es un elemento no restringido de la matriz de datos. El número de grados de libertad para un modelo propuesto se calcula como

$$gl = 1/2[(p+q)(p+q+1)] - t$$

Donde:
***p*** = número de indicadores endógenos
***q*** = el número de indicadores exógenos
***t*** = número de coeficientes estimados en el modelo propuesto

La primera parte de la ecuación calcula el tamaño no redundante de la matriz de covarianza o correlación (es decir, la mitad inferior o superior de la matriz más

la diagonal). Entonces cada coeficiente estimado gasta un grado de libertad. La principal diferencia entre los grados de libertad utilizados en **SEM** comparado con otras técnicas multivariantes es que el número de parámetros estimados se compara con el número de elementos de la matriz de datos, NO con el tamaño muestra!. En **SEM** el tamaño muestral se utiliza para estimar el error de muestreo, pero no afecta a los grados de libertad.

## Normas para la identificación

Aunque no existe una regla aislada que establezca la identificación de un modelo, el investigador dispone de varias *normas* o heurística. Las 2 normas más básicas son: las condiciones de orden y rango. Las condiciones de orden afirman que los grados de libertad del modelo deben ser mayores o iguales a cero. Esto corresponde a lo que hemos denominado como modelo identificado o modelo sobreidentificado. Un modelo identificado tiene exactamente cero grados de libertad. Aunque esto ofrece un ajuste perfecto del modelo, la solución no tiene interés puesto que no se puede generalizar. Un modelo sobreidentificado es el objetivo de todos los modelos de ecuaciones estructurales. Tiene más información en la matriz de datos que el número de parámetros a estimar, lo que significa que tiene un número positivo de grados de libertad. Al igual que en otras técnicas multivariantes, el investigador se esfuerza por conseguir un ajuste aceptable con el mayor grado de libertad posible. Esto asegura que el modelo es tan generalizable como sea posible. Un modelo que no llega a cumplir la condición de orden se conoce como un modelo infraestimado. Este modelo tiene grados de libertad negativos, lo que significa que se intentan estimar más parámetros de lo que permite la información disponible. El modelo no puede ser estimado hasta que algunos parámetros sean fijos o restringidos. La condición de orden es una condición de identificación necesaria, pero no suficiente. El modelo también debe cumplir la condición de rango, lo que exige que el investigador determine algebraicamente si cada parámetro se identifica (estima) especialmente. Pero incluso, para los modelos más sencillos, se trata de un ejercicio muy complejo para ser considerado directamente por el investigador. En su lugar, hay heurística disponible. Primero está la norma de las 3 medidas, que evalúa que cualquier constructo **con tres o** más indicadores siempre estará identificado. También existe la *norma del modelo recursivo*, que dice que los modelos recursivos con constructos identificados (norma de las tres medidas) siempre estarán identificados . Un modelo recursivo NO tiene relaciones no recursivas o recíprocas en el modelo estructural .

## Diagnóstico de los problemas de identificación

Los programas de ecuaciones estructurales también realizan contrastes para diagnosticar problemas de identificación. **LISREL** tiene un test sencillo de identificación durante el proceso de estimación para examinar la matriz de información, mientras que **EQS** tiene el test de rango de Wald (1943). Aunque estos contrastes identifican la mayoría de los problemas de identificación, pueden no evaluar la unicidad de cada parámetro estimado, como exige la condición de rango.

El investigador puede llevar a cabo contrastes cuando la ecuación se identifica para ver si los resultados son inestables debido al nivel de identificación. En

primer lugar, el modelo puede ser reestimado varias veces, cada una de ellas con un valor de partida diferente. El investigador puede especificar un valor inicial para cualquier parámetro estimado, un punto de partida del proceso de estimación. Si el valor de partida no está disponible, el programa lo calcula automáticamente por diversos métodos. Si los resultados no convergen en el mismo punto para diversos valores de partida, la identificación debería examinarse más a fondo. El segundo test evalúa el efecto de la identificación sobre un único coeficiente, que consiste en estimar en primer lugar el modelo y a continuación obtener la estimación de coeficiente. Posteriormente, se fija el coeficiente a su valor estimado y se reestima la ecuación. Si el ajuste conjunto del modelo varía significativamente, esto indica problemas de identificación. Otro enfoque es observar los posibles síntomas de un problema de identificación. Estos incluyen:

- Errores estándar muy elevados para uno o más coeficientes,
- La incapacidad del programa para invertir la matriz de información,
- Estimaciones muy poco razonables o estimaciones imposibles tales como varianzas de error negativas, o
- Elevadas correlaciones **(± 0,90 o superiores**) entre los coeficientes estimados.

## Causas y soluciones para los problemas de identificación

Si se localiza un problema de identificación, el investigador debería buscar 3 posibles causas:

- Un mayor número de coeficientes estimados relativos al número de correlaciones o cavaríanzas, indicada por el reducido número de grados de libertad, similar al problema de sobreajuste de los datos encontrado en otras técnicas multivariantes;
- El uso de efectos recíprocos (flechas causales de dos sentidos entre dos constructos); o
- Fallo en la fijación de la escala de un constructo.

Discutiremos este procedimiento más adelante en nuestro análisis de los datos del ejemplo.

La única solución para un problema de identificación es definir más restricciones para el modelo ,esto es, eliminar algunos de los coeficientes estimados. El investigador debería seguir un proceso estructurado, *añadiendo gradualmente más restricciones (eliminando secuencias del diagrama de trayectorias)* hasta que el problema esté solucionado. Al hacerlo, el investigador está intentando conseguir un modelo sobreidentificado que tiene grados de libertad suficientes con las cuales evaluar, si es posible, la cantidad de error de medida y de muestreo y ofrecer mejores estimaciones de las verdaderas relaciones casuales. Para llegar a este fin se sugiere onstruir un modelo teórico *con el mínimo número de coeficientes* (incógnitas) que puedan ser justificados. Si se *encuentran problemas de identificación,* procederemos con las soluciones en este orden DE acciones sugeridas:

- FIjar las varianzas de error de medida de los constructos si es posible,
- Fijar algunos coeficientes estructurales que sean conocidos con confiabilidad, y

- Eliminar las variables problemáticas.

Si los problemas de identificación todavía existen, el investigador debe reformular el modelo teórico para ofrecer más constructos relativos al número de relaciones causales examinadas.

## Paso 6
## Evaluación de los criterios de calidad de ajuste

El primer paso de la evaluación de los resultados es una inspección inicial de las estimaciones infractoras. Una vez que el modelo está establecido como para ofrecer estimaciones aceptables, a continuación debe evaluarse la calidad del ajuste a diversos niveles: en primer lugar para el modelo conjunto y a continuación para los modelos estructurales y de medida por separado.

### Estimaciones infractoras

En primer lugar se examinan los resultados buscando estimaciones infractoras. Se trata de coe-ficientes estimados tanto en los modelos de medida como los estructurales que exceden los límites aceptables. Los ejemplos más normales de estimaciones infractoras son:

- **Varianzas de error negativas o varianzas de error no significativas para cualquier constructo,**
- **Coeficientes estandarizados que sobrepasan o están muy cerca de 1.0 o**
- **Errores estándar muy elevados asociados con cualquier coeficiente estimado.** Si se encuentran estimaciones infractoras, el investigador debería resolver en primer lugar cada caso antes de evaluar cualquier resultado específico del modelo, en la medida en que cambios en una parte del modelo puedan tener efectos significativos sobre otros resultados. Se han utilizado varios enfoques para la resolución de estos problemas en la discusión de los problemas de identificación. Si se corrigen los problemas de identificación y aún así, los problemas todavía se mantienen, existen otras soluciones. En el caso de varianzas de error negativas (también conocidas como casos de Heywood), una posibilidad es fijar las varianzas de los errores infractores a un valor positivo **muy pequeño (.005)** . Aunque este remedio cumple los requisitos prácticos del proceso de estimación, sólo enmascara el problema subyacente y debe ser considerado cuando se interpreten los resultados. Si las correlaciones de la **solución estandarizada exceden de 1.0**, o dos estimaciones están altamente correlacionadas, entonces el investigador deberá considerar la eliminación de uno de los constructos o deber asegurarse que se ha establecido entre los constructos una verdadera validez discriminante. En muchos casos, tales situaciones son el resultado de modelos ateóricos, establecidos sin la suficiente justificación teórica o modificados solamente a partir de consideraciones empíricas.

### Ajuste global del modelo

Una vez que el investigador ha establecido que no existen estimaciones infractoras, el siguiente paso es evaluar el ajuste global del modelo con una o más medidas de calidad del ajuste. La calidad del ajuste mide la correspondencia entre la matriz de entrada real u observada (covarianza o correlación) con la que se predice mediante el modelo propuesto.

Al desarrollar cualquier modelo estadístico, el investigador debe ser precavido con el *sobreajuste* del modelo a los datos. Al discutir la regresión, mostramos que

deberían mantenerse ciertas relaciones (**quizá 5 a 1**, ver Mejìa-Trejo, 2017c) entre el número de coeficientes estimados y el número de encuestados. Esta relación debería mantenerse también en **SEM**. **El** investigador debe esforzarse en tener un gran número de grados de libertad, si el resto no cambia. Al hacerlo así, el modelo consigue parsimonia, el objetivo de un mejor o mayor ajuste del modelo para cada coeficiente estimado. Cuanto mejor ajuste pueda conseguirse con pocos coeficientes, mejor será el contraste del modelo y más confianza podemos tener en que los resultados no sean producto del sobreajuste de los datos.

Las medidas de calidad del ajuste son de tres tipos:

- **Medidas absolutas del ajuste,**
- **Medidas del ajuste incremental, o**
- **Medidas de ajuste de parsimonia.**

Las medidas absolutas del ajuste evalúan sólo el ajuste global del modelo (tanto los modelos de medida como los estructurales colectivamente), sin ajuste para el grado de sobreajuste que pudiera ocurrir. Las medidas del ajuste incremental comparan el modelo propuesto con otro modelo especificado por el investigador. Finalmente, las medidas de ajuste de parsimonia ajustan las medidas de ajuste para ofrecer una comparación entre modelos con diferentes números de coeficientes estimados, siendo el propósito determinar la cantidad del ajuste conseguido por cada coeficiente estimado. El investigador se enfrenta con la cuestión de qué medidas elegir. No surge ninguna medida aislada o conjunto de medidas como las únicas medidas necesarias. Como **SEM** ha evolucionado en los últimos años, las medidas de calidad del ajuste se han estado desarrollando continuamente y se han propuesto medidas adicionales]. Se anima al investigador a buscar y emplear una o más medidas de cada tipo. La evaluación de la calidad del ajuste de un modelo es más un proceso relativo que un criterio absoluto. La aplicación de diversas medidas del ajuste múltiple permitirá al investigador ganar en consenso tanto para los tipos de medidas como para la aceptación del modelo propuesto. Un nivel aceptable de calidad del ajuste conjunto NO garantiza que todos los constructos cumplan los requisitos del ajuste del modelo, ni que el modelo estructural puede ser mantenido con certeza absoluta. El investigador deberá evaluar cada una de estas áreas separadamente para confirmar su cumplimiento de los requisitos o como medio de identificar problemas potenciales que afectan a la calidad del ajuste conjunto del modelo.

## Ajuste del modelo de medida

Una vez que se ha evaluado el ajuste del modelo conjunto, podemos evaluar la unidimensionalidad y la confiabilidad de la medida de cada constructo. La unidimensionalidad es un supuesto que subyace al cálculo de la confiabilidad y se demuestra cuando los indicadores de un constructo tienen un ajuste aceptable sobre un modelo de un único factor (unidimensional). El uso de medidas de confiabilidad, tales como la *alfa de Cronbach* NO asegura la unidimensionalidad sino que en su lugar supone que existe. Se anima al investigador a que re-alice contrastes de unidimensionalidad sobre todos los constructos de indicadores múltiples antes de evaluar su confiabilidad. El siguiente paso es examinar las ponderaciones estimadas y evaluar la significación estadística de cada una. Si la significación estadística no se

consigue, el investigador puede desear eliminar el indicador o intentar transformarlo para un mejor ajuste del constructo.

**Confiabilidad compuesta**

Antes de realizar el examen de las ponderaciones de cada indicador, se utiliza una medida fundamental para la evaluación del modelo de medida, la confiabilidad compuesta de cada constructo. La confiabilidad es una medida de la consistencia interna de los indicadores del constructo, que representa el grado en que éstos *indican* el constructo común latente (no observado). Medidas más confiables ofrecen al investigador una mayor confianza de que todos los indicadores individuales son consistentes en sus medidas. Un valor umbral comúnmente aceptado para aceptar la hipótesis de fiabilidad es .7, aunque no se trate de un estándar absoluto, y los valores por debajo de .7 se han demostrado aceptables si la investigación tiene naturaleza exploratoria. Debemos tener en cuenta, sin embargo, que la confiabilidad no asegura validez. La validez es la medida en que los indicadores miden *con precisión* lo que se supone que están midiendo. Por ejemplo, pueden ser confiables diversas medidas de cómo y por qué los consumidores compran productos, pero el investigador puede asumir erróneamente que miden la lealtad a la marca cuando de hecho son indicadores de intenciones de compra. En este caso, los indicadores son un conjunto confiable de medidas pero una medida inválida de lealtad a la marca. El supuesto de validez descansa en la especificación del investigador de indicadores para un constructo latente.

La confiabilidad y la varianza extraída (véase la sección siguiente) de un constructo latente deben calcularse por separado para cada constructo con indicadores múltiples en el modelo. Aunque **LISREL** y **EQS** no los calculan directamente, se proporciona toda la información necesaria. (Ver más en Capítulo 3).

**Varianza extraída**

Otra medida de confabilidad es la medida de varianza extraída. Esta medida refleja la cantidad total de la varianza de los indicadores tenida en cuenta por el constructo latente. Los mayores valores de la varianza extraída se producen cuando los indicadores son verdaderamente representativos del constructo latente. La varianza extraída es una medida complementaria del valor de la confiabilidad del constructo. Esta medida es muy parecida a la medida de la fiabilidad pero difiere en que las ponderaciones estandarizadas se elevan al cuadrado antes de sumarlas. En general, se sugiere que el valor de la varianza extraída debería exceder de .5 para un constructo. Los ejemplos reales de los cálculos tanto para la medida de la fiabilidad como la varianza extraída, se ofrecen en el ejemplo del análisis factorial confirmatorio de este mismo capítulo. (Ver más en Capítulo 3 y 4).

**Ajuste del modelo estructural**

El examen más obvio del modelo estructural consiste en la significación de los coeficientes estimados. Los métodos de modelización de ecuaciones estructurales ofrecen no sólo coeficientes estimados sino también errores estándar y valores *t* calculados para cada coeficiente. Si podemos especificar el nivel de significación que juzgamos apropiado (**es decir .05**), entonces cada

coeficiente estimado puede ser contrastado por la significación estadística (es decir, que sea diferente de cero) para las relaciones causales supuestas. Sin embargo, dadas las propiedades estadísticas del **MLE** y sus características (tamaños de muestra reducidos), se anima al investigador a ser conservador en la especificación del nivel de significación, **eligiendo niveles pequeños ( .025 o .01) en lugar del tradicional nivel de .05.** La selección de un valor crítico también depende de la justificación teórica de las relaciones propuestas. Si se supone una relación positiva o negativa, se puede emplear un test de significación con una sola cola. Sin embargo, si el investigador no puede preespecificar la dirección de la relación, entonces se necesita utilizar un test de dos colas. La diferencia está en los valores críticos de la *t* utilizados para evaluar la significación. Por ejemplo, para el nivel de significación de **.05**, el valor crítico para un test de una sola cola es de **1.645**, pero aumenta a **1.96** para un test de dos colas. Por tanto, el investigador puede detectar con mayor precisión las diferencias si puede utilizarse una teoría más fuerte en la especificación del modelo como medida de la ecuación estructural, se calcula un coeficiente conjunto de determinación *($R^2$),* similar al que se encuentra en la regresión múltiple. Aunque no pueda llevarse a cabo ningún test de significación estadística, ofrece una medida de ajuste relativa para cada ecuación estructural. Los resultados de **SEM** pueden verse afectados por la multicolinealidad, al igual que en la regresión. Aquí el investigador debe tener cuidado de las correlaciones entre las estimaciones de los constructos en los resultados de SEM. Si aparece un valor muy elevado, entonces debe llevarse a cabo una acción correctiva. Esta acción puede incluir la destrucción de un constructo o la reformulación de relaciones causales. Aunque no se ha fijado ningún límite que defina lo que se con-sideran como correlaciones elevadas, los valores que exceden .9 siempre deberían ser examinados, y muchas veces las correlaciones que exceden .8 pueden ser indicadoras de problemas.

## Comparación de modelos anidados o rivales

Las estrategias de modelización más habituales son: estrategia de desarrollo de modelo y modelos rivales consisten en la comparación de los resultados del modelo para determinar el modelo mejor ajustado de un conjunto de modelos. En una estrategia de desarrollo del modelo, el investigador comienza con un modelo inicial y sigue con una serie de re-especificaciones del modelo, con las que cada vez se espera mejorar el ajuste del modelo mientras se mantenga la concordancia con la teoría subyacente.Para ayudar a la comparación entre modelos, se han desarrollado una gran cantidad de medidas para evaluar el ajuste del modelo. Una clase de medidas evalúa el ajuste del modelo conjunto en términos absolutos, ofreciendo una medida específica del ajuste. Una conclusión sobre estas medidas es que no tienen en cuenta el número de relaciones utilizadas en la obtención del ajuste del modelo. Para medir la parsimonia del modelo, se han propuesto una serie de medidas de ajuste de parsimonia. Su objetivo es determinado por el ajuste por coeficiente, dado que el ajuste absoluto siempre mejorará a medida que se añaden coeficientes. Anderson y Gerbing (1988) propusieron un procedimiento exhaustivo para este propósito, en el que se especifican una serie de modelos

rivales. Puede mostrarse que las diferencias entre modelos son simplemente la diferencia entre los valores de la Chi-Cuadrado para los diferentes modelos. A continuación se puede contrastar la significación de la diferencia de la Chi-Cuadrado con los grados de libertad apropiados, esto es, la diferencia en el número de coeficientes estimados para los dos modelos. El único requisito es que el número de constructos o indicadores sea el mismo, de tal forma que el modelo nulo sea el mismo para ambos modelos (es decir, son modelos anidados). El efecto de añadir o destruir una o más relaciones causales se puede contrastar también de esta forma haciendo comparaciones entre modelos con y sin las relaciones. Si los modelos no se hacen anidados (tener un número diferente de indicadores o constructos), el investigador debe basarse en las medidas de ajuste de parsimonia descritas previamente, en la medida en que el test de la Chi-Cuadrado no es apropiado para este caso

## Paso 7
## Interpretación y modificación del modelo

Una vez que el modelo se considera aceptable, el investigador debe examinar en primer lugar los resultados y su correspondencia con la teoría propuesta. ¿Están corroboradas y son estadísticamente significativas las principales relaciones de la teoría? ¿Añaden los modelos rivales mayor perspectiva sobre las formulaciones alternativas de la teoría como para que puedan ser tenidas en cuenta? ¿Están todas las relaciones en la dirección supuesta (positiva o negativa)? Todas estas cuestiones y muchas más pueden ser contestadas a partir de los resultados empíricos. En el proceso de respuesta a estas tres cuestiones, el investigador puede tener necesidad de considerar dos supuestos de interpretación: el uso de las soluciones estandarizadas frente a las no estandarizadas y la re- especificación del modelo.

### Soluciones estandarizadas frente a las no estandarizadas

Un aspecto de la evaluación de una relación estimada es la valoración del tamaño del parámetro. Al igual que en otras técnicas multivariantes, tales como la regresión múltiple, existe una diferencia notable en las soluciones estandarizadas y no estandarizadas en términos de su interpretación y uso. En los modelos de ecuaciones estructurales, los coeficientes estandarizados tienen todos igual varianza y un **valor máximo de 1.0,** por tanto, efectos tamaño muy aproximados, como se mostró con las ponderaciones beta en la regresión. Los coeficientes cercanos a cero tienen poco efecto sustantivo, si es que tienen alguno, mientras que un aumento en el valor corresponde a un aumento de la importancia en las relaciones causales. Los coeficientes estandarizados son útiles en la determinación de la importancia relativa, pero son específicos de una muestra y no son comparables entre las muestras. Los coeficientes sin estandarizar corresponden a las ponderaciones de la regresión en una regresión múltiple en la que se expresan en términos de la escala del constructo, en este caso su varianza. Esto hace comparables a estos coeficientes para las muestras y retiene sus efectos escala. Dado que la escala varía para cada constructo, sin embargo, la comparación entre coeficientes es más dificil que con los coeficientes estandarizados

### Re-especificación del modelo

Una vez que la interpretación del modelo se ha completado, el investigador

normalmente bus-ca métodos para mejorar el ajuste del modelo y/o su correspondencia con la teoría subyacente. En tales casos, puede iniciar la re-especificación del modelo, el proceso de añadir o eliminar los parámetros estimados del modelo original. Antes de tratar algunos enfoques para identificar la modificación del modelo, aconsejamos al investigador hacer tales modificaciones con cuidado y sólo después de obtener justificación teórica para lo que se considera empíricamente deseable. Las modificaciones del modelo original deberían hacerse sólo después de una consideración deliberada. Si se hacen las modificaciones, el modelo debería tener validación cruzada (es decir. estimado sobre un conjunto distinto de datos) antes de que el modelo modificado sea aceptado.

## Un proceso de re-especificación del modelo

Antes de identificar cualquier posible re-especificación del modelo, el investigador deberá clasificar todas las relaciones (estimadas o no) en una de dos categorías: teóricas o empíricas. Las relaciones teóricas son esenciales a la teoría subyacente y no pueden ser modificadas. Están fuera de los límites de la reespecificación. La categoría empírica contiene relaciones que se añaden para mejorar el ajuste del modelo. Estas pueden ser re-especificadas. El objetivo es determinar un conjunto de modelos teóricos anidados, donde el conjunto de modelos se contempla como un conjunto de diferentes niveles de parsimonia para la misma teoría subyacente. De esta forma, el modelo se convierte en una serie de modelos rivales que llevan a varios niveles de corroboración para la teoría.

## Indicadores empíricos de posibles reespecificaciones

¿Dónde puede buscar el investigador la mejora de los modelos? La primera indicación viene del examen de los residuos de la matriz de las predicciones de la covarianza y correlación. Los residuos estandarizados (también denominados residuos normalizados) representan las diferencias entre la matriz de covarianza o co- rrelación observada y la matriz de covarianza o correlación estimada. Con las últimas versiones de **LISREL** (versión 7 y superiores), se ha mejorado el cálculo de los residuos y el patrón de evalua- ción de los residuos «significativos» ha cambiado (el umbral anterior estaba en ::!:: **2,0**). Los valores residuales mayores que ::!:: **2.58** se consideran ahora estadísticamente significativos al **nivel de. 05**. Los residuos significativos indican un error de predicción sustancial para un par de indicadores (es decir, una de las covarianzas o correlaciones de los datos de entrada originales).

Un residuo estandarizado indica sólo que existe diferencia pero no arroja luz acerca de cómo puede ser reducida. El investigador debe identificar el remedio mediante la suma o modificación de las relaciones causales.

Otro apoyo en la evaluación del ajuste de un modelo especificado consiste en los índices de modificación, que se calculan para cada relación no estimada. El valor del *índice de modificación* corresponde aproximadamente a la reducción en la Chi-Cuadrado que se produciría si el coeficiente fuera estimado. Un valor de **3.84** o superior sugiere que se obtiene una reducción estadísticamente significativa en la Chi-Cuadrado cuando se estima el coeficiente. Aunque los índices de modificación pueden ser útiles en la evaluación del impacto de modificaciones basadas teóricamente, el

investigador nunca hace cambios de modelo basándose sólo en índices de modificación. Este enfoque ateórico es completamente contrario al *espíritu* de la técnica y debería ser evitado en todos los casos. La modificación del modelo debería tener una justificación teórica antes de ser considerada e incluso entonces el investigador debería ser escéptico acerca de los cambios. Además del *índice de modificación*, **LISREL** y **EQS** también ofrecen un parámetro de cambioesperado, que hace referencia a la magnitud y la dirección de cada parámetro fijo (no estimado). Este parámetro difiere del índice de modificación en que no indica el cambio en el ajuste del modelo conjunto **($x^2$ )** y en su lugar representa el cambio en el valor del parámetro efectivo. **EQS** ofrece el multiplicador de Lagrange y el estadístico de Wald, que evalúa el efecto de liberar un conjunto de parámetros simultáneamente. Una vez hechas las modificaciones en el modelo, **el investigador debe volver al paso 4 del proceso de 7 pasos y reevaluar los modelos modificados**. Si se anticipan modificaciones amplias del modelo, los datos deberían dividirse en dos muestras, una de ellas ofreciendo la base para la estimación y modificación del modelo, y la otra validación para el modelo final.

**Una recapitulación del proceso de siete pasos**

La modelización de ecuaciones estructurales ofrece al investigador más flexibilidad que cualquier otro método multivariante discutido. Pero junto con esta flexibilidad viene el potencial uso inapropiado del modelo. Un asunto primordial en cualquier aplicación de **SEM**, es que deber tener una fundamentación teórica incondicional, apoyada por un test confirmatorio robusto para una serie de relaciones casuales. Sin embargo, cuando se aplica el método de forma exploratoria, el investigador se enfrenta con la muy elevada probabilidad de no conseguir resultados mediante la búsqueda o de datos y las relaciones identificadas tendrán poca generalización mediante la simple extrapolación de las relaciones específicas de los datos de la muestra que se están estudiando.

Fuente: Jöreskog (1993); Hair (et al.,1999); Byrne 2006 con adaptación propia.

# Formulación de estructuras de covarianza y medias

Los parámetros centrales en los modelos de ecuaciones estructurales que se centran en el análisis de las estructuras de covarianza son los coeficientes de regresión y las varianzas y covarianzas de las variables independientes**.** Sin embargo, dado que los datos de la muestra comprenden solo puntajes observados, es necesario que haya una evaluación de mecanismo interno, mediante el cual los datos se transponen a los parámetros del modelo. Esta tarea se realiza a través de un modelo matemático que representa todo el sistema de variables. Dichos sistemas de representación pueden y deben variar con cada programa de computadora **SEM**; el mecanismo utilizado por el programa **EQS** se discute más adelante. Al igual que con cualquier forma de comunicación, uno debe primero entender el lenguaje antes de poder entender el

mensaje transmitido; así es en la comprensión de la especificación de los modelos **SEM.**

# Notación

**EQS** considera que todas las variables caen en una de dos categorías: variables medidas (observadas) o variables no medidas (no observadas). Todas las variables medidas se designan como **V** y constituyen los datos *reales* de un estudio. Todas las demás variables son hipotéticas y representan la red estructural del fenómeno bajo investigación. Aunque conceptualmente es innecesario, tiene sentido en la práctica diferenciar las variables no medidas, debido a:

**a.** *El propio constructo latente* (considerada generalmente como un factor en **EQS**), designada como **F**;

**b.** *Un residual asociado* con la medición de cada variable observada (**V**), designada como **E**; y

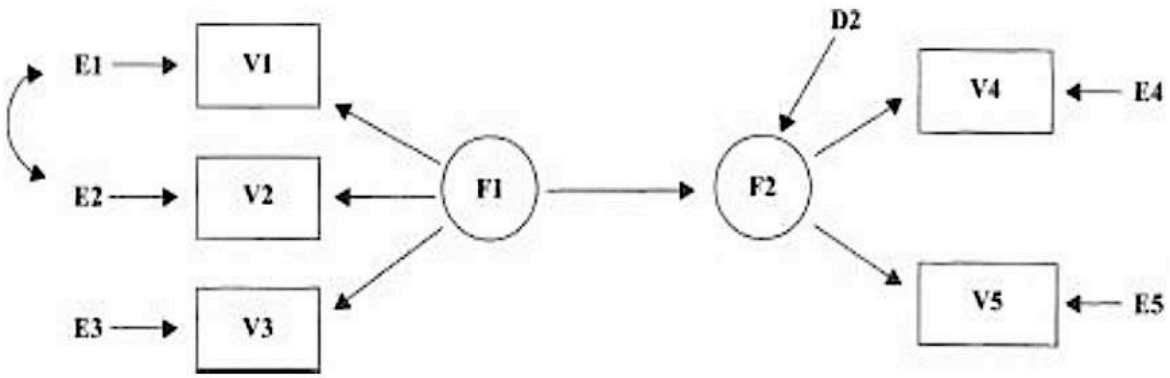
**c.** *Un residuo asociado con la predicción de cada factor, designado como **D***.

Los términos residuales a menudo se conocen como *disturbios* (*disturbance*), que es la terminología utilizada en el programa **EQS**. El término disturbio se usa en lugar de residual a lo largo del resto de este libro. Finalmente, para simplificar, todas las **E** y los **D** están numerados para corresponderse con las **V** y las **F** con las que están asociados, respectivamente.

# Diagrama de trayectoria

Para comprender completamente cómo funciona el sistema de símbolos y etiquetado, veamos la **Figura 7.3**.

**Figura 7.3. EQS. Modelo general de ecuación estructural**




Fuente: Jöreskog (1993); Hair (et al.,1999); Byrne 2006 con adaptación propia.

Como se observa, es muy similar a lo planteado desde la **Figura 7.1.** agregando los sólo, comentarios que llevan en el número que acompaña a cada variable etiquetada en el modelo, ya que:

**a.EQS** automáticamente numera cada variable observada de acuerdo con su ubicación de entrada de datos. Como tal, la primera variable en el conjunto de datos se designaría **V1**, el segundo **V2**, y así sucesivamente.

**b.**El valor numérico asociado con cada **error (E)** y el término de perturbación (**D**) son consistentes con su relación observada **(V**) y no observada (**F**) variables, respectivamente. Por lo tanto, aunque solo hay un término de perturbación en este modelo, se denomina **D2** en lugar de **D1.**

# Representación del sistema Bentler-Weeks

Como se discutió anteriormente, dado que los datos de muestra constituyen puntajes observados solamente, cada programa **SEM** requiere algunos medios para transponer estos puntajes en parámetros del modelo. Esto se logra a través de un modelo matemático que represente a todo el sistema de variables. En **EQS**, el modelo matemático deriva del trabajo de Bentler y Weeks (1979, 1980). Lo valioso de dicho modelo, es que todas las variables del modelo, pueden clasificarse ya sea como variables dependientes o independientes. Cualquier variable que tenga una flecha unidireccional dirigida a ella representa una variable dependiente; si no hay una flecha unidireccional dirigida a él, una variable se considera independiente. Como es habitual, las variables dependientes se explican en términos de otras variables en el modelo, mientras que las variables independientes sirven como variables explicativas.

No es tan habitual, sin embargo, la conceptualización de Bentler-Weeks (1979, 1980). sobre lo que es una variable dependiente o independiente. De hecho, la interpretación de este concepto es mucho más amplio de lo común. Según Bentler y Weeks (1979, 1980), cualquier variable que no sea una variable dependiente se considera automáticamente una variable independiente, sin importar de si se trata de una puntuación observada, un factor no observado o un término de perturbación. Por

ejemplo, en la **Figura 7.3**, las variables dependientes**,** son: **V1, V2, V3, V4, V5** y **F2**; las variables independientes, son: **E1, E2, E3, E4, E5, F1** y **D2**. Una variable dependiente, entonces, es cualquier variable que se puede expresar como una función de regresión estructural de otras variables Por lo tanto, para cada variable dependiente, esta función de regresión se puede resumir en la forma de una ecuación Al igual que con las ecuaciones basadas en las variables que se muestran en la **Figura 1.1**, cada función de regresión modelado en la **Figura 7.3** se puede traducir a ecuaciones específicas de **EQS**. Como se hará evidente en las aplicaciones ilustradas más adelante, estas ecuaciones sirven esencialmente para definir el modelo para el programa. los Las ecuaciones relacionadas son las siguientes:

$$F2 = F1 + D2$$
$$V1 = F1 + E1$$
$$V2 = F1 + E2$$
$$V3 = F1 + E3$$
$$V4 = F2 + E4$$
$$V5 = F2 + E5$$

Está claro que las flechas unidireccionales vinculan a los factores con las variables observadas y el **Factor 1** con el **Factor 2** representan coeficientes de regresión. Sin embargo, la explicación con respecto a la vinculación en términos de la perturbación o disturbios asociados a sus variables a través de las flechas de acceso unidireccional, pueden ser algo menos obvias**.** Aunque estas flechas también simbolizan los coeficientes de regresión, sus trayectorias están implícitas en la predicción de una variable a otra; así, se consideran conocidos y, por lo tanto, se fijan **a 1.0**. Por ejemplo, en el lenguaje simple regresión, la predicción de **V1** a partir de **F1** se puede escribir como:

$$V1 = b_n F1 + E1$$

Donde $b_n$ representa el peso beta desconocido asociado con el predictor **Fi** y **Ei** representa un error en esta predicción**.** Tenga en cuenta que no hay un peso beta asociado con el término de error, lo que indica que no se debe estimar. Por implicación entonces, el peso beta para $E_i$ se considera conocido y se fija arbitrariamente a 1 (la varianza de $E_1$ necesita ser fijado a **1** porque tanto el coeficiente de regresión como la varianza no pueden ser estimados simultáneamente). De manera similar, el la predicción de $F_2$ a partir de $F_1$ se puede escribir como $F_2 = b_{12}F_1 + D_2$, donde $D_2$ representa un error en la predicción, aunque esto implica la predicción de un factor de otro,(mientras que en la predicción anterior, la ecuación implicaba la predicción de una variable observada a partir de un factor (por lo tanto, la distinción entre la términos **E y D**). Finalmente, un **corolario** importante del modelo de Bentler-Weeks es que las varianzas de las variables dependientes o sus covarianzas con otras variables nunca son parámetros del modelo; más bien, siguen siendo explicado por esos parámetros. Por el contrario, las varianzas y covarianzas de las variables independientes son parámetros importantes que necesitan ser estimados.

# EQS. Uso del programa

Ahora que se comprende la notación básica de **EQS**, es hora de ver:

**a.** Cómo se pueden vincular estos símbolos juntos para formar declaraciones y párrafos del programa admisibles utilizados en la construcción de un archivo de entrada **SEM**, y

**b.** Cómo ejecutar estos archivos de entrada para que el modelo especificado pueda ser probado.

El objetivo principal aquí es presentar una visión multifacética y exhaustiva de las capacidades del programas; en consecuencia, este capítulo es necesariamente extenso. Sin embargo, después de trabajar, a través de este material, se espera que tenga una mejor apreciación para el programa banco de construcción de modelos y los recursos de prueba de modelo que le esperan al usar **EQS.** Comenzamos esta visita guiada de la programación de **EQS** inspeccionando primero todos los requisitos y la mayoría de los opcionales componentes del archivo de entrada. Para una visión más completa de cómo se combinan estos componentes, se revisará su aplicación con respecto a tres modelos diferentes (es decir, **CFA** de primer orden, segundo orden **CFA** y **SEM** completo). Se presenta el importante concepto de identificación de modelo (o estadístico). Posteriormente, se aborda la cuestión de cómo crear un archivo de entrada **EQS**. Aquí, se ofrece tres opciones:

-Manual;

-interactivo, usando la opción de comando **BUILD_EQS**, y

-Gráfico**,** usando la opción **DIAGRAMMER.**

Continuando con la siguiente etapa, preparamos la salida y análisis de datos

Finalmente, se presenta un resumen de posibles mensajes de error que pueden ocurrir en el uso de **EQS.**

# EQS. Componentes del archivo de entrada

Un archivo de entrada **EQS** describe el modelo bajo estudio. Se compone de varias declaraciones agrupadas dentro de párrafos según ciertas reglas. En esta subsección, trataremos sobre instrucciones y datos básicos que rigen la construcción de archivos de entrada **EQS,** luego se examinarán los componentes básicos de un archivo de entrada **EQS** y, finalmente, se revisará el métodos manual para formular y ejecutar archivos de entrada **EQS** relacionados con tres modelos **SEM** diferentes.

# EQS. Reglas básicas en la creación de archivos de entrada

Existen una serie de instrucciones básicas, las cuales resumimos como sigue:

## *Keywords. (Palabras clave)*

Como en la mayoría de los programas de computadora, **EQS** usa un sistema de palabras clave que el programa interpreta como comandos básicos**.** Las palabras clave principales en **EQS** son los nombres utilizados para identificar párrafos

particulares que comprende el archivo de entrada (por ejemplo, el párrafo /**SPECIFICATIONS**). Las palabras clave secundarias representan cada declaración dentro de un párrafo; funcionan como subcomandos. Como tales, se refieren a un aspecto particular de el párrafo, lo que permite al usuario elegir una de varias opciones relacionadas con este componente.
Por ejemplo, un aspecto del párrafo /**SPECIFICATIONS** involucra el **método de** estimación. De la opciones disponibles, el usuario puede desear emplear una estimación generalizada de mínimos cuadrados (*generalized least squares estimation*)**,** que se establecería como **ME = GLS**, donde **ME** es la forma abreviada de la palabra clave **METHOD**, un subcomando que indica el método de estimación que se utilizará, y **GLS** es la palabra clave de tres letras que indica que la estimación generalizada de mínimos cuadrados (*generalized least squares estimation*) es el método de estimación que se utilizará.
Dos reglas principales rigen el uso de las párrafos en las palabras clave:

**a.** *Siempre deben ir precedidas de una barra oblicua* (por ejemplo, **/SPECIFICATIONS**), y

**b.** *La línea en la que aparece la palabra clave no debe contener ninguna otra entrada información*. Un uso opcional del párrafo en las palabras clave es que pueden abreviarse a las tres primeras letras (por ejemplo, **/SPE**). Sin embargo, el usuario puede preferir más de tres letras porque da como resultado un *palabra clave significativa,* y esta elección es bastante aceptable. Por ejemplo, abreviatura **/SPECIFICATIONS a /SPEC** parece ser una mejor opción.

## Descriptive Statements

Toda la información que describe el modelo y los datos debe expresarse con *declaraciones específicas* escritas dentro del párrafo apropiado. Un punto y coma **(;)** debe separar cada instrucción. Aunque no es necesario, a menudo es útil para construir archivos de entrada de tal manera que los haga fáciles de leer. Una forma de hacerlo es comenzar *párrafos de palabras clave en la columna 1***,** con todas las demás declaraciones dentro del párrafo con *sangría* de algunos caracteres. Finalmente, todas las declaraciones de entrada deben especificarse *en no más de 80 columnas por línea*. En ocasiones, es posible que desee **insertar un comentario** recordatorio en el archivo de entrada con respecto a algún aspecto de los datos o análisis. **EQS** le permite incluir tales comentarios en cualquier línea del archivo de entrada por medio de un *signo de exclamación* **(!),** que debe preceder al comentario. El programa ignora todo el material a la derecha del símbolo de exclamación.

## File Editors

Para que se ejecute **EQS**, el archivo de entrada debe ser un archivo simple que no contenga caracteres de control ocultos**.** Dado que la mayoría de los procesadores de texto (por ejemplo, ***Word***) usan dichos símbolos en la formulación del texto, los caracteres invisibles primero deben eliminarse antes de que **EQS** pueda leerlos. Sin embargo, esta restricción es fácil direccionado al guardar el archivo de entrada en formato **ASCII**, que **EQS** puede leer.

## Data Input

Los datos en forma de una matriz de covarianza o correlación se pueden incrustar dentro del archivo de entrada o residir en un archivo externo. Los datos de insumo siempre deben residir en un archivo separado. Detalles específicos sobre la integración de los archivos de datos con el archivo de entrada **EQS** se abordan en la siguiente subsección.

## Basic Components of the EQS Input File

Como se discutió anteriormente, la estructura básica de un archivo de entrada **EQS** comprende una serie de declaraciones agrupadas dentro de párrafos, cada uno de los cuales se introduce por medio de una palabra clave precedida por una barra oblicua **(/).** Sin embargo, no todos los párrafos son necesarios. Así, serán revisadas las características más importantes relacionados con seis párrafos básicos aunque existen más que se refieren más bien, a aplicaciones particulares del modelo en diseño. Ver Manual **EQS 6** (Bentler, 2005) o la guía del usuario (Bentler Wu, 2002), ambas disponibles a través de la ventana de Ayuda (**Help**) del programa.

## /TITLE (Optional)

Aunque este párrafo es opcional (es decir, no se requiere para ejecutar un trabajo de **EQS**), no solo es altamente recomendable usarlo ya que incluye una generosa cantidad de información en el título. Es muy común que la documentación al inicio sencilla y fácil de interpretar, con el tiempo *pierda su sentido*. El uso libre de la información del título, es particularmente útil cuando se implementan varias ejecuciones **EQS** para un conjunto de datos determinado. **EQS** permite tantas líneas como desee en este título sección. Un ejemplo de un párrafo **/TITLE** es el siguiente:

```
/TITLE
CFA of BDI-French Version
Initial Model
```

Fuente: EQS 6. con datos y adaptación propia

## /SPECIFICATIONS (SPEC; Obligatorio)

Este párrafo define tanto los datos que se analizarán como el método de análisis que se utilizará. En particular, detalla:

**a.**Información relacionada con los datos (ubicación, forma de matriz);
**b.**La cantidad de casos y aportes variables; (c) el método de estimación deseado; y
**c.**Otra información que pueda ser necesaria para asegurar que **EQS**, corra

Hay tres factores importantes con respecto al párrafo **/SPEC**:

**1.**Aunque la información de entrada se puede colocar en cualquier orden, un punto y coma (**;**)debe delinear cada operando;
**2.**Información relacionada con el tamaño de muestra y el número de variables debe ser siempre provisto; y

**3.**Las palabras clave del subcomando se pueden abreviar para formato de tres y dos letras.

Ahora damos un vistazo general a algunos aspectos de este párrafo.

## *Data (DA)*

El uso de esta palabra clave se rige por dos condiciones:

1.Los datos que se analizarán existen en algún archivo externo, y

2.El entorno informático implica un sistema interactivo. Como tal, la instrucción se leería: **DA = 'C: \ EQS61 \ Files \ FRBDI.dat**;

Donde :

**C: EQS61 \ Files** representa la ubicación de los datos y

**FRBDI.dat** representa el nombre del archivo de datos, los cuales están encerrados dentro de citas. Como se explicó anteriormente, estos datos pueden estar en forma de **matriz** en bruto, de correlación o de covarianza.

## *Variables (VAR)*

El número indicado aquí debe representar el número total de variables en el conjunto de datos; no representa la cantidad de variables que se analizarán en una ejecución específica de **EQS.** El programa ordena automáticamente qué variables incluir en los análisis en su lectura posterior de los párrafos: **/EQUATIONS y /VARIANCES**.

## *Cases (CAS)*

Este término define el número de sujetos que comprenden los datos de muestra. El número debe representar todos los casos, independientemente que cualquiera de ellos pueda eliminarse más tarde. Tales eliminaciones, ocurren según lo especifique el usuario así como el procedimiento analítico de **EQS**, nunca altera realmente el conjunto de datos original, por lo tanto, la el número de casos siempre permanece igual.

Si los datos de entrada están en forma de una matriz o correlación o covarianza, es fundamental que el número de casos sea correcto porque **EQS** no tiene forma de verificar su precisión; un número incorrecto dará lugar a estadísticas incorrectas. Sin embargo, si los datos de entrada comprenden puntajes brutos (*raw scores*), este número no necesita ser específico; **EQS** registra automáticamente el número de casos en el archivo, que se usa en todos los cálculos. Cualquier discrepancia entre el número especificado y el número real de casos se observa en el salida.

## Method (ME)

El usuario **EQS**, puede elegir alguno de los 9 métodos de estimación:

- *Maximum Likelihood***.** (**ML**.Máxima verosimilitud),
- *Least Squares.* (**LS.** Mínimos cuadrados),
- *Generalized Least Squares.*(**GLS**. Mínimos cuadrados generalizados),
- *Elliptical LS.* (**ELS**.Elíptica **LS**),
- *Elliptical GLS.(***EGLS**. Elíptica **GLS**),
- *Elliptical Reweighted***.(ERLS**. Elíptica ponderada **LS**),

- *Heterogeneous Kurtosis GLS.* (**HGKLS.** Curtosis heterogénea **GLS**),
- *Heterogeneous Kurtosis RLS.* (**HKRLS.** Curtosis heterogénea **RLS**) y
- *Arbitrary distribution GLS.* (**AGLS).**

Si no se especifica algún método, el programa automáticamente usa la estimación de **ML**; en otras palabras, **ML** representa el método predeterminado. (Para detalles relacionados a estos métodos de estimación, ver el manual de Bentler, 2005).

Una característica especial del programa **EQS** es que también permite la especificación de varias estimaciones métodos en un único envío de trabajo (por ejemplo, **ME = ML, AGLS**;). Se puede especificar un máximo de dos métodos cuando **AGLS** está incluido y un máximo de tres métodos de lo contrario.

Los métodos *elíptico* y de *distribución arbitraria* siempre van precedidos automáticamente por sus contrapartidas teóricas normales.

Más específicamente, los métodos **ELS, EGLS y ERLS** están precedidos por los métodos teóricos normales de **LS, GLS, y ML**, respectivamente; **AGLS** está precedido por **LS**. No obstante, el usuario puede anular estos métodos predeterminados simplemente especificando otro método válido. (Para más detalles con respecto a la especificación y uso apropiado de estos métodos, ver Bentler, 2005).

Finalmente, una característica extremadamente valiosa exclusiva del programa **EQS** es la disponibilidad de estadística robusta (*robust statistics*) asociada con cualquier método de estimación seleccionado, excepto **AGLS**.

Por ejemplo, al especificar **ME = ML, ROBUST**, la salida proporciona una estadística *Chi-Cuadrada robusta* ($\chi^2$) llamada estadística escalada de *Satorra-Bentler* (S- Bx *; Satorra Bentler, 1988, 1994) y errores estándar robustos (Bentler Dijkstra, 1985), ambos de los cuales han sido corregidos para falta de normalidad en muestras grandes.

El **S-B$\chi^2$** se ha mostrado como el cálculo más aproximado a $\chi^2$ que la prueba usual estadística y los errores estándar robustos para ser corregidos en muestras grandes a pesar del hecho de que los supuestos de distribución con respecto a las variables sean incorrectos (Bentler, 2005). Aunque estas robustas estadísticas son computacionalmente exigentes, se ha demostrado que tienen un mejor rendimiento que las estadísticas no corregidas donde el supuesto de la normalidad no se sostiene y es mejor que la asintótica **GLS** sin distribución (***AGLS.****Asymptotic Distribution-Free*) en todas las muestras excepto en las más grandes (Chou, Bentler, Satorra, 1991; Hu, Bentler, Kano, 1992). Sin embargo, una advertencia importante sobre el uso de estadísticas robusta es que pueden ser calculadas solo a partir de datos sin procesar (*raw data*).

## *Analysis. (ANAL).*

Esta palabra clave describe el tipo de matriz que se analizará si se trata de algo más que una matriz de covarianza, que es la predeterminada. Si los análisis se basan en la matriz de covarianza, pero los datos se ingresaron como una matriz de correlación, las desviaciones estándar se deben agregar a la entrada. Esto se logra proporcionando un párrafo separado introducido con la palabra clave /**STANDARD DEVIATONS (o /STA)** y luego listar las desviaciones estándar en la línea siguiente,

dejando un espacio en blanco entre cada una; más de una línea puede usarse si es necesario. Se deben recordar tres instrucciones al usar el párrafo /STA:

**1.** Debe haber exactamente el mismo número de desviaciones estándar que variables en el que se lee la **matriz de correlación**,

**2.** Las desviaciones estándar deben estar en el mismo orden que las variables en el matriz de entrada, y,

**3.** No se usa punto y coma **(;)** para finalizar esta sección de entrada.

## *Matrix (MA).*

Esta palabra clave describe la forma de los datos de entrada, es decir, si es una correlación, covarianza o matriz de datos sin procesar (*raw data matrix*). Como se señaló anteriormente, siempre se supone que los datos brutos (*raw data*) residen en un archivo externo; las matrices de datos de covarianza o correlación se pueden incrustar en el archivo de entrada (es decir, como un archivo interno) o residen en un archivo separado (externo).

Cuando las matrices de covarianza (o correlación) residen como un archivo interno, los datos se especifican en un párrafo separado, etiquetado **/MATRIX (MAT).** Si, por otro lado, residen en un archivo externo, su ubicación se anota a través de la palabra clave **DATA (DA).** Una vez que **EQS** ha localizado el archivo de datos, necesita saber cómo leer los datos. Si los datos están en formato libre, que es el predeterminado, **EQS** puede leer el archivo siempre que cada elemento en la matriz esté separado por al menos un espacio en blanco. Si los datos no están en formato libre, esta información se proporciona por medio de una declaración **Fortran**, que debe ser precedida inmediatamente por la palabra clave **FORMAT (FO).**

Finalmente, si el usuario desea que el análisis no sea basado en la matriz de covarianza (que es el valor predeterminado), la matriz que se analizará se debe anotar utilizando la palabra clave **ANALYSIS (ANAL).** Ahora, intentemos todo esto y observemos dos párrafos de especificación que describen una matriz de datos de entrada. Vea los siguientes casos:

• **Caso 1.** La matriz de datos como un archivo interno

```
/SPECIFICATIONS
CASE=250; VAR=4; ME=ML; MA=COR;
/MATRIX
1.00
 .34 1.00
 .55 .27 1.00
 .48 .33 .63 1.00
/STANDARD DEVIATIONS
1.09 .59 .98 1.10
```

Fuente: EQS 6. con datos y adaptación propia

**Nota:**

• La entrada de datos nunca requiere un punto y coma (**;**). Por lo tanto, no hay **(;)** en los párrafos la **/MATRIX** y **/STANDARD DEVIATIONS**.

•Aunque los datos de entrada estén en forma de matriz de correlación, la especificación de las desviaciones estándar permite que los análisis se basen en la matriz de covarianza.

•**Caso 2.** La matriz de datos como un archivo externo:

```
/SPECIFICATIONS
CASE=250;VAR=4; ME=ML, ROBUST; MA=COV;
DA='C:\EQS61\Files\FRBDI.dat'; FO=(4F2.0);
```

Fuente: EQS 6. con datos y adaptación propia

Además de los subcomandos del párrafo **/SPEC** que se muestran aquí, otros son opcionales o usados con aplicaciones particulares.

## /LABELS (Opcional)

Este párrafo se puede usar para identificar los nombres de las variables observadas **(Vs)** y/o latentes **(Fs)** en el modelo. Las etiquetas pueden tener de 1 a 8 caracteres de longitud y están asignadas solo a las variables de tipo **V** y **F**. Las variables observadas deben numerarse de acuerdo con su posición en el conjunto de datos. Por lo tanto, la especificación que **V5 = MATH** indica que la quinta variable en la matriz de datos debe etiquetarse como **MATH.** Aunque *la numeración de variables latentes es arbitraria*, debe ser secuenciada lógicamente dentro del contexto del modelo. Porque **EQS** asigna automáticamente $V_1$ a la primera variable observada, **V2** a la segunda variable, y así en adelante, estas designaciones se utilizan como etiquetas si el usuario no proporciona nombres para las variables; del mismo modo a las variables latentes se les asignan automáticamente las etiquetas **F1**, **F2**, etc. Si el usuario proporciona etiquetas para solo algunas de las variables, estos nombres anulan las etiquetas predeterminadas (por ejemplo, **V1**).

## /EQUATIONS (EQU; Obligatorio)

Esta palabra clave señala información de especificación con respecto al modelo en estudio. Específicamente, **el** párrafo define cada trayectoria de regresión en el modelo. Por medio de una serie de declaraciones de ecuación, el párrafo **/EQU** especifica todos los vínculos entre las variables independientes y dependientes (y viceversa) e identifica aquellos parámetros a restringir (es decir, a cero, 1.0 o algún otro valor) y aquellos que se estiman libremente. Antes de completar esta parte del archivo de entrada, es fundamental construir un diagrama de trayectoria de su modelo en que :

**a.**Todas las variables observadas y latentes están claramente etiquetadas;
**b.**Todas las trayectoria de regresión estructural son especificadas;
**c.**Se especifican todos los términos de error y perturbación, junto con sus trayectorias de regresión relacionadas;
**d.**Todas las covarianzas hipotéticas (entre las variables independientes); y
**e.**Todos los parámetros a ser estimados (incluidas las varianzas de factores) se identifican por medio de un asterisco **(*)**. Una vez que obtenga esta visual

representación de su modelo, es fácil completar todas las ecuaciones en el párrafo **/EQU** simplemente leyendo fuera del diagrama de trayectoria.
Nota: Con **EQS** ahora hay un párrafo recientemente establecido: **/MODEL** que permite al usuario expresar las especificaciones del modelo de una manera más simplificada y manera concisa. En adición, las variables ***E*** y ***D*** son generadas automáticamente por el programa. Cuando se utiliza este párrafo, reemplaza a la palabra clave **/EQU**, también como las palabras clave **/VAR** y **/COV**.

# EQS. Cómo escribir ecuaciones

Como se señaló, la finalización de estas ecuaciones se lleva a cabo como sigue:
-Para cada variable dependiente (es decir, cualquier variable que tenga una flecha apuntando hacia ella), escriba una ecuación que resuma el impacto directo sobre ella de otras variables en el modelo. Por lo tanto, siempre habrá tantas ecuaciones en el párrafo **/EQUATIONS** como variables dependientes en el modelo probado. La variable dependiente siempre estará en el lado izquierdo de la ecuación, con todas las variables independientes (es decir, *explicativas*) que aparecen a la derecha (recordemos que en **EQS,** los términos de variables dependientes e independientes, se definen en el contexto del modelo de Bentler-Weeks). Finalmente, una vez Las ecuaciones se han formado, asegúrese de insertar asteriscos **(*)** junto a todos los parámetros que se estimarán. **EQS** proporciona estimaciones solo para los parámetros con asterisco **(*);** todos los demás se consideran fijos.
Se debe tomar en cuenta el uso de los valores de inicio, que se refieren al punto en el cual un programa comienza un proceso iterativo para establecer estimaciones de parámetros. Los usuarios pueden permitir que el programa suministre estos valores (**EQS** usa el valor predeterminados tales como *0.0 para covarianzas*) o proporcionar sus propios valores. Los valores de inicio proporcionados por el usuario representan una mejor estimación de cuál será el valor esperado de una estimación de parámetro en particular. Estos valores de mejor estimación son incluídos en la ecuación, actuando como modificadores de los parámetros a estimar; siempre precediendo al asterisco **(*)**. Aunque no es necesario especificar los valores de inicio para la mayoría de los trabajos de **EQS,** a menudo facilitan el proceso iterativo. cuando los modelos complejos están bajo estudio; tales modelos pueden generar problemas de no convergencia si el inicio los valores proporcionados por **EQS** son *inadecuados***.** Sin embargo, para el usuario de **EQS,** el problema de tratar de determinar los valores de inicio apropiados se han ido para siempre como resultado de una característica ingeniosa llamada la opción **RETEST**, que genera automáticamente estos valores de inicio y luego le permite al usuario editar el archivo de entrada en consecuencia. (Para una discusión más extensa de los valores de inicio, vea Bollen,1989). Algunos párrafos para examinar, son los casos:

• **Caso 1.** Ecuaciones sin valores de inicio:

```
/EQUATIONS
VI = FI + EI;
F3 = *F1 + *F2 + D3;
```

Fuente: EQS 6. con datos y adaptación propia

• **Caso 2.** Ecuaciones con valores de inicio:

```
/EQUATIONS
VI = .9 FI + EI;
F3 = .6 *F1 + -.2*F2 + D3;
```

Fuente: EQS 6. con datos y adaptación propia

**Nota:**

• El valor de inicio asociado con **F1** se fija en **.**9 por elección.

• Los valores de inicio asociados con **E1** y **D3** no tienen ningún valor especificado; estos parámetros son considerados conocidos y fijados a 1.00 por defecto del programa.

# EQS. /VARIANCES (VAR; Obligatorio)

Este párrafo especifica el estado de las varianzas relacionadas con las variables independientes en el modelo. Como tal, cada varianza debe identificarse ya sea como un parámetro fijo en el modelo, o como uno que se estima libremente. (Recuerde que las varianzas para las variables dependientes nunca se especifican, independientemente de si son fijas o libres).

Como en el párrafo **/EQUATIONS**, las varianzas a ser estimadas se identifican con un asterisco **(*)**.Veamos un par de casos de el párrafo **/VAR.**

• **Caso 1 . Varianzas sin valores de inicio**

```
/VARIANCES
F1, F2 = *;
F1 to F3 = *;
F4-F6 = *;
```

Fuente: EQS 6. con datos y adaptación propia

**Nota:**

• Las listas consecutivas de variables se pueden especificar usando ***to*** o un *guión* **(-).** Como tal, **F1 to F3** indica que se deben estimar las varianzas de **F1, F2** y **F3.** Del

mismo modo, **F4**, a través de **F6** especifica que las varianzas de **F4, F5 y F6** deben ser estimadas.

- Las etiquetas de variables (como se define en el párrafo **/LABEL**) se pueden usar en lugar de las etiquetas **V** o **F.** Sin embargo, *dado que las etiquetas con menos de 8 caracteres pueden causar un error, se recomienda que, en este caso, se use el guión **(-)** en lugar de la convención **to**.*

- **Caso 2. Varianzas con valores de inicio**

```
/VARIANCES
F1, F2 = .3*;
F1 to F3 = .4*;
F4-F6 = .6*;
```

Fuente: EQS 6. con datos y adaptación propia

# EQS. /COVARIANCES (COV; Opcional)

Por razones obvias, este párrafo es necesario solo cuando se especifican covarianzas en el modelo. Así, se usa para especificar tanto covarianzas fijas distintas de cero como libres entre las variables independientes. Sin embargo, cualquier variable involucrada en una covarianza también debe tener su varianza especificada en el párrafo **/VAR**; en consecuencia, las covarianzas tampoco se pueden especificar para las variables dependientes **(***si desea obtener dos variables dependientes covariantes, una opción es especificar su covarianza entre sus términos de perturbación porque estos los residuos son siempre variables independientes en un modelo***)**. Al especificar un covarianza, el par de variables se establece y se separa por una coma, como se ejemplifica a continuación:

```
/COVARIANCES
E1, E3 = *;
F1 to F3 = *;
F4-F6 = *;
```

Fuente: EQS 6. con datos y adaptación propia

**Nota:**

- El uso de ***to*** y **(-)** indica que las covarianzas relacionadas con todos los posibles pares de variables, *deben ser estimadas*. Por ejemplo, la especificación de **F1 to F3 = *;"** indica que los parámetros estimados son **F1, F2; F1, F3; F2, F3**.

El último párrafo de todos los archivos de entrada de **EQS** debe ser **/END**. Esta palabra clave marca la terminación del programa entrada.

# EQS.Casos de ejemplos de archivos de entrada

Para proporcionar una vista más completa de cómo las diversas secciones del archivo de entrada **EQS** se relacionan con el diagrama de trayectoria de un modelo en particular, ahora revisamos tres modelos simples aunque diferentes:

**a.***Un modelo **CFA** de primer orden* (**Figura 7.4**),
**b.***Un modelo **CFA** de segundo orden* (**Figura 7.5**) y
**c.***Un modelo **SEM** completo* (**Figura 7.6**).

Sin embargo, debido a limitaciones de espacio, solo se muestra el archivo de entrada completo para la **Figura 7.4**; los párrafos de **/SPECIFICATIONS** y **/LABELS** están ausentes para los dos modelos restantes. Aunque parte de los mismos archivos de entrada aparecen más adelante en el capítulo, se presentan en su forma original según lo estructurado por la opción **BUILD_EQS**. Por lo tanto, los archivos presentados en esta subsección proporcionan una idea de cómo hacer que el archivo sea más conciso. Para obtener el máximo beneficio de esta subsección, estudie cada figura mientras se refiere a su declaración de entrada respectiva.

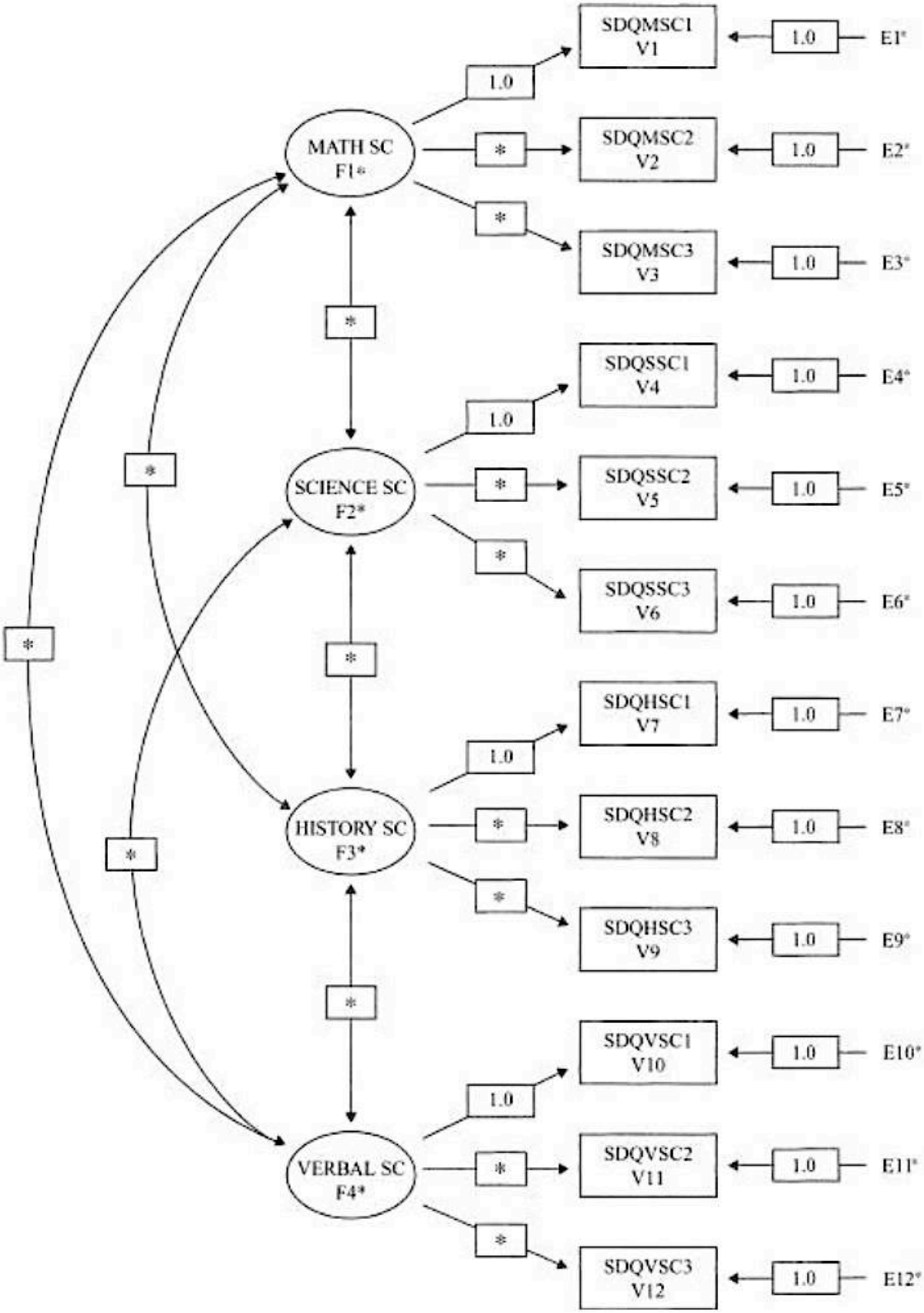
## *Modelo de primer orden*

Aunque este archivo de entrada parece ser bastante claro, tres características son dignas de hacer notar (Partiremos de observar y analizar la **Figura 7.4)**:

**1.**La convención ***to*** se ha utilizado en los párrafos **/VARIANCES** y **/COVARIANCES,** indicando que las varianzas de todas las variables independientes (**F1-F4**) deben estimarse libremente así como sus covarianzas.

**2.**Recuerde que una variable independiente puede tener o ya sea su trayectoria o su varianza estimada pero NO ambas. Por lo tanto, debido a que solo estamos interesados en las varianzas de los términos de error, tenga en cuenta que en el párrafo **/EQUATIONS**, sus ponderaciones beta relacionadas sean asignadas automáticamente por el programa a *valor de 1.0*. Aunque estos valores no se muestran en el archivo de entrada (vea **EI-E12**), son visibles en la **Figura 7.4**.

**Figura 7.4. Modelo CFA de primer orden con notación EQS y valores asignados de inicio**



```
/TITLE
Example 1st-order CFA Input File
Initial Model
/SPECIFICATIONS
DA=6C: \EQS61\FILES\CFASC.DAT; FO=(12F1 .0); CASE=361; VAR=12; ME=ML; MA=RAW;
/LABELS
V1=SDQMSC1;        V2=SDQMSC2;        V3=SDQMSC3;
V4=SDQSSC1;        V5=SDQSSC2;        V6-SDQSSC3;
V7=SDQHSC 1;       V8=SDQHSC3;        V9=SDQHSC3;
V10=SDQVSC1;       V11=SDQVSC2;       V12=SDQVSC3;
F1=MSC; F2=SSC; F3=HSC; F4=VSC;

/EQUATIONS
VI =  F1 + E1;
V2 = *F1 + E2;
V3 = *F1 + E3;
        V4 =  F2 + E4;
        V5 = *F2 + E5;
        V6 = *F2 + E6;
                V7 =  F3 + E7;
                V8 = *F3 + E8;
                V9 = *F3 + E9;
                        V10 =  F4 + E10;
                        V11 = *F4 + Ell;
                        V12 = *F4 + E12;
/VARIANCES
F1 to F4 = *;
E1 to E12 = *;
/COVARIANCES
F1 to F4 = *;
/END
```

Fuente: Byrne (2006)

**3**.Finalmente, tenga en cuenta que el primer indicador de medición para cada factor (**V1, V4, V7, V10**) se ha especificado como *fijo* (es decir, no hay asterisco junto a estos parámetros). Como se muestra en la **Figura 7.4**, estos parámetros son *limitados a un valor igual a 1.0*; **EQS** asigna automáticamente este valor a estos parámetros fijos, por lo tanto, el valor de *1.0 no aparece en el archivo de entrada.* Dos puntos deben hacerse a este respecto: **(a)** los valores asignados a estos parámetros *no necesitan ser 1.0*, aunque se puede asignar cualquier número a estos parámetros, *el valor de 1.0* ha sido típicamente la asignación de elección; y **(b)** el parámetro restringido no necesita ser limitado a la primera variable indicadora; cualquiera de un conjunto de *parámetros congenéricos* puede ser elegido (por ejemplo, como se muestra en la **Figura 7.4. SDQMSC1, SDQMSC2 y SDQMSC3** sirven todos como medidas de la variable **MATHSC,** por lo tanto, representan un *conjunto congenérico* de variables indicadoras).

Más allá estas notaciones técnicas, sin embargo, el punto más importante con respecto a las *cargas factoriales fijas* es que abordan el problema de la *identificación del modelo o model identification* (también denominado *identificación estadística o statitiscal identification*).
Lo fundamental a saber, es si un modelo al estar estadísticamente identificado comprender el número de estimaciones de parámetros en el modelo. Por lo tanto, una advertencia muy importante al trabajar con modelos de ecuaciones estructurales es contar siempre con el número de parámetros libremente estimados antes de ejecutar los análisis.
Como un prerrequisito a la discusión de la identificación del modelo, entonces, vamos a contar la cantidad de parámetros a estimar para el modelo de la **Figura 7.4.** en la que encontramos que hay 12 coeficientes de regresión (cargas factoriales) y 16 varianzas (12 varianzas de error, 4 varianzas de factor) y 6 covarianzas de factor. Los *1s* asignados a uno de cada conjunto de parámetros de trayectoria de regresión representan un *valor fijo de 1.00*; como tal, estos los parámetros No deben ser estimados. En total, entonces, hay 30 parámetros para ser estimados para el modelo **CFA** representado en la **Figura 7.4.** A continuación, se presenta una breve discusión del importante concepto de identificación de modelos.

## *El concepto de identificación de modelo.*

La identificación de modelos es un tema complejo que es difícil de explicar en términos no técnicos. Aunque una explicación completa del principio de identificación excede el alcance de este libro, no es crítico para la comprensión y el uso por parte del lector, lo que sigue es una breve explicación *no matemática* de la idea básica que subyace a este concepto. Esencialmente, la llamada *t- rule* se aborda, una de las varias pruebas asociadas con la identificación. En términos generales, el problema de la identificación se centra en si existe un conjunto único de parámetros consistente con los datos. Esta pregunta influye directamente en la transposición de la matriz de varianza-covarianza de variables observadas (los datos) en los parámetros estructurales del modelo en estudio. Si es determinada una solución única para los valores de los parámetros estructurales, el modelo se considera identificado. Como consecuencia, los parámetros se consideran estimables y el modelo, por lo tanto, comprobable. Por otra parte si un modelo no puede identificarse, indica que los parámetros están sujetos a arbitrariedad, lo que implica que diferentes valores de parámetros definen el mismo modelo. Si es el caso, el logro de la coherencia de las estimaciones para todos los parámetros no son posibles y, por lo tanto, el modelo no puede evaluarse empíricamente**.** Por medio de un ejemplo simple, el proceso sería conceptualmente similar a tratar de determinar valores únicos para **X e Y** cuando la única información disponible es que **X + Y = 15**. Generalizando este ejemplo a la estructura del análisis de covarianza, el problema de identificación del modelo se centra en la medida en que se puede inferir un conjunto único de valores para los parámetros desconocidos de una matriz de covarianza dada de las variables analizadas reproducidas por el modelo.
Los modelos estructurales pueden ser simplemente identificados (*just-identified*), sobreidentificados (*overidentified*) o no identificados (*underidentified*). Así:

**a.** *Un modelo identificado* (*just-identified*),es uno en el que hay una correspondencia uno a uno entre los datos y los parámetros estructurales. Es decir, el número de varianzas y covarianzas de datos es igual a la cantidad de parámetros a estimar. Sin embargo, a pesar de la capacidad del modelo de ofrecer una solución única para todos los parámetros, este modelo no es científicamente interesante porque no tiene grados de libertad y, por lo tanto, nunca puede ser rechazado.

**b.** *Un modelo sobreidentificado* (*overidentified*) es aquel en el que el número de parámetros a estimar es menor que el número de puntos de datos (es decir, varianzas y covarianzas de las variables observadas). Esta situación resulta en grados positivos de libertad que permiten el rechazo del modelo, por lo que se hace de uso científico. El objetivo en **SEM** es el especificar un modelo tal que cumpla con el *criterio de sobreidentificación.*

**c.** *Finalmente, un modelo no identificado* (*underidentified*) es uno en el que el número de parámetros a estimar excede el número de varianzas y covarianzas (es decir, puntos de datos). Como tal, el modelo contiene información insuficiente (datos de entrada) para lograr un solución determinada de estimación de parámetros; es decir, es posible un número infinito de soluciones para un modelo no identificado.

Al revisar el modelo **CFA** en la **Figura 7.4**, determinemos con cuántos puntos de datos hay que trabajar (es decir, ¿cuánta información tenemos con respecto a nuestros datos?). Como se señaló anteriormente, estos puntos de datos constituyen las varianzas y covarianzas de las variables observadas; con las variables p, hay p (p + 1) / 2 de tales elementos. Dado que hay 12 variables observadas, esto significa que hay 12 (12 + 1) / 2 = 78 puntos de datos.

Antes de esta discusión de identificación, determinamos un total de 30 parámetros desconocidos. Así, con 78 puntos de datos y 30 parámetros para estimar, tenemos un modelo sobreidentificado con 48 grados de libertad. Sin embargo, es importante observar que la especificación de un modelo sobreidentificado es necesaria, pero no condición suficiente para resolver el problema de identificación. De hecho, la imposición de restricciones en particular, los parámetros, a veces pueden ser beneficiosos para ayudar al investigador a alcanzar un modelo sobreidentificado

## *El problema de la escala adecuada*

Vinculado con el problema de la identificación está el requerimiento de que cada variable latente debe tener su escala determinada. Este requisito surge porque las variables latentes no se observan y, por lo tanto, no tienen escala métrica definida. Este requisito se puede lograr de una de dos maneras:

**a.** Mediante el enfoque con que se incia la presente obra de determinar: *dimensionalidad, confiabilidad y validez* al constructo para su prueba **EFA** y **CFA**

**b.** Ligar la especificación del modelo de medición, mediante el cual *la variable latente* no medida se mapea en su variable del indicador observado relacionado. Este requisito de escalamiento se cumple restringiendo a un valor distinto de cero (*típicamente 1.0*), un parámetro de carga de factorial en cada conjunto *congenérico* de cargas. Esta restricción vale para variables latentes tanto independientes como dependientes. Al revisar la **Figura 7.4,** esto significa que para una de los 3

trayectorias de regresión que conducen desde cada factor (**SC**) a un conjunto de indicadores observados, algunos valores fijos deben ser especificados; este parámetro fijo se denomina variable de referencia (*reference variable*).
Con respecto al modelo en **Figura 7.4**, por ejemplo, la escala se estableció limitando a un *valor de 1.0*, el primer parámetro en cada conjunto de variables observadas. Con una mejor idea de los aspectos importantes de la especificación de un modelo **CFA** en general, con la especificación que se utiliza del programa **EQS** en particular, y las nociones básicas asociadas con la identificación del modelo, es posible analizar los siguientes modelos de entrada.

## Modelo de segundo orden

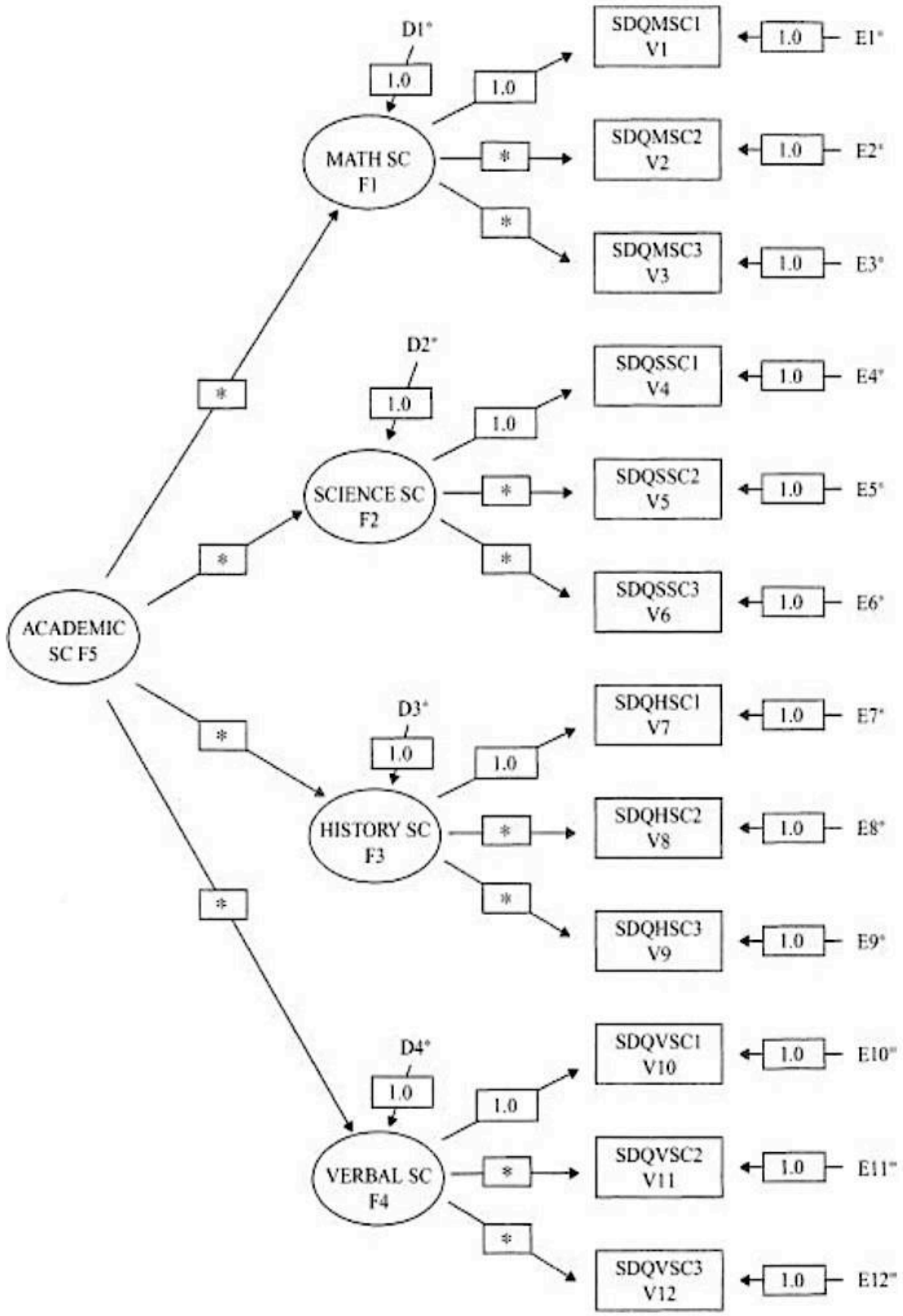
Este modelo, apoya la noción de que las variables **MATH SC** (matemáticas de auto-concepto o *math self-concept*), **SCIENCE SC** (ciencia de autoconcepto o *science self-concept*), **HISTORY SC** (historia de auto-concepto o *history self-concept*) y los **VERBAL SC** (autoconceptos verbales o *verbal self-concept)* son causados por el constructo de **ACADEMIC SC** (autoconcepto académico o *academic self-concept*) global de orden superior. Partiremos de observar y analizar la **Figura 7.5**
Por lo tanto, aunque la estructura de primer orden de este el modelo sigue siendo básicamente el mismo que en la **Figura 7.4**, se debe tener en cuenta cuatro características importantes:

**1.** Primero, en contraste con el ejemplo **CFA** de primer orden en el que todas las ecuaciones en el párrafo **/EQUATIONS** representaban las trayectorias de regresión entre todas las variables observadas y sus factores subyacentes relacionados, las ecuaciones especificadas para este el modelo de segundo orden también incluyen trayectorias de regresión entre factores. Específicamente, estos parámetros representan el impacto de **F5** (una variable independiente) en **F1, F2**, **F3 y F4**.

**2.** Dado que la estimación de todas las cargas factoriales de orden más alto son típicamente de interés en modelos de segundo orden, la varianza del único el factor de orden superior, **F5**, *se ha limitado a 1.0*, como se especifica en el párrafo **/ VARIANCES**. (Alternativamente, si la varianza de **F5** hubiera sido de interés, una de las cargas de orden superior tendría que estar *limitada a 1.0* (u otra distinta de cero).Esta la restricción, por supuesto, aborda el problema de la identificación del modelo discutido anteriormente.

**3.** En tercer lugar, dado que se supone que la variable **ACADEMIC SC** causa cada uno de los 4 factores de primer orden, **F1** a **F4** ahora representan variables dependientes en el sentido de Bentler-Weeks. Como tal**, MATH SC, SCIENCE SC, HISTORY SC** son las hipótesis predichas por el **ACADEMIC SC** académico, pero con algún grado de error, que es capturado por el término de perturbación asociado con cada uno de estos factores. Por lo tanto, en el párrafo **/VARIANCES**, tenga en cuenta también que las varianzas de las perturbaciones (las **D's**) se designan como estimadas libremente. Relacionado, sus caminos son restringido automáticamente a *1.0* por el programa, como se muestra en la **Figura 7.5**

**Figura 7.5. Modelo hipotético CFA de segundo orden con notación EQS y valores asignados de inicio**



```
/TITLE
Example 2nd-order CFA Input File
Initial Model
/EQUATIONS
V1 =  FI +E1;
V2 = *F1 + E2;
V3 = *F1 + E3;
        V4 = F2 + E4;
        *F2 + E5;
        *F2 + E6;
                V7 =  F3 + E7;
                V8 = *F3 + E8;
                V9 = *F3 + E9;
                        V10 =  F4 + E10;
                        V11 = *F4 + E11;
                        V12 = *F4 + E12 ;
FI = *F5 + D1;
  F2 = *F5 + D2;
    F3 = *F5 + D3;
      F4 = *F5 + D4;
/VARIANCES
F5 = 1.0;
D1 to D4 = *;
El toE12 = *;
/END
```

Fuente: Byrne (2006)

**4.** Finalmente**,** en los modelos de segundo orden, cualquier covarianza supuesta entre los factores de primer orden se explica por el (los) factor (es) de orden superior. En consecuencia, observe la ausencia de flechas de doble punta que unen los cuatro factores de primer orden en el diagrama de trayectoria así como la ausencia de un párrafo **/COVARIANCES** en el archivo de entrada.

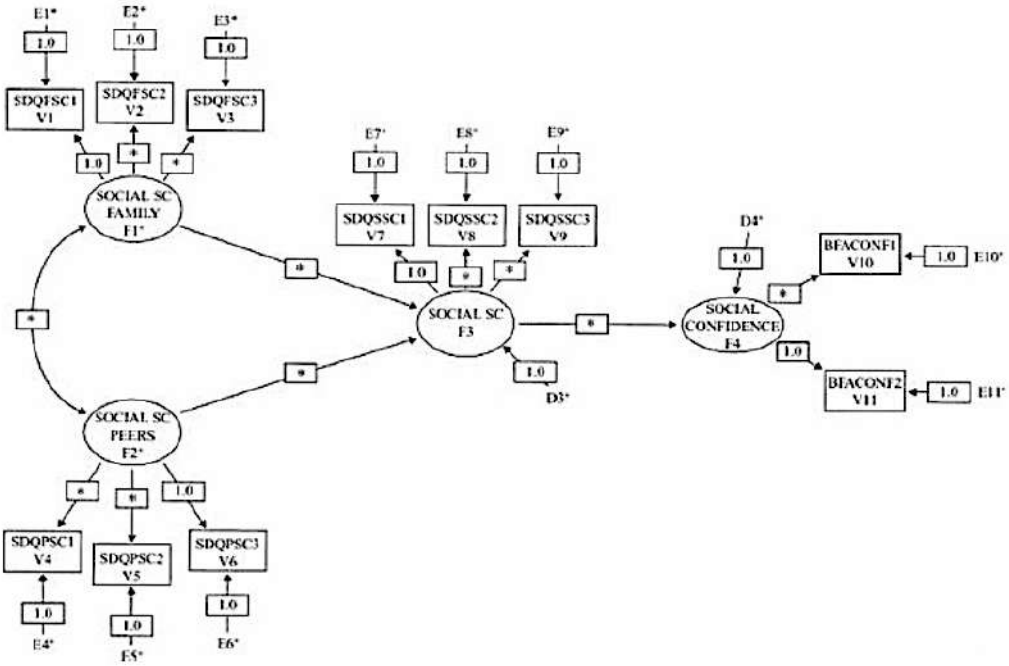
## *SEM. Modelo completo*

Al revisar tanto el diagrama de trayectoria así como el archivo de entrada para este modelo completo de **SEM**, tenga en cuenta tres particularidades (Partiremos de observar y analizar la **Figura 7.6)**:

**1.** De los cuatro factores que componen este modelo, solo **F1 y F2** son variables independientes (en el sentido de Bentler-Weeks); todos los demás son variables dependientes en el modelo. En consecuencia, solo las varianzas para **F1 y F2**, así como sus covarianzas, se pueden estimar como se especifica en el archivo de entrada.

**2.** En segundo lugar, como con el *modelo de segundo orden* que se muestra en la **Figura 7.5**, una vez más, las ecuaciones de regresión involucran conjuntos de factores. En el presente caso, estas ecuaciones se especifican solo para **F3** y **F4** porque cada una se explica por otros factores en el modelo.

**Figura 7.6. Modelo hipotético SEM con notación EQS y valores asignados de inicio**




```
/TITLE
Example Full SEM Input File
Initial Model
/EQUATIONS
V1 =  F1 +E1;
V2 = *F1 + E2;
V3 = *F1 + E3;
        V4 =  F2 + E4;
        *F2 + E5;
        *F2 + E6;
                V7 =  F3 + E7;
                V8 = *F3 + E8;
                V9 = *F3 + E9;
                        V10 =  F4 + E10;
                        V11 = *F4 + E11;
F3 = *F1 + *F2 + D3;
F4 = *F3 + D4;
/VARIANCES
F1 = *; F2 = *;
D3 = *;D4 = *;
El to E11 = *;
/COVARIANCES
F1, F2 = *;
/END
```

Fuente: Byrne (2006)

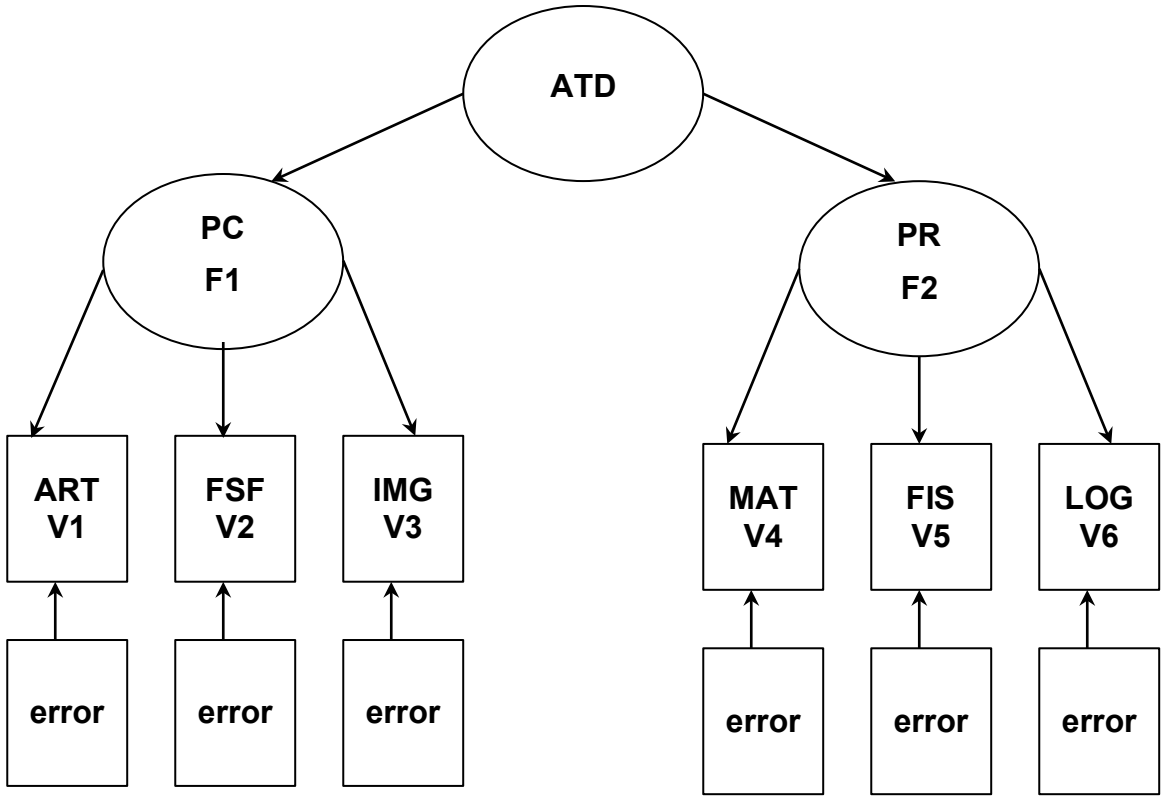
**3.** En tercer lugar, dado el estado de la variable dependiente de **F3** y **F4** en el modelo, tenga en cuenta que la varianza de ninguno de ellos está especificado en el párrafo **/VARIANCES;** relacionado a esto, tenga en cuenta la especificación de las varianzas asociadas con sus respectivas perturbaciones, **D3 y D4**, respectivamente. Finalmente, aunque **F3** parece funcionar tanto como una variable

independiente así como una dependiente, esto no es así. Una vez que una variable se define como variable dependiente en un modelo, mantiene esa designación en todos los análisis relacionados con la hipótesis modelo.

# Modelo de primer orden. Análisis de caso de estudio 1.

Recordar que el **CFA** es un método estadístico multivariable cuyo propósito principal es definir la estructura subyacente en una matriz de datos.
Su objetivo principal es seleccionar factores que expliquen las interrelaciones entre las variables. Las variables originales juegan el papel de variables dependientes que se explican por factores comunes y únicos, que no son observables.
En conclusión: sirve para Medir Modelos Teóricos**.** Ver ejemplo del modelo del diagrama de taryectoria de la **Figura 7.7.**

**Figura 7.7. Diagrama de trayectoria de la asertividad en la toma de decisiones (ATD)**




Fuente: propia.

## *Problema*

Suponga que en una investigación, se han recogido las notas de **275** ejecutivos de la industria del software son calificados en 6 áreas: Arte (**ART**); Filosofía (**FSF**);

Imaginación (**IMG**); Matemáticas (**MAT**); Física (**FIS**); Lógica (**LOG**). El investigador no posee ninguna hipótesis al respecto de si la asertividad en la toma de decisiones (**ATD**) es unidimensional o bidimensional (pensamiento creativo y pensamiento racional). Suponiendo que las notas miden la (**ATD**), tiene la siguiente hipótesis:
*La **ATD** de los ejecutivos de la industria del software es explicada mediante las notas del **ART**; **FSF** e **IMG** que componen al **PC**, así como las notas de **MAT; FIS** y **LOG** que componen al **PR.***
Dentro de la encuesta de captura de datos, se tienen entre otros: ventas, número de empleados (ambos como indicadores de crecimiento), años de funcionamiento en la empresa y en el puesto de dirección (se analizan empresas jóvenes y/o maduras de 10 a más años, así como sus ejecutivos), si es o no familiar (para distinguir las pymes de las que no), género (estudios de género pertinentes), edad de los ejecutivos *(¿gerentes maduros menos emprendedores a los jóvenes?)*, grado de estudios del directivo, entre otros relevantes. Aunque el modelo es general, dichas preguntas de segmentación ya aportan datos importantes, incluso por cada una de los factores por separado.Con esto, es factible utilizar diversos factores parciales *parcelar* en modelos lineales, es decir, sólo con dos factores. Siguiendo la fórmula :

**Fórmula 7.1. Ecuación para determinar elementos de una matriz**

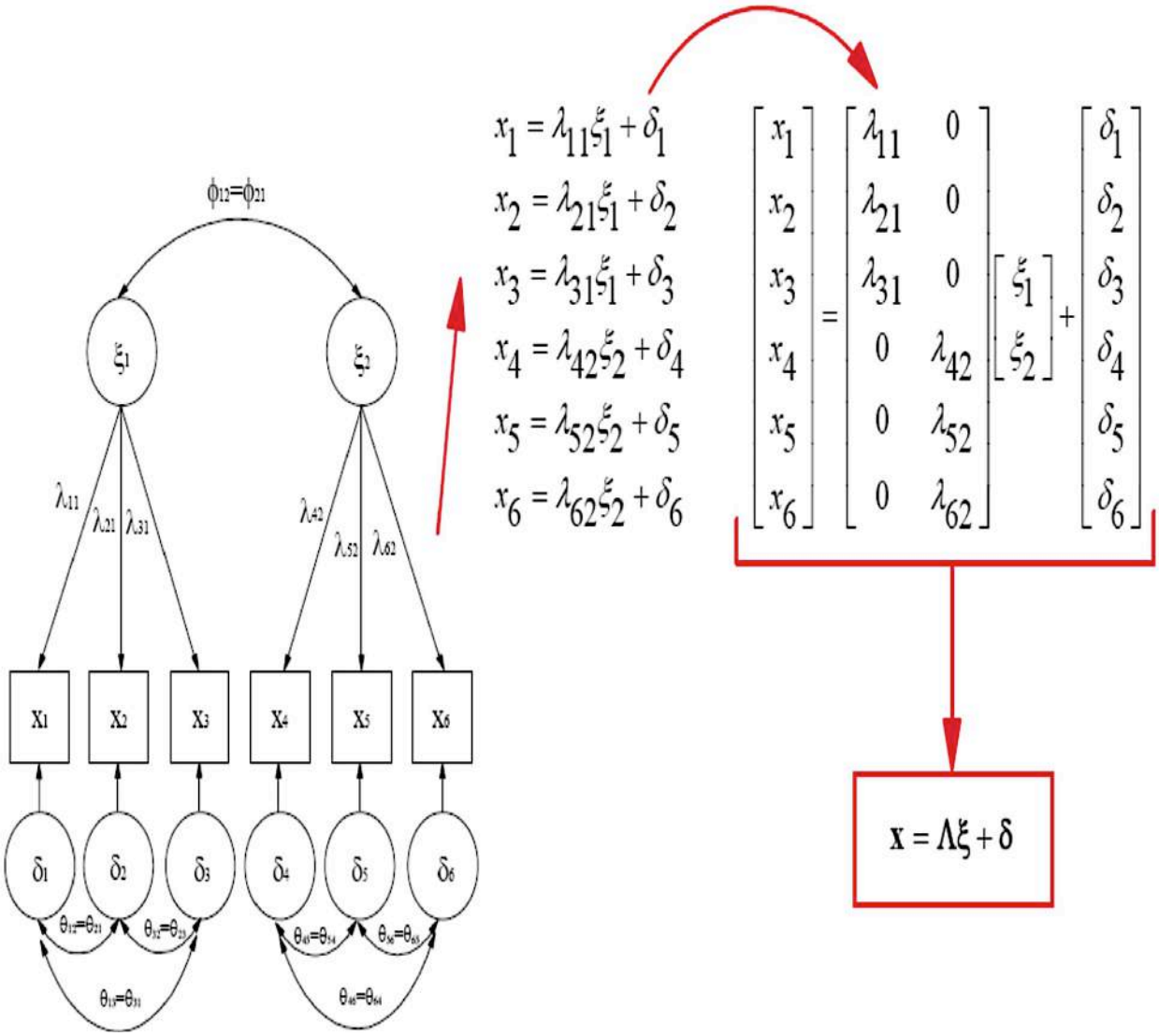
$$N = \frac{n(n-1)}{2}$$

Donde :
**n=** Número de factores
**N=** cantidad de relaciones lineales
Eventualmente, sería posible hacer artículos que describan de constructos lineales esto es,*parcelar*, del modelo para explicar cualquier fenómeno relacionado, hasta lograr la integración de todo el modelo completo sugerido. Lo importante es el modelo, no tanto los datos. (Es posible conservar bases de datos con más de 4 años de antigüedad aún útiles para la investigación). Cabe señalar, *que en **CFA** NO se tienen hipótesis (modelo de medición) ya que se requiere ajustar (confiabilidad y validez) el modelo*. Hecho lo anterior, ya es es posible determinar las hipótesis y probarlas con la parte de modelo estructural. Al momento, es más práctico trabajar de **CFA** es con matrices de covarianza que de correlación.

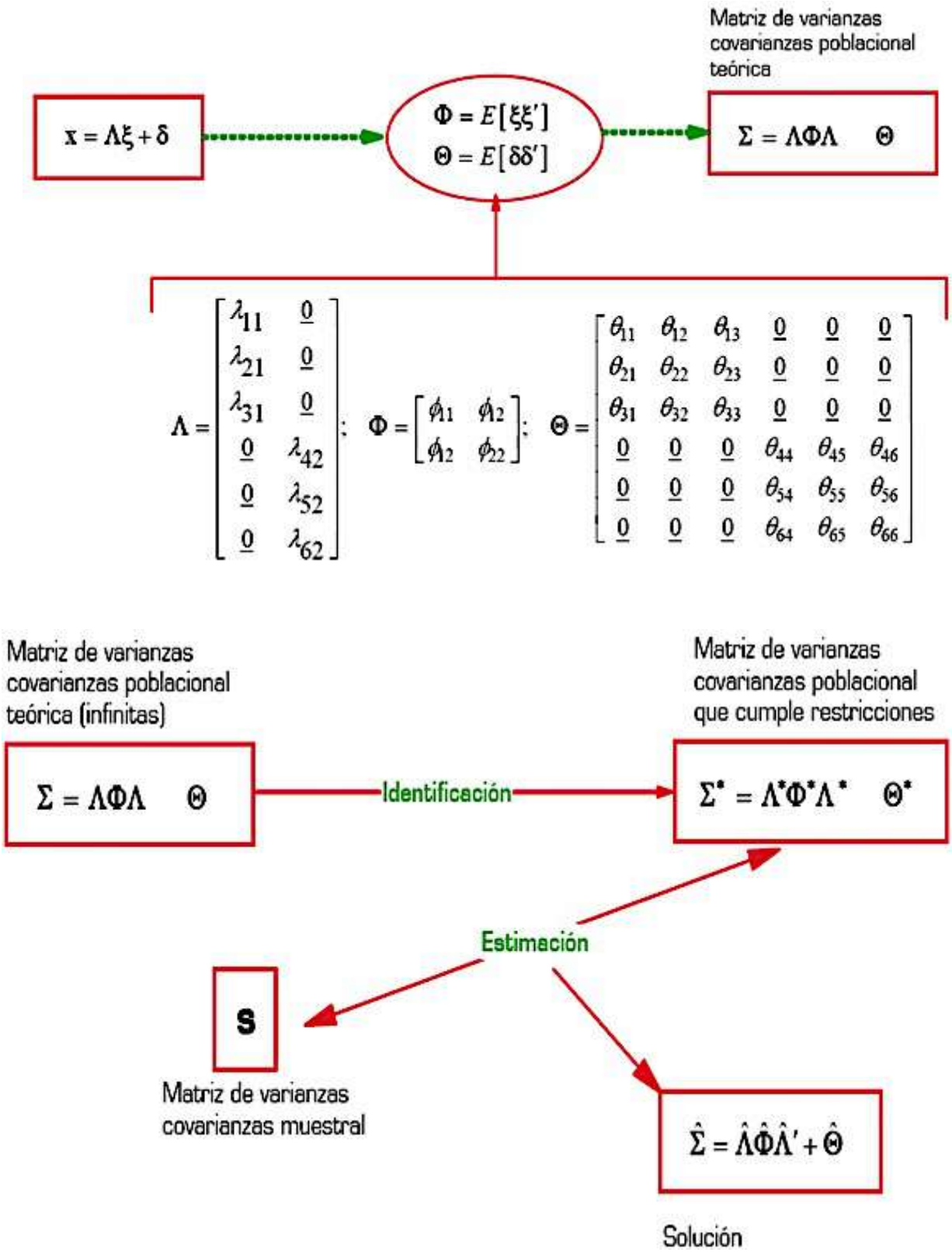
## Formulación de matriz

Aunque está más allá de los alcances de esta obra la explicación detallada, una formulación matemática del **CFA** se muestra en la **Figura 7.8**

**Figura 7.8. CFA formulación de matriz**




Fuente: Jöreskog (1993); Hair (et al.,1999); Byrne 2006 con adaptación propia.




Fuente: Jöreskog (1993); Hair (et al.,1999); Byrne 2006 con adaptación propia.

## Condiciones de aplicabilidad

A este momento se deben de confirmar las condiciones de aplicabilidad, tales como::

- Las variables manifiestas (los ítems de las distintas escalas) deben ser variables métricas.
- Debe existir la Normalidad, Homoscedasticidad y Linealidad.
- Las relaciones entre las variables deben ser lineales y aditivas.
- No debe existir problema de multicolinealidad (asegurar variabilidad, evitar resultados sesgados).
- El modelo debe ser sobreidentificado.
- Debe existir un número mínimo de observaciones (Hair et al., 1999 = 100 y Bagozzi y Yi, 1988 =200), o bien 5 observaciones por cada parámetro.
- Deben existir al menos tres indicadores (variables manifiestas) por cada variable latente (Factor o Dimensión).
- Pueden existir un máximo de 20 a 30 indicadores como ideal.

Por lo que una representación matemática del **CFA** bajo análisis, se muestra en la **Figura 7.9**

Donde:

-Variables observadas o manifiestas o indicadores, son **V1** a **V6: ART, FSF,IMG, MAT,FIS,LOG** (representadas dentro de un cuadrado).

-Variables latentes o no observables o subyacentes, como:

Factores comunes: (compartidos por más de una variable observada): Pensamiento Creativo (**PC**) y Pensamiento Racional (**PR**).

Factores específicos o errores: aleatorios que pueden haberse producido en la medida de una variable observada (representadas con un cuadrado).

-Covarianzas entre las variables latentes: representadas con flecha de doble punta.

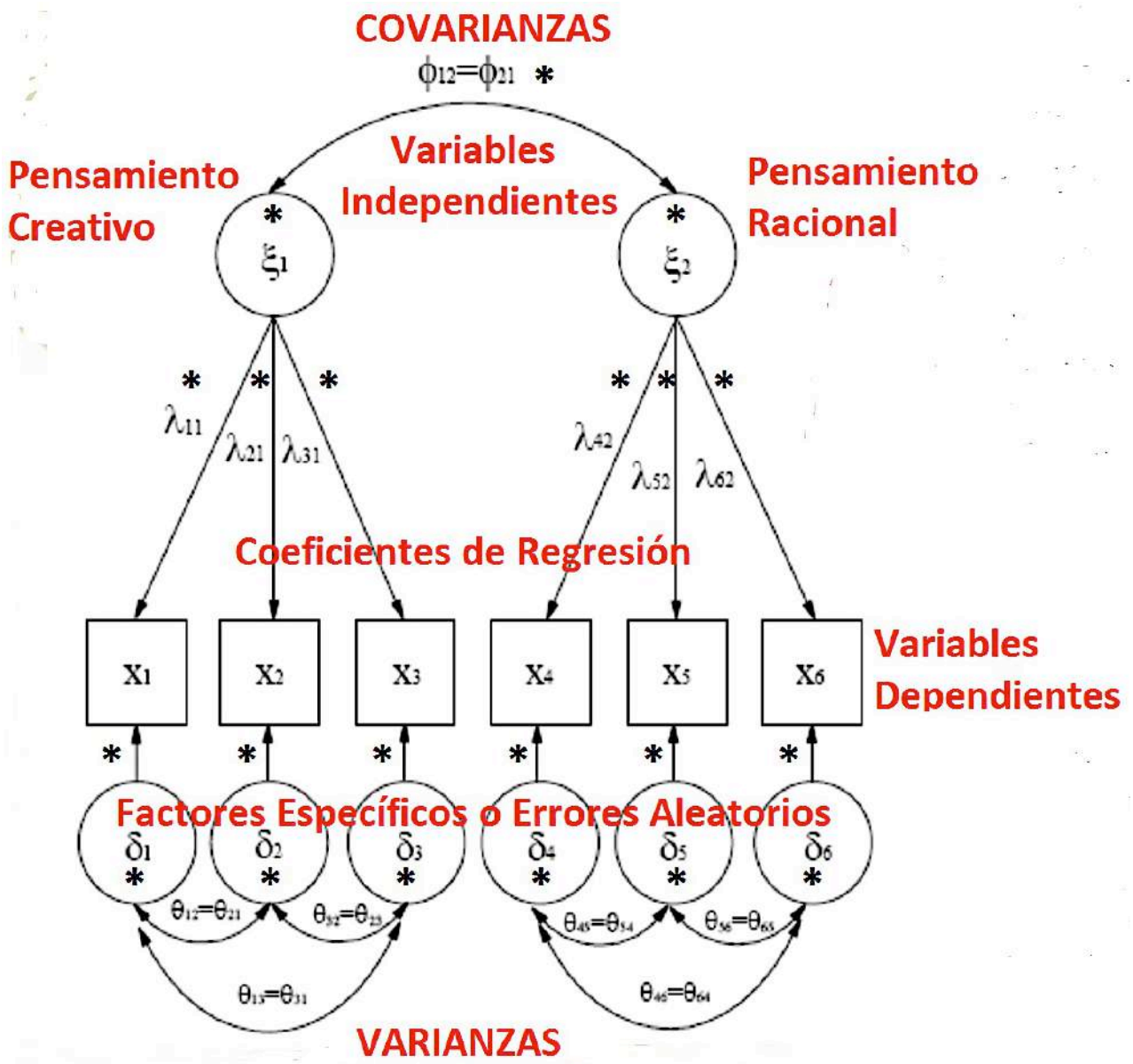
-Coeficientes de regresión, trayectoria, carga factorial: representadas con flecha de una sola punta.

La realización de un **CFA,** sugiere realizar los siguientes pasos (ver **Figura 7.9**):

1. Especifique un término de error por cada variable dependiente.
2. Sólo las variables independientes, pueden tener covarianza (y/o varianza. Punto 4).
3. Las variables independientes no tienen término de error.
4. Se deben estimar las varianzas para cada variable independiente del modelo.
5. Se deben estimar las covarianzas para cada par de variables independientes.
   **Nota:** valores >.8 de covarianza No son deseables ya que indicaría que las factores tienden a medir lo mismo; sin embargo <.1 No son deseables ya que indicaría que los factores serían interdependientes, porque juntas deben medir un mismo constructo.
6. No se estiman las covarianzas de los términos de error (éstos son también variables independientes), dado que lo interesante es la covarianza entre factores, no entre errores.
7. Hay que estimar los coeficientes de regresión.

El simbolo de asterisco (*) es para indicar que el software realice la estimación del valor de la relación.

**Figura 7.9. CFA aproximación matemática del diagrama de trayectoria del estudio de caso**




Fuente: Jöreskog (1993); Hair (et al.,1999); Byrne 2006 con adaptación propia.

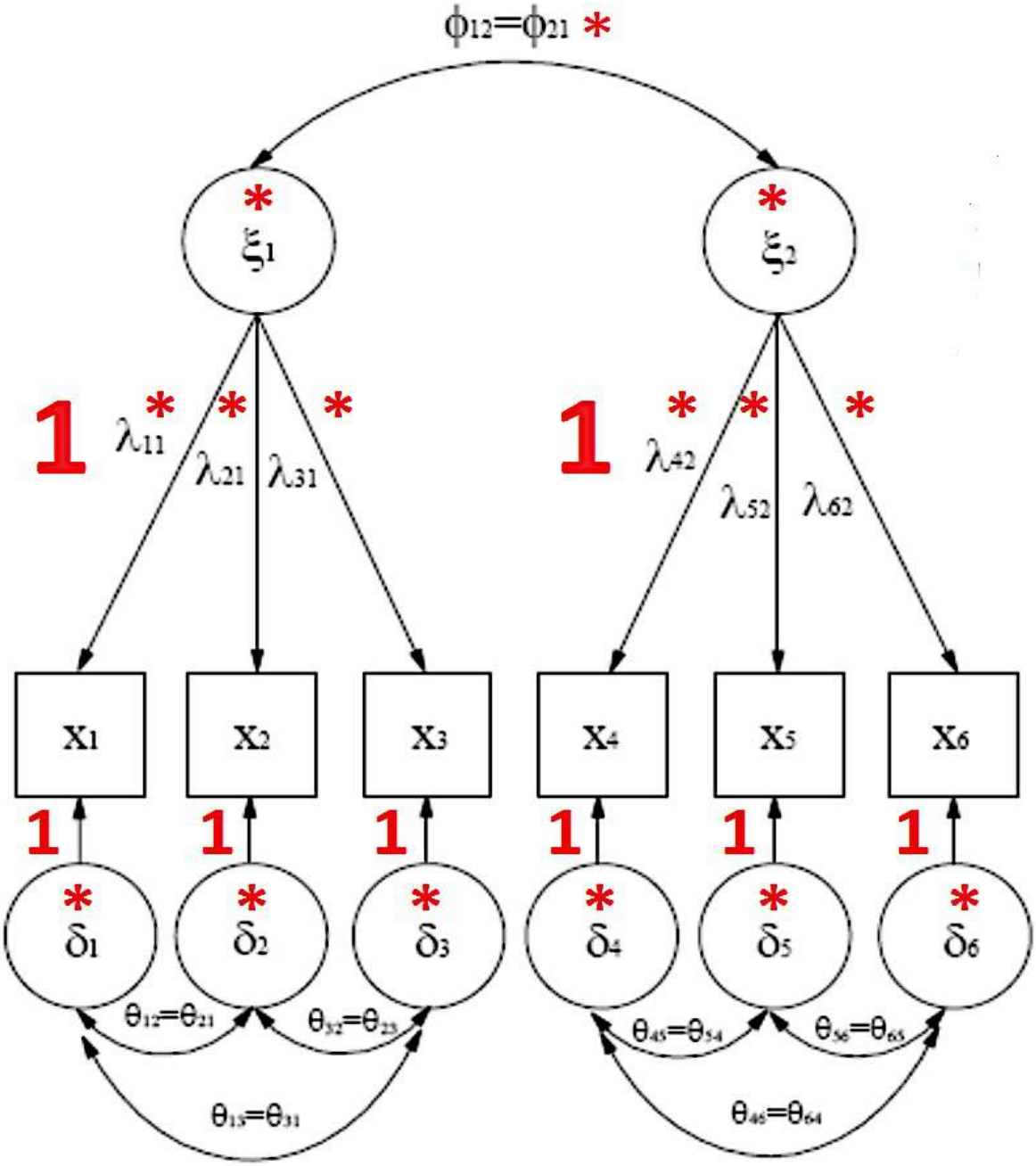
## *Identificación*

Se plantea la pregunta: ¿pueden los parámetros del modelo ser determinados de manera única?. Se debe establecer una escala para los factores comunes (esto es, evitar indeterminación entre la varianza y las cargas factoriales), mediante:

- Fijar la carga factorial a 1 (al menos 1 variable de cada factor debe tener varianza 1),
- Fijar la varianza a 1 **(**1 a todos los errores…No ambas**)**
- Fijar arbitrariamente el coeficiente de regresión del término de error a 1

- Comparar el número de datos conel número de parámetros que han de estimarse
- Datos varianzas-covarianzas muestrales: **q(q+1)/2**
- Parámetros a estimar (Vea la **Figura 7.10**)
  - *Varianzas de los factores comunes
  - *Coeficientes de regresión entre valores observados y factores comunes
  - *Varianzas de los errores

**Figura 7.10. Identificación**




Fuente: Jöreskog (1993); Hair (et al.,1999); Byrne 2006 con adaptación propia.

## Estimación

Se puede realizar mediante la consideración de los métodos discutidos en el apartado *Method (ME)*:

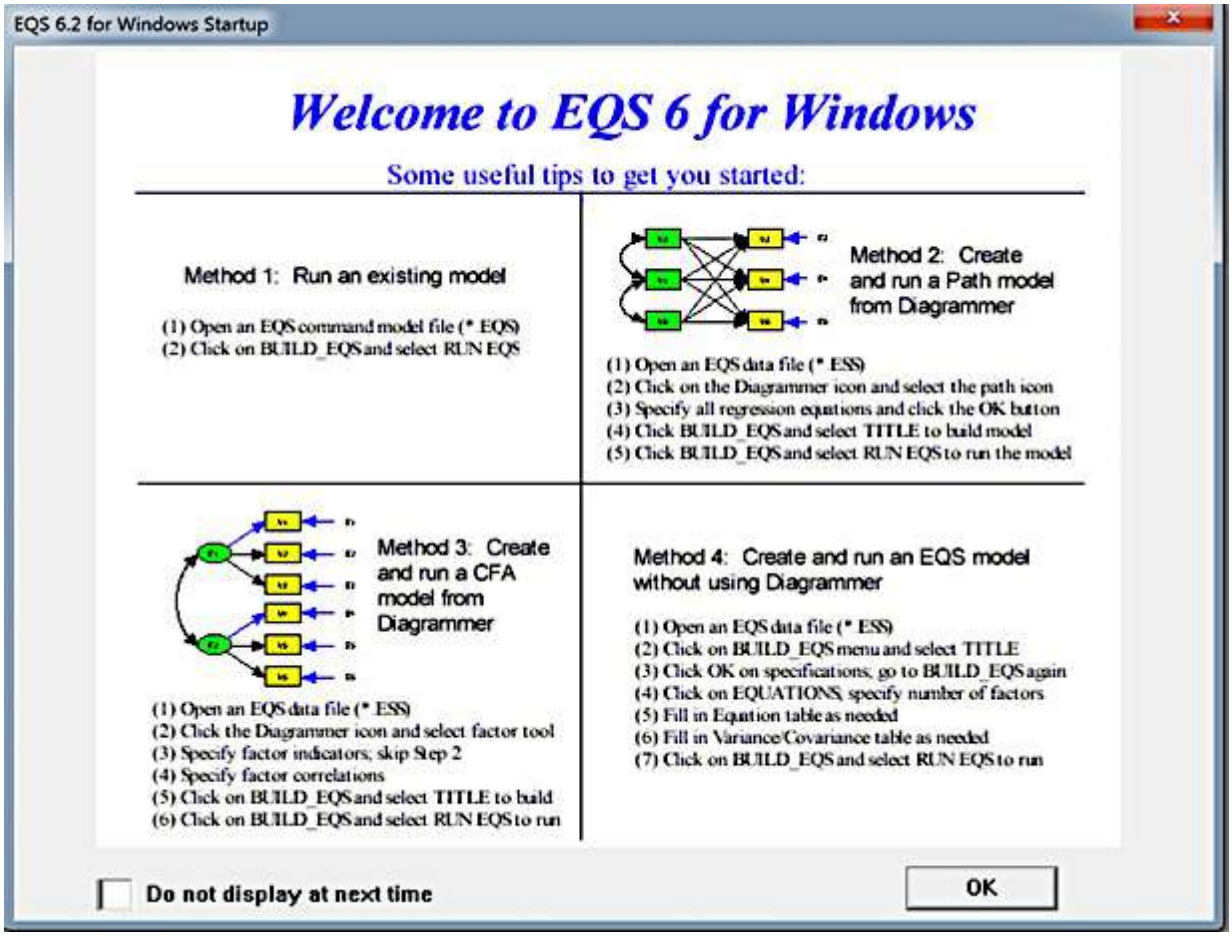
- Mínimos cuadrados no ponderados.
- Mínimos cuadrados generalizados.
- Teoría de la distribución asintótica
- Con libre distribución asintótica.
- Máxima verosimilitud (el más utilizado).

De este método, de acuerdo a Ullman (1996), se recomienda:

- **ML** y **GLS** siempre que sea lograble la normalidad e independencia.
- Si no se cumple la normalidad, **ML** y **GLS** sólo deberá realizarse para tamaños muestrales muy grandes (mayores a 2500)
- Si no se cumple la normalidad **USL**, aunque no existan contrastes estadísticos y los estimadores dependen de la escala de medida de las variables observadas

**Nota**: Bentler en su versión de **EQS 6** estandariza los datos para conservar la normalidad. Ver **Figura 7.11**

**Figura 7.11**. **Presentación de pantalla EQS 6**




Fuente: EQS 6

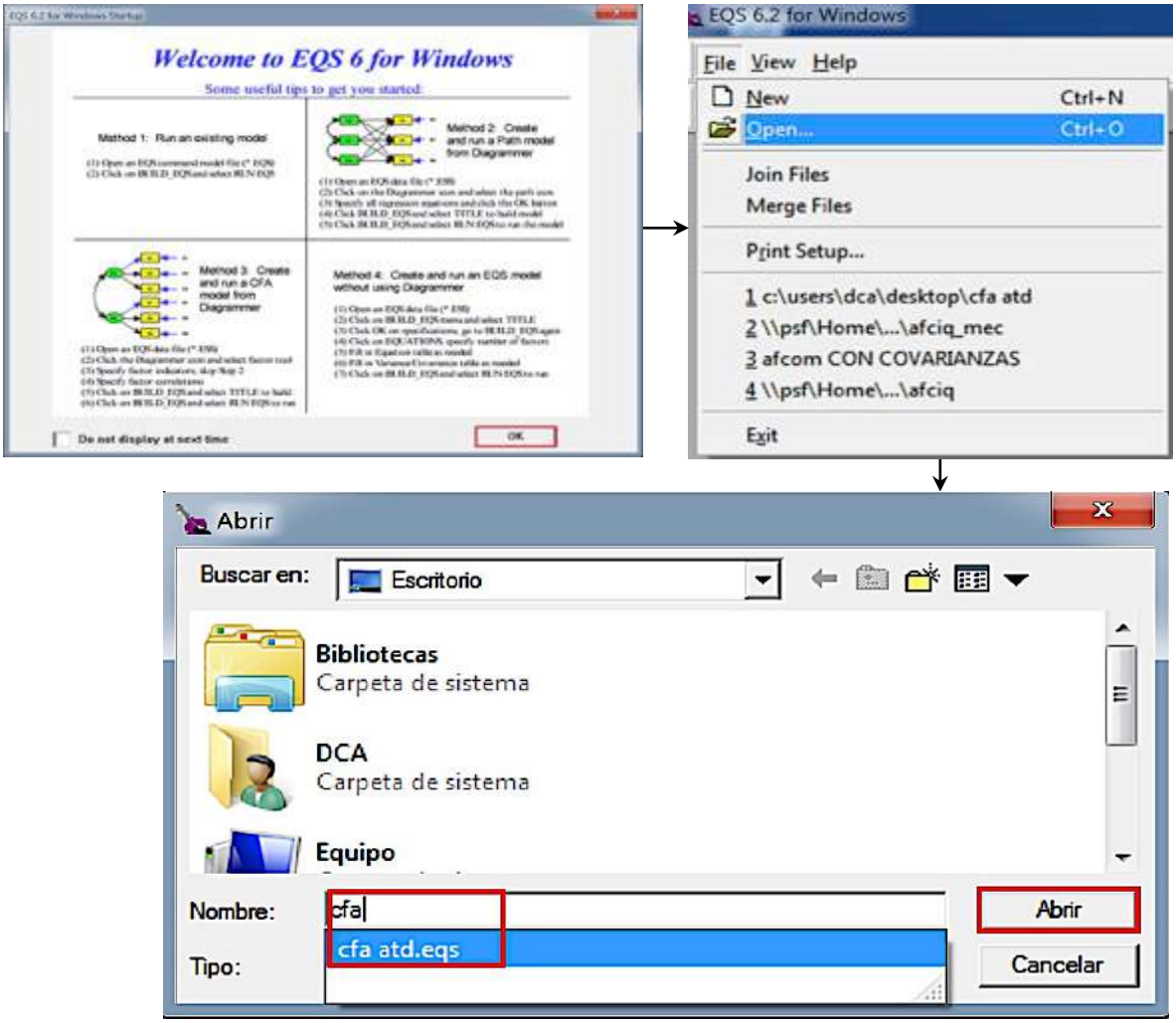
## Entrada de datos al programa

A continuación y a fin de entender el **CFA** aplicando **EQS**, realizaremos un análisis al programa nombrado como CFA ATD.eqs, y se documenta en el título como : *CFA ATD . CFA EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES* mediante la siguiente serie de comandos:

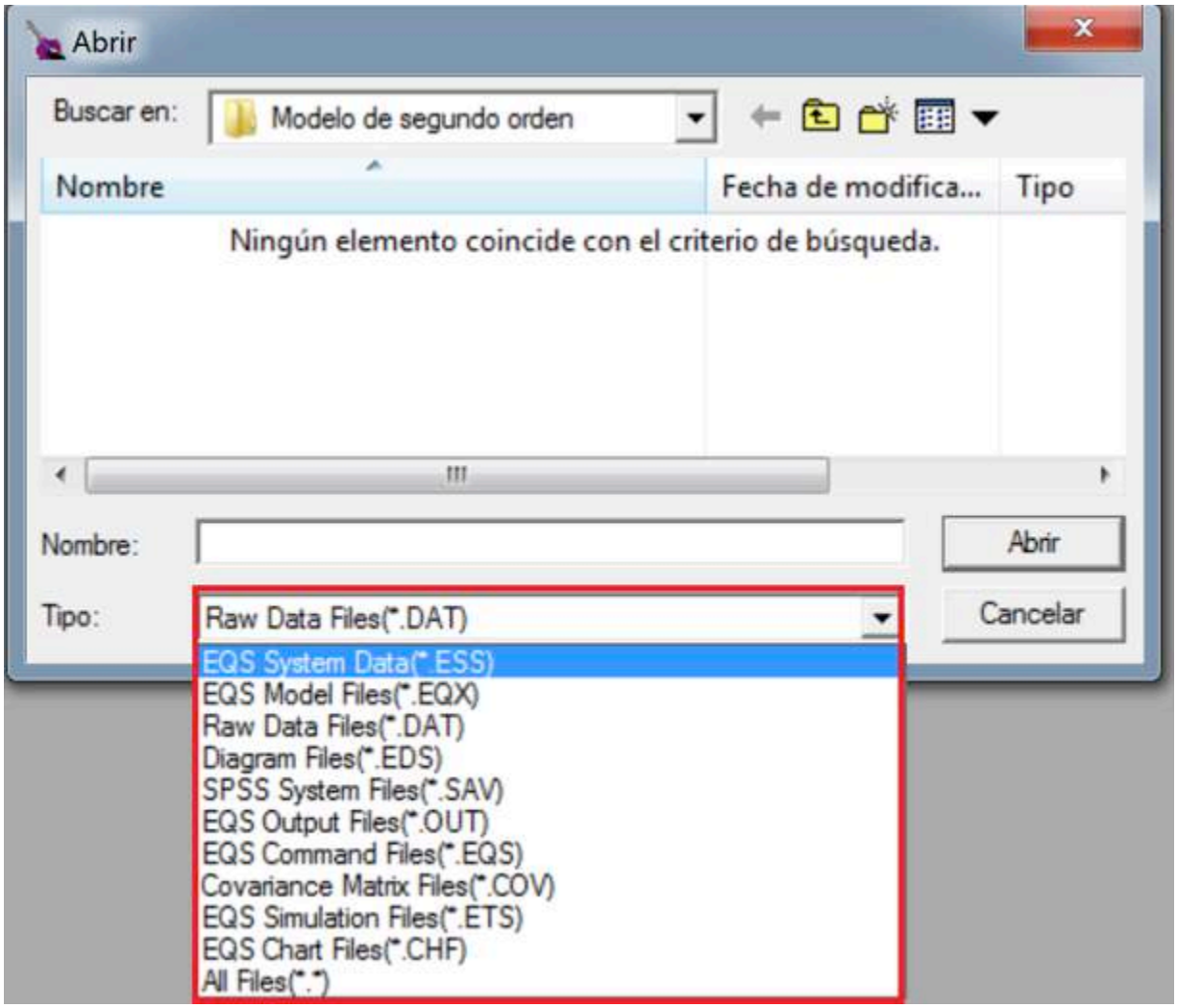
**Icono EQS 6 for windows->OK->File->Open->Nombre: cfa atd.eqs->Tipo: EQS System Data (*ESS)->Abrir.**

Observe que el programa **EQS**, ofrece una serie de formatos de archivos con los cuales opera, siendo los más utilizados los **EQS Command Files. (*.EQS)**, con los que se configura el archivo de inicio de programa; **EQS Output Files (*.OUT)**, que genera el archivo de datos de salida para análisis; **SPSS System files (*.SAV),** que permite visualizar el contenido delas bases de datos de **SPSS**; **EQS System Data (*.ESS),** utilizado para ingresar datos externos, entre otros. Ver **Figura 7.12.**

**Figura 7.12. EQS.Secuencia de comandos para apertura de programa de entrada de datos.Análisis CFA.**



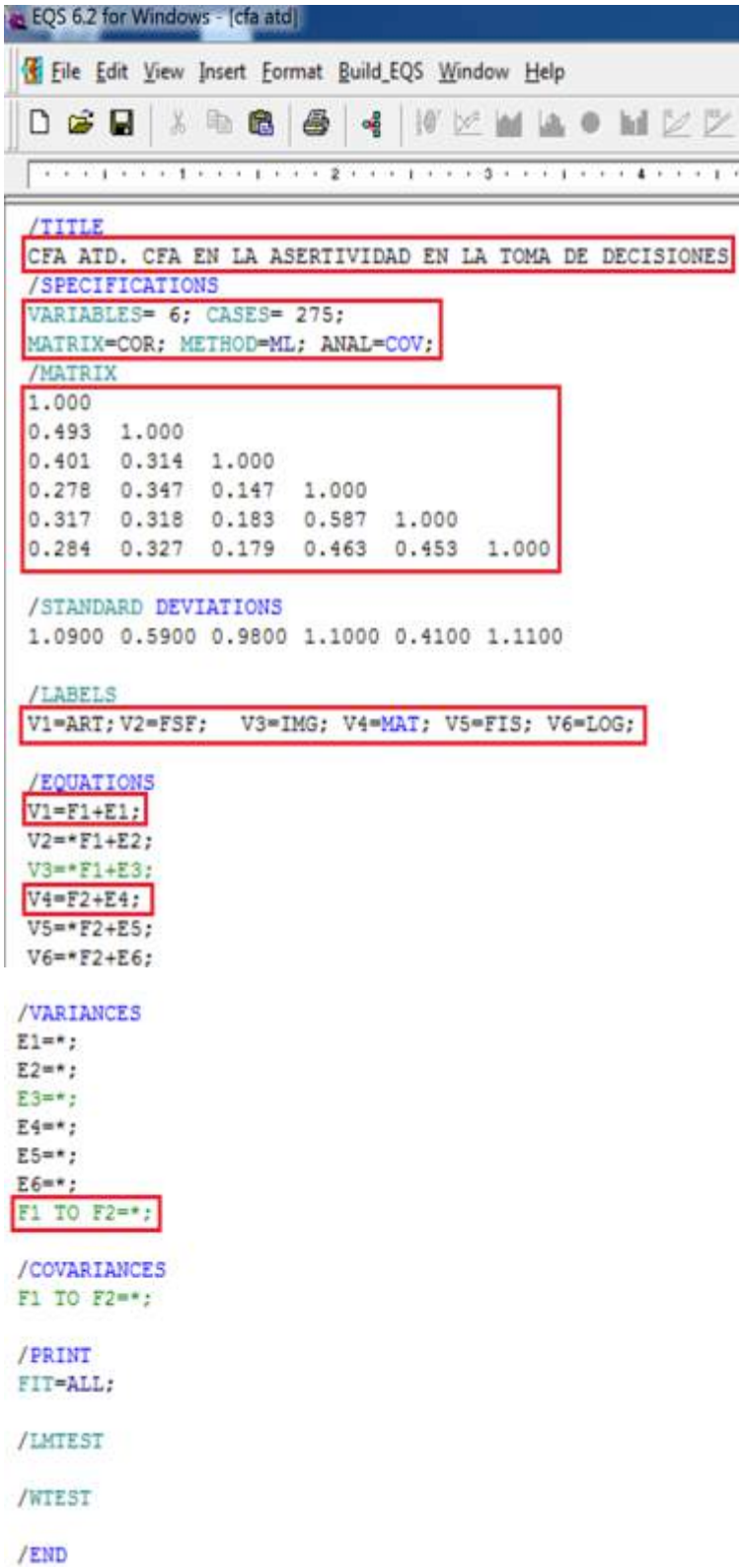

Fuente: EQS 6. con datos y adaptación propia



Fuente: EQS 6. con datos y adaptación propia

En el que se tiene codificada: la matriz de covarianzas de datos del estudio de caso en el cual, se aprecian las instrucciones a nivel del título del proyecto, las especificaciones de la base de datos, el método de análisis, la matriz de covarianzas y las desviaciones estándar consideradas, entre otros datos relevantes. Observe que si tiene 6 factores observables, en 275 casos, analizados con matriz de correlación, método de máxima verosimilitud (**ML.***Maximum Likelihood*) y análisis por covarianza. La cantidad de elementos en la matriz aplicando la **Fórmula 7.1** es de **15**. Este dato esta reportado en las magnitudes de la matriz de covarianzas de datos. **Ver Figura 7.13.**

**Figura 7.13. EQS Codificación de entrada de datos caso de estudio 1**




Fuente: EQS 6. con datos y adaptación propia

De acuerdo a la codificación mostrada en la **Figura 7.12**, se tiene de manera relevante las siguientes condiciones de aplicabilidad cumplidas:

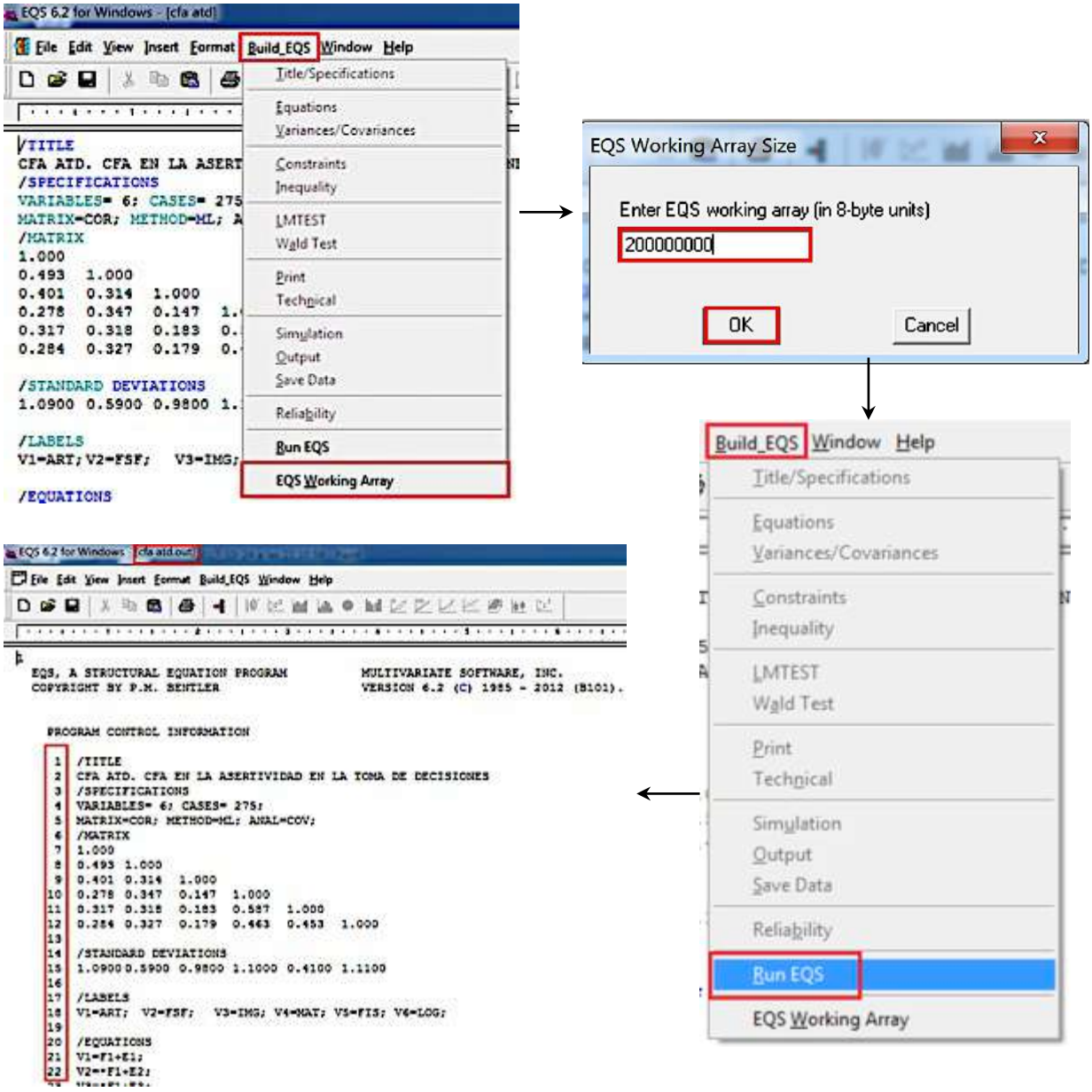
- Observe el párrafo **/TITLE**, donde se dan detalles del nombre del archivo y lo que hace. En nuestro caso: **CFA ATD. Asertividad en la toma de decisiones.**
- Observe el párrafo **/SPECIFICATIONS,** donde se muestran las instrucciones que declaran las 6 variables observables bajo estudio; los 275 casos de los datos de ejecutivos a analizar; con una matriz de correlación (**COR**), bajo el método de máxima verosimilitud (**ML**); con análisis por covarianza (**COV**). Se muestra la matriz de los 15 casos de enlaces que arrojan las 6 variables (**Fórmula 7.1**), integrada en el programa de entrada de datos **EQS**.
- Observe el párrafo **/STANDARD DEVIATIONS** con las desviaciones estándar propuestas
- Observe el párrafo **/LABELS** con las etiquetas de las **6** variables que describen al modelo y su equivalente en **EQS** (variables **VX, que describen el modelo teórico**)
- Observe el párrafo **/EQUATIONS,** en la parte de ecuaciones se eligió **V1=F1+E1,** SIN ASTERISCO **(*)**, con el fin de afirmarlo con varianza=1 y cumplir los requerimientos (así también, se replicó en el caso de **V4**); **V2** ya tiene * para que se calcule la relación de la carga factorial de **V2=*F1+E2** hacia **F1** (es decir, que tanto explica **V2** a **F1**); **como los errores no tienen asterisco, se afirman por tanto con** varianza= 1; ídem con **V3** y así consecutivamente.
- Observe el párrafo **/VARIANCES,** estas cumplen la condición en errores y en factores. Es factible la instrucción **E1 to E6=*;** , pero no se sugiere en errores ya que limita el análisis por eliminación y por lo tanto el ajuste del modelo. En factores no hay problema ya que éstos No se deben de eliminar (es por ello que se justifica **F1 to FX**).
- Observe el párrafo **/COVARIANCES,** donde se afirma que sólo las variables independientes covarían (**F1 to F2=* o F1=*; F2=*;** es lo mismo)
- Observe el párrafo **/PRINT FIT ALL;**, donde se indica que se requiere imprimir todos los índices de ajuste para su análisis.
- Observe el párrafo **/LMTEST**, que significa se genere el *test de Lagrange*, el cual es o univariado o multivariado. El primero indica que tanto aporta cada una de las variables a su correspondiente factor. El segundo, indica qué variable tiene que medir en qué factor para mejorar el ajuste. Es decir, que cambios se deben realizar en las variables para que midan factores distintos para que se mejore el modelo. No significa finalmente, que la **V1** sea parte o mida a **F2**, sino que su contribución disminuye el error y mejora el ajuste del modelo.
- Observe el párrafo **/WTEST,** *o test de Wald***,** que implica determinar *que relación entre factores que mejora del ajuste modelo.*
- Observe el párrafo **/END**, término de la programación.

## *EQS. Cómo ejecutar el programa*

Una vez definidas las ecuaciones del modelo teórico y cumplidas las condiciones de aplicabilidad, se corre el programa mediante la siguiente secuencia de comandos:

**Build EQS->EQS Working Array->Enter EQS working array (in 8-bits units): cambiar de 2000000 a 200000000 (se asegura capacidad de memoria)->OK->Build EQS->Run->Generación de archivo de salida cfa atd.out.** Ver **Figura 7.14**

**Figura 7.14 EQS.Secuencia de comandos para ejecutar programa de entrada de datos**




Fuente: EQS 6. con datos y adaptación propia

## EQS. Análisis de resultados

Como se observa, el archivo de salida generado **cfa atd.out,** muestra el mismo contenido de instrucciones que el archivo de entrada **cfa atd.eqs**, sólo que con las líneas de comando numeradas y adicionalmente, con los reportes de resultados como producto del cálculo **EQS**. Por ejemplo, las secciones de reporte de las **Figuras 7.15 y 7.16,** que muestran los arreglos de datos que cumplen con la distribución esperada.

**Figura 7.15. EQS. Reporte matriz de covarianza analizada con CFA**

```
DETERMINANT OF INPUT MATRIX IS   0.23735D-01

PARAMETER ESTIMATES APPEAR IN ORDER,
NO SPECIAL PROBLEMS WERE ENCOUNTERED DURING OPTIMIZATION.


RESIDUAL COVARIANCE MATRIX  (S-SIGMA) :
```

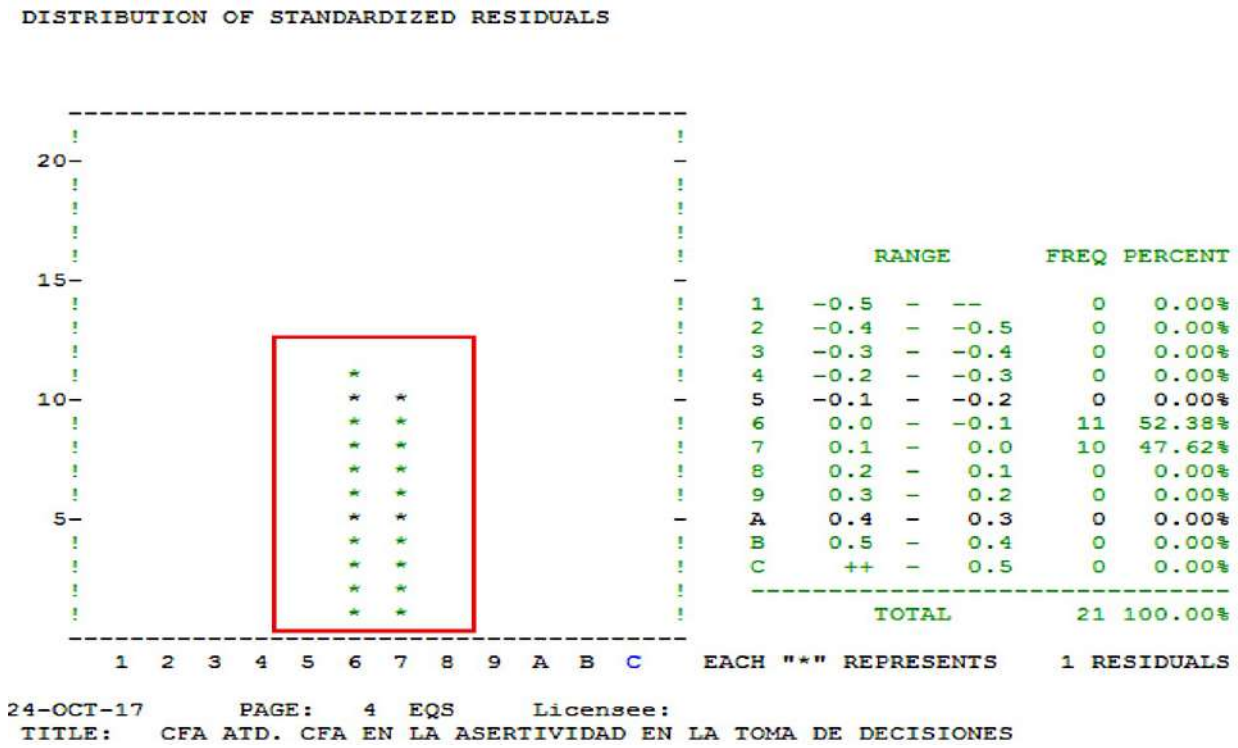
| | | ART V1 | FSF V2 | IMG V3 | MAT V4 | FIS V5 |
|---|---|---|---|---|---|---|
| ART | V1 | 0.000 | | | | |
| FSF | V2 | -0.007 | 0.000 | | | |
| IMG | V3 | 0.044 | -0.014 | 0.000 | | |
| MAT | V4 | -0.055 | 0.028 | -0.076 | 0.000 | |
| FIS | V5 | -0.003 | 0.003 | -0.014 | 0.004 | 0.000 |
| LOG | V6 | 0.027 | 0.053 | 0.003 | -0.004 | -0.006 |

| | | LOG V6 |
|---|---|---|
| LOG | V6 | 0.000 |

```
                    AVERAGE ABSOLUTE RESIDUAL =          0.0163
          AVERAGE OFF-DIAGONAL ABSOLUTE RESIDUAL =       0.0228
```

Fuente: EQS 6. con datos y adaptación propia

**Figura 7.16. EQS. Reporte residuos estandarizados con CFA**




Fuente:EQS 6, con datos y adaptación propia

El gráfico de la **Figura 7.17.** donde se observa, que los datos están adecuadamente distribuídos, ya que no existen desviaciones estándar ni a la izquierda ni a la derecha del intervalo.

**Figura 7.17.EQS. Gráfico distribución estandarizada de los residuos con CFA**

24-OCT-17 PAGE: 3 EQS Licensee:
TITLE: CFA ATD. CFA EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES

MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

LARGEST STANDARDIZED RESIDUALS:

| NO. | PARAMETER | ESTIMATE | NO. | PARAMETER | ESTIMATE |
|---|---|---|---|---|---|
| 1 | V6, V2 | 0.081 | 11 | V2, V1 | -0.011 |
| 2 | V4, V3 | -0.070 | 12 | V5, V4 | 0.010 |
| 3 | V4, V1 | -0.046 | 13 | V5, V1 | -0.007 |
| 4 | V4, V2 | 0.043 | 14 | V6, V4 | -0.004 |
| 5 | V3, V1 | 0.041 | 15 | V6, V3 | 0.003 |
| 6 | V5, V3 | -0.035 | 16 | V2, V2 | 0.000 |
| 7 | V3, V2 | -0.024 | 17 | V3, V3 | 0.000 |
| 8 | V6, V1 | 0.022 | 18 | V6, V6 | 0.000 |
| 9 | V6, V5 | -0.014 | 19 | V5, V5 | 0.000 |
| 10 | V5, V2 | 0.013 | 20 | V4, V4 | 0.000 |

Fuente:EQS 6, con datos y adaptación propia

Si existiera el caso, es indicativo de que algunas variables no están aportando al ajuste del modelo. Se debe tener mucho cuidado a la eliminación de variables para justificar el modelo de estudio (*eliminar 10% al 20% de variables*).
Por otro lado, es importante analizar la sección de reporte de índices de ajuste (*fit indices*), los cuales en reporte muestran entre otros resultados relevantes (Ver **Figura 7.18** , donde lo encuadrado en rojo, es la referencia a analizar):

- El método de máxima verosimilitud **ML**
- Se observa que p= .35579 > p = .05 o sea aprox. un *36% de error*, es decir, la probabilidad de que el modelo se equivoque, es muy alta; esto es importante para ya no considerarlo.

-Los índices de ajuste (*fit indices*) más utilizados son los primeros 3:

- Índice de ajuste normalizado o Bentler-Bonnet Normed Fit Index, **NFI**
- Índice de ajuste no normalizado o Bentler-Bonnet Non Normed Fit Index, **NNFI.**
- Índice de ajuste comparativo o Comparative Fit Index, **CFI.**
- Error medio cuadrático de aproximación del error o *Root Mean-Square Error of Approximation*, **RMSEA.**

Si el promedio de los tres primeros se encuentre entre .8-.89 el modelo se está ajustado estadístico (Hair, et al. 1999; Bagozzi y Yi, 1988); si es >.9, tiene ajuste excelente; **RMSEA** =<.08.
En conclusión, al momento el *modelo está ajustado pero con un alto error.*

**Figura 7.18. EQS. Reporte índices de ajuste con CFA**

```
GOODNESS OF FIT SUMMARY FOR METHOD = ML

INDEPENDENCE MODEL CHI-SQUARE        =     392.818 ON    15 DEGREES OF FREEDOM

INDEPENDENCE AIC =     362.818   INDEPENDENCE CAIC =     293.566
       MODEL AIC =      -7.158          MODEL CAIC =     -44.092

CHI-SQUARE =        8.842 BASED ON       8 DEGREES OF FREEDOM
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS      0.35579

THE NORMAL THEORY RLS CHI-SQUARE FOR THIS ML SOLUTION IS          9.156.

FIT INDICES
-----------
BENTLER-BONETT     NORMED FIT INDEX =     0.977
BENTLER-BONETT NON-NORMED FIT INDEX =     0.996
COMPARATIVE FIT INDEX (CFI)         =     0.998
BOLLEN'S          (IFI) FIT INDEX   =     0.998
MCDONALD'S        (MFI) FIT INDEX   =     0.998
JORESKOG-SORBOM'S  GFI  FIT INDEX   =     0.989
JORESKOG-SORBOM'S AGFI  FIT INDEX   =     0.971
ROOT MEAN-SQUARE RESIDUAL (RMR)     =     0.027
STANDARDIZED RMR                    =     0.031
ROOT MEAN-SQUARE ERROR OF APPROXIMATION (RMSEA)    =     0.020
90% CONFIDENCE INTERVAL OF RMSEA  (       0.000,         0.075)
```

Fuente: EQS. 6. con datos y adaptación propia

Cabe preguntarse: *¿en qué medida la matriz de varianzas-covarianzas predicha se ajusta a la muestra?*.Esto se puede resolver si se aplicacn los siguientes criterios:

- *Matriz residual de covarianzas.* Todos los residuos deberían estar cercanos a **0** y simétricamente distribuídos
- *Chi-Cuadrada no significativa.* Contrasta la hipótesis nula de que la matriz de varianzas-covarianzas estimada no es cualquier matriz positiva definida, sino que está condicionada al modelo.
    - No se debe rechazar esta hipótesis.
    - La prueba es muy sensible al tamaño de la muestra (tendencia a rechazo en muestras muy grandes)
- Considerar los estadísticos:

**NFI; NNFI; CFI; IFI; MFI; GFI; AGFI** *cercanos c/u a .9 (Nunca es >= 1.* Esto representarìa multicolinealidad y no es posible que midan lo mismo), y *la Raíz Cuadrática Media Residual Estandarizada* **(SRMR.***Standardized Root-Mean Square Residual***)** *cercano a .05* (**SRMR,** se utiliza cuando en primera instancia el **RMSEA=0** que implicaría *alta multicolinealidad entre las variables del modelo y tres indices de ajuste al menos sean >.9.* Esto genera la necesidad de aclarar el ajuste del modelo).

- Que el proceso de estimación haga convergencia.

Continuando el análisis, se prosigue a revisar la sección de reporte del *alfa de Cronbach* (*reliability coeffiecients*) que reporta la confiabilidad del modelo. Se destaca que lo importante a revisar es la confiabilidad y validez de la escala NO del modelo. Ver **Figura 7.19.**

**Figura 7.19. EQS. Reporte confiabilidad alfa de Cronbach del modelo con CFA**

```
RELIABILITY COEFFICIENTS
------------------------
CRONBACH'S ALPHA                    =      0.712
RELIABILITY COEFFICIENT RHO         =      0.777
```

Fuente: EQS. 6. con datos y adaptación propia.

Otro reporte a observar es la sección de iteraciones (*iterative summary*) que reporta la cantidad de iteraciones que hace el proceso hasta que hace convergencia en cero, en este caso se reportan 5 iteraciones. Ver **Figura 7.20**

**Figura 7.20. EQS. Iteraciones del sistema con CFA**

ITERATIVE SUMMARY

| ITERATION | PARAMETER ABS CHANGE | ALPHA | FUNCTION |
|---|---|---|---|
| 1 | 0.147746 | 1.00000 | 0.04518 |
| 2 | 0.025186 | 1.00000 | 0.03297 |
| 3 | 0.006181 | 1.00000 | 0.03233 |
| 4 | 0.001730 | 1.00000 | 0.03228 |
| 5 | 0.000515 | 1.00000 | 0.03227 |

Fuente: EQS. 6. con datos y adaptación propia.

Continuando el análisis de resultados que reporta **EQS,** se observa la sección de reporte de medición de las ecuaciones (*measurement equations with standard errors and tests statistics significant at the 5% are marked with @*). Aquí se reportan el valor de las ecuaciones NO estandarizadas , con la varianza de cada una de las variables NO estandarizadas. Así, *donde hay @ la p es significativa*. Ver **Figura 7.21**

**Figura 7.21. EQS. Reporte medición de las ecuaciones con CFA**

MEASUREMENT EQUATIONS WITH STANDARD ERRORS AND TEST STATISTICS
STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.

VARIANZA 1 DADO QUE NO SE LES ASIGNO ASTERISCO (*) COMO DISEÑO ORIGINAL

ART =V1 = 1.000 F1 + 1.000 E1

VALOR DE LA VARIANZA DE V2 RESPECTO A F1

TODAS LAS VARIANZAS DE LOS ERRORES TIENEN VALOR 1

FSF =V2 = .509*F1 + 1.000 E2

ERROR DE LA VARIANZA DE V2 .068

VALOR t DE STUDENT ES SIGNIFICATIVA AL .05 7.467@

IMG =V3 = .604*F1 + 1.000 E3
.096
6.320@

MAT =V4 = 1.000 F2 + 1.000 E4

FIS =V5 = .373*F2 + 1.000 E5
.039
9.467@

LOG =V6 = .817*F2 + 1.000 E6
.096
8.552@

Fuente: EQS 6. con datos y adaptación propia

De este reporte se debe analizar que todas las variables que presentan varianzas a sus respectivos factores presenten valores *t de Student* con una significancia (**@**) al *.05* para continuar el análisis, situación que es el caso.
Interpretando el modelo, puede requerir de preguntarse en este punto, sobre los parámetros estimados:
-¿Toman valores adecuados teóricamente?
-¿Existen correlaciones superiores a la unidad?
-¿Existen cargas factoriales estandarizadas fuera del intervalo (-1,+1)?
-¿Son los residuos estandarizados normalmente grandes o pequeños?
-¿Hay estimaciones negativas de las varianzas?
-¿Son significativos?

```
FSF    =V2   =     .509*F1     + 1.000 E2
                   .068
                  7.467@
```

Donde:
*Valor t (coeficiente/error estándar)*
t> 1.96 valor significativo al .05 o 5%
t> 2.56 valor significativo al .01 o 1%
**Así, la asignación de asteriscos (*) de error/confiabilidad/ valor de t, es:**
***** p<.01= 1% error o 99% confiabilidad, cuando valor de t>2.56**
**** p<.05= 5% error o 95% confiabilidad, cuando valor de t=1.96<=2.55**
*** p<.10=10% error o 90% confiabilidad, cuando valor de t=1.65<=1.95**
Revisando los reportes emitidos por **EQS**, encontramos la sección de reporte de varianzas de variables independientes (*variances of independent variables*) entre factores y entre las variables y sus errores.
Cabe preguntarse:*¿Existe Varianza negativa?.* Ver **Figuras 7.22 y 7.23.**

**Figura 7.22. EQS. Reporte de varianzas entre factores con CFA**

```
24-OCT-17      PAGE:   6  EQS     Licensee:
 TITLE:   CFA ATD. CFA EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

 VARIANCES OF INDEPENDENT VARIABLES
 ----------------------------------
 STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.

                 V                              F
                ---                            ---
                              I F1  -   F1                .636*I
                              I                           .117 I
                              I                          5.444@I
                              I                               I
                              I F2  -   F2                .698*I
                              I                           .112 I
                              I                          6.244@I
                              I                               I
```

Fuente: EQS. 6. con datos y adaptación propia

**Figura 7.23. EQS. Reporte de varianzas entre las variables y sus errores con CFA**

```
24-OCT-17        PAGE:   7  EQS      Licensee:
 TITLE:    CFA ATD. CFA EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

 VARIANCES OF INDEPENDENT VARIABLES
 ----------------------------------
 STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.

                  E                                D
                 ---                              ---
E1  -  ART                .552*I                                  I
                          .088 I                                  I
                         6.255@I                                  I
                              I                                   I
E2  -  FSF                .183*I                                  I
                          .025 I                                  I
                         7.296@I                                  I
                              I                                   I
E3  -  IMG                .728*I                                  I
                          .071 I                                  I
                        10.281@I                                  I
                              I                                   I
E4  -  MAT                .512*I                                  I
                          .075 I                                  I
                         6.828@I                                  I
                              I                                   I
E5  -  FIS                .071*I                                  I
                          .010 I                                  I
                         6.807@I                                  I
                              I                                   I
E6  -  LOG                .767*I                                  I
                          .079 I                                  I
                         9.655@I                                  I
                              I                                   I
```

Fuente: EQS. 6. con datos y adaptación propia

Ambos reportes validan lo analizado hasta el momento. La sección de reporte que es muy importante analizar sus resultados es la de la covarianza entre las variables independientes (*covariances among independent variables*). Ver **Figura 7.24.**

**Figura 7.24. EQS. Reporte covarianzas entre las variables independientes con CFA**

```
24-OCT-17        PAGE:   8  EQS      Licensee:
 TITLE:    CFA ATD. CFA EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

 COVARIANCES AMONG INDEPENDENT VARIABLES
 ---------------------------------------
 STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.

                  V                                  F
                 ---                                ---
                      I F2  -   F2                      .388*I
                      I F1  -   F1                      .068 I
                      I                                5.712@I
                      I                                      I
```

SE LEE: LA COVARIANZA ENTRE F1 y F2 ES DE .388 CON UN ERROR DE .068 Y UN VALOR t DE STUDENT SIGNIFICATIVO AL .05

Fuente: EQS. 6. con datos y adaptación propia

Este reporte es muy importante considerar su análisis, ya que a partir del mismo se determina la *validez* del modelo. Se recuerda que:
-El valor de la covarianza entre factores *nunca debe ser 1*, ya que entonces los factores miden lo mismo;
-El valor de la covarianza *No debe ser > .8* ya que puede tener problemas debido a que la *covarianza es sumada con 2 veces el error,* lo que tiende a ser *1* y por lo tanto, generar problemas de validez (*colinealidad*). Ver apartado de cálculo de la validez discriminante. Test del intervalo de confianza.
-El valor de la covarianza *No debe ser < .1* ya que los factores serían *interdependientes* (lo que significa que nada habría que ver entre ellas; sin relación alguna en el modelo)
En nuestro caso (.388*), No existe el 1 y No es >.8* cumple el requerimiento hasta el momento para verificar su validez.
Al seguir analizando los reportes emitidos por **EQS,** encontramos la sección de solución estandarizada (*standardized solution*), es decir, *son resultados que ya se ha eliminado el error a cada una de las variables* que intervienen en el modelo. Lo visto hasta el momento no se encontraba estandarizado.
Cabe preguntarse: *¿Estimaciones estandarizadas >1 como valor absoluto?.* Ver **Figura 7.25.**

**Figura 7.25. EQS. Reporte solución estandarizada con CFA**

```
24-OCT-17      PAGE:   9  EQS     Licensee:
 TITLE:   CFA ATD. CFA EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)


 STANDARDIZED SOLUTION:                                          R-SQUARED
 CARGAS FACTORIALES ESTANDARIZADOS DE LAS 6  VARIABLES RESPECTO A
 LOS FACTORES Y LOS ERRORES CON R CUADRADA DE C/U

  ART   =V1  =   .732 F1    + .682 E1                              .535
  FSF   =V2  =   .688*F1    + .725 E2                              .474
  IMG   =V3  =   .492*F1    + .871 E3                              .242
  MAT   =V4  =   .759 F2    + .651 E4                              .577
  FIS   =V5  =   .760*F2    + .650 E5                              .578
  LOG   =V6  =   .615*F2    + .789 E6                              .378
```

Fuente: EQS. 6. con datos y adaptación propia

Antes de finalizar el método, se reporta la sección de correlaciones entre variables independientes (*correlations among independent variables*).
Cabe preguntarse: *¿Existen Correlaciones >1?. No puede ser porque sería el mismo factor. Deseable cargas factoriales de factores >.6.* Ver **Figura 7.26.**

**Figura 7.26. EQS. Reporte correlaciones entre variables independientes con CFA**

```
24-OCT-17      PAGE:  10  EQS     Licensee:
 TITLE:   CFA ATD. CFA EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

 CORRELATIONS AMONG INDEPENDENT VARIABLES
 ----------------------------------------

                V                              F
               ---                            ---
                          I F2  -   F2                .582*I
                          I F1  -   F1                     I
                          I                                I
```

VERIFICAR QUE EL VALOR SEA <1

```
-----------------------------------------------------------------------------
                          E N D   O F   M E T H O D
-----------------------------------------------------------------------------
```

Fuente: EQS. 6. con datos y adaptación propia

Hasta aquí, fin de los reportes del método. Sin embargo, continuamos con los análisis de los resultados de pruebas específicas como lo es la *prueba de Wald* (*Wald test–for dropping parameters*) considerando parámetros multivariantes (*multivariate Wald test by simultaneous process*).Ver **Figura 7.27.**

**Figura 7.27. EQS. Reporte test de Wald con CFA**

```
24-OCT-17      PAGE:  11  EQS     Licensee:
 TITLE:   CFA ATD. CFA EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

 WALD TEST (FOR DROPPING PARAMETERS)
 MULTIVARIATE WALD TEST BY SIMULTANEOUS PROCESS

        CUMULATIVE MULTIVARIATE STATISTICS              UNIVARIATE INCREMENT
        ----------------------------------              --------------------

 STEP  PARAMETER   CHI-SQUARE  D.F.  PROBABILITY       CHI-SQUARE  PROBABILITY
 ---- -----------  ----------  ----  -----------       ----------  -----------

    ************
         NONE OF THE FREE PARAMETERS IS DROPPED IN THIS PROCESS.
```

| NO | CODE | PARAMETER | CHI-SQUARE | PROB. | HANCOCK 8 DF PROB. | PARAM. CHANGE | STANDARDIZED CHANGE | PREDICTED RMSEA | PREDICTED CFI |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 12 | V2,F2 | 4.741 | 0.029 | 0.785 | 0.174 | 0.354 | 99.999 | 1.000 |
| 2 | 2 12 | V6,F1 | 2.098 | 0.148 | 0.978 | 0.181 | 0.205 | 99.999 | 1.000 |
| 3 | 2 12 | V3,F2 | 1.675 | 0.196 | 0.989 | -0.142 | -0.174 | 0.009 | 1.000 |
| 4 | 2 12 | V1,F2 | 1.125 | 0.289 | 0.997 | -0.167 | -0.184 | 0.019 | 0.998 |
| 5 | 2 12 | V4,F1 | 1.000 | 0.317 | 0.998 | -0.134 | -0.152 | 0.021 | 0.998 |
| 6 | 2 12 | V5,F1 | 0.069 | 0.793 | 1.000 | -0.013 | -0.040 | 0.030 | 0.995 |
| 7 | 2 0 | V4,F2 | 0.000 | 1.000 | 1.000 | 0.000 | 0.000 | 0.031 | 0.995 |
| 8 | 2 0 | V1,F1 | 0.000 | 1.000 | 1.000 | 0.000 | 0.000 | 0.031 | 0.995 |

```
*** NOTE *** IF PREDICTED RMSEA COULD NOT BE CALCULATED, 99.999 IS PRINTED.
             IF PREDICTED  CFI  COULD NOT BE CALCULATED,  9.999 IS PRINTED.
```

Fuente: EQS. 6. con datos y adaptación propia

El resultado: NONE OF THE FREE *PARAMETERS IS DROPPED IN THIS PROCESS* de la prueba, nos indica que no hay ningún parámetro que se pueda considerar en esta relación, dado que sólo hay dos factores que generan covarianza; requiere para su demostración al menos tres factores. Aún así, es correcto medir con dos factores la asertividad en la toma de decisiones (**ATD**).

Verificando los reportes **EQS,** nos encontramos con la sección del *multiplicador de Lagrange* (*Lagrange multiplier test -for adding parameters-*), para datos univariados y multivariados. Ver **Figuras 7.28 y 7.29.**

**Figura 7.28. EQS. Reporte multiplicador de Lagrange. Datos univariados con CFA**

```
24-OCT-17        PAGE:  12  EQS      Licensee:
 TITLE:    CFA ATD. CFA EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)


 LAGRANGE MULTIPLIER TEST (FOR ADDING PARAMETERS)
```

SE DESCARTA ESTA SECCIÓN PARA ANÁLISIS EN LA MODLALIDAD UNIVARIADA. LO IMPORTANTE ES LA MODALIDAD MULTIVARIADA

```
 ORDERED UNIVARIATE TEST STATISTICS:
                                                  HANCOCK            STANDAR-
                                  CHI-             8 DF    PARAM.    DIZED     PREDICTED
  NO  CODE    PARAMETER          SQUARE   PROB.   PROB.   CHANGE    CHANGE   RMSEA   CFI
  --  ----    ---------          ------   -----   -----   ------   -------   -----------
   1  2 12        V2,F2            4.741   0.029   0.785    0.174     0.354  99.999 1.000
   2  2 12        V6,F1            2.098   0.148   0.978    0.181     0.205  99.999 1.000
   3  2 12        V3,F2            1.675   0.196   0.989   -0.142    -0.174   0.009 1.000
   4  2 12        V1,F2            1.125   0.289   0.997   -0.167    -0.184   0.019 0.998
   5  2 12        V4,F1            1.000   0.317   0.998   -0.134    -0.152   0.021 0.998
   6  2 12        V5,F1            0.069   0.793   1.000   -0.013    -0.040   0.030 0.995
   7  2  0        V4,F2            0.000   1.000   1.000    0.000     0.000   0.031 0.995
   8  2  0        V1,F1            0.000   1.000   1.000    0.000     0.000   0.031 0.995

*** NOTE *** IF PREDICTED RMSEA COULD NOT BE CALCULATED, 99.999 IS PRINTED.
             IF PREDICTED  CFI  COULD NOT BE CALCULATED,  9.999 IS PRINTED.
```

Fuente: EQS. 6. con datos y adaptación propia

Se descarta análisis ya que lo importante son los multivariados. Ver **Figura 7.29**

**Figura 7.29. EQS. Reporte multiplicador de Lagrange. Datos multivariados con CFA**

```
24-OCT-17        PAGE:  13  EQS      Licensee:
 TITLE:    CFA ATD. CFA EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

 MULTIVARIATE LAGRANGE MULTIPLIER TEST BY SIMULTANEOUS PROCESS IN STAGE  1

 PARAMETER SETS (SUBMATRICES) ACTIVE AT THIS STAGE ARE:

    PVV PFV PFF PDD GVV GVF GFV GFF BVF BFF


      CUMULATIVE MULTIVARIATE STATISTICS      UNIVARIATE INCREMENT
      ----------------------------------      ----------------------------
```

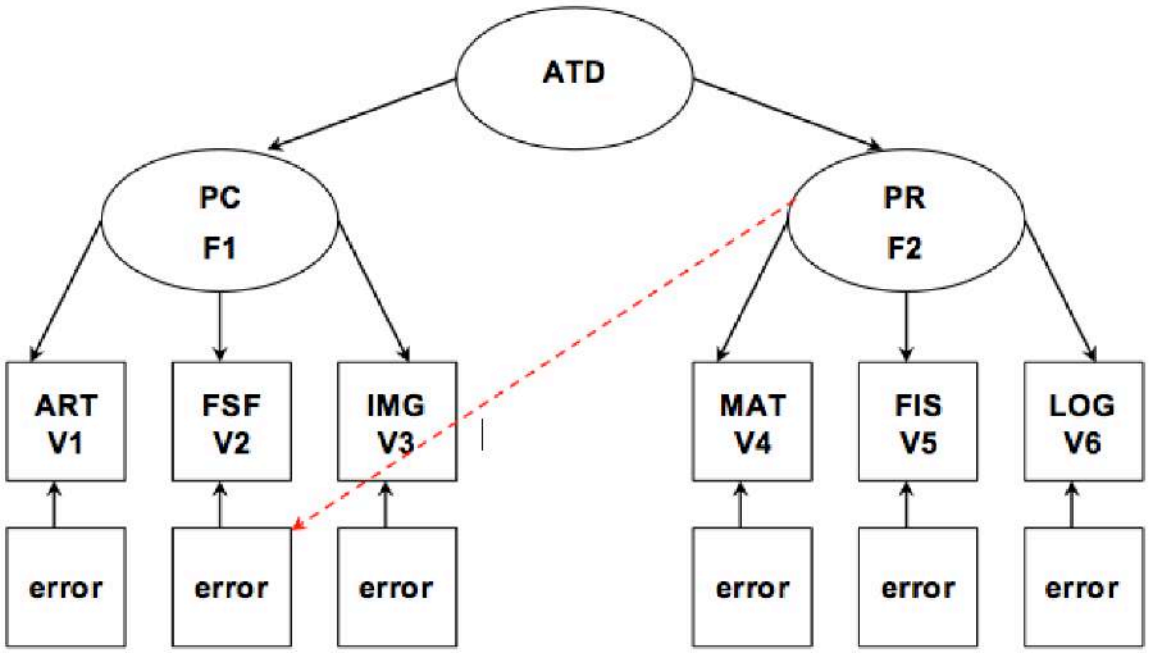
SE LEE: SI V2 MIDE A F2 ENTONCES CHI CUADRADO SE INCREMENTA EN 4.741

```
                                                                HANCOCK'S
                                CHI-                   CHI-     SEQUENTIAL   PREDICTED
 STEP    PARAMETER              SQUARE   D.F.  PROB.   SQUARE  PROB.  D.F.  PROB.  RMSEA   CFI
 ----    ---------              ------   ----  -----   ------  -----  ----  -----  -----------
    1       V2,F2                4.741      1  0.029    4.741  0.029     8  0.785  99.999 1.000

*** NOTE *** IF PREDICTED RMSEA COULD NOT BE CALCULATED, 99.999 IS PRINTED.
             IF PREDICTED  CFI  COULD NOT BE CALCULATED,  9.999 IS PRINTED.
```

Fuente: EQS. 6. con datos y adaptación propia

La lectura de que **V2** mide a **F2** con un incremento de Chi-Cuadrada, trae como consecuencia un cambio de incidencia variable-factor, como se indica en la **Figura 7.30**

**Figura 7.30. Diagrama de trayectoria con cambio de contribución de variable-factor**




Fuente: propia.

## Discusión

La discusión debe dar una clara exposición de los constructos del modelo y sus implicaciones en función del conocimiento, línea de investigación, incluso sesgo e interés del investigador, de la teoría del modelo vs. la realidad y explicar coincidencias y diferencias con pleno conocimiento del sector abordado. Las posibles causas de resultados diferentes se deben explicar a fin de que el modelo se depure y se haga universal. Hacer observaciones para estudios futuros.

El resultado de nuestro caso de estudio, nos sugiere que para mejorar el modelo, el Pensamiento Racional (**PR**), es la causa de la Filosofía (**FSF**) y no del Pensamiento Creativo (**PC**). Es decir, **V2** al medir a **F2** (**PR**), con el aumento de Chi-Cuadrado en vez de **F1,** implica que el modelo sólo a nivel estadístico requiere de este reacomodo, sin embargo *No tiene implicaciones teóricas* ya que el modelo fue previamente validado a novel de constructo en un marco teórico, por lo que es muy probable que la nueva relación No tenga sustento como ajuste teórico. Puede hacerse la pregunta ¿Puede la Filosofía (**FSF**) medir la Asertividad de la Toma de Decisiones (**ATD**) en la parte de su Pensamiento Racional (**PR**)?. Esto exigiría nuevamente una revisión teórica del modelo. Se debe recordar también, que a estas alturas, *No es el objetivo*

*mejorar el modelo* ya que incluso, el modelo *cumple con las condiciones de los índices de ajuste (fit indices, >.9)*, mostrados en la **Figura 7.18**.
Asumiendo que el modelo está correctamente respaldado por el marco teórico y con el fin de ser más rigurosos en el análisis, se recurre a una de las condiciones de aplicabilidad emitidas por Bagozzi y Yi (1988), las cuáles, para ser tomadas en cuenta, *deben ser cada una >.6*. Analizando la solución estandarizada de la **Figura 7.25,** observamos, que esta condición no cumple para la variable. Ver **Figura 7.31** .

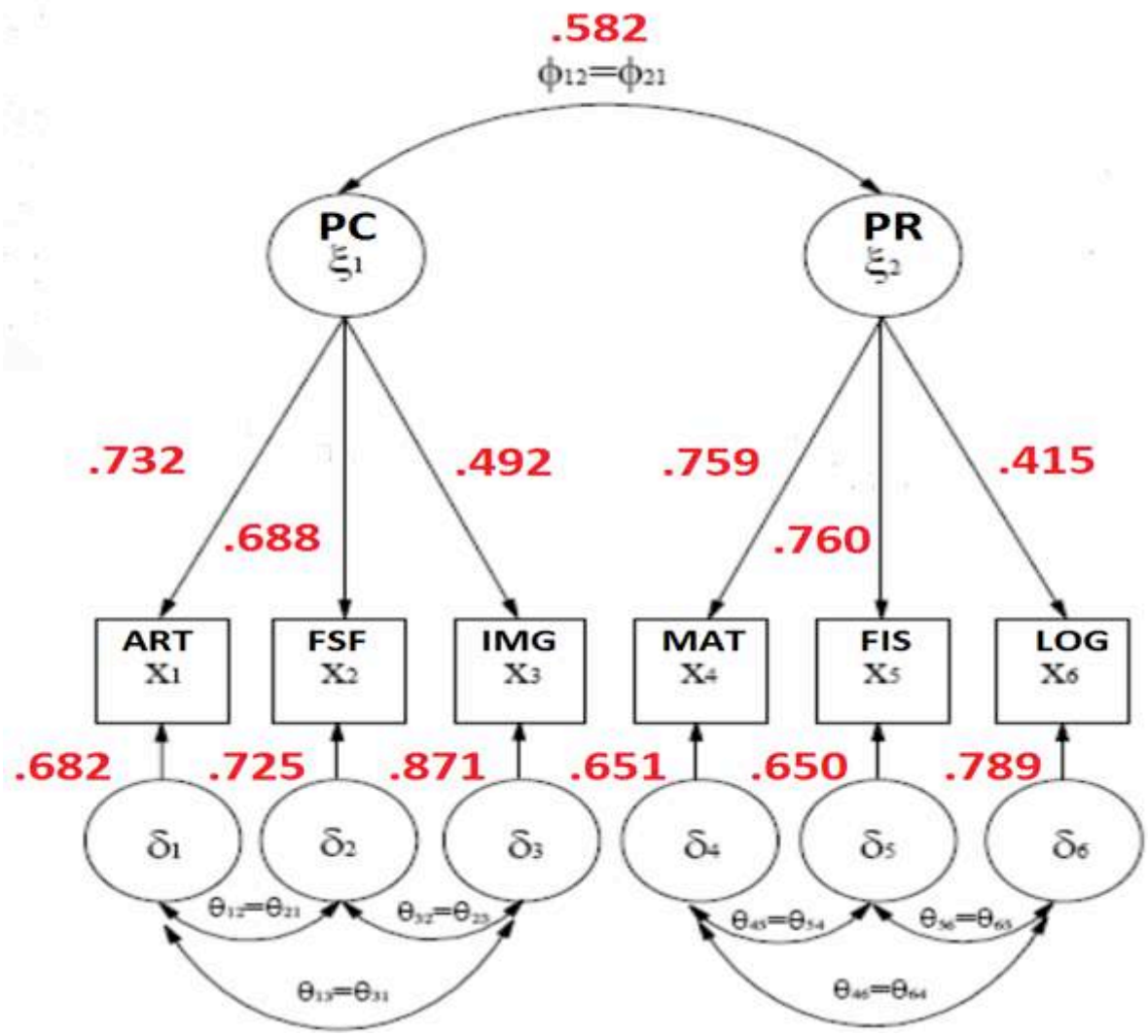
**Figura 7.31. EQS. Reporte solución estandarizada con CFA**

```
24-OCT-17      PAGE:   9  EQS     Licensee:
 TITLE:   CFA ATD. CFA EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)


 STANDARDIZED SOLUTION:                                                R-SQUARED


  ART   =V1  =   .732 F1    + .682 E1                                     .535
  FSF   =V2  =   .688*F1    + .725 E2                                     .474
  IMG   =V3  =   .492*F1    + .871 E3                                     .242
  MAT   =V4  =   .759 F2    + .651 E4                                     .577
  FIS   =V5  =   .760*F2    + .650 E5                                     .578
  LOG   =V6  =   .615*F2    + .789 E6                                     .378
```

Fuente: EQS. 6. con datos y adaptación propia

En esta figura, se aprecia que la variable Imaginación (**IMG**), no cumple debidamente con la aportación al factor **F1**, Pensamiento Creativo (**PC**) ya que su carga factorial: .492<.6 *En este caso no cumple con Bagozzi y Yi (1988).*
Recuerde que la interpretación de las cargas factoriales de cada uno de los factores (no de manera conjunta) inciden en un porcentaje directo en el factor de análisis. Por ejemplo, se lee: *que la variable Arte (**ART**), explica en un 73.2% el factor de Pensamiento Creativo (**PC**); que la variable Filosofía (**FSF**), explica en un **68.8%** el factor de Pensamiento Creativo (**PC**); que la variable Imaginación (**IMG**) explica en un 49.2% el factor de Pensamiento Creativo (**PC**)*; etc.
Así y de acuerdo a Bagozzi y YI (1988) Imaginación (**IMG**), *no debe estar en el modelo.*
Debe considerarse también a Hair (et al.,1999), que indica que el promedio de las *cargas factoriales deben ser >.6*. Si fuera así, *se debe eliminar la más baja (irrelevancia)*. En este caso: (.732+.688+.492)/3= .637 *por lo que cumple con Hair (et al. 1999).*
Al momento, como *diagrama de trayectoria*, se tiene la **Figura 7.32**

**Figura 7.32. Diagrama de trayectoria con los resultados del caso de estudio con solución estandarizada**




Fuente: Jöreskog (1993); Hair (et al.,1999); Byrne 2006 con adaptación propia.

## *Reespecificación del modelo*

Como se observa, lo logrado hasta el momento implica añadir o eliminar relaciones entre las variables del modelo motivado por 2 razones:

*-Mejorar el ajuste*

*-Contrastar alguna hipótesis teórica*

Sin embargo, se recomienda incorporar o eliminar relaciones cuando éstas tengan un soporte teórico que justifique la escala. *La mejora del ajuste No debe ser el fin en sí mismo.*

Lo anterior, es posible realizarlo mediante:

- *El multiplicador de Lagrange (Lagrange Multiplier) para añadir relaciones, parámetro que conduce a un mayor incremento de la Chi-Cuadrada*
- *Prueba de Wald (Wald Test)*

```
/LMTEST

/WTEST

/END
```

En nuestro caso No aplicó. Importante referirse a autores acreditados para basar su decisión como lo mostrado en la discusión.Ver **Figuras 7.27 y 7.29**

## *Confiabilidad*

En esta etapa, se debe verificar:

- *Consistencia.* Coherencia interna de los ítems (¿existen correlaciones entre ellos?). ¿Realmente ítems, variables y factores del modelo miden lo planeado?
- Si los ítems están fuertemente correlacionados entre sí, significa que miden una misma variable latente (los causa), pero no que ésta variable latente que miden, sea el constructo que se pretende medir (validez).
- La confiabilidad como condición necesaria pero *No suficiente para la validez*. Así, *una escala puede tener confiabilidad pero No validez, sin embargo, Jamás una escala tendrá validez Sino tiene antes, confiabilidad.*

-Su medición, suele realizarse por una de 3 vías de explicacion de la varianza (Bagozzi y Yi, 1988):

- *Coeficiente alfa de Cronbach* ***> .7***
- *Indice de Fiabilidad Compuesta (**IFC**)* ***> .7***
- *Índice de Varianza Extraída (**IVE**)* ***> .5***

- Desde el punto de vista de análisis estructural, se sugiere cumplir con los 3, para lograr confiabilidad de la escala (*No de los datos*)
- El punto de partida para construir los indicadores de confiabilidad: la matriz de varianzas-covarianzas entre los ítems de la escala.

Recordar que *el alfa de Cronbach,* es el método más utilizado para medir la confiabilidad cuando ésta se entiende como *consistencia interna***,** analiza la homogeneidad de las cuestiones planteadas de forma simultánea y tiene la ventaja de que para su cálculo, sólo se necesita realizar una medición con un grupo de personas. El proceso consiste en medir la confiabilidad de una suma simple de medias (ítems) paralelas o *tau-equivalentes* y esla medida de todos los coeficientes de correlación resultantes de correlacionar dos mitades de una escala de todas las formas posibles. Vea nuevamente la **Ecuación 3.9**

**Ecuación 3.9. Alfa de Cronbach por correlación**

$$\alpha = \frac{k\rho}{1+(k-1)\rho}$$

Donde:
**ρ** es, como ya hemos indicado, la media de los coeficientes de correlación entre todos los ítems que conforman la escala.

El coeficiente *alfa de Cronbach* debe variar entre *0 y 1*; así, conforme se incrementa su valor,*menor es la varianza*. Así, se tienen dos interpretaciones del valor de *alfa de Cronbach:*

- *Es la correlación existente entre la escala y todas las posibles escalas que, midiendo el mismo concepto, tengan igual número de ítems.*

*o*

- *Es la correlación entre la puntuación que una persona (o personas) obtiene en una escala (valor observado) y la puntuación que debería tener, considerando todos los ítems del universo que definen el concepto.*

El punto de inicio, es: *la matriz de varianzas-covarianzas*. Si *se estandariza*, se obtiene la *matriz de correlaciones***.** De esta manera, se afirma que la varianza total de la escala es la suma de los elementos de C **(ver Matriz 3.1**)

**Matriz 3.1. Varianzas-Covarianzas**

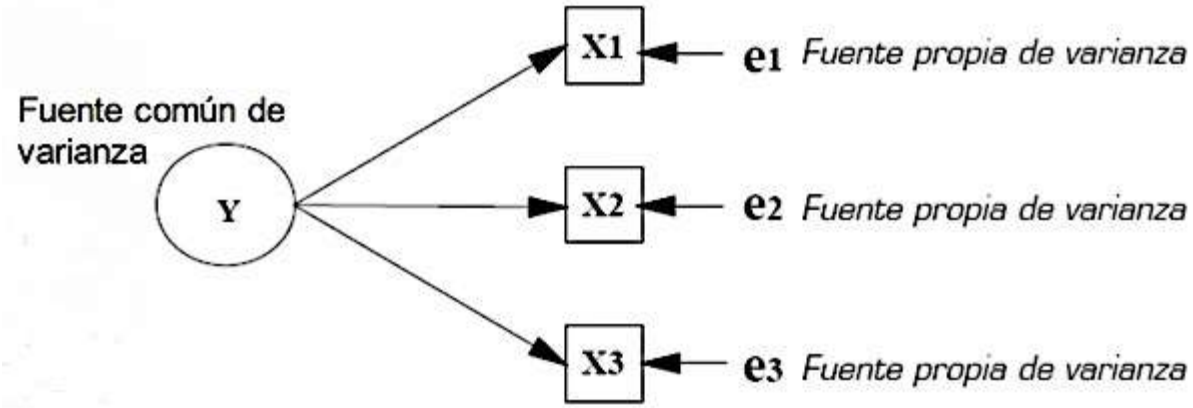
$$C=\begin{pmatrix} \boxed{\sigma_1^2} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \boxed{\sigma_2^2} & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \boxed{\sigma_3^2} & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \boxed{\sigma_4^2} \end{pmatrix}$$

Siendo así: *Varianza total= Varianza común+Varianza específica*

- *Varianza común.* Se trata de la varianza en los ítems provocada por la variable latente (que los ítems comparten entre sí). Si varía la variable latente, varían los ítems $\mathbf{X_i}$ de la escala.
- **Vari***anza específica, única***.** Varianza causada por los errores de medida de cada ítem en particular en la diagonal principal de la matriz.

Observe la **Figura 7.33.**

**Figura 7.33. Comportamiento de la varianza**




Fuente: Jöreskog (1993); Hair (et al.,1999); Byrne 2006 con adaptación propia.

La variable independiente (*fuente común de la varianza*) y los errores (*fuente propia de varianza de cada variable dependiente*) afecta a las dependientes.
*Alfa de Cronbach*= la parte de la varianza total que es atribuible a la variable latente (*varianza común*), donde:

$$\alpha = \frac{\sigma_y^2 - \sum \sigma_i^2}{\sigma_y^2} = 1 - \frac{\sum \sigma_i^2}{\sigma_y^2}$$

**VARIANZA TOTAL.**
**Suma de los elementos de la Matriz de Varianzas-Covarianzas**

**VARIANZA ESPECÍFICA.**
**Suma de los elementos de la Diagonal de la Matriz de Varianzas-Covarianzas**

Fuente: Jöreskog (1993); Hair (et al.,1999); Byrne 2006 con adaptación propia.

La expresión anterior, debe ponderarse debido a que el número de elementos necesarios para calcular la *varianza total* (todos los de la matriz: $k^2$) *no es el mismo que los necesarios para calcular la varianza específica* (la diagonal: **k**).
Por lo tanto, en el numerador hay $\mathbf{k^2-k}$ elementos y en el denominador se tienen k elementos, por lo que para que el ratio exprese las magnitudes relativas, más que el número de casos, corregimos por $\mathbf{k^2/-k^2}$ o lo que es lo mismo **k/k-1**. Así, la formula matemática que queda para el calculo del *alfa de Cronbach,* es mostrado en la **Ecuación 3.8.** Si se requiere expresar el *alfa de Cronbach* en función de las *varianzas-covarianzas*, se tiene la **Ecuación 3.9**

**Ecuación 3.9. Alfa de Cronbach por varianza-covarianza**

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum \sigma_i^2}{\sigma_y^2}\right)$$

Si se requiere expresar *el alfa de Cronbach en función de las correlaciones y no de las varianzas-covarianzas*, se obtiene así la siguiente formulación matemática (Crocker y Algina, 1986) , se tiene la **Ecuación 3.10.**

**Ecuación 3.10. Alfa de Cronbach por correlación**

$$\alpha = \frac{k\rho}{1 + (k-1)\rho}$$

Donde:
p es la media de las correlaciones entre ítems.

Ahora bien, los valores aceptables del *alfa de Cronbach* :
*Nunnally (1987):* >= *.7*
Se debe tener especial atención al número de ítems de la escala y la fase de construcción en la que se encuentra, ya que se sugiere, dependiendo autor a considerar, se tiene como recomendación:

- *Alfa de Cronbach, escala en construcción: .6 a .8*
- *Alfa de Cronbach escala en uso o ajuste normal: .8 a .9*
- *Alfa de Cronbach escala existente> .9 es altamente ajustada*

## *Cálculo IFC*

Como notas importantes, se tiene:
-Permite tener en cuenta todos los constructos implicados en la escala y no un análisis uno a uno como *Cronbach.*
*-Requiere de la realizacion de un* ***CFA.***
*-Se calcula para cada factor implicado (Fornell y Larcker, 1981)*. Ver **Ecuación 3.11**

**Ecuación 3.11. Índice de la fiabilidad compuesta**

$$IFC_i = \frac{\left(\sum_j L_{ij}\right)^2}{\left(\sum_j L_{ij}\right)^2 + \sum_j Var\left(E_{ij}\right)}$$

Fuente: Fornell y Larcker (1981)

Donde:
**L** es la carga factorial estandarizada de cada uno de los **j** indicadores que cargan sobre el factor
**Var ($E_{ij}$)** es la varianza del término de error que se calcula como:

$$Var\left(E_{ij}\right) = 1 - L_{ij}^{\ 2}$$

Se debe calcular ya que el software no lo reporta. Siguiendo la **Tabla 3.14** para el problema del caso de estudio, tenemos la **Tabla 7.5**

Así, los cálculos necesarios para obtener el **IFC** se ilustran en la **Tabla 3.14**:

**Tabla 7.5. Recopilación de información que calcula el IFC de ATD del caso de estudio con CFA**

| Constructo e indicadores | Carga Factorial Estandarizada $L_{ij}$ | Varianza del Término de Error $Var(E_{ij}) = 1 - (L_{ij})^2$ | Índice de Fiabilidad Compuesta IFC >.7 $\sum(L_{ij})^2/[\sum(L_{ij})^2 + \sum Var(E_{ij})]$ |
|---|---|---|---|
| **Pensamiento Creativo (PC)** | | | |
| $V_1$=ART | .732 | .464 | **.676** |
| $V_2$=FSF | .688 | .527 | |
| $V_3$=IMG | .492 | .758 | |
| Total | 1.912 | 1.749 | |
| **Pensamiento Racional (PR)** | | | |
| $V_4$=MAT | .759 | .424 | **.691** |
| $V_5$=FIS | .760 | .422 | |
| $V_6$=LOG | .415 | .828 | |
| Total | 1.934 | 1.674 | |

Fuente: EQS 6. con datos y adaptación propios.

Observe que en ambos casos es < .7 . Considerando el apartado de *confiabilidad* y anteriores *No cumple con los requerimientos la escala***.** Observe que con los promedios de las cargas factoriales >.6 enunciados por Hair et al., (1999), *SI cumple con el requerimiento de confiabilidad*, sin embargo, de acuerdo a los criterios expuestos por Bagozzi y Yi (1988) de cada carga factorial debe ser >.7 y con la

comprobación de la **Tabla 7.5,** *se concluye que la escala diseñada No cumple los requerimientos de confiabilidad.* Observe que las variables que aportan escasamente a la explicaciòn del modelo de **ATD**, son **IMG** y **LOG.**
*Acción sugerida: eliminar **V3** y realizar de nueva cuenta análisis.*

## Redefiniendo el modelo

Al realizar la acciòn sugerida, se renombra el archivo de entrada de datos de programa como: **cfa atd sv3.eqs** y se documenta en el título como : *CFA ATD SV3. CFA EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES SIN V3***.** Para eliminar la acción de **V3,** a ésta se le antepone, en todos los lugares que se mencione así como su error con el signo ( **!** ). Ver **Figura 7.34**

**Figura 7.34. EQS Codificación de entrada de datos caso de estudio 1 con CFA sin V3**

```
COPYRIGHT BY P.M. BENTLER                    VERSION 6.2 (C) 1985 - 2012 (B101).

  PROGRAM CONTROL INFORMATION

  1  /TITLE
  2  CFA ATD SV3. CFA EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES SIN V3
  3  /SPECIFICATIONS
  4  VARIABLES= 6; CASES= 275;
  5  MATRIX=COR; METHOD=ML; ANAL=COV;
  6  /MATRIX
  7  1.000
  8  0.493  1.000
  9  0.401  0.314   1.000
 10  0.278  0.347   0.147   1.000
 11  0.317  0.318   0.183   0.587   1.000
 12  0.284  0.327   0.179   0.463   0.453   1.000
 13
 14  /STANDARD DEVIATIONS
 15  1.0900 0.5900  0.9800  1.1000  0.4100  1.1100
 16
 17  /LABELS
 18  V1=ART;  V2=FSF;   V3=IMG; V4=MAT; V5=FIS; V6=LOG;
 19
 20  /EQUATIONS
 21  V1=F1+E1;
 22  V2=*F1+E2;
 23  !V3=*F1+E3;
 24  V4=F2+E4;
 25  V5=*F2+E5;
 26  V6=*F2+E6;
 27
 28  /VARIANCES
 29  E1=*;
 30  E2=*;
 31  !E3=*;
```

Fuente: EQS. 6. con datos y adaptación propia

Ejecute nuevamente el programa siguiendo las instrucciones del apartado *de cómo ejecutar programa***.** De los nuevos resultados, observe que la secciones de reporte de matriz de covarianza analizada, de residuos estadarizados y el gráfico de distribución esrandarizada de los residuos, permanecen con resultados esperados, sin mayores complicaciones. Sin embargo, observe la sección de reporte de índices de ajuste (*fit indices*). Ver **Figura 7.35**

**Figura 7.35. EQS. Reporte índices de ajuste con CFA sin V3**

```
26-OCT-17      PAGE:   4  EQS     Licensee:
 TITLE:   CFA ATD SV3. CFA EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)



 GOODNESS OF FIT SUMMARY FOR METHOD = ML

 INDEPENDENCE MODEL CHI-SQUARE        =     338.050 ON    10 DEGREES OF FREEDOM

 INDEPENDENCE AIC =     318.050   INDEPENDENCE CAIC =     271.882
        MODEL AIC =      -4.438          MODEL CAIC =     -22.905

 CHI-SQUARE =        3.562 BASED ON       4 DEGREES OF FREEDOM
 PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS      0.46855

 THE NORMAL THEORY RLS CHI-SQUARE FOR THIS ML SOLUTION IS         3.559.

 FIT INDICES
 -----------
 BENTLER-BONETT     NORMED FIT INDEX =     0.989
 BENTLER-BONETT NON-NORMED FIT INDEX =     1.003
 COMPARATIVE FIT INDEX (CFI)         =     1.000
 BOLLEN'S          (IFI) FIT INDEX   =     1.001
 MCDONALD'S        (MFI) FIT INDEX   =     1.001
 JORESKOG-SORBOM'S  GFI  FIT INDEX   =     0.995
 JORESKOG-SORBOM'S AGFI  FIT INDEX   =     0.981
 ROOT MEAN-SQUARE RESIDUAL (RMR)     =     0.017
 STANDARDIZED RMR                    =     0.019
 ROOT MEAN-SQUARE ERROR OF APPROXIMATION (RMSEA)     =     0.000
 90% CONFIDENCE INTERVAL OF RMSEA  (       0.000,          0.087)
```

Fuente: EQS. 6. con datos y adaptación propia

En este reporte se encuentran resultados atípicos, por ejemplo:

- *Nunca debe considerarse índices de ajuste >=1***.** Por lo tanto se descartan los índices: **NNFI, CFI**; Bollen´s; Mcdonald´s.
- *El **RMSEA**=0.000* **(***debe ser: <.08<>0.00***)**; dado que hay índices de ajuste *>=1* esto validaría el hecho de que hay *una alta multicolinealidad* de las variables por *la supresión de **V3.***

Byrne (2006) y Bentler (1993) *condicionan a tomar tres índices*, generalmente los primeros 3; como en este caso hay 2 que no cumplen, queda en libertad del investigador basarse en los que Sí cumplan los requerimientos. En este caso, se tiene la alternativa de utilizar **SRMR=** 0.019 *ya que cumple con estar <.5*, con la posibilidad de utilizar los índices de ajuste **NFI, GFI, AGFI** >.9 de alto ajuste**.** Se deberá soportar por marco teórico su uso. (**Nota:** *los indices de ajuste normalizados siempre reportan valores < a los no normalizados ya que estan ajustad*os).
Todo lo anterior da a pauta a explicar a mayor detalle las condiciones del modelo dentro de un artículo a publicar.
Asumiendo, que todo va en acuerdo a lo esperado, se continúa revisando las secciones de reporte como la mostrada en la **Figura 7.36**

**Figura 7.36. EQS. Reporte ecuaciones de medida con CFA sin V3**

```
26-OCT-17        PAGE:   5  EQS      Licensee:
 TITLE:   CFA ATD SV3. CFA EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)


 MEASUREMENT EQUATIONS WITH STANDARD ERRORS AND TEST STATISTICS
 STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.


  ART    =V1   =  1.000 F1     +1.000 E1


  FSF    =V2   =   .610*F1     +1.000 E2
                   .099
                  6.139@

  MAT    =V4   =  1.000 F2     +1.000 E4


  FIS    =V5   =   .368*F2     +1.000 E5
                   .039
                  9.518@

  LOG    =V6   =   .811*F2     +1.000 E6
                   .095
                  8.582@
```

Fuente: EQS 6. con datos y adaptación propia

Se debe preguntar: *¿son significativos los resultados?* Si la repuesta es SI, se continúa con el análisis de la sección de reporte de la **Figura 7.37.**

**Figura 7.37. EQS. Reporte covarianzas entre las variables independientes con CFA**

```
26-OCT-17      PAGE:   8  EQS     Licensee:
 TITLE:   CFA ATD SV3. CFA EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

 COVARIANCES AMONG INDEPENDENT VARIABLES
 ---------------------------------------
 STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.

                V                              F
               ---                            ---
                          I F2  -   F2                .370*I
                          I F1  -   F1                .070 I
                          I                          5.292@I
                          I                                I
```

Fuente: EQS 6. con datos y adaptación propia

*¿La covarianza entre las variables independientes es >=1?* NO ya que es de .37
**Nota**: *La covarianza NUNCA debe ser >=.8*
¿Es significativa al .5? Sí, ya que tienen @ asignada. Se asegura alta probabilidad de *Validez* y se continúa con el análisis de la sección de reporte de la **Figura 7.38**

**Figura 7.38. EQS. Reporte solución estandarizada con CFA**

```
26-OCT-17      PAGE:   9  EQS     Licensee:
 TITLE:   CFA ATD SV3. CFA EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

 STANDARDIZED SOLUTION:                                          R-SQUARED


  ART    =V1   =   .661 F1    + .750 E1                            .437
  FSF    =V2   =   .745*F1    + .667 E2                            .556
  MAT    =V4   =   .765 F2    + .644 E4                            .585
  FIS    =V5   =   .755*F2    + .656 E5                            .569
  LOG    =V6   =   .615*F2    + .789 E6                            .378
```

Fuente: EQS 6. con datos y adaptación propia

*¿Cargas factoriales >.6?* Sí, se continúa con el cálculo de la fiabilidad compuesta.

## *Recálculo de IFC y cálculo IVE*

En la **Tabla 7.6**, se muestra el concentrado de datos para cálculo

**Tabla 7.6. Recopilación de información que calcula el IFC de ATD del caso de estudio CFA sin V3**

| Constructo e indicadores | Carga Factorial Estandarizada $L_{ij}$ | Varianza del Término de Error $Var(E_{ij}) = 1 - (L_{ij})^2$ | Índice de Fiabilidad Compuesta IFC >.7 $\sum(L_{ij})^2/[\sum(L_{ij})^2 + \sum Var(E_{ij})]$ |
|---|---|---|---|
| **Pensamiento Creativo (PC)** | | | |
| $V_1$=ART | .661 | .563 | **.662** |
| $V_2$=FSF | .745 | .445 | |
| Total | 1.406 | 1.008 | |
| **Pensamiento Racional (PR)** | | | |
| $V_4$=MAT | .765 | .415 | **.757** |
| $V_5$=FIS | .755 | .430 | |
| $V_6$=LOG | .615 | .622 | |
| Total | 2.135 | 1.467 | |

Fuente: EQS 6. con datos y adaptación propios.

Dado que el **IFC** >.6 y partiendo de la **Tabla 3.15** previa, tenemos para nuestro caso la **Tabla 7.7**, como sigue:

**Tabla 7.7. Recopilación de información que calcula el IVE de ATD del caso de estudio CFA sin V3**

| Constructo e indicadores | Cuadrado de la Carga Factorial Estandarizada $(L_{ij})^2$ | Varianza del Término de Error $Var(E_{ij}) = 1 - (L_{ij})^2$ | Indice de Varianza Compuesta Extraída IVE >= .5 $\sum(L_{ij})^2/[\sum(L_{ij})^2 + \sum Var(E_{ij})]$ |
|---|---|---|---|
| **Pensamiento Creativo (PC)** | | | |
| $V_1$=ART | .437 | .563 | **.496** |
| $V_2$=FSF | .555 | .445 | |
| Total | .992 | 1.008 | |
| **Pensamiento Racional (PR)** | | | |
| $V_4$=MAT | .585 | .415 | **.511** |
| $V_5$=FIS | .570 | .430 | |
| $V_6$=LOG | .378 | .622 | |
| Total | 1.533 | 1.467 | |

Fuente: EQS 6. con datos y adaptación propios.

Recuerde que la confiablidad y la validez se encuentran estrechamente relacionadas, desempeñando papeles complementarios. La *confiabilidad trata acerca de la consistencia , exactitud y capacidad de predicción de los hallazgos. La validez, está*

*relacionada con el problema de estamos realmente midiendo lo que creemos,* por lo que es un asunto más amplio y complejo que la confiabilidad.

## Determinando la validez

Recuerde lo citado en la **secciones 4.1, 4.2 y 4.3** en que la validez de una escala, sucede cuando la variable latente que están midiendo los indicadores, es realmente lo que se quiere medir. También representa el grado en que un instrumento mide el concepto bajo estudio (Bohrnstedt,1976).

Recuerde que existen en resumen, los siguientes tipos de validez más utilizados:

- *Validez de contenido.* Grado en que una escala recoge todas las dimensiones teóricas del concepto que se quiere medir. *(Se refleja en las cargas factoriales >.6)*
- *Validez de constructo o concepto.* Se tienen para considerar:
  - ***-Convergente****.* Los indicadores utilizados para *medir un mismo concepto,* están fuertemente correlacionados entre sí. *(Sucede sólo entre los ítems de cada factor del constructo; es decirlos ítems de un factor sólo se relacionan para dicho factor, o sea, convergen)*
  - ***-Discriminante****.* Los indicadores utilizados para *medir conceptos distintos* están escasamente correlacionados entre sí.(*Sucede entre al menos 2 o más factores o constructos, midiendo cuestiones diferentes con un mínimo de relación)*
- *Validez nomológica.* El constructo medido por los indicadores propuestos guarda relaciones con otros conceptos teóricamente previsibles, (concurrente o predictiva). En otras palabras, *¿el modelo teórico inicial es mejor que el ajustado? Se espera que No. Así, si el modelo teórico Sí es mejor que el ajustado, se dice que No hay validez nomológica*.

Sin embargo, *si el modelo teórico no es mejor al ajustado, se dice que SI hay validez nomológica. Esto es, que el modelo ajustado está más cercano a la explicación de la realidad que el modelo teórico inicial*.

En nuestro caso, el suprimir la **V3** hizo un mejora a los índices del modelo teórico original, por lo que SI se obtuvo validez nomológica.

*Cuando el modelo teórico coincide con la realidad, No existe validez nomológica.* Al suprimir variables, por baja explicación, se está ajustando el modelo mejorándolo respecto al teórico (SI hay validez nomológica), sin embargo*, no se recomienda suprimir más del 10%. Lo ideal es que No exista validez nomológica (modelo teórico=modelo ajustado= coincide con la realidad). Cuando existe validez nomológica la teoría No coincide con la realidad, por lo que se deberá buscar explicación de la realidad del porqué no se consideró en la teoría. En nuestro caso, la **V3** suprimida no implica que la teoría esté errónea, sino que en la realidad, no se está llevando a cabo su práctica.*

Esta validez podría medirse haciendo la diferencia de la Chi-Cuadrada del modelo teórico y ajustado y el resultado buscar los grados de libertad en qué nivel de error se estarían manejando los ajustes y a que nivel de significatividad. La validez nomológica, simple y llanamente se demuestra al suprimir o no las variables del modelo.

En cuanto a la *validez discriminante*, recuerde lo citado en el apartado *validez*

*discriminante (discriminant validity)***,** *de los test más utilizados:*
*-Test de intervalos de confianza (Anderson y Gerbing, 1988)*
*-Test de la varianza extraída (Fornell y Larcker, 1981 )*
*-Test de diferencias de las Chi-Cuadrada (Anderson y Gerbing, 1988).*
Con un solo test que soporte el modleo teórico, basta para demostrar que existe validez discriminante.*Se estila hacer los dos primeros.*

## *EQS.Cálculo de la validez discriminante. Test del intervalo de confianza*

Siguiendo el caso y considerando que *es validez de los constructos* (*No de los ítems*), se observa que se tienen 2 constructos: **F1** y **F2.** Recuerde que de la **Figura 7.37,**

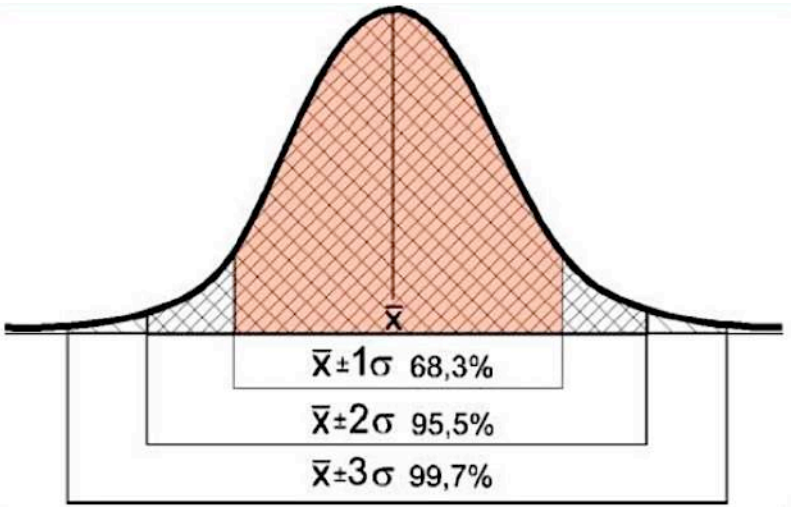
**Figura 7.37. EQS. Reporte covarianzas entre las variables independientes**

```
26-OCT-17       PAGE:   8  EQS     Licensee:
 TITLE:   CFA ATD SV3. CFA EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

 COVARIANCES AMONG INDEPENDENT VARIABLES
 ---------------------------------------
 STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.

                V                          F
               ---                        ---
                          I F2  -   F2            .370*I
                          I F1  -   F1            .070 I
                          I                      5.292@I
                          I                            I
```

Fuente: EQS 6. con datos y adaptación propia

La covarianza de las variables independientes No debe tener **1**, para que tenga validez discriminante. Se debe realizar el cálculo de la siguiente manera:
En **EQS** se codifica **: * F1-F2**
Intervalo límite inferior = .370 – 2 (.070) = .230
Intervalo límite superior= **.**370 + 2 (.070) = .510
El 2 se refiere al nivel de confiabilidad de 95.5% al que se desea trabajar, como el factor más común utilizado en las ciencias económico-administrativas. **Vea Figura 7.38**

**Figura 7.38. Campana de Gauss y la desviaciòn estándar**




Fuente Levin y Rubin (2004)

Del resultado, se lee: *entre el límite inferior y superior entre los factores* **F1** y **F2***, No debe existir la unidad (1) ya que los factores estarían midiendo lo mismo***.**
En nuestro caso (.230 - .510), por lo tanto *Sí cumple con las condiciones del test, hay validez discriminante.*
En otras palabras, *los factores miden cuestiones diferentes*.
**Nota:** *la covarianza Nunca debe ser >=.8* ya que el cálculo de los límites inferior y superior, al multiplicar por **2** el error, forzarían a que el resultado sea **1**.
Si fueran más factores se deben considerar las diversas combinaciones y cumplir con el requerimiento. Ejemplo: 5 Factores
En **EQS**, se codIfica **: F1 to F5= *;**
Realizando 10 combinanciones.
F1-F2; F1-F3; F1-F4; F1-F5; F2-F3; F2-F4; F2-F5; F3-F4; F3-F5; F4-F5 y cumplir las condiciones de los límites.

## *Cálculo de la validez discriminante. Test de la varianza extraída.*

Tomando en cuenta la seccion de reporte covarianzas entre las variables independientes (*covariances among the independent variables*) de la **Figura 7.37,** se debe determinar la covarianza al cuadrado = $(.370)^2$ = .137**.**
La lectura, para nuestro caso de estudio,es:
-*¿Es el cuadrado de la covarianza de las variables independientes F1-F2 es menor (<) a los **IVEs** calculados (***Tabla 7.7***)?*
o
-*¿Los valores calculados de los* **IVEs** *(***Tabla 7.7***), son mayores (>) al cuadrado de la covarianza de las variables independientes* **F1-F2***?*
Así, de la **Tabla 7.7**, tenenmos los valores de **IVE** como $\mathbf{IVE_{PC}}$=.496 **e** $\mathbf{IVE_{PR}}$= .511.
Por lo tanto *Sí cumple con las condiciones del test, hay validez discriminante.*

## *Reportando resultados: consistencia interna y validez convergente*

En este punto, es importante tener una referencia del acomodo de los datos, como se sugiere en el ejemplo de la **Figura 7.39**

**Figura 7.39. Reporte tabla de consistencia interna y validez convergente del modelo teórico sin V3**

| Factor | Ítem V | Variable | Carga Factorial >0.6 (a) | Valor Robusto t | Factor Promedio de Carga Factorial | Fiabilidad de la Escala | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | Alfa de Cronbach >=0.7 (b) | IFC >0.7 (b) | IVE >0.5 (c) |
| F1 (PC) | 1 | ART | .661*** | 1.000a | 0.703 | - | **.662** | **.496** |
| | 2 | FSF | .745*** | 6.139 | | | | |
| F2 (PR) | 4 | MAT | .765*** | 1.000a | 0.712 | - | **.757** | **.511** |
| | 5 | FIS | .755*** | 9.518 | | | | |
| | 6 | LOG | .615*** | 8.582 | | | | |

Resultados: Chi-Cuadrada= 3.562; gl= 4; p < .468; NFI = .989; GFI = .995; AGFI = .981; SRMR = .019

*Conclusion: las relaciones entre las variables y los factores tienen un ajuste adecuado a los datos ya que se presentan con una alta probabilidad de ocurrencia del 99% (significancia .01,marcado con ***). Sin embargo, el p= .468 es muy alto, o sea, representa un error del 46.8% (debe ser <.05). Esto es, no hay error en el ajuste del modelo, pero SI existe error en los índices (varios son >1).Aunqe existe suficiente evidencia de validación convergente y fiabilidad , que justifica la fiabilidad interna de la escala aplicada (Nunnally y Bernstein, 1994; Hair et al., 1999), los datos reportados, se deben considerar de forma muy cuidadosa ya que presentan error.*

Notas:
*** Parámetros constreñidos a ese valor en el proceso de identificación = **p** < 0.001
(a) De acuerdo a Bagozzi y Yi, 1988.
(b) De acuerdo a Hair et al. 1999.
(c) Índice de Varianza Extraída (**IVE**), de acuerdo a Fornell y Larcker, 1981.

Fuente: EQS 6 y SPSS 21 IBM, con datos y adaptación propios.

**Donde las columnas significan:**

- *Factores,* listado de los involucrados en el estudio. En nuestro caso Pensamiento Creativo: **F1 (PC)** y Pensamiento Racional: **F2 (PR)**
- *Ítem,* consecutivo de las variables (**V**) mostradas, como : **V1; V2; V3,** etc.
- *Variable,* en donde se ubica las siglas que representa cada una de las variables no suprimidas: **V1; V2, V4, V5, V6**, etc. En nuestro caso: **ART; FSF; MAT; FIS; LOG**
- *Carga Factorial > .6 (a)* (Ver **Figura 7.38**) donde sólo se reportan los factores no suprimidos y con carga factorial >.6**.** Es posible que el investigador utilice escalas previas que si a su juicio, por el marco teórico, deban asignarse más ítems, es válido

realizarlo siempre y cuando, se mencione que la nueva escala es una *adaptación* de la previa más la añadida como propuesta del investigador dadas las condiciones de la realidad en la que se aplica contra las que originalmente la escala se creó.En esta etapa es donde se valida o no la pertinencia de los nuevos ítems basados en el marco teórico planteado.

• *Valor Robusto t* (ver **Figura 7.36**, que se lee por cada variable: *covarianza; error y valor de t significativo a .5), donde se apreciará que hay 1 ítem de cada factor con coeficiente 1.000a que implica que tiene varianza 1, como condición requerida mínima de que una variable tenga dicho valor de la varianza, para que se realice el cálculo por factor como parte del proceso de identificación; esto implica que no habrá valor de t dado que no se puso asterisco (*) en el arranque del software EQS del programa de entrada y por lo tanto, no reportará el valor de t.*

Debemos recordar que:

• *Un nivel del 95% de confianza incluirá el 47.5%* del área que se encuentra a ambos lados de la media de la distribución de muestreo (vea **Figura 7.40**) a 0.475. Descubrimos que *0.475 del área bajo la curva normal está contenida entre la media y un punto situado a 1.96 errores estándar a la derecha de la media*.

Por consiguiente, sabemos que *(2)(0.475)= 0.95 del área está localizada entre +- 1.96 errores estándar de la media .* En consecuencia, *95% del área está contenida entre +- 1.96 errores estándar de la media. (p=.05)*

**-***Un nivel de confianza del 99% incluiría 49.5%* del área que se encuentra a cualquier lado de la media de la distribución de muestreo Si observamos la **Figura 7.40,** veremos *que 0.495 del área bajo la curva normal está localizada entre la media y un punto que se encuentra a +- 2.58 errores estándar de la media. En consecuencia, 99% del área está contenida entre +- 2.58 errores estándar de la media. (p=.01)*

**Figura 7.40. Áreas bajo la curva de distribución de probabilidad normal estándar entre la media y los valores positivos de z**

**Ejemplo:**

Para encontrar el área bajo la curva entre la media y un punto que está a 2.24 desviaciones estándar a la derecha de la media, busque el valor que se encuentra a la altura del renglón correspondiente a 2.2 y en la columna del 0.04; 0.4875 del área bajo la curva se encuentra entre la media y un valor de *z* de 2.24.

| z | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.0000 | 0.0040 | 0.0080 | 0.0120 | 0.0160 | 0.0199 | 0.0239 | 0.0279 | 0.0319 | 0.0359 |
| 0.1 | 0.0398 | 0.0438 | 0.0478 | 0.0517 | 0.0557 | 0.0596 | 0.0636 | 0.0675 | 0.0714 | 0.0753 |
| 0.2 | 0.0793 | 0.0832 | 0.0871 | 0.0910 | 0.0948 | 0.0987 | 0.1026 | 0.1064 | 0.1103 | 0.1141 |
| 0.3 | 0.1179 | 0.1217 | 0.1255 | 0.1293 | 0.1331 | 0.1368 | 0.1406 | 0.1443 | 0.1480 | 0.1517 |
| 0.4 | 0.1554 | 0.1591 | 0.1628 | 0.1664 | 0.1700 | 0.1736 | 0.1772 | 0.1808 | 0.1844 | 0.1879 |
| 0.5 | 0.1915 | 0.1950 | 0.1985 | 0.2019 | 0.2054 | 0.2088 | 0.2123 | 0.2157 | 0.2190 | 0.2224 |
| 0.6 | 0.2257 | 0.2291 | 0.2324 | 0.2357 | 0.2389 | 0.2422 | 0.2454 | 0.2486 | 0.2517 | 0.2549 |
| 0.7 | 0.2580 | 0.2611 | 0.2642 | 0.2673 | 0.2704 | 0.2734 | 0.2764 | 0.2794 | 0.2823 | 0.2852 |
| 0.8 | 0.2881 | 0.2910 | 0.2939 | 0.2967 | 0.2995 | 0.3023 | 0.3051 | 0.3078 | 0.3106 | 0.3133 |
| 0.9 | 0.3159 | 0.3186 | 0.3212 | 0.3238 | 0.3264 | 0.3289 | 0.3315 | 0.3340 | 0.3365 | 0.3389 |
| 1.0 | 0.3413 | 0.3438 | 0.3461 | 0.3485 | 0.3508 | 0.3531 | 0.3554 | 0.3577 | 0.3599 | 0.3621 |
| 1.1 | 0.3643 | 0.3665 | 0.3686 | 0.3708 | 0.3729 | 0.3749 | 0.3770 | 0.3790 | 0.3810 | 0.3830 |
| 1.2 | 0.3849 | 0.3869 | 0.3888 | 0.3907 | 0.3925 | 0.3944 | 0.3962 | 0.3980 | 0.3997 | 0.4015 |
| 1.3 | 0.4032 | 0.4049 | 0.4066 | 0.4082 | 0.4099 | 0.4115 | 0.4131 | 0.4147 | 0.4162 | 0.4177 |
| 1.4 | 0.4192 | 0.4207 | 0.4222 | 0.4236 | 0.4251 | 0.4265 | 0.4279 | 0.4292 | 0.4306 | 0.4319 |
| 1.5 | 0.4332 | 0.4345 | 0.4357 | 0.4370 | 0.4382 | 0.4394 | 0.4406 | 0.4418 | 0.4429 | 0.4441 |
| 1.6 | 0.4452 | 0.4463 | 0.4474 | 0.4484 | 0.4495 | 0.4505 | 0.4515 | 0.4525 | 0.4535 | 0.4545 |
| 1.7 | 0.4554 | 0.4564 | 0.4573 | 0.4582 | 0.4591 | 0.4599 | 0.4608 | 0.4616 | 0.4625 | 0.4633 |
| 1.8 | 0.4641 | 0.4649 | 0.4656 | 0.4664 | 0.4671 | 0.4678 | 0.4686 | 0.4693 | 0.4699 | 0.4706 |
| 1.9 | 0.4713 | 0.4719 | 0.4726 | 0.4732 | 0.4738 | 0.4744 | 0.4750 | 0.4756 | 0.4761 | 0.4767 |
| 2.0 | 0.4772 | 0.4778 | 0.4783 | 0.4788 | 0.4793 | 0.4798 | 0.4803 | 0.4808 | 0.4812 | 0.4817 |
| 2.1 | 0.4821 | 0.4826 | 0.4830 | 0.4834 | 0.4838 | 0.4842 | 0.4846 | 0.4850 | 0.4854 | 0.4857 |
| 2.2 | 0.4861 | 0.4864 | 0.4868 | 0.4871 | 0.4875 | 0.4878 | 0.4881 | 0.4884 | 0.4887 | 0.4890 |
| 2.3 | 0.4893 | 0.4896 | 0.4898 | 0.4901 | 0.4904 | 0.4906 | 0.4909 | 0.4911 | 0.4913 | 0.4916 |
| 2.4 | 0.4918 | 0.4920 | 0.4922 | 0.4925 | 0.4927 | 0.4929 | 0.4931 | 0.4932 | 0.4934 | 0.4936 |
| 2.5 | 0.4938 | 0.4940 | 0.4941 | 0.4943 | 0.4945 | 0.4946 | 0.4948 | 0.4949 | 0.4951 | 0.4952 |
| 2.6 | 0.4953 | 0.4955 | 0.4956 | 0.4957 | 0.4959 | 0.4960 | 0.4961 | 0.4962 | 0.4963 | 0.4964 |
| 2.7 | 0.4965 | 0.4966 | 0.4967 | 0.4968 | 0.4969 | 0.4970 | 0.4971 | 0.4972 | 0.4973 | 0.4974 |
| 2.8 | 0.4974 | 0.4975 | 0.4976 | 0.4977 | 0.4977 | 0.4978 | 0.4979 | 0.4979 | 0.4980 | 0.4981 |
| 2.9 | 0.4981 | 6.4982 | 0.4982 | 0.4983 | 0.4984 | 0.4984 | 0.4985 | 0.4985 | 0.4986 | 0.4986 |
| 3.0 | 0.4987 | 0.4987 | 0.4987 | 0.4988 | 0.4988 | 0.4989 | 0.4989 | 0.4989 | 0.4990 | 0.4990 |

Fuente: Levin y Rubin (2004)

Con lo anterior, es posible afirmar que *si los valores de t (ver sección reporte ecuaciones de medida o measurement equations,* ***Figura 7.36****) son mayores (>) 2.58 errores estándar de la media, entonces se anotarán tres asteriscos (***) por carga factorial, lo cual indica que p<.0001. Así se tienen una seguridad que las cargas factoriales aportan sus valores a un 99% de confiabilidad*

**-***Alfa de Cronbach,* no es posible calcularla directamente dado que se insertaron los datos de nuestro ejemplo a partir de una matriz de correlación (ver **Figura 7.13**). Sería factible calcularla separadamente si los datos estuvieran contenidos en una base de datos (como **SPSS**); sin embargo, no se tuvieron. Lo que se ha calculado con **EQS** hasta aquí es el *alfa de Cronbach***=** .706 **(ver Figura 7.41)** *de todo el modelo general,* sin embargo, lo que se requiere es realizarlo por factor, para comprobar la confiabilidad de la escala. Se tiene confianza de que el reporte de confiabilidad por factor tienda a ser superior a .7. (Se debe calcular teniendo el desglose de la base de datos en **SPSS** y comprobar)

**Figura 7.41. EQS. Reporte índices de ajuste CFA sin V3**

```
INDEPENDENCE AIC =      318.050   INDEPENDENCE CAIC =      271.882
       MODEL AIC =       -4.438          MODEL CAIC =      -22.905

CHI-SQUARE =        3.562 BASED ON       4 DEGREES OF FREEDOM
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS      0.46855

THE NORMAL THEORY RLS CHI-SQUARE FOR THIS ML SOLUTION IS          3.559.

FIT INDICES
-----------
BENTLER-BONETT      NORMED FIT INDEX =      0.989
BENTLER-BONETT NON-NORMED FIT INDEX =      1.003
COMPARATIVE FIT INDEX (CFI)         =      1.000
BOLLEN'S          (IFI) FIT INDEX   =      1.001
MCDONALD'S        (MFI) FIT INDEX   =      1.001
JORESKOG-SORBOM'S  GFI  FIT INDEX   =      0.995
JORESKOG-SORBOM'S AGFI  FIT INDEX   =      0.981
ROOT MEAN-SQUARE RESIDUAL (RMR)     =      0.017
STANDARDIZED RMR                    =      0.019
ROOT MEAN-SQUARE ERROR OF APPROXIMATION (RMSEA)    =      0.000
90% CONFIDENCE INTERVAL OF RMSEA  (       0.000,          0.087)

RELIABILITY COEFFICIENTS
------------------------
CRONBACH'S ALPHA                    =      0.706
RELIABILITY COEFFICIENT RHO         =      0.771
```

Fuente: EQS 6. con datos y adaptación propia

- **IFC.** Se toman los datos de la **Figura 7.6**
- **IVE.** Se toman los datos de la **Figura 7.7**

Hasta este momento, tomando en cuenta la **Figura 7.39** afirmamos que para **maximizar** la destreza de la asertividad en la toma de decisiones (**ATD**) en los ejecutivos de la industria del software, es necesario capacitarlos más en Filosofía (**FSF**) y Artes (**ART**) que en técnicas de Imaginación (**IMG**), para que se fomente el Pensamiento Creativo (**PC**) y que esto se debe complementar con una capacitación mayor en Matemáticas (**MAT**) Física (**FIS**) que en Lógica (**LOG**) para reforzar su Pensamiento Racional (**PR**). Todo lo anterior basados en las cargas factoriales encontradas con una probabilidad de equivocación de 1/100 o 99% de que lo afirmado ocurra (ver las ***)

Donde los renglones por debajo de las cargas factoriales, significan:

- ***S-Bχ2*** *Indice Satorra Bentler Chi-Cuadrada,* es el recomendado cuando se tienen más de dos factores para generar datos robustos. Sin embargo nuestro ejemplo no corresponde, por lo que es válido reportar la Chi-Cuadrada, con sus grados de libertad. Así, se explica la lectura reportada en la base de la **Figura 7.39**:
- ***Resultados: Chi-Cuadrada****= 3.562;* ***gl****= 4;* ***p*** *< .468;* ***NFI*** *= .989;* ***GFI*** *= .995;* ***AGFI*** *= .981;* ***SRMR*** *= .019*

• ***Conclusión:*** Las relaciones entre las variables y los factores tienen un ajuste adecuado a los datos. Sin embargo, el p= .468 es muy alto, o sea, representa un error del 46.8% (debe ser <.05). Esto es, no hay error en el ajuste del modelo, perosí jay error en los índices (varios son >1).Aunqe existe suficiente evidencia de validación convergente y fiabilidad , que justifica la fiabilidad interna de la escala aplicada (Nunnally y Bernstein, 1994; Hair et al., 1999), los datos se deben considerar de forma muy cuidadosa ya que presentan error.

De no haber tenido el problema p y la falta de *alfa de Cronbach* por constructo, una posible redacción de los resultados de la **Figura 7.39**, sería:
*…Por otro lado, para evaluar la fiabilidad y validez de las escalas de medida se realizó un Análisis Factorial Confirmatorio (**CFA**) utilizando el método de máxima verosimilitud con el software **EQS 6** (Bentler, 2005; Brown, 2006; Byrne, 2006). Asimismo, la fiabilidad de las escalas de medida se evaluó a partir del coeficiente alfa de Cronbach y el índice de fiabilidad compuesta (**IFC**) (Bagozzi y Yi, 1988). Todos los valores de la escala excedieron el nivel recomendado de **.70** para el alfa de Cronbach y el **IFC** que proporciona una evidencia de fiabilidad y justifica la fiabilidad interna de las escalas (Nunally y Bernstein, 1994; Hair et al., 1995).*
*Los resultados de la aplicación del **AFC** se presentan en la Tabla 1 y sugiere que el modelo de medición proporciona un buen ajuste de los datos (**S-BX2** = XXX.YYY**; df =** XX**; p =** 0.000**; NFI** >. 9XXX **; NNFI** >.9XXX**; CFI**>.9XXX**; y RMSEA** <.08XXX). Como evidencia de la validez convergente, los resultados del **CFA** indican que todos los ítems de los factores relacionados son significativos (p < 0.001) y el tamaño de todas las cargas factoriales estandarizadas son superiores a 0.60 (Bagozzi y Yi, 1988).*
*La Tabla xxx muestra una alta consistencia interna de los constructos, en cada caso, la alfa de Cronbach excede el valor de .70 recomendado por Nunnally y Bernstein (1994). El **IFC** representa la varianza extraída entre el grupo de variables observadas y el constructo fundamental (Fornell y Larcker, 1981). Generalmente un **IFC** superior a .60 es considerado como deseable (Bagozzi y Yi, 1988), en este estudio este valor es superior. El índice de la varianza extraída (**IVE**) fue calculado para cada par de constructos, resultando un **IVE** superior a **.5** en todos los factores (Fornell y Larcker, 1981)…*

## Reporte de resultados: validez discriminante del modelo teórico

Con el fin de facilitar la presentacion de resultados de la validez discriminante, se sugiere utilizar lo mostrado en la **Figura 7.42**

**Figura 7.42. Reporte tabla de validez discriminante del modelo teórico**

| Variable | F1 Pensamiento Creativo (PR) | F2 Pensamiento Racional (PR) |
|---|---|---|
| F1 Pensamiento Creativo (PR) | **.496** | **.137** |
| F2 Pensamiento Racional (PR) | **.230 -. 510** | **.511** |

Fuente: EQS 6. con datos y adaptación propios.

Esta tabla de reporte, se recopila de la siguiente manera:

- La diagonal es representada por las **IVE** calculadas (Ver **Figura 7.39** )
- Los datos por debajo de la diagonal, son los límites superior-inferior de los Factores **F1-F2.** (Ver Cálculo de la validez discriminante. Test del intervalo de confianza)

-Los datos por arriba de la diagonal, son las covarianzas al cuadrado entre las variables independientes (Ver Cálculo de la validez discriminante. Test de la varianza extraída)

La tabla de reporte, se debe de interpretar de la siguiente manera:

- *Los datos por debajo de la diagonal, no deben contener 1*. Por lo tanto, cumple con validez convergente y validez discriminante por test del intervalo de confianza.
- *Los datos por arriba de la diagonal*. **IVE** debe ser mayor a los datos capturados por arriba de la columna y también a los datos que se capturen a la derecha. Por lo tanto, cumple con validez discriminante por test de la varianza extraída.

Una posible redacción de los resultados de la **Figura 7.39**, sería:

*Por lo que respecta a la evidencia de la validez discriminante, la medición se proporciona en dos formas que se pueden apreciar en la Tabla yyy. Primero, con un intervalo del 95% de confidencialidad, ninguno de los elementos individuales de los factores latentes de la matriz de correlación, contiene el valor 1.0 (Anderson y Gerbing, 1988). Segundo, la varianza extraída entre cada par de constructos es superior que su correspondiente **IVE** (Fornell y Larcker, 1981). Con base a estos criterios, se puede concluir que las distintas mediciones realizadas en este estudio demuestran suficiente evidencia de fiabilidad y validez convergente y discriminante.*

## Conclusiones.

- La Asertividad en la toma de Decisiones (**ATD**) de los ejecutivos de la industria del software, tiene una mayor explicación en el conocimiento del Arte (**ART**) y de la Filosofía (**FSF**) en su componente de Pensamiento Creativo (**PC**) que de la imaginación (**IMG**). Por lo tanto, invierta menos técnicas de la Imaginación (**IMG**) e invierta más en Arte (**ART**) y Filosofía (**FSF**).

•La Asertividad en la toma de Decisiones (**ATD**) de los ejecutivos de la industria del software, tiene una mayor explicación en las Matemáticas (**MAT**) y en la Física (**FIS**) que en la Lógica (**LOG**) en su componente de Pensamiento Racional (**PR**). Por lo tanto, invierta menos en las técnicas de la Lógica (**LOG**), e invierta más en Matemáticas (**MAT**) y Física (**FIS**).

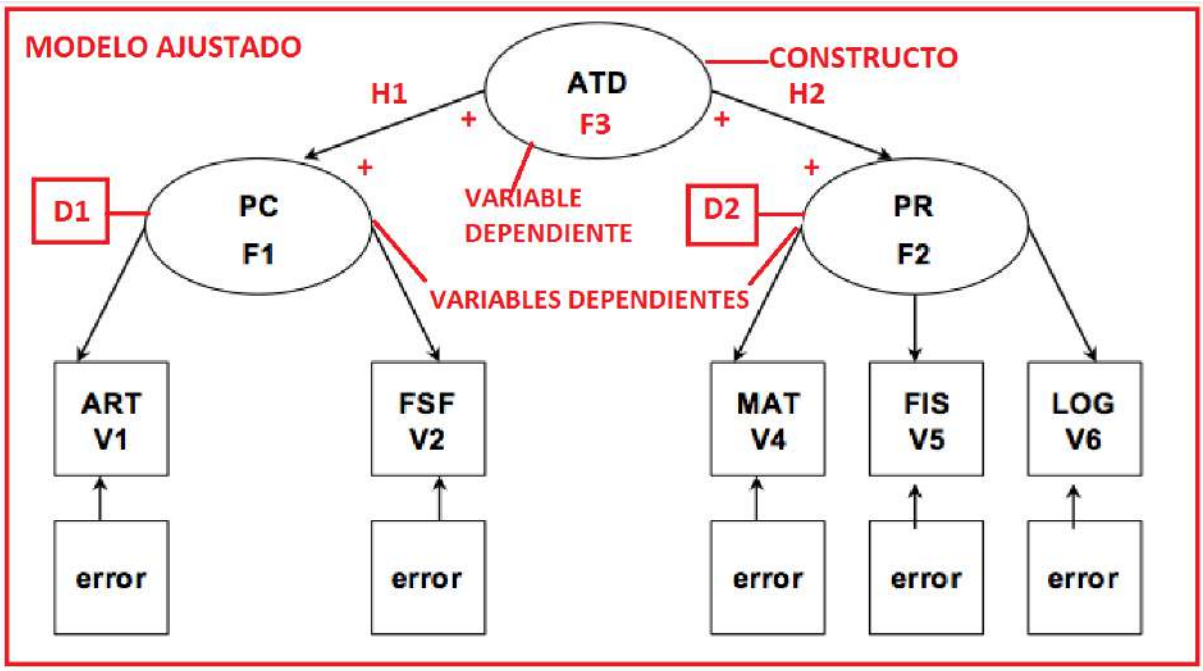
•Por lo tanto, la hipótesis: *la **ATD** de los ejecutivos de la industria del software es explicada mediante las notas del **ART**; **FSF** e **IMG** que componen al **PC**, así como las notas de **MAT; FIS** y **LOG** que componen al **PR** SI es explicada*.

Dado que existe confiabilidad y validez, se procede a realizar el análisis estructural ***(SEM.**Structural Equation Modeling**)***

## Calculando la ecuación estructural (SEM)

Recuerde que en el **CFA** se realizan análisis de confiabilidad y validez a través de correlaciones entre los ítems y los factores del constructo. Con el análisis estructural, y teniendo como base el **CFA,** ya se tiene la capacidad de realizar prueba de hipótesis en el constructo, el cual es denominado también, **modelo ajustado.** Así, partiendo de los resultados mostrados en la **Figura 7.43**, tenemos:

**Figura 7.43. Constructo o modelo ajustado de la asertividad en la toma de decisiones (ATD) para el análisis estructural**



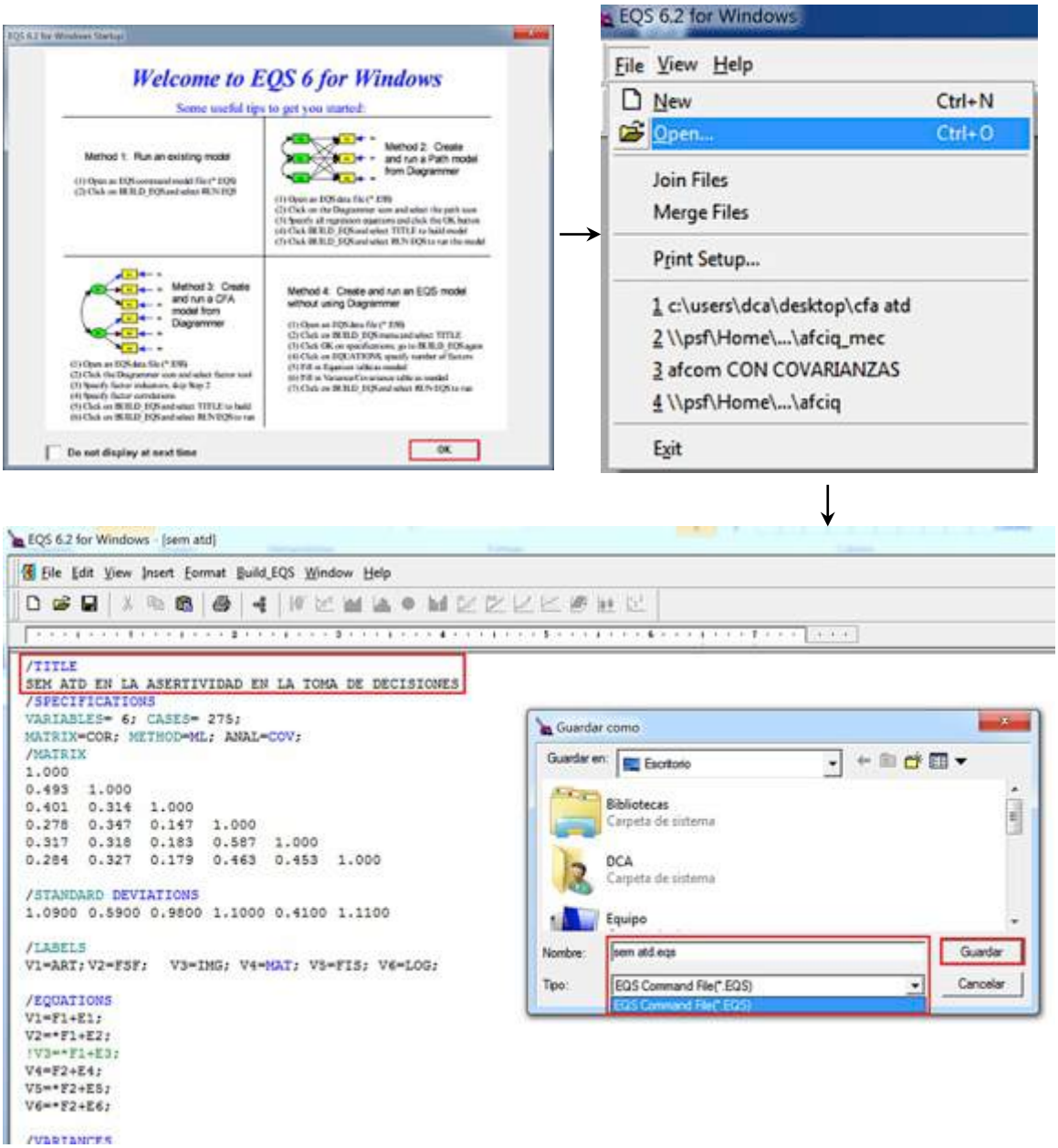

Fuente: propia.

Así, es posible plantear las siguientes hipótesis:
**H1:** *A mayor **ATD**, mayor **PC***
**H2:** *A mayor **ATD**, mayor **PR***

Para verificarlo, se requiere entrar al programa de entrada de datos de **EQS** y renombrar el archivo como **sem atd.eqs**; se sugiere asignar el título : *SEM ATD EN LA ASERTIVIDAD PAR ALA TOMA DE DECISIONES,* mediante la siguiente serie de comandos:

**Icono EQS 6 for windows->OK->File->Open->Nombre: sem atd.eqs->Tipo: EQS System Data (*ESS)->Abrir->cambio título->Guardar.** Ver **Figura 7.44**

**Figura 7.44. EQS.Secuencia de comandos para apertura de programa de entrada de datos. Análisis SEM.**




Fuente: EQS 6. con datos y adaptación propia

Así, se deberá realizar el ingreso del nuevo modelo estructural donde se debeán declarar las relaciones del modelo ajustado, tomando en cuenta los factores **F1** (**PC**)

y **F2** (**PR**) hacia el nuevo factor o constructo a medir **F3** (**ATD**); también, se deberán incluir los efectos provocados por los disturbios (**D1** y **D2**) dado que los factores **F1** (**PC**) y **F2** (**PR**) son tratados ahora como *variables dependientes* (los errores quedan implicitos en cada uno de los ítems que los conforman), que afectan a **F3** (variable independiente)**,** que ya se puede medir en el tratamiento de **SEM** (en **CFA,** No era factible medirlo).

**Nota:** recuerde que una ecuacion **SEM** de 1er orden Sí se puede medir dado que los factores tienen los ítems; por lo tanto, en una ecuacion **SEM** de segundo orden (**ATD**, se convierte en un factor de este tipo) dependerá su medición de *manera indirecta* a través de los factores primer orden. Lo anterior genera expresiones que se deben integrar al programa de entrada en el párrafo **/EQUATIONS**, tales como: **F1=*F3+D1;F2=*F3+D2 ,** por lo que se lee: *calcula el valor de **F1** respecto a **F3**, considerando el error generado por **F1**.*

Por lo anterior, recordar que sólo tienen varianza las variables independientes, y por lo tanto al revisar dicho estatus de las nuevas variables dependientes **F1(PC)** y **F2 (PR)** *éstas No varían o covarían*, por lo que se concluye que ya no debe considerarse sus varianzas en el programa de entrada **EQS;** *se deben suprimir***.** Al tener sólo un factor independiente **F3 (ATD)** y no tener otro factor para covariar, *se SUPRIME* los contenidos en el párrafo **/COVARIANCES** .Ver **Figura 7.45.**

**Figura 7.45. EQS. Ingreso de expresiones SEM al programa de entrada en los párrafos: /EQUATIONS /VARIANCES y /COVARIANCES**

```
/EQUATIONS
V1=F1+E1;
V2=*F1+E2;
!V3=*F1+E3;
V4=F2+E4;
V5=*F2+E5;
V6=*F2+E6;

F1=*F3+D1;
F2=*F3+D2;


/VARIANCES
E1=*;
E2=*;
!E3=*;
E4=*;
E5=*;
E6=*;
!F1 TO F2=*;SE CONVIRTIERON EN VARIABLES DEPENDIENTES POR LO TANTO YA NO VARIAN...
!SÓLO LO HACEN LAS VARIABLES INDEPENDIENTES EN PARES
!F3 NO VARIA YA QUE ES LA UNICA QUE EXISTE...
!D1 y D2 SE AGREGAN PORQUE SON ERRORES DE FACTORES DEPENDIENTES
D1 TO D2=*;

/COVARIANCES
!F1 TO F2=*; SE ANULA; F3 NO SE CONSIDERA PORQUE ES VARIABLE INDEPENDIENTE ÚNICA

/PRINT
FIT=ALL;
```

Fuente: EQS 6. con datos y adaptación propia

Se guardan el archivo y se corre conforme a lo indicado en la **Figura 7.14**

## Análisis de resultados

Verifique los resultados más relevantes de las secciones de reporte, tales como la **Figura 7.46.**

**Figura 7.46.EQS. Gráfico distribución estandarizada de los residuos con SEM**

```
DISTRIBUTION OF STANDARDIZED RESIDUALS

  ----------------------------------------
  !                                      !
 20-                                     -
  !                                      !
  !                                      !
  !                                      !
  !                                      !
  !                                      !          RANGE      FREQ PERCENT
 15-                                     -
  !                                      !    1   -0.5  -  --      0   0.00%
  !                                      !    2   -0.4  -  -0.5    0   0.00%
  !                                      !    3   -0.3  -  -0.4    0   0.00%
  !                                      !    4   -0.2  -  -0.3    0   0.00%
 10-                                     -    5   -0.1  -  -0.2    0   0.00%
  !                                      !    6    0.0  -  -0.1    7  46.67%
  !                    *                 !    7    0.1  -   0.0    8  53.33%
  !                 *  *                 !    8    0.2  -   0.1    0   0.00%
  !                 *  *                 !    9    0.3  -   0.2    0   0.00%
  5-                *  *                 -    A    0.4  -   0.3    0   0.00%
  !                 *  *                 !    B    0.5  -   0.4    0   0.00%
  !                 *  *                 !    C     ++  -   0.5    0   0.00%
  !                 *  *                 !    ------------------------------
  !                 *  *                 !          TOTAL        15 100.00%
  ----------------------------------------
     1  2  3  4  5  6  7  8  9  A  B  C    EACH "*" REPRESENTS    1 RESIDUALS
```

Fuente: EQS 6. con datos y adaptación propia

Donde se observa, nuevamente que los datos están adecuadamente distribuídos, ya que no existen desviaciones estándar ni a la izquierda ni a la derecha del intervalo. Esto es un indicativo de que el modelo conserva su ajuste.

Continuar revisando hasta verificar la sección de reporte de ajuste de índices (*fit indices*) de la **Figura 7.47.**

**Figura 7.47. EQS. Reporte índices de ajuste con SEM**

```
INDEPENDENCE AIC =     318.050   INDEPENDENCE CAIC =     271.882
       MODEL AIC =      -0.438          MODEL CAIC =      -9.672

CHI-SQUARE =        3.562 BASED ON       2 DEGREES OF FREEDOM
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS      0.16849

THE NORMAL THEORY RLS CHI-SQUARE FOR THIS ML SOLUTION IS          3.559.

FIT INDICES
-----------
BENTLER-BONETT     NORMED FIT INDEX =     0.989
BENTLER-BONETT NON-NORMED FIT INDEX =     0.976
COMPARATIVE FIT INDEX (CFI)         =     0.995
BOLLEN'S          (IFI) FIT INDEX   =     0.995
MCDONALD'S        (MFI) FIT INDEX   =     0.997
JORESKOG-SORBOM'S  GFI  FIT INDEX   =     0.995
JORESKOG-SORBOM'S AGFI  FIT INDEX   =     0.961
ROOT MEAN-SQUARE RESIDUAL (RMR)     =     0.017
STANDARDIZED RMR                    =     0.019
ROOT MEAN-SQUARE ERROR OF APPROXIMATION (RMSEA)    =     0.053
90% CONFIDENCE INTERVAL OF RMSEA  (       0.000,          0.142)

RELIABILITY COEFFICIENTS
------------------------
CRONBACH'S ALPHA                    =     0.706
```

Fuente: EQS 6. con datos y adaptación propia

Observe que al haber realizado todos los ajustes **SEM** al ingresar en el programa de entrada **EQS** de factores y errores así como la supresión-agregados de varianzas y supresión de covarianzas se modifica la distribución de la infomación a analizar, lo que trae como consecuencia una mejora del modelo, respecto al de **CFA,** en sus índices de ajuste >.9 <1.0 (ya no aparecen los 1), **RMSEA** <.08. Lo cual se lee: ***Chi-Cuadrada****= 3.562; gl=2; p= .16849;* ***NFI****=.989;* ***NNFI****=.976;* ***CFI****=.995;* ***RMSEA****=.053"*

Se continúa la revisión hasta llegar a la sección de reporte mostrado en la **Figura 7.48.**

**Figura 7.48. EQS. Reporte medición de las ecuaciones con SEM**

```
MEASUREMENT EQUATIONS WITH STANDARD ERRORS AND TEST STATISTICS
STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.

ART   =V1  =        1.000 F1    +      1.000 E1

FSF   =V2  =         .610*F1    +      1.000 E2
                     .099
                    6.139@

MAT   =V4  =        1.000 F2    +      1.000 E4

FIS   =V5  =         .368*F2    +      1.000 E5
                     .039
                    9.518@

LOG   =V6  =         .811*F2    +      1.000 E6
                     .095
                    8.583@
```

Fuente: EQS 6. con datos y adaptación propios.

*Las tres variables del modelo muestran altos valores de t significativos al .05*

Continúe revisando y observe la sección de reporte de la **Figura 7.49**

**Figura 7.49. EQS. Ecuaciones de constructo con errores estándar con SEM**

```
28-OCT-17        PAGE:   8  EQS      Licensee:
 TITLE:   SEM ATD EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)


 CONSTRUCT EQUATIONS WITH STANDARD ERRORS AND TEST STATISTICS
 STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.


   F1    =F1   =          .663*F3     +        1.000 D1
                           .159
                          4.164@

   F2    =F2   =          .694*F3     +        1.000 D2
                           .000
                        1.0E+38@
```

Fuente: EQS 6. con datos y adaptación propia

El cual se lee: *ésta es la medición F1-F3 con una covarianza de F1 en F3 de .663, con un error de .159 así como un valor de t alto y significativo al .05.*
Del segundo caso, se tiene la siguiente lectura: *ésta es la medición F2-F3 con una covarianza de F2 en F3 de .694, con un error de .000 así como un valor extremadamente alto de t y significativo al .05*,
Analizando más, observamos la sección del reporte de la **Figura 7.50**

**Figura 7.50. EQS. Reporte de varianzas entre las variables independientes con SEM**

```
28-OCT-17        PAGE:   9  EQS      Licensee:
 TITLE:   SEM ATD EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

 VARIANCES OF INDEPENDENT VARIABLES
 ----------------------------------
 STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.

                V                                F
                ---                             ---
                              I F3  -   F3                .805*I
                              I                           .230 I
                              I                          3.504@I
                              I                                I
```

Fuente: EQS 6. con datos y adaptación propios.

Revisar que la varianza *No contenga 1*; como es así, continuar con la revisión.
Se localiza la sección de reporte mostrada en la **Figura 7.51**

**Figura 7.51. Reporte de entre las variables independientes y factores con SEM**

```
28-OCT-17      PAGE:  10  EQS     Licensee:
 TITLE:   SEM ATD EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

 VARIANCES OF INDEPENDENT VARIABLES
 ----------------------------------
 STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.

                 E                                D
                ---                              ---
E1  -  ART                .668*I D1  -   F1                .166*I
                          .097 I                           .000 I
                         6.897@I                       1.0E+38@I
                              I                               I
E2  -  FSF                .155*I D2  -   F2                .320*I
                          .032 I                           .100 I
                         4.839@I                        3.206@I
                              I                               I
E4  -  MAT                .502*I                              I
                          .075 I                              I
                         6.718@I                              I
                              I                               I
E5  -  FIS                .072*I                              I
                          .010 I                              I
                         6.996@I                              I
                              I                               I
E6  -  LOG                .766*I                              I
                          .079 I                              I
                         9.665@I                              I
                              I                               I
```

Fuente: EQS 6. con datos y adaptación propia

Este reporte, *sólo es informativo* y se prosigue con la revisión, localizando la sección de reporte mostrada en la **Figura 7.52.**

**Figura 7.52. EQS. Reporte solución estandarizada con SEM**

```
28-OCT-17      PAGE:  11  EQS     Licensee:
 TITLE:   SEM ATD EN LA ASERTIVIDAD EN LA TOMA DE DECISIONES

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)


 STANDARDIZED SOLUTION:                                          R-SQUARED


  ART   =V1  =  .661 F1    + .750 E1                                .438
  FSF   =V2  =  .745*F1    + .667 E2                                .555
  MAT   =V4  =  .765 F2    + .644 E4                                .585
  FIS   =V5  =  .754*F2    + .656 E5                                .569
  LOG   =V6  =  .615*F2    + .789 E6                                .378
   F1   =F1  =  .825*F3    + .565 D1                                .680
   F2   =F2  =  .740*F3    + .672 D2                                .548
```

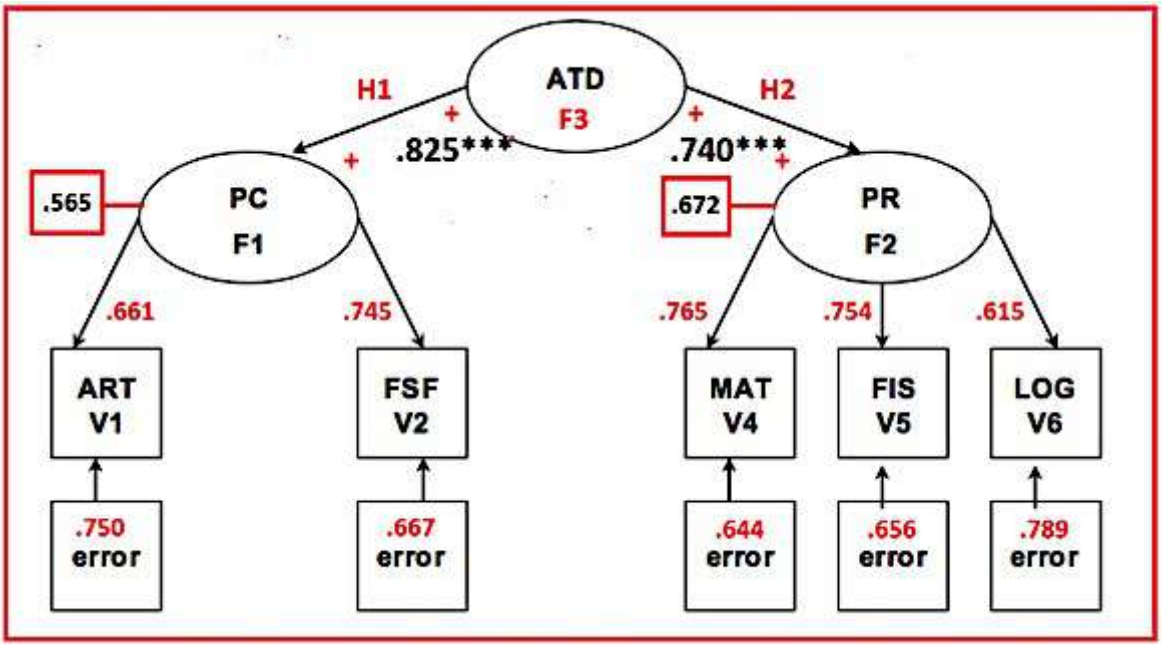
Fuente: EQS 6. con datos y adaptación propia

De esta figura, se aprecian las cargas factoriales que inciden de las variables a los factores y entre factores.
A fin de determinar la significatividad de la relación de Factores **F1-F3** y **F2-F3**,se procede a revisar la **Figura 7.49** y las consideraciones mostradas de la **Figura 7.40**, de las que se procede a realizar, lo siguiente:
-Para la relación **F1-F3** y de la **Figura 7.49,** el valor de t= 4.164> 2.58 (**Figura 7.40**), por lo tanto se asignan : tres asteriscos (***).
-Para la relación **F2-F3** y de la **Figura 7.49,** el valor de t= 1.0E+38@, representa un valor extremadamente alto que se deberá calcular a partir de calcular *el promedio de los valores t* de todas las variables relacionadas con **F2** (ver **Figura 7.48**), de tal forma que:
Valor de t de **F2-F3**= Promedio de las variables relacionadas con **F2= Promedio (V5-F2+ V6-F2)** (ver **Figura 7.48**)
Valor de t de **F2-F3=** (9.518+8.583)/2= 9.050> 2.58 (**Figura 7.40**), por lo tanto se asignan : tres asteriscos (***)

## *Discusión*

Lo anterior, nos permitirá hacer el gráfico correspondiente al diagrama de trayectorias para responder a las hipótesis planteadas como se muestra en la **Figura 7.53.**

**Figura 7.53. Constructo o modelo ajustado de la asertividad en la toma de decisiones (ATD) para el análisis estructural con valores de carga factorial SEM**




Fuente: propia

El resultado del modelo estructural, nos indica que: *para medir la **ATD**, aporta más en su explicación, el **PC** que el **PR***. Si se decide hacer inversiones al respecto, equivaldría a decir, sólo a manera ejemplificativa y guardando las debidas proporciones, que:

*Del 100% de recursos, invertir aprox. un 60% a **PC** y 40% a **PR**. De lo recursos del 60% de **PC** y tomando en cuenta la **Figura 7.53** asignar un **60%** a **FSF** y un 40% a **ART.** Del 40% de **PR**, asignar 30% a **MAT**, 30% a **FIS** y 40% de **LOG;** la $r^2$ mostradas indican una alta correlación entre cada una las variables y factores del modelo*.

Todo esto depende dl marco teórico planteado sobre la trealidad estudiada así como de la experiencia del investigador para evidenciarlo. Es interesante reflexionar que los estudios **SEM,** basados en múltiples regresiones prrmiten hacer este tipo de planteamientos incluyendo el estimado del error, que no son posibles determinar si se hicieran por separado para posteriormente conjuntarlas.

## Escribiendo la comprobación de hipótesis

Con el fin de reportar la comprobaciòn de hipótesis basados en los resultados mostrados en la **Figura 7.53**, se sugiere seguir, el formato mostrado en la **Figura 7.54.** más que los gráficos por cuestiones de espacio, en el escrito de una artículo. (*En algunas editoriales, solicitan entre 20-30 paginas; arial 10, 1, 1.5 o doble espacio o 6000-8000 palabras y/o 150 palabras de resumen, 15 referencias con opción dejar correo para reenvío de las restantes*)

**Figura 7.54. Reporte tabla del modelo de ecuaciones estructurales**

| Hipótesis | Relación Estructural | Coeficiente estandarizado | Valor t Robusto |
|---|---|---|---|
| **H1:** *A mayor **ATD**, mayor **PC*** | **ATD->PC** | .825*** | 4.164 |
| **H2:** *A mayor **ATD**, mayor **PR*** | **ATD->PR** | .740*** | 9.050 |
| Chi-Cuadrada= 3.562; gl=2; p= .16849; NFI=.989; NNFI=.976; CFI=.995; RMSEA=.053. Nota: aún hay problemas con p =.16849>.05 | | | |

Fuente: EQS 6. con datos y adaptación propia

## *Conclusión*

Una posible redacción de explicación de la **Figura 7.54**, de tener p<0.05, sería*:*

*Con respecto a la hipótesis **H1**, los resultados obtenidos (Coeficiente estandarizado = 0..825, p < 0.001), indican que la **ATD** tiene efectos significativos en el **PC** de los ejecutivos de la industria del software. En cuanto a la hipótesis H2, los resultados obtenidos (Coeficiente estandarizado = 0.740, p < 0.001), indican la **ATD** tiene efectos significativos en el **PR** de los ejecutivos de la industria del software. En conclusión se puede comprobar que los dos factores analizados tienen significancia (p < 0.001).*

La discusión tendrá que versar sobre cómo justificar los hallazgos encontrados empíricamente, el marco teórico y la realidad estudiada.

# Modelo de segundo orden. Análisis de caso de estudio 2.

Tomemos como ejemplo el concepto de *orientación al mercado* y la escala desarrollada por Narver y Slater (1990) (en Aldás-Manzano y Maldonado-Guzmán,2016) para su medición (ver **Figura 7.55**). *La orientación al mercado es latente en cuanto que no es manifiesto, esto es, directamente observable*. Además la orientación al mercado es también *variable* en cuanto que no es constante, es decir, no es fija en el tiempo. Una escala de medición de la orientación al mercado, buscará estimar la magnitud de esta variable latente en el momento y entorno geográfico actual, lo que denominamos, puntuación verdadera.

**Figura 7.55. Escala orientación al mercado de Narver y Slater (1990)**

**ORIENTACIÓN AL CLIENTE**

X1. Nos preocupamos por responder a las exigencias de los clientes.
X2. Las acciones de mi empresa van dirigidas a que el cliente obtenga más por el mismo precio.
X3. Comprendemos las necesidades de los clientes.
X4. Nos fijamos objetivos de satisfacción del cliente.
X5. Medimos el grado de satisfacción del cliente.
X6. Ofrecemos servicios post-venta.

**ORIENTACIÓN A LA COMPETENCIA**

X7. Poseemos información sobre la cuota de mercado de la competencia.
X8. El personal de ventas regularmente comparte información con nuestro negocio en relación a la estrategia de los competidores.
X9. Damos una respuesta rápida a las acciones de la competencia.
X10 La alta dirección efectúa análisis de las estrategias de la competencia.
X11 Vemos como ventajas competitivas las oportunidades de mercado

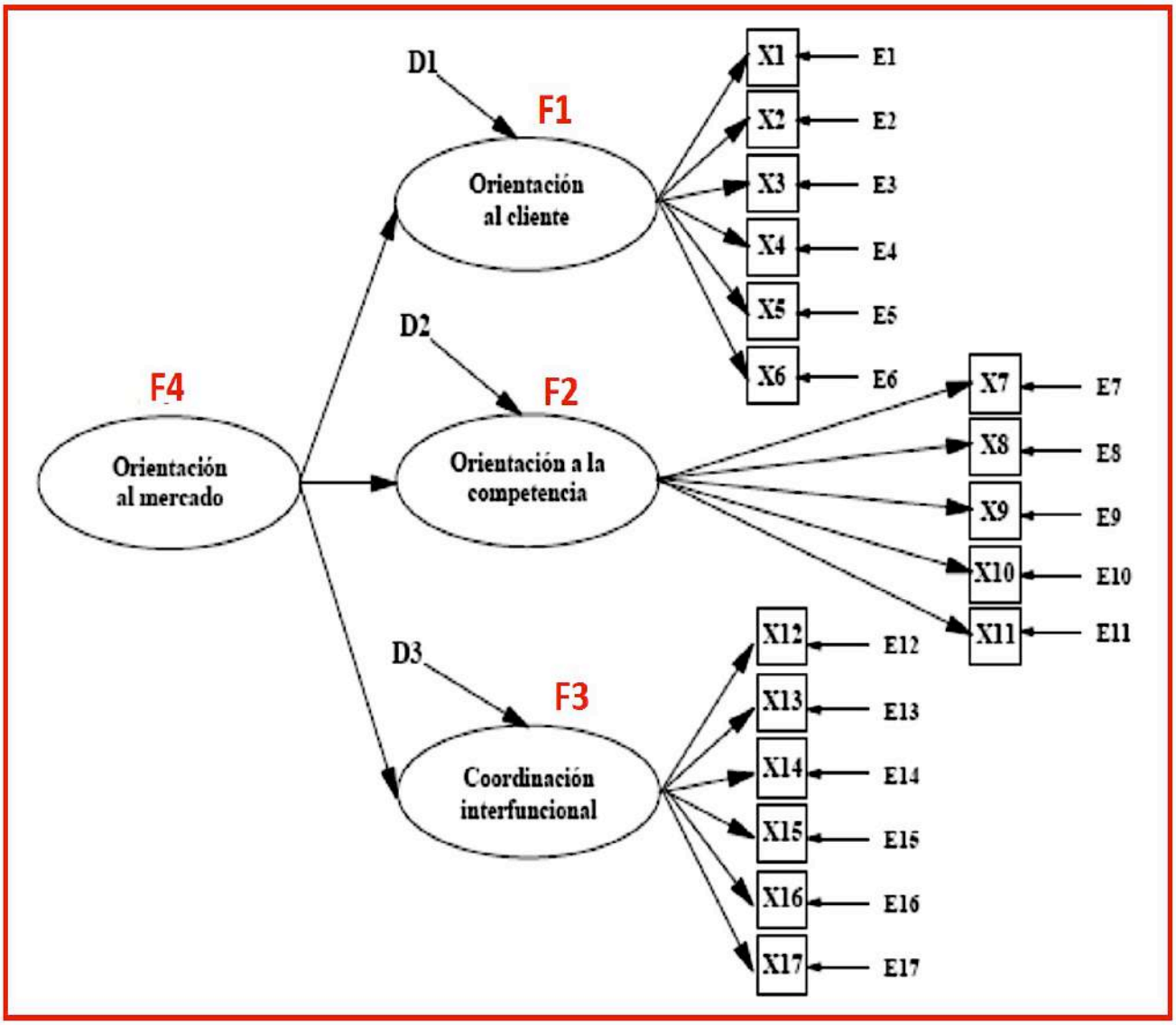
**COORDINACIÓN INTERFUNCIONAL**

X12. Los distintos departamentos tienen contacto con los clientes.
X13. Los departamentos comparten entre sí información relevante relativa a clientes o mercados.
X14. Los distintos departamentos integran sus funciones en el marco de una estrategia única general.
X15. Los distintos departamentos buscan la satisfacción de los clientes.
X16. Todos los departamentos realizan su contribución para que el cliente reciba más sin pagar más.
X17. Los departamentos de la empresa efectúan acciones de manera conjunta.

Fuente: Narver y Slater (1990) en Aldás-Manzano y Maldonado-Guzmán (2016).

La noción de *variable latente*, implica que debe existir cierta relación entre ella y los ítems que componen la escala. La variable latente debe ser, si la escala se ha desarrollado correctamente, la causa de la puntuación obtenida en los distintos ítems de la escala. Existirá una correlación entre la puntuación dada al ítem
**X1** *nos preocupamos de responder a las exigencias de nuestros clientes* y la orientación al mercado de la empresa que responde, aunque no pueda ser calculado al no ser este concepto directamente medible. La forma más común de representar las relaciones entre los distintos ítems de una escala y el concepto que miden, son los *diagramas de trayectoria*. **Ver Figura 7.56.**

**Figura 7.56. CFA. Diagrama de trayectoria del constructo orientación al mercado (Narver y Slater, 1990)**




Fuente: Narver y Slater (1990); Byrne (2006) con adaptación propia

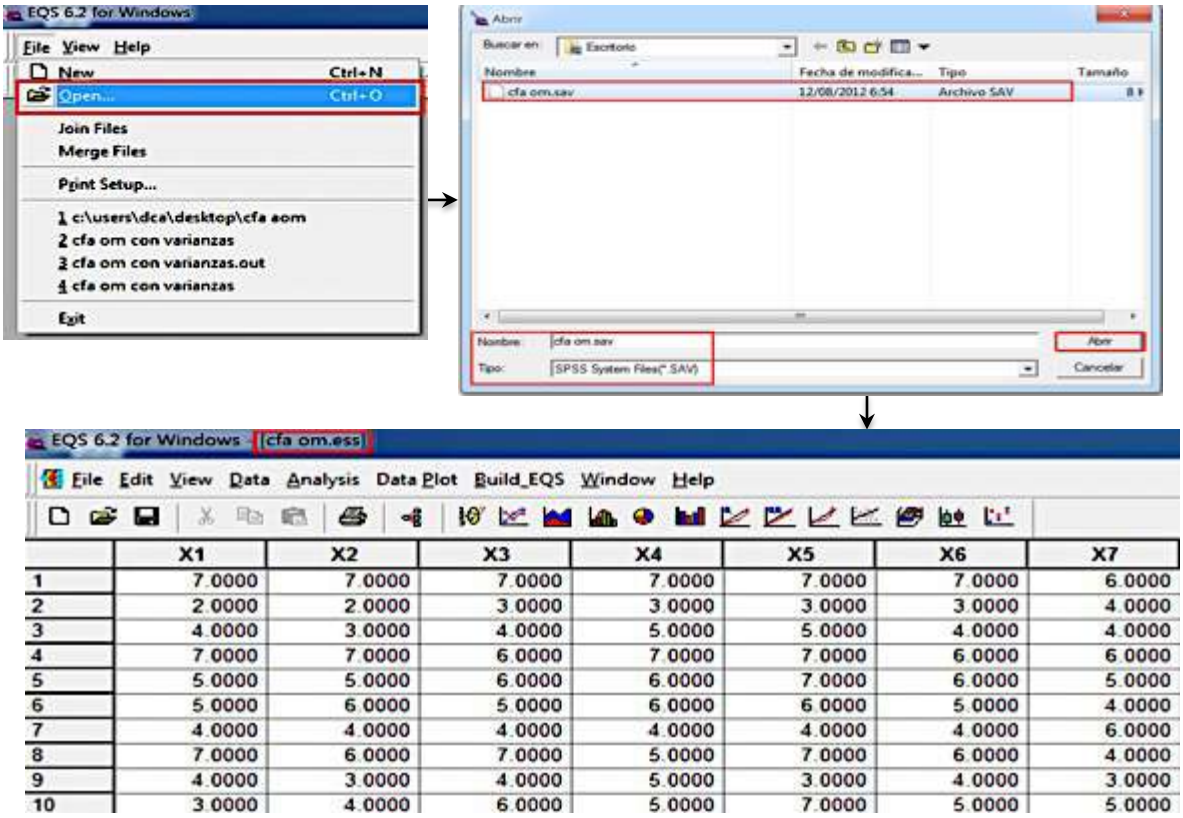
Donde :
**F1.** Orientación al cliente; **F2.** Orientación a la competencia; **F3.** Coordinación interfuncional; **F4.** Orientación al mercado.

Recuerde que en un **CFA** de *segundo orden*, los factores *de primer orden* al ser observables se miden de manera directa mientras que los de segundo orden se miden de manera indirecta**.** Dado lo anterior, al ser el factor de segundo orden una variable independiente (orientación al mercado. **F4**) no se considera con errores, sin embargo los de primer orden como orientación al cliente (**F1**); Orientación a la competencia (**F2**) y Coordinación interfuncional (**F3**), tendrán cada uno su asignación de error (**D1** a **D3**).

## *EQS. Entrada de datos al programa*

Se sugiere seguir lo recomendado en las **Figuras 7.11,** accesando el archivo de entrada de programa (***.EQS**), llamado **cfa om.sav (**formato **SPSS)**, el cual, al ser abierto, cambia su formato original ***.SAV** a ***.EES** (**cfa aom.ess**) para su despliegue y manipulación**.** Ver **Figura 7.57.**

**Figura 7.57. EQS. Selección de archivo externo para caso de estudio 2. Cambio de formato SPSS (*.SAV) a EQS (*.ESS)**

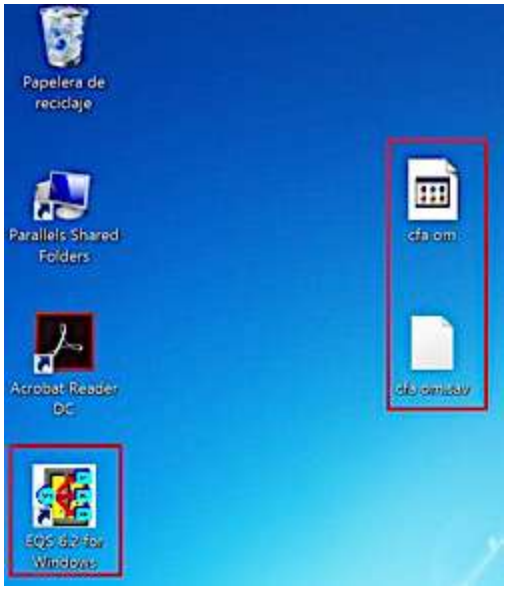



| | X1 | X2 | X3 | X4 | X5 | X6 | X7 |
|---|---|---|---|---|---|---|---|
| 1 | 7.0000 | 7.0000 | 7.0000 | 7.0000 | 7.0000 | 7.0000 | 6.0000 |
| 2 | 2.0000 | 2.0000 | 3.0000 | 3.0000 | 3.0000 | 3.0000 | 4.0000 |
| 3 | 4.0000 | 3.0000 | 4.0000 | 5.0000 | 5.0000 | 4.0000 | 4.0000 |
| 4 | 7.0000 | 7.0000 | 6.0000 | 7.0000 | 7.0000 | 6.0000 | 6.0000 |
| 5 | 5.0000 | 5.0000 | 6.0000 | 6.0000 | 7.0000 | 6.0000 | 5.0000 |
| 6 | 5.0000 | 6.0000 | 5.0000 | 6.0000 | 6.0000 | 5.0000 | 4.0000 |
| 7 | 4.0000 | 4.0000 | 4.0000 | 4.0000 | 4.0000 | 4.0000 | 6.0000 |
| 8 | 7.0000 | 6.0000 | 7.0000 | 5.0000 | 7.0000 | 6.0000 | 4.0000 |
| 9 | 4.0000 | 3.0000 | 4.0000 | 5.0000 | 3.0000 | 4.0000 | 3.0000 |
| 10 | 3.0000 | 4.0000 | 6.0000 | 5.0000 | 7.0000 | 5.0000 | 5.0000 |

Fuente: EQS 6. con datos y adaptación propia

Así también es posible verificar el archivo de programa de entrada **EQS**, del que resaltamos el nombre del archivo en el párrafo **/TITLE:** *CFA ORIENTACIÓN AL MERCADO DE SEGUNDO ORDEN.*
En el párrafo: **/SPECIFICATIONS,** la declarativa de analizar 17 variables con 300 casos.
**Nota: EQS**, para el caso de sistema operativo **WINDOWS** debe especificarse, en un archivo externo, la trayectoria del archivo de programa de entrada; por ejemplo: **DATAFILE='E://cfa om.eqs'.** Si es **IOS**, sólo se menciona el nombre del archivo dentro de la computadora a *nivel de escritorio con los archivos apareciendo en el emulador de Windows* (***Parallels,*** en nuestro caso: **DATAFILE='cfa om.eqs'**). Ver **Figura 7.58.**

**Figura 7.58. Vista de archivos desde emulador Windows (parallels) manejo de EQS con IOS**




Fuente: EQS 6. con datos y adaptación propios.

Otra instrucción es **MATRIX=RAW**, que indica que la matriz de datos será generada por el mismo software. La instrucción **METHOD=ML** que indica que se aplicará máxima verosimilitud (**ML*.Maximum Likelihood***), la instrucción **ROBUST** en la que se da el indicativo que los datos tienen filtros para garantizar mínimos errores (no se recomienda trabajar con los datos originales) y finalmente, la instrucción **ANALYSIS= COV**, sobre la aplicación de covarianzas (es mejor a la matriz de correlaciones).
En el párrafo **/EQUATIONS**, se tienen las ecuaciones de todas y cada una de las variables con sus errores señalando con la supresión de *asterisco* **(*)** *al menos una variable que tenga varianza 1*.Observe en particular las expresiones que involucran a los constructos o factores que identifican al modelo de segundo orden: **F1=*F4+D1; F2=*F4+D2; F3=*F4+D3,** en las que al incluir al **F4** en las mismas, se da la idea de que *se correrá al mismo tiempo* **CFA con SEM,** situación que es válida y que facilita el análisis, dado lo complejo del modelo.
**Nota:** *¿qué sucedería si se introdujera otro constructo de segundo orden, por*

*ejemplo innovación (**F8**) con **tres factores** como variables dependientes como proceso (**F5**), producto (**F6**) y tecnología (**F7**), que se relacionen con el modelo de orientación de mercado (**F4**)?* Se relacionaría con una flecha de **F4-F8**, se generaría hipótesis inicial de *a mayor orientación del mercado (**F4**), mayor innovación (**F8**).* Se asignarían **D5** a **D7** para calcular a **F5** a **F7.** Para su análisis dentro de **CFA EQS**, se deberían agregar además, la relación de **F5-F8;F6-F8;F7-F8**, así como los ítems que involucran a cada factor, numerados a partir del 18 en adelante (17 contenidos previamente hasta orientación de mercado (**F4**). Hay modelos de *tres factores de segundo orden*, por ejemplo competitividad (**F12**) con factores de primer orden (**F9-F10-F11**) con **D9** a **D11** con tres factores, muy complejos; se generarían hipótesis tales como (dependiendo el sentido de las flechas): *a mayor orientación de mercado mayor competitividad y a mayor innovación, mayor competitividad*. Por lo anterior, *se recomienda hacer estudios de primer orden, a partir de calcular promedios (sólo en regresiones lineales)* para calcular a **F4**; sin embargo, resta información importante.
Por cierto, un ítem sólo abona a un factor de primer orden y de forma indirecta a un factor de segundo orden; *nunca un ítem se propone a dos factores* al mismo tiempo. Cabe anotar que si se llegara el caso de tener más *de 1,000* muestras bajo estudio, en regresión lineal sería mucho más sensible que en un estudio de segundo orden, es decir, todas las relaciones son altamente significativas. En los estudios factoriales esto no afecta ya que realiza cientos de regresiones lineales al mismo tiempo. En el párrafo de **/VARIANCES** se asigna a **F4=1;** dado que no se puede medir la variable independiente **F4**, y **D1 to D3**=*, por su posicionamiento como parte de las **variables dependientes** de los factores **F1 a F3.** No se incluye el párrafo **/COVARIANCES** (ya que sólo correen como variables independientes).Es importante hacer notar que tanto los nombres de la base de datos proveniente se **SPSS** (**.SAV**), tanto el *nombre de variables y su aparición sea estrictamente como corresponde al modelo teórico***,** con lectura de izquierda a derecha, ya que esto es un a condición necesaria para que el modelo **EQS**, corra correctamente y/o subdividir las bases de datos. En otras palabras, debe haber correspondencia en nombre y orden de aparición de las variables de la base de datos, con el modelo teórico, con el fin de que **EQS** ejecute correctamente las instrucciones con las variables correspondientes y produzca los resultados esperados. **Ver Figura 7.59**

**Figura 7.59. EQS Codificación de entrada de datos caso de estudio 2**

```
/TITLE
CFA ORIENTACIÓN AL MERCADO DE SEGUNDO ORDEN.
NOMBRE DEL ARCHIVO: cfa om.eqs
/SPECIFICATIONS
VARIABLES=17; CASES=300; DATAFILE='cfa om.ess';
MATRIX=RAW; METHOD=ML, ROBUST; ANALYSIS=COV;

/EQUATIONS
V1=F1+E1;
V2=*F1+E2;
V3=*F1+E3;
V4=*F1+E4;
V5=*F1+E5;
V6=*F1+E6;

V7=F2+E7;
V8=*F2+E8;
V9=*F2+E9;
V10=*F2+E10;
V11=*F2+E11;

V12=F3+E12;
V13=*F3+E13;
V14=*F3+E14;
V15=*F3+E15;
V16=*F3+E16;
V17=*F3+E17;

F1=*F4+D1;
F2=*F4+D2;
F3=*F4+D3;
```

```
/VARIANCES
E1=*;
E2=*;
E3=*;
E4=*;
E5=*;
E6=*;
E7=*;
E8=*;
E9=*;
E10=*;
E11=*;
E12=*;
E13=*;
E14=*;
E15=*;
E16=*;
E17=*;

F4=1;
D1 TO D3=*;

/LMTEST

/PRINT
FIT=ALL;

/END
```

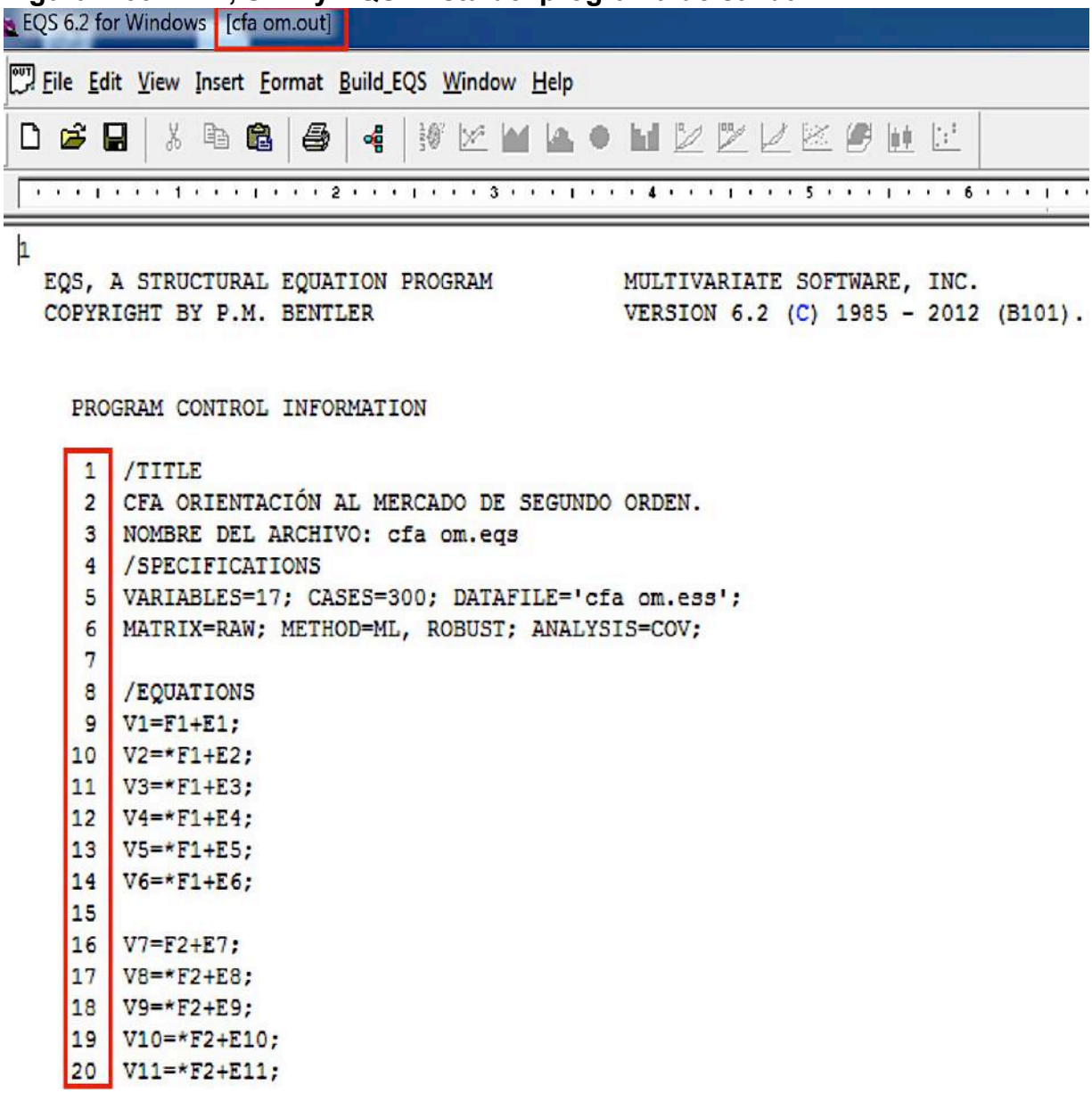
Fuente: EQS 6. con datos y adaptación propia

Realizado lo anterior, ejecute el programa **EQS**, siguiendo lo mostrado en las instrucciones del apartado de *cómo ejecutar programa.*

## *Análisis de resultados*

Se procede a revisar y analizar las diversas secciones de reporte que se presentan en el archivo: **cfa om.out.** Observe que contienen los datos del programa de entrada (**cfa om.eqs**) entrada a diferencia de que se encuentran numeradas de forma consecutiva y se anexa el reporte de resultados por secciones en el archivo de salida: **cfa om.out.** Ver **Figura 7.60**

**Figura 7.60. CFA, SEM y EQS. Vista del programa de salida**




Fuente: EQS 6. con datos y adaptación propia

Así, la primera sección de reporte, describe a los estadísticos univariados (*univariate statistics*), que muestra la media, la asimetría, la curtosis, y la desviación estandar variable por variable; asimismo, se presenta el resultado de la curtosis multivariados (*multivariate kurtosis*), que por cierto, a través del coeficiente de Mardia (*Mardia´s coefficient*) mientras más grande se presente, da el indicativo de *problemas de normalidad*. Sin embargo, esto no es de preocupar ya que el sistema **EQS,** resuelve y aporta los resultados a nivel de *solución normalizada o estandarizada* correspondiente y a consideracion del investigador, en el proceso de generación de la matriz de datos a analizar. Así también se reportan los números de casos y estimaciones. Ver **Figura 7.61**

**Figura 7.61. CFA, SEM y EQS. Reporte de estadísticos univariados, curtosis multivariada y curtosis elíptica para el caso de estudio 2.**

```
30-OCT-17        PAGE:   3  EQS     Licensee:
 TITLE:   CFA ORIENTACIÓN AL MERCADO DE SEGUNDO ORDEN.


 SAMPLE STATISTICS BASED ON COMPLETE CASES

                    UNIVARIATE STATISTICS
                    ---------------------

VARIABLE          X1          X2          X3          X4          X5
                 V1          V2          V3          V4          V5

MEAN            5.4300      5.1267      5.4200      4.9800      5.1233
SKEWNESS (G1)  -1.0011     -0.7650     -0.8937     -0.5964     -0.7291
KURTOSIS (G2)   1.0529      0.0403      0.9292     -0.1302      0.3503
STANDARD DEV.   1.3801      1.4690      1.3023      1.5560      1.4840

                  MULTIVARIATE KURTOSIS
                  ---------------------

MARDIA'S COEFFICIENT (G2,P) =     104.7620
NORMALIZED ESTIMATE =              35.6959

                ELLIPTICAL THEORY KURTOSIS ESTIMATES
                ------------------------------------

MARDIA-BASED KAPPA =       0.3243 MEAN SCALED UNIVARIATE KURTOSIS =      0.0334

MARDIA-BASED KAPPA IS USED IN COMPUTATION. KAPPA=           0.3243

CASE NUMBERS WITH LARGEST CONTRIBUTION TO NORMALIZED MULTIVARIATE KURTOSIS:
---------------------------------------------------------------------------

CASE NUMBER        158         172         191         206         223

ESTIMATE        997.7348    889.0390   1818.8817    701.6374   1114.6302
```

Fuente: EQS 6. con datos y adaptación propia

La siguiente sección de reporte, llamada matriz de covarianza (*covariance matrix*), en la que se observan las 17 variables analizadas de los 300 casos recopilados en la base de datos, que se construye en el proceso **EQS**. Ver **Figura 7.62.**

**Figura 7.62. CFA, SEM y EQS. Matriz de covarianza para el caso de estudio**

29-OCT-17 PAGE: 4 EQS Licensee:
TITLE: CFA ORIENTACIÓN AL MERCADO DE SEGUNDO ORDEN.

COVARIANCE MATRIX TO BE ANALYZED: 17 VARIABLES (SELECTED FROM 17 VARIABLES)
BASED ON 300 CASES.

| | | X1 | X2 | X3 | X4 | X5 |
|---|---|---|---|---|---|---|
| | | V1 | V2 | V3 | V4 | V5 |
| X1 | V1 | 1.905 | | | | |
| X2 | V2 | 1.256 | 2.158 | | | |
| X3 | V3 | 1.096 | 1.171 | 1.696 | | |
| X4 | V4 | 0.708 | 1.518 | 0.821 | 2.421 | |
| X5 | V5 | 0.559 | 1.078 | 0.811 | 1.350 | 2.202 |
| X6 | V6 | 0.807 | 1.311 | 0.942 | 1.511 | 1.345 |
| X7 | V7 | 0.749 | 1.108 | 0.819 | 1.129 | 0.996 |
| X8 | V8 | 0.670 | 1.214 | 0.775 | 1.406 | 1.102 |
| X9 | V9 | 0.626 | 1.239 | 0.922 | 1.552 | 1.294 |
| X10 | V10 | 0.612 | 1.245 | 0.757 | 1.440 | 1.104 |
| X11 | V11 | 0.448 | 0.728 | 0.857 | 0.858 | 0.752 |
| X12 | V12 | 0.508 | 0.970 | 0.603 | 1.074 | 0.914 |
| X13 | V13 | 0.645 | 1.298 | 0.844 | 1.323 | 1.161 |
| X14 | V14 | 0.338 | 0.883 | 0.570 | 1.085 | 0.897 |
| X15 | V15 | 0.543 | 1.175 | 0.535 | 1.392 | 0.904 |
| X16 | V16 | 0.360 | 0.903 | 0.515 | 1.116 | 0.947 |
| X17 | V17 | 0.459 | 1.017 | 0.421 | 1.191 | 1.068 |

Fuente: EQS 6. con datos y adaptación propia

Se observa la subsección de reporte estructural de Bentler-Weeks, que presenta la cantidad de variables dependientes e independientes, parámetros libres y parámetros *fijos <>0*. Ver **Figura 7.63** .

**Figura 7.63. CFA, SEM y EQS. Reporte estructural de Bentler-Weeks.**

```
BENTLER-WEEKS STRUCTURAL REPRESENTATION:

NUMBER OF DEPENDENT VARIABLES = 20
    DEPENDENT V'S :     1    2    3    4    5    6    7    8    9   10
    DEPENDENT V'S :    11   12   13   14   15   16   17
    DEPENDENT F'S :     1    2    3

NUMBER OF INDEPENDENT VARIABLES = 21
    INDEPENDENT F'S :     4
    INDEPENDENT E'S :     1    2    3    4    5    6    7    8    9   10
    INDEPENDENT E'S :    11   12   13   14   15   16   17
    INDEPENDENT D'S :     1    2    3

NUMBER OF FREE PARAMETERS =  37
NUMBER OF FIXED NONZERO PARAMETERS =  24

*** WARNING MESSAGES ABOVE, IF ANY, REFER TO THE MODEL PROVIDED.
    CALCULATIONS FOR INDEPENDENCE MODEL NOW BEGIN.

*** WARNING MESSAGES ABOVE, IF ANY, REFER TO INDEPENDENCE MODEL.
    CALCULATIONS FOR USER'S MODEL NOW BEGIN.


 3RD STAGE OF COMPUTATION REQUIRED    300104 WORDS OF MEMORY.
 PROGRAM ALLOCATED 200000000 WORDS
```

Fuente: EQS 6. con datos y adaptación propia.

Se observa la sección de reporte de matriz de covarianzas residuales (*residual covariance matrix*) y estandarizada (*standardized residual matrix*). Sólo son informativos.Ver **Figura 7.64.**

**Figura 7.64. CFA, SEM y EQS. Matriz de covarianza residual generada para el caso de estudio 2**

RESIDUAL COVARIANCE MATRIX (S-SIGMA) :

| | | X1 V1 | X2 V2 | X3 V3 | X4 V4 | X5 V5 |
|---|---|---|---|---|---|---|
| X1 | V1 | 0.000 | | | | |
| X2 | V2 | 0.354 | 0.000 | | | |
| X3 | V3 | 0.475 | 0.179 | 0.000 | | |
| X4 | V4 | -0.235 | 0.014 | -0.215 | 0.000 | |
| X5 | V5 | -0.217 | -0.160 | -0.042 | 0.057 | 0.000 |
| X6 | V6 | -0.075 | -0.096 | -0.027 | 0.040 | 0.135 |
| X7 | V7 | 0.031 | -0.039 | 0.029 | -0.070 | 0.010 |
| X8 | V8 | -0.139 | -0.077 | -0.113 | 0.057 | -0.008 |
| X9 | V9 | -0.264 | -0.182 | -0.057 | 0.067 | 0.072 |
| X10 | V10 | -0.135 | 0.053 | -0.064 | 0.195 | 0.079 |
| X11 | V11 | -0.037 | -0.045 | 0.324 | 0.050 | 0.087 |
| X12 | V12 | -0.090 | 0.017 | -0.054 | 0.078 | 0.094 |
| X13 | V13 | -0.125 | 0.070 | -0.002 | 0.039 | 0.104 |
| X14 | V14 | -0.271 | -0.089 | -0.100 | 0.070 | 0.061 |
| X15 | V15 | -0.163 | 0.047 | -0.242 | 0.213 | -0.066 |
| X16 | V16 | -0.245 | -0.063 | -0.150 | 0.107 | 0.116 |
| X17 | V17 | -0.176 | 0.004 | -0.277 | 0.132 | 0.196 |

Fuente: EQS 6. con datos y adaptación propia

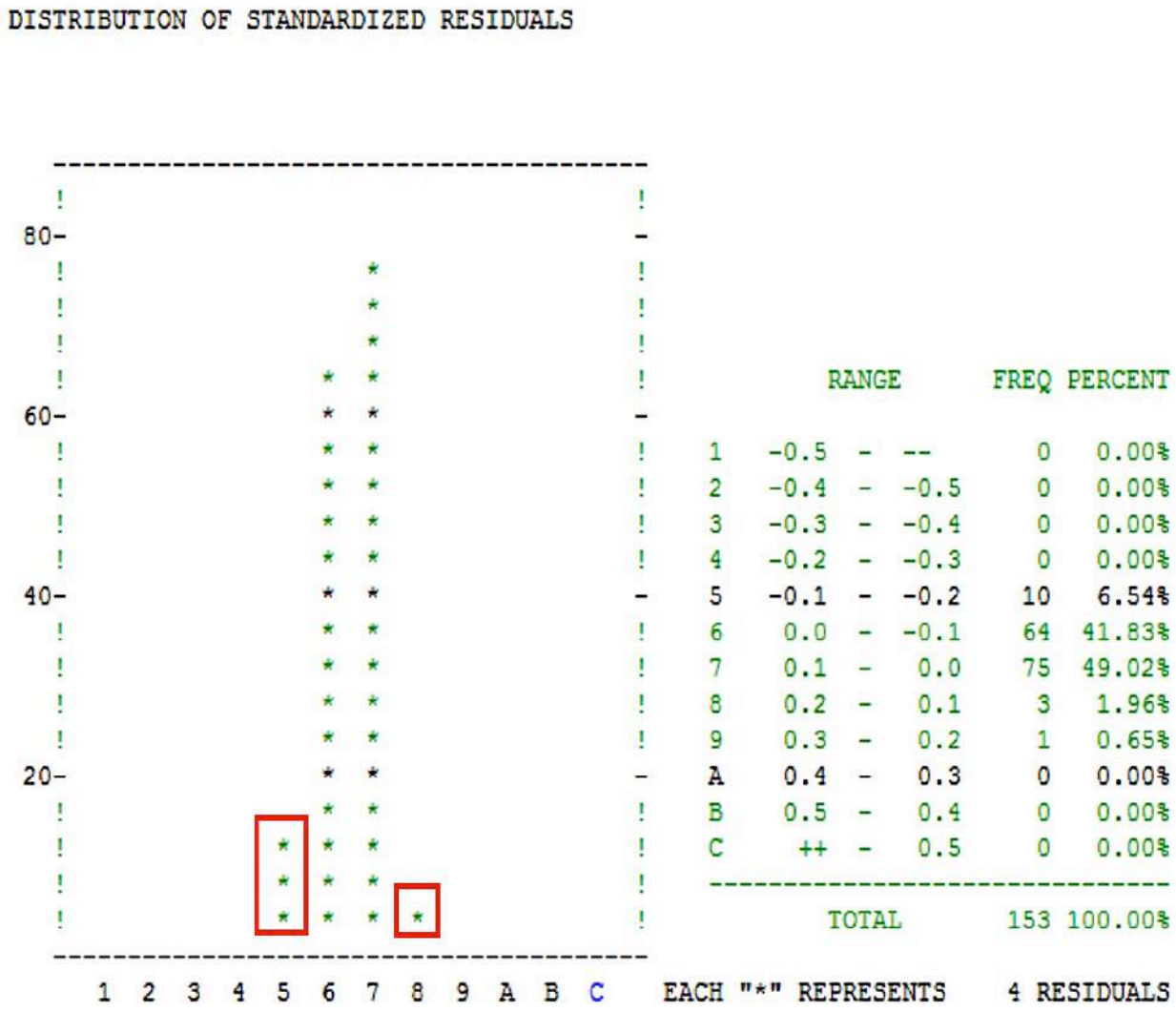
STANDARDIZED RESIDUAL MATRIX:

| | | X1 V1 | X2 V2 | X3 V3 | X4 V4 | X5 V5 |
|---|---|---|---|---|---|---|
| X1 | V1 | 0.000 | | | | |
| X2 | V2 | 0.175 | 0.000 | | | |
| X3 | V3 | 0.264 | 0.094 | 0.000 | | |
| X4 | V4 | -0.109 | 0.006 | -0.106 | 0.000 | |
| X5 | V5 | -0.106 | -0.073 | -0.022 | 0.025 | 0.000 |
| X6 | V6 | -0.037 | -0.045 | -0.014 | 0.018 | 0.062 |
| X7 | V7 | 0.015 | -0.018 | 0.015 | -0.029 | 0.004 |
| X8 | V8 | -0.064 | -0.033 | -0.055 | 0.023 | -0.004 |
| X9 | V9 | -0.113 | -0.073 | -0.026 | 0.025 | 0.029 |
| X10 | V10 | -0.063 | 0.023 | -0.032 | 0.081 | 0.034 |
| X11 | V11 | -0.020 | -0.023 | 0.181 | 0.023 | 0.042 |
| X12 | V12 | -0.046 | 0.008 | -0.029 | 0.036 | 0.045 |
| X13 | V13 | -0.058 | 0.031 | -0.001 | 0.016 | 0.045 |
| X14 | V14 | -0.146 | -0.045 | -0.057 | 0.033 | 0.030 |
| X15 | V15 | -0.079 | 0.021 | -0.123 | 0.091 | -0.030 |
| X16 | V16 | -0.133 | -0.032 | -0.087 | 0.051 | 0.059 |
| X17 | V17 | -0.091 | 0.002 | -0.151 | 0.061 | 0.094 |

Fuente: EQS 6. con datos y adaptación propia

Al revisar la sección de reporte del gráfico de distribución estandarizada, observamos que hay tres datos desviados a la izquierda de la media y uno desviado a la derecha de la media. Ver **Figura 7.65**

**Figura 7.65.CFA, SEM y EQS. Gráfico distribución estandarizada caso de estudio 2.**




Fuente: EQS 6. con datos y adaptación propia

La sección de reporte de índices de ajuste (*fit indices*) reporta que se usó el método de máxima verosimilitud (**ML.***Maximum Likelihood*) para proponer la solución normal. Así también, se observa que el resultado de Chi-Cuadrada= 554.577, con gl= 116, valor de p<=0.01 (*altamente significativo*). Así también, los índices de ajuste **NFI**=.853; **NNFI**=.859 **y CFI=.880**, cumplen si se considerara que esta escala ha sido muy utilizada de manera efectiva. Recordar que para casos de nuevas escalas, los índices de ajuste deben ser >.9. Por otro lado, el **RMSEA=** 0.112 > 0.08, *NO cumple el requerimiento***.** Por último, el *alfa de Cronbach*=.943, no es de interés a nuestro estudio ya que es el del modelo en general, **EQS**, no lo reporta a nivel de factor o

variable de manera individual ,que es lo que se necesita reportar. **Ver Figura 7.66.**

**Figura 7.66.CFA, SEM y EQS. Reporte índices de ajuste con valores normales para el caso de estudio 2**

```
GOODNESS OF FIT SUMMARY FOR METHOD = ML

INDEPENDENCE MODEL CHI-SQUARE        =    3778.176 ON   136 DEGREES OF FREEDOM

INDEPENDENCE AIC =    3506.176   INDEPENDENCE CAIC =    2866.461
       MODEL AIC =     322.577          MODEL CAIC =    -223.061

CHI-SQUARE =      554.577 BASED ON     116 DEGREES OF FREEDOM
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS      0.00000

THE NORMAL THEORY RLS CHI-SQUARE FOR THIS ML SOLUTION IS        577.028.

FIT INDICES
-----------
BENTLER-BONETT     NORMED FIT INDEX =     0.853
BENTLER-BONETT NON-NORMED FIT INDEX =     0.859
COMPARATIVE FIT INDEX (CFI)         =     0.880
BOLLEN'S          (IFI) FIT INDEX   =     0.880
MCDONALD'S        (MFI) FIT INDEX   =     0.481
JORESKOG-SORBOM'S  GFI  FIT INDEX   =     0.815
JORESKOG-SORBOM'S AGFI  FIT INDEX   =     0.756
ROOT MEAN-SQUARE RESIDUAL (RMR)     =     0.122
STANDARDIZED RMR                    =     0.060
ROOT MEAN-SQUARE ERROR OF APPROXIMATION (RMSEA)    =     0.112
90% CONFIDENCE INTERVAL OF RMSEA  (       0.103,         0.122)

RELIABILITY COEFFICIENTS
------------------------
CRONBACH'S ALPHA                    =     0.943
RELIABILITY COEFFICIENT RHO         =     0.956
```

Fuente: EQS 6. con datos y adaptación propia

Más adelante se encuentra la sección de reporte de índices de ajuste (*fit indices*), con el método robusto (*robust*).Esta sección es la que se sugiere tomar en cuenta para los análisis ya que presentan mínimos errores. Así, por ejemplo, el valor *Chi-Cuadrada de Satorra-Bentler***=** 369.7671**;** gl=116; valor p<.01 ((Nota: 369.77671/116=3.18>2 Tendencia a que los índices de ajuste sean <.9), *SI cumple con el requerimiento de significatividad*.

En el reporte, los índices de ajuste **NFI**=.851**; NNFI**=.874 **y CFI=**.892, a pesar de ser >.8, *el promedio*= (.851+.874+.892)/3=.872<.9, no cumple para la creación de una escala nueva. Si embargo, **RMSEA**=.086>.08, Sí cumple como indicador. Cabe señalar, que en la sección de método robusto (*robust*), No se encuentran presentes los índices de ajuste **GFI,AGFI;SRMR** (sólo en los normales).

Se considera así, dado que *Satorra y Bentler*, creadores de **EQS** y no consideran otros (que aparecen por ejemplo en **LISREL**, donde Jöreskog y Sorbom son los creadores), Ver **Figura 7.67.**

**Figura 7.67.CFA, SEM y EQS. Reporte índices de ajuste valores robustos para el caso de estudio 2.**

```
GOODNESS OF FIT SUMMARY FOR METHOD = ROBUST

ROBUST INDEPENDENCE MODEL CHI-SQUARE =    2488.294 ON   136 DEGREES OF FREEDOM

INDEPENDENCE AIC =    2216.294   INDEPENDENCE CAIC =    1576.580
       MODEL AIC =     137.767          MODEL CAIC =    -407.872

SATORRA-BENTLER SCALED CHI-SQUARE =    369.7671 ON     116 DEGREES OF FREEDOM
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS      0.00000

MEAN- AND VARIANCE-ADJUSTED CHI-SQUARE              =      91.624 ON    29 D.F.
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS      0.00000

MEAN-SCALED AND SKEWNESS-ADJUSTED STATISTIC         =      21.143 ON     7 D.F.
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS      0.00356

RESIDUAL-BASED TEST STATISTIC                       =   531.896
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS      0.00000

YUAN-BENTLER RESIDUAL-BASED TEST STATISTIC          =   190.995
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS      0.00001

YUAN-BENTLER RESIDUAL-BASED F-STATISTIC     =      2.822
DEGREES OF FREEDOM  =                        116,   184
PROBABILITY VALUE FOR THE F-STATISTIC IS        0.00000
```

```
FIT INDICES
-----------
BENTLER-BONETT     NORMED FIT INDEX =     0.851
BENTLER-BONETT NON-NORMED FIT INDEX =     0.874
COMPARATIVE FIT INDEX (CFI)         =     0.892
BOLLEN'S          (IFI) FIT INDEX   =     0.893
MCDONALD'S        (MFI) FIT INDEX   =     0.655
ROOT MEAN-SQUARE ERROR OF APPROXIMATION (RMSEA)    =     0.086
90% CONFIDENCE INTERVAL OF RMSEA  (       0.076,         0.095)
```

Fuente: EQS 6. con datos y adaptación propia

Como ya es de su conocimiento, las siguientes secciones de reporte, se deberán revisar que cumplan con sus condiciones de aplicabilidad. Por ejemplo, en la sección de medición de ecuaciones (*measurement equations*) en la parte de variables, se debe revisar que **l***os valores de t sean significativas*. *En las 17 variables cumple*. Ver **Figura 7.68**

**Figura 7.68. CFA, SEM y EQS. Medición de ecuaciones de medición para el caso de estudio 2**

```
MEASUREMENT EQUATIONS WITH STANDARD ERRORS AND TEST STATISTICS
STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.
(ROBUST STATISTICS IN PARENTHESES)

X1     =V1  =   1.000 F1     + 1.000 E1

X2     =V2  =   1.596*F1     + 1.000 E2
                 .166
                9.626@
             (   .218)
             (  7.320@

X3     =V3  =   1.099*F1     + 1.000 E3
                 .132
                8.335@
             (   .132)
             (  8.304@

X4     =V4  =   1.668*F1     + 1.000 E4
                 .174
                9.561@
             (   .283)
             (  5.903@
```

Fuente: EQS 6. con datos y adaptación propia

Al revisar la sección de reporte de ecuaciones de constructo (*construct equations*), en su relación entre los factores, debe cumplir con *valores t* significativos. En este caso cumple en la relación de **F1-F4; F2-F4; F3-F4.** Ver **Figura 7.69.**

**Figura 7.69. CFA, SEM y EQS. Medición de ecuaciones de constructo para el caso de estudio 2**

```
30-OCT-17        PAGE:  10  EQS      Licensee:
 TITLE:   CFA ORIENTACIÓN AL MERCADO DE SEGUNDO ORDEN.

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)


 CONSTRUCT EQUATIONS WITH STANDARD ERRORS AND TEST STATISTICS
 STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.
 (ROBUST STATISTICS IN PARENTHESES)


   F1    =F1   =     .675*F4     + 1.000 D1
                     .074
                    9.181@
                (    .097)
                (   6.971@

   F2    =F2   =    1.064*F4     + 1.000 D2
                     .077
                   13.821@
                (    .077)
                ( 13.802@

   F3    =F3   =     .885*F4     + 1.000 D3
                     .072
                   12.270@
                (    .077)
                ( 11.475@
```

Fuente: EQS 6. con datos y adaptación propia

Un dato importante, normalmente a verificar es la covarianza entre variables independientes (*covariance among independent variables*), que en este caso de estudio *No se presentan ya que es sólo existe una variable independiente **F4**.*
Avanzando en el análisis de los reportes, encontramos el de soluciones estandarizadas (*standardized solutions*), del cual, se realiza la pregunta:
*¿Son todas las cargas factoriales >.6?* . Para nuestro caso de estudio, se observa que *No cumple en* **V1**= .545 **y** $\mathbf{r^2}$= .297, lo que significa que este ítem aporta muy poco al modelo de *orientación al mercado***.** Además, si se observan las cargas

factoriales **F1**=.898**; F2**=.858**; F3**=.868**,** y sus $\mathbf{r^2}$ de**:** .807; .736; .753, son muy similares, no tienen una diferencia notable, lo que se traduce que *las 3 orientaciones son igual de importantes para explicar el modelo de orientación al mercado*. Ver **Figura 7.70**

**Figura 7.70. CFA, SEM y EQS. EQS. Reporte solución estandarizada**

```
30-OCT-17      PAGE:  14  EQS     Licensee:
 TITLE:   CFA ORIENTACIÓN AL MERCADO DE SEGUNDO ORDEN.

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)


 STANDARDIZED SOLUTION:                                          R-SQUARED


 X1    =V1  =   .545 F1    + .839 E1                                .297
 X2    =V2  =   .817*F1    + .577 E2                                .667
 X3    =V3  =   .634*F1    + .773 E3                                .403
 X4    =V4  =   .806*F1    + .592 E4                                .649
 X5    =V5  =   .695*F1    + .719 E5                                .484
 X6    =V6  =   .806*F1    + .592 E6                                .649
 X7    =V7  =   .819 F2    + .574 E7                                .671
 X8    =V8  =   .880*F2    + .475 E8                                .774
 X9    =V9  =   .908*F2    + .419 E9                                .824
 X10   =V10 =   .833*F2    + .554 E10                               .693
 X11   =V11 =   .609*F2    + .793 E11                               .371
 X12   =V12 =   .727 F3    + .687 E12                               .529
 X13   =V13 =   .846*F3    + .533 E13                               .716
 X14   =V14 =   .771*F3    + .637 E14                               .595
 X15   =V15 =   .802*F3    + .597 E15                               .643
 X16   =V16 =   .776*F3    + .631 E16                               .602
 X17   =V17 =   .772*F3    + .635 E17                               .596
  F1   =F1  =   .898*F4    + .440 D1                                .807
  F2   =F2  =   .858*F4    + .514 D2                                .736
  F3   =F3  =   .868*F4    + .497 D3                                .753
```

Fuente: EQS 6. con datos y adaptación propia

Sobre la sección de reporte de multiplicadores de Lagrange multivariado (*multivariate Lagrange multiplier*), tenemos las siguientes lecturas. *que* **V10***, debiera impactar en* **F3;** que **V1** a **F4; V3** a **F3; V7** a **F3; V2** a **F2; V7** a **F4** , de hacerlo, se incrementarían las Chi-Cuadradas en: 18.555; 16.399; 12.466; etc.

Así, se debe buscar *No aumentar las Chi-Cuadradas, sino disminuirlas, ya que disminuye el error, por lo que aumentan los índices de ajuste. Se debe buscar disminuir el error, es decir, la Chi-Cuadrada*.
Cabe hacer notar que las sugerencias del reporte multiplicador de multivariado, son sólo estadísticas, *No proceden si sugieren orientar a otro factor diferente del modelo*. Ver **Figura 7.71.**

**Figura 7.71. CFA, SEM y EQS. Test de los multiplicadores de Lagrange**

```
30-OCT-17      PAGE:  16  EQS     Licensee:
 TITLE:   CFA ORIENTACIÓN AL MERCADO DE SEGUNDO ORDEN.

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)


 MULTIVARIATE LAGRANGE MULTIPLIER TEST BY SIMULTANEOUS PROCESS IN STAGE  1

 PARAMETER SETS (SUBMATRICES) ACTIVE AT THIS STAGE ARE:

    PVV PFV PFF PDD GVV GVF GFV GFF BVF BFF
```

| | CUMULATIVE MULTIVARIATE STATISTICS | | | | UNIVARIATE INCREMENT | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| STEP | PARAMETER | CHI-SQUARE | D.F. | PROB. | CHI-SQUARE | PROB. | HANCOCK'S SEQUENTIAL D.F. | PROB. | ROBUST PREDICTED RMSEA | CFI |
| 1 | V10,F3 | 18.555 | 1 | 0.000 | 18.555 | 0.000 | 116 | 1.000 | 0.083 | 0.900 |
| 2 | V1,F4 | 34.954 | 2 | 0.000 | 16.399 | 0.000 | 115 | 1.000 | 0.080 | 0.906 |
| 3 | V3,F3 | 47.420 | 3 | 0.000 | 12.466 | 0.000 | 114 | 1.000 | 0.079 | 0.911 |
| 4 | V7,F3 | 58.171 | 4 | 0.000 | 10.751 | 0.001 | 113 | 1.000 | 0.077 | 0.915 |
| 5 | V2,F2 | 66.529 | 5 | 0.000 | 8.358 | 0.004 | 112 | 1.000 | 0.076 | 0.918 |
| 6 | V7,F4 | 70.665 | 6 | 0.000 | 4.137 | 0.042 | 111 | 1.000 | 0.076 | 0.920 |

```
*** NOTE *** IF PREDICTED RMSEA COULD NOT BE CALCULATED, 99.999 IS PRINTED.
             IF PREDICTED  CFI  COULD NOT BE CALCULATED,  9.999 IS PRINTED.
```

Fuente: EQS 6. con datos y adaptación propia

## Eliminando V1

Así y dado que la variable **V1**= .545 (**Figura 7.72)** aporta escasamente al modelo y de que se requiere además, disminuir Chi-Cuadrada, para que disminuya el error, se sugiere retirar del modelo **EQS** a la variable **V1**; cambiar el *asterisco* **(*)** a **V2** *para conservar los requerimientos de análisis de EQS de que una de las variables de cada factor debe tener varianza 1.* Ver **Figura 7.72**

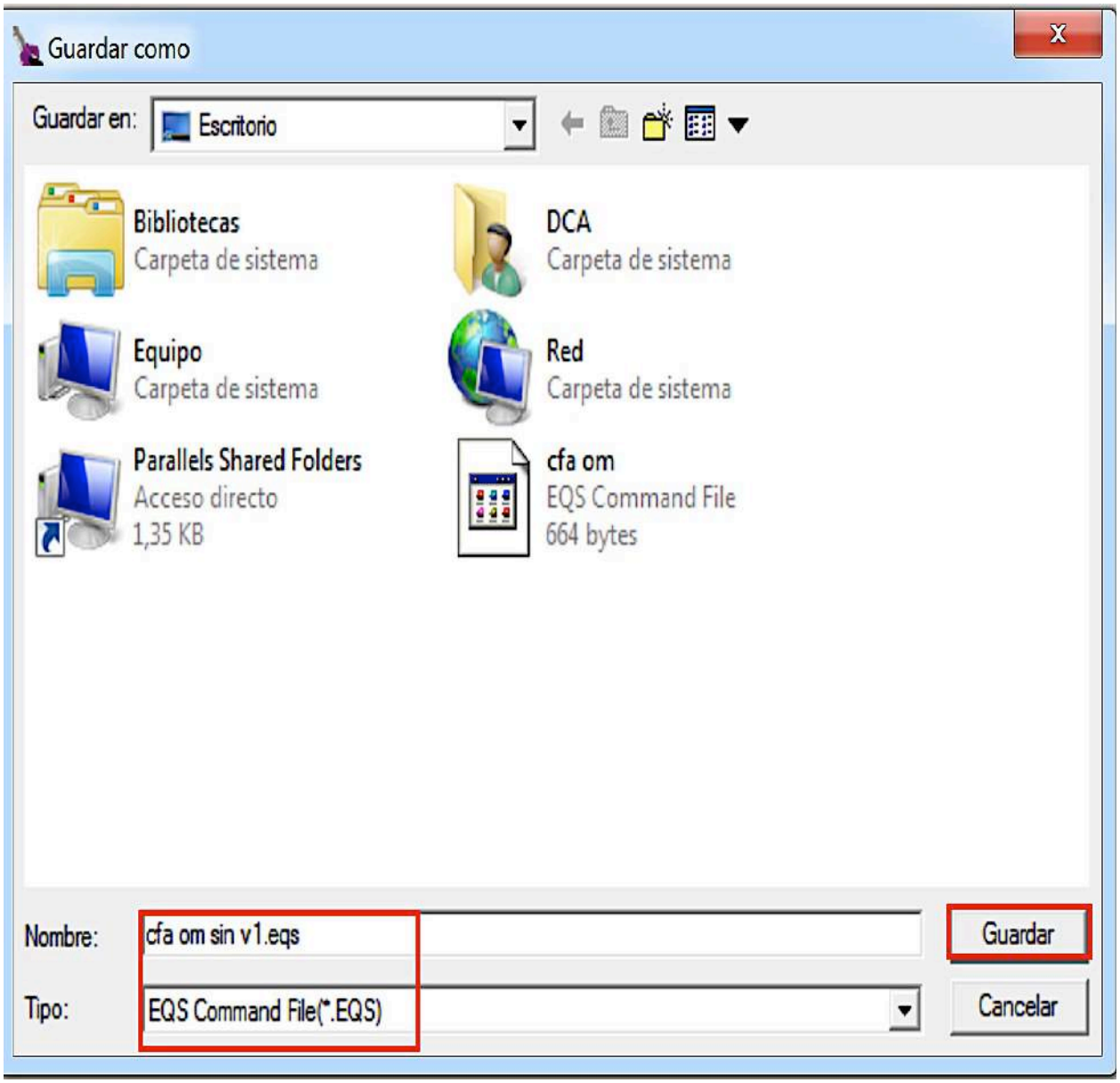
**Figura 7.72. CFA, SEM y EQS Recodificación de entrada de datos caso de estudio 2. Sin V1.**

```
/TITLE
CFA ORIENTACIÓN AL MERCADO DE SEGUNDO ORDEN.SIN V1
NOMBRE DEL ARCHIVO: cfa om sin v1.eqs
/SPECIFICATIONS
VARIABLES=17; CASES=300; DATAFILE='cfa om.ess';
MATRIX=RAW; METHOD=ML, ROBUST; ANALYSIS=COV;

/EQUATIONS
!V1=F1+E1;
V2= F1+E2;
V3=*F1+E3;
V4=*F1+E4;
V5=*F1+E5;
V6=*F1+E6;

V7=F2+E7;
V8=*F2+E8;
V9=*F2+E9;
V10=*F2+E10;
V11=*F2+E11;

V12=F3+E12;
V13=*F3+E13;
V14=*F3+E14;
V15=*F3+E15;
V16=*F3+E16;
V17=*F3+E17;

F1=*F4+D1;
F2=*F4+D2;
F3=*F4+D3;

/VARIANCES
!E1=*;
```

Fuente: EQS 6, con datos y adaptación propia

y realizar la ejecución del nuevo modelo, para analizar los resultados. En nuestro caso se renombra el archivo **eqs** como: **cfa om sin v1.eqs.** Ver **Figura 7.73**

**Figura 7.73.CFA, SEM y EQS. Renombrando archivo de entrada caso de estudio 2. Sin V1**




Fuente: EQS 6, con datos y adaptación propia

Se guardan el archivo y se corre conforme a lo indicado en la **Figura 7.14**

Se procede a revisar rápidamente las secciones de reporte de resultados como se ve en la **Figura 7.74**

**Figura 7.74. CFA, SEM y EQS. Proceso de revisión rápida de resultados caso estudio 2. Sin V1.**

```
EQS, A STRUCTURAL EQUATION PROGRAM          MULTIVARIATE SOFTWARE, INC.
COPYRIGHT BY P.M. BENTLER                   VERSION 6.2 (C) 1985 - 2012 (B101).

 PROGRAM CONTROL INFORMATION

  1  /TITLE
  2  CFA ORIENTACIÓN AL MERCADO DE SEGUNDO ORDEN.SIN V1
  3  NOMBRE DEL ARCHIVO: cfa om sin v1.eqs
  4  /SPECIFICATIONS
  5  VARIABLES=17; CASES=300; DATAFILE='cfa om.ess';
  6  MATRIX=RAW; METHOD=ML, ROBUST; ANALYSIS=COV;
  7
  8  /EQUATIONS
  9  !V1=F1+E1;
 10  V2= F1+E2;
 11  V3=*F1+E3;
 12  V4=*F1+E4;
 13  V5=*F1+E5;
 14  V6=*F1+E6;
 15
 16  V7=F2+E7;
 17  V8=*F2+E8;
 18  V9=*F2+E9;
 19  V10=*F2+E10;
 20  V11=*F2+E11;
 21
 22  V12=F3+E12;
 23  V13=*F3+E13;
 24  V14=*F3+E14;
 25  V15=*F3+E15;
 26  V16=*F3+E16;
 27  V17=*F3+E17;
```

```
DISTRIBUTION OF STANDARDIZED RESIDUALS

 80-
 60-
 40-
 20-
   1 2 3 4 5 6 7 8 9 A B C   EACH "*" REPRESENTS   4 RESIDUALS

                    RANGE        FREQ PERCENT
  1  -0.5  -  --       0   0.00%
  2  -0.4  -  -0.5     0   0.00%
  3  -0.3  -  -0.4     0   0.00%
  4  -0.2  -  -0.3     0   0.00%
  5  -0.1  -  -0.2     3   2.21%
  6   0.0  -  -0.1    68  50.00%
  7   0.1  -   0.0    62  45.59%
  8   0.2  -   0.1     3   2.21%
  9   0.3  -   0.2     0   0.00%
  A   0.4  -   0.3     0   0.00%
  B   0.5  -   0.4     0   0.00%
  C    ++  -   0.5     0   0.00%
                TOTAL     136 100.00%
```

```
GOODNESS OF FIT SUMMARY FOR METHOD = ROBUST

ROBUST INDEPENDENCE MODEL CHI-SQUARE =    2344.224 ON   120 DEGREES OF FREEDOM

INDEPENDENCE AIC =    2104.224   INDEPENDENCE CAIC =    1539.770
       MODEL AIC =      84.115          MODEL CAIC =    -390.967

SATORRA-BENTLER SCALED CHI-SQUARE =    286.1146 ON     101 DEGREES OF FREEDOM
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS      0.00000

FIT INDICES
-----------
BENTLER-BONETT      NORMED FIT INDEX =      0.878
BENTLER-BONETT NON-NORMED FIT INDEX =       0.901
COMPARATIVE FIT INDEX (CFI)          =      0.917
BOLLEN'S            (IFI) FIT INDEX  =      0.917
MCDONALD'S          (MFI) FIT INDEX  =      0.735
ROOT MEAN-SQUARE ERROR OF APPROXIMATION (RMSEA)     =      0.078
90% CONFIDENCE INTERVAL OF RMSEA  (       0.068,          0.089)
```

Fuente: EQS 6. con datos y adaptación propia

CONSTRUCT EQUATIONS WITH STANDARD ERRORS AND TEST STATISTICS
STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.
(ROBUST STATISTICS IN PARENTHESES)

```
F1   =F1  =   1.059*F4    + 1.000 D1
               .076
             13.890@
           (   .077)
           ( 13.662@ )

F2   =F2  =   1.060*F4    + 1.000 D2
               .077
             13.814@
           (   .076)
           ( 14.010@ )

F3   =F3  =    .887*F4    + 1.000 D3
               .072
             12.319@
           (   .077)
           ( 11.505@ )
```

MEASUREMENT EQUATIONS WITH STANDARD ERRORS AND TEST STATISTICS
STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.
(ROBUST STATISTICS IN PARENTHESES)

```
X2   =V2  =   1.000 F1    + 1.000 E2

X3   =V3  =    .673*F1    + 1.000 E3
               .064
             10.444@
           (   .055)
           ( 7.934@ )

X4   =V4  =   1.109*F1    + 1.000 E4
               .073
             15.241@
           (   .075)
           ( 14.699@ )

X5   =V5  =    .914*F1    + 1.000 E5
               .071
             12.788@
           (   .055)
           ( 16.753@ )
```

30-OCT-17 PAGE: 14 EQS Licensee:
TITLE: CFA ORIENTACIÓN AL MERCADO DE SEGUNDO ORDEN.SIN V1

MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

STANDARDIZED SOLUTION:

| | | | | | R-SQUARED |
|---|---|---|---|---|---|
| X2 | =V2 | = | .786 F1 | + .618 E2 | .618 |
| X3 | =V3 | = | .597*F1 | + .802 E3 | .357 |
| X4 | =V4 | = | .823*F1 | + .568 E4 | .677 |
| X5 | =V5 | = | .711*F1 | + .703 E5 | .506 |
| X6 | =V6 | = | .813*F1 | + .583 E6 | .660 |
| X7 | =V7 | = | .818 F2 | + .576 E7 | .668 |
| X8 | =V8 | = | .879*F2 | + .476 E8 | .773 |
| X9 | =V9 | = | .909*F2 | + .417 E9 | .826 |
| X10 | =V10 | = | .833*F2 | + .553 E10 | .694 |
| X11 | =V11 | = | .609*F2 | + .793 E11 | .370 |
| X12 | =V12 | = | .726 F3 | + .687 E12 | .528 |
| X13 | =V13 | = | .845*F3 | + .535 E13 | .714 |
| X14 | =V14 | = | .772*F3 | + .635 E14 | .596 |
| X15 | =V15 | = | .802*F3 | + .598 E15 | .643 |
| X16 | =V16 | = | .777*F3 | + .630 E16 | .604 |
| X17 | =V17 | = | .773*F3 | + .634 E17 | .598 |
| F1 | =F1 | = | .917*F4 | + .400 D1 | .840 |
| F2 | =F2 | = | .856*F4 | + .517 D2 | .733 |
| F3 | =F3 | = | .870*F4 | + .493 D3 | .757 |

LAGRANGE MULTIPLIER TEST (FOR ADDING PARAMETERS)

ORDERED UNIVARIATE TEST STATISTICS:

| NO | CODE | PARAMETER | CHI-SQUARE | PROB. | HANCOCK 101 DF PROB. | PARAM. CHANGE | STANDAR-DIZED CHANGE | ROBUST PREDICTED RMSEA | CFI |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 20 | V10,F3 | 18.491 | 0.000 | 1.000 | 0.406 | 0.257 | 0.075 | 0.925 |
| 2 | 2 20 | V7,F3 | 16.678 | 0.000 | 1.000 | -0.387 | -0.250 | 0.075 | 0.924 |
| 3 | 2 12 | V10,F4 | 13.646 | 0.000 | 1.000 | 0.564 | 0.364 | 0.076 | 0.922 |
| 4 | 2 20 | V3,F3 | 6.572 | 0.010 | 1.000 | -0.328 | -0.247 | 0.077 | 0.919 |
| 5 | 2 12 | V7,F4 | 6.472 | 0.011 | 1.000 | -0.389 | -0.257 | 0.078 | 0.919 |
| 6 | 2 20 | V10,F1 | 6.460 | 0.011 | 1.000 | 0.241 | 0.135 | 0.078 | 0.919 |
| 7 | 2 20 | V4,F3 | 3.669 | 0.055 | 1.000 | 0.250 | 0.158 | 0.078 | 0.918 |
| 8 | 2 20 | V13,F2 | 3.019 | 0.082 | 1.000 | 0.138 | 0.072 | 0.078 | 0.918 |
| 9 | 2 12 | V3,F4 | 2.828 | 0.093 | 1.000 | -0.483 | -0.371 | 0.078 | 0.918 |
| 10 | 2 20 | V6,F3 | 2.524 | 0.112 | 1.000 | -0.195 | -0.132 | 0.078 | 0.917 |
| 11 | 2 12 | V4,F4 | 2.478 | 0.115 | 1.000 | 0.479 | 0.308 | 0.078 | 0.917 |

Fuente: EQS 6. con datos y adaptación propia

## Discusión rápida de resultados, sin V1.

De los reportes emitidos, podemos observar y concluir de cada sección de reporte lo siguiente:

- *Gráfico de distribución de residuos* (*distribution of standardized residuals*). *Cumple.*
- *índices de ajuste* (*fit indices*) robustos (*robust*),valores a analizar:

Chi-Cuadrada de Satorra-Bentler= 286.1146; gl= 101 (**Nota:** 286.1146/101=2.83>2 Tendencia a que alguno de los índices de ajuste sea <.9); valor p<0.01**.** *Cumple.*
**NFI**= .878; **NNFI**= .901 ; **CFI**= .917. *No Cumple en* **NFI. RMSEA**=.078<.08. *Cumple.*

- Ecuaciones de medición (*measurment equations),*valores de t (@) =.05.*Cumple.*
- Ecuaciones de constructo (*construct equations*), valores de t (@)= .05.*Cumple.*
- Varianzas de las variables independientes (*variances of independent variables*), valores de t (@)= .05. *Cumple.*
- Solución estandarizada (*standardized solutions*), *¿todas las cargas factoriales son >.6?.* En **V3**=.597 . Recuerde que Hair (et. al, 1999) afirma que el promedio de todas las cargas factoriales >*.6* vs. Bagozzi y Yi (1988) que cada carga debe ser factorial debe ser >.6. Tomando en cuenta este último criterio, *NO cumple.*
- Prueba del multiplicador de Lagrange ***(****Lagrange multiplier test*), se observa la importancia del ajuste estadístico del reacomodo de variables y no se justifica hacer reacomodo de variables, como sugiere la prueba del multiplicador de Lagrange, cambiar **V3** a **F3** , *No se debe proceder si se hizo un EFA donde* ***V3*** *abona a* ***F1****, a menos que tenga un sustento teórico y evidencia suficiente para explicarlo. Por lo tanto, considerar la posible supresión de variables, que tengan un Chi-Cuadrado bajo para bajar el error y aumentar los índices de ajuste, como en nuestro caso No cumple en* ***NFI*** *(.878)*.

Así, se percibe que si se anula **V3** con su *error* **E3**, se obtendría una mejora del índice de ajuste. Por otro lado, recuerde que sólo se recomienda *suprimir un máximo del 10% de los indicadores,* ya que esto da aún la posibilidad de explicar el modelo sin mayores problemas, de lo contrario, se estaría obteniendo un *modelo ateórico*.

## Eliminando V3

Así y dado que la variable **V3**= .597 (**Figura 7.76)** aporta escasamente al modelo y de que se requiere además, disminuir *Chi-Cuadrada*, para que disminuya el **error**, se sugiere retirar del modelo **EQS** a la variable **V3;** *se conserva el asterisco* **(*)** a **V2** para conservar los requerimientos de análisis de **EQS** de que una de las variables de cada factor debe tener varianza 1. Ver **Figura 7.75.**

**Figura 7.75. CFA, SEM y EQS Recodificación de entrada de datos caso de estudio 2. Sin V1 y V3**

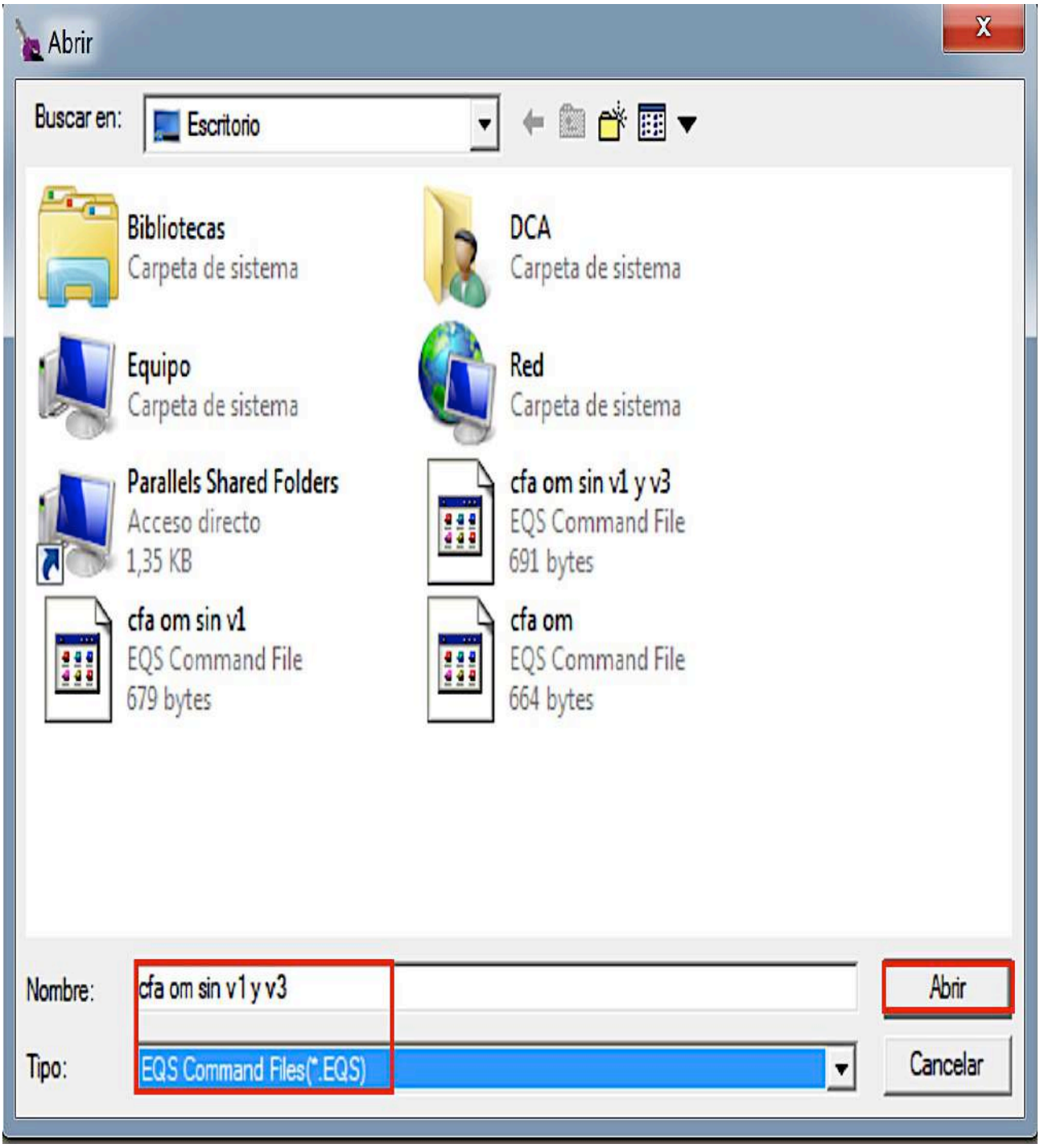
```
/TITLE
CFA ORIENTACIÓN AL MERCADO DE SEGUNDO ORDEN.SIN V1 y V3
NOMBRE DEL ARCHIVO: cfa om sin v1 y v3.eqs
/SPECIFICATIONS
VARIABLES=17; CASES=300; DATAFILE='cfa om.ess';
MATRIX=RAW; METHOD=ML, ROBUST; ANALYSIS=COV;

/EQUATIONS
!V1=F1+E1;
V2= F1+E2;
!V3=*F1+E3;
V4=*F1+E4;
V5=*F1+E5;
V6=*F1+E6;

V7=F2+E7;
V8=*F2+E8;
V9=*F2+E9;
V10=*F2+E10;
V11=*F2+E11;

V12=F3+E12;
V13=*F3+E13;
V14=*F3+E14;
V15=*F3+E15;
V16=*F3+E16;
V17=*F3+E17;

F1=*F4+D1;
F2=*F4+D2;
F3=*F4+D3;

/VARIANCES
!E1=*;
!E2=*;
```

```
/VARIANCES
E1=*;
E2=*;
E3=*;
E4=*;
E5=*;
E6=*;
E7=*;
E8=*;
E9=*;
E10=*;
E11=*;
E12=*;
E13=*;
E14=*;
E15=*;
E16=*;
E17=*;

F4=1;
D1 TO D3=*;

/LMTEST

/PRINT
FIT=ALL;

/END
```

Fuente: EQS 6. con datos y adaptación propios.

Y realizar la ejecución del nuevo modelo, para analizar los resultados. En nuestro caso se renombra el archivo **eqs** como: **cfa om sin v1** y **v3.eqs. Ver Figura 7.76**

**Figura 7.76.CFA, SEM y EQS. Renombrando archivo de entrada caso de estudio 2. Sin V1 y V3**



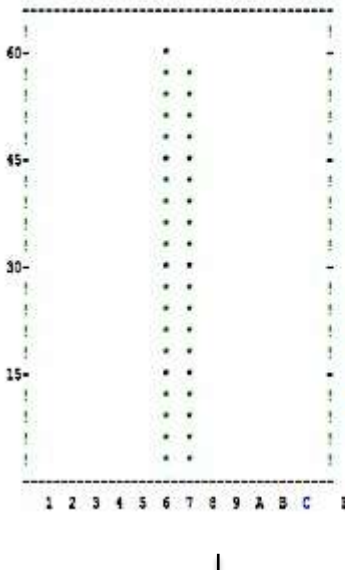

Fuente: EQS 6. con datos y adaptación propia

Se guardan el archivo y se corre conforme a lo indicado en la **Figura 7.14**
Se procede a revisar rápidamente las secciones de reporte de resultados como se ve en la **Figura 7.77**

**Figura 7.77. CFA, SEM y EQS. Proceso de revisión rápida de resultados caso estudio 2. Sin V1 y V3.**

```
/TITLE
CFA ORIENTACIÓN AL MERCADO DE SEGUNDO ORDEN.SIN V1 y V3
NOMBRE DEL ARCHIVO: cfa cm sin v1 y v3.eqs
/SPECIFICATIONS
VARIABLES=17; CASES=300; DATAFILE='cfa cm.ess';
MATRIX=RAW; METHOD=ML, ROBUST; ANALYSIS=COV;

/EQUATIONS
!V1=F1+E1;
V2= F1+E2;
!V3=*F1+E3;
V4=*F1+E4;
V5=*F1+E5;
V6=*F1+E6;

V7=F2+E7;
V8=*F2+E8;
V9=*F2+E9;
V10=*F2+E10;
V11=*F2+E11;

V12=F3+E12;
V13=*F3+E13;
V14=*F3+E14;
V15=*F3+E15;
V16=*F3+E16;
V17=*F3+E17;

F1=*F4+D1;
F2=*F4+D2;
F3=*F4+D3;

/VARIANCES
!E1=*;
!E2=*;
```

DISTRIBUTION OF STANDARDIZED RESIDUALS



| | RANGE | FREQ | PERCENT |
|---|---|---|---|
| 1 | -0.5 - -- | 0 | 0.00% |
| 2 | -0.4 - -0.5 | 0 | 0.00% |
| 3 | -0.3 - -0.4 | 0 | 0.00% |
| 4 | -0.2 - -0.3 | 0 | 0.00% |
| 5 | -0.1 - -0.2 | 1 | 0.83% |
| 6 | 0.0 - -0.1 | 61 | 50.83% |
| 7 | 0.1 - 0.0 | 57 | 47.50% |
| 8 | 0.2 - 0.1 | 1 | 0.83% |
| 9 | 0.3 - 0.2 | 0 | 0.00% |
| A | 0.4 - 0.3 | 0 | 0.00% |
| B | 0.5 - 0.4 | 0 | 0.00% |
| C | ++ - 0.5 | 0 | 0.00% |
| | TOTAL | 120 | 100.00% |

1 2 3 4 5 6 7 8 9 A B C  EACH "*" REPRESENTS 3 RESIDUALS

```
GOODNESS OF FIT SUMMARY FOR METHOD = ROBUST

ROBUST INDEPENDENCE MODEL CHI-SQUARE =    2196.220 ON   105 DEGREES OF FREEDOM

INDEPENDENCE AIC =    1986.220   INDEPENDENCE CAIC =    1492.323
       MODEL AIC =      35.128          MODEL CAIC =    -369.398

SATORRA-BENTLER SCALED CHI-SQUARE =    207.1276 ON      86 DEGREES OF FREEDOM
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS      0.00000


FIT INDICES
-----------
BENTLER-BONETT      NORMED FIT INDEX =      0.906
BENTLER-BONETT NON-NORMED FIT INDEX =      0.929
COMPARATIVE FIT INDEX (CFI)          =      0.942
BOLLEN'S          (IFI) FIT INDEX    =      0.943
MCDONALD'S        (MFI) FIT INDEX    =      0.817
ROOT MEAN-SQUARE ERROR OF APPROXIMATION (RMSEA)     =      0.069
90% CONFIDENCE INTERVAL OF RMSEA  (       0.057,          0.080)
```

Fuente: EQS 6. con datos y adaptación propia

MEASUREMENT EQUATIONS WITH STANDARD ERRORS AND TEST STATISTICS
STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.
(ROBUST STATISTICS IN PARENTHESES)

```
X2   =V2  =      1.000 F1   +     1.000 E2

X4   =V4  =      1.164*F1   +     1.000 E4
                  .078
                14.879@
          (       .081)
          (    14.284@

X5   =V5  =       .943*F1   +     1.000 E5
                  .076
                12.410@
          (       .083)
          (    10.125@
```

CONSTRUCT EQUATIONS WITH STANDARD ERRORS AND TEST STATISTICS
STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.
(ROBUST STATISTICS IN PARENTHESES)

```
F1   =F1  =      1.036*F4   +     1.000 D1
                  .077
                13.492@
          (       .078)
          (    13.244@

F2   =F2  =      1.052*F4   +     1.000 D2
                  .077
                13.722@
          (       .075)
          (    13.981@

F3   =F3  =       .891*F4   +     1.000 D3
                  .072
                12.369@
          (       .077)
          (    11.605@
```

STANDARDIZED SOLUTION:

| | | | | R-SQUARED |
|---|---|---|---|---|
| X2 | =V2 = | .763 F1 | + .646 E2 | .583 |
| X4 | =V4 = | .839*F1 | + .544 E4 | .704 |
| X5 | =V5 = | .713*F1 | + .702 E5 | .508 |
| X6 | =V6 = | .809*F1 | + .588 E6 | .654 |
| X7 | =V7 = | .817 F2 | + .577 E7 | .667 |
| X8 | =V8 = | .880*F2 | + .475 E8 | .774 |
| X9 | =V9 = | .909*F2 | + .417 E9 | .826 |
| X10 | =V10 = | .834*F2 | + .553 E10 | .695 |
| X11 | =V11 = | .606*F2 | + .795 E11 | .368 |
| X12 | =V12 = | .725 F3 | + .688 E12 | .526 |
| X13 | =V13 = | .842*F3 | + .539 E13 | .710 |
| X14 | =V14 = | .772*F3 | + .636 E14 | .595 |
| X15 | =V15 = | .803*F3 | + .596 E15 | .644 |
| X16 | =V16 = | .778*F3 | + .628 E16 | .605 |
| X17 | =V17 = | .776*F3 | + .631 E17 | .602 |
| F1 | =F1 = | .924*F4 | + .383 D1 | .854 |
| F2 | =F2 = | .850*F4 | + .527 D2 | .723 |
| F3 | =F3 = | .876*F4 | + .483 D3 | .767 |

MULTIVARIATE LAGRANGE MULTIPLIER TEST BY SIMULTANEOUS PROCESS IN STAGE 1

PARAMETER SETS (SUBMATRICES) ACTIVE AT THIS STAGE ARE:

PVV PFV PFF PDD GVV GVF GFV GFF BVF BFF

| CUMULATIVE MULTIVARIATE STATISTICS | | | | | UNIVARIATE INCREMENT | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| STEP | PARAMETER | CHI-SQUARE | D.F. | PROB. | CHI-SQUARE | PROB. | HANCOCK'S SEQUENTIAL D.F. | HANCOCK'S SEQUENTIAL PROB. | ROBUST PREDICTED RMSEA | ROBUST PREDICTED CFI |
| 1 | V10,F3 | 18.638 | 1 | 0.000 | 18.638 | 0.000 | 86 | 1.000 | 0.064 | 0.951 |
| 2 | V7,F3 | 29.246 | 2 | 0.000 | 10.607 | 0.001 | 85 | 1.000 | 0.061 | 0.955 |
| 3 | V6,F3 | 33.486 | 3 | 0.000 | 4.241 | 0.039 | 84 | 1.000 | 0.060 | 0.957 |

Fuente: EQS 6. con datos y adaptación propia

## Discusión rápida de resultados, sin V1 y V3

De los reportes emitidos, podemos observar y concluir de cada sección de reporte lo siguiente:

- Gráfico de distribución de residuos (*distribution of standardized residuals*). *Cumple.*
- índices de ajuste (*fit indices*) robustos (*robust*),valores a analizar:

*Chi-Cuadrada de Satorra-Bentler*= 207.1276; gl= 86 (**Nota:** 2017.1276/86=0.2727<2 Tendencia a que los índices de ajuste SI sean .9); valor p<0.01**.** *Cumple.*
**NFI=** .906; **NNFI**= .929 ; **CFI**= .942. Cumple >.9**. RMSEA**=.069<.08. *Cumple.*

- Ecuaciones de medición (*measurment equations),*valores de t (@) =.05.*Cumple.*
- Ecuaciones de constructo (*construct equations*), valores de t (@)= .05.*Cumple.*
- Varianzas de las variables independientes (*variances of independent variables*), valores de t (@)= .05. *Cumple.*
- Solución estandarizada (*standardized solutions*), *¿todas las cargas factoriales son >.6?. Bagozzi y Yi (1988) que cada carga debe ser factorial debe ser >.6. Cumple.*
- *Ya No es necesario verificar la prueba del multiplicador de Lagrange (Lagrange multiplier test), ya que cumplen los requisitos anteriores*. Con estos resultados, ya estamos, ten posibilidades de hacer aprobar o desaprobar, 3 hipótesis de manera directa, tales como:

**H1:** *A mayor orientación al mercado, mayor orientación al cliente.*
**H2:** *A mayor orientación al mercado mayor orientación a la competencia.*
**H3:** *A mayor orientación al mercado mayor coordinación interfuncional.*

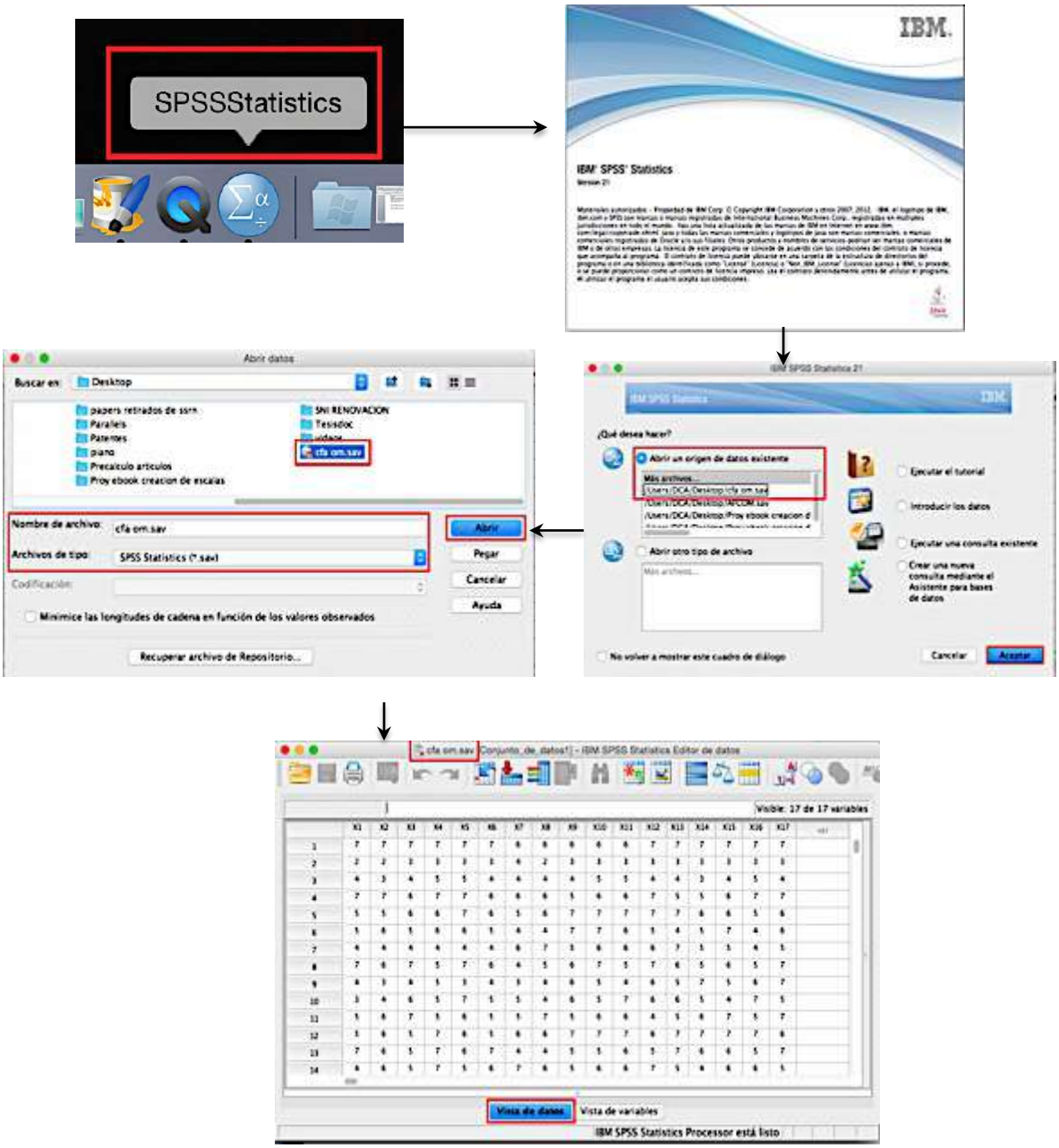
La respuesta a los 3 es afirmativa en un valor correspondiente a (.924); (.850); y (.876). Lo anterior significa también, que el empresario con orientación a mercado, debe enfatizar más su atención a la orientación al cliente, seguido de la orientación de infraestructura y finalmente, en la orientación a la competencia para optimizar sus recursos con mejores resultados.
Con los datos ajustados, se procederá a calcular la confiabilidad de cada constructo con el *alfa de Cronbach*, para posteriormente compararlo con el índice de fiabilidad compuesta **IFC.** Si cumplen el requisito , se realizará la prueba de validez calculando el índice de varianza extraída (**IVE**) correspondiente

## Cálculo del alfa de Cronbach, sin V1 y V3

Dado que se requiere el *alfa de Cronbach*, por constructo (*No por modelo, que es lo que se reporta en **EQS***), requerimos calcularlo.Este cálculo debará proceder de la base de datos **SPSS**, en nuestro caso **cfa om.sav** y proceder únicamente con las variables que quedaron del modelos estructura, o sea, *No incluir las variables* **V1** *ni* **V3.** Es importante considerar que si está utilizando **SPSS** soportado con **IOS** o **Windows** *(con un emulador como Parallels)***,** ubique su archivo **.SAV**, de acuerdo al sistema operativo con el cual va a trabajar. En nuestro caso es de forma directa en sistema operativo **IOS.**
Así, la secuencia de comandos para ubicar y abrir el archivo **cfa om.sav**, es:
**Teclear en ícono SPSS desde aplicaciones o barra de herramientas-> ¿Qué desea hacer? abrir un archivo de datos existente:… cfa om.sav->Aceptar-**

**>Nombre de archivo: cfa om.sav; Archivos de tipo: SPSS Statistics (*.sav)->Abrir ->Despliegue de datos en modo: Vista de datos. Ver Figura 7.78.**

**Figura 7.78. Secuencia de comandos apertura de archivo base de datos SPSS, caso de estudio 2.**

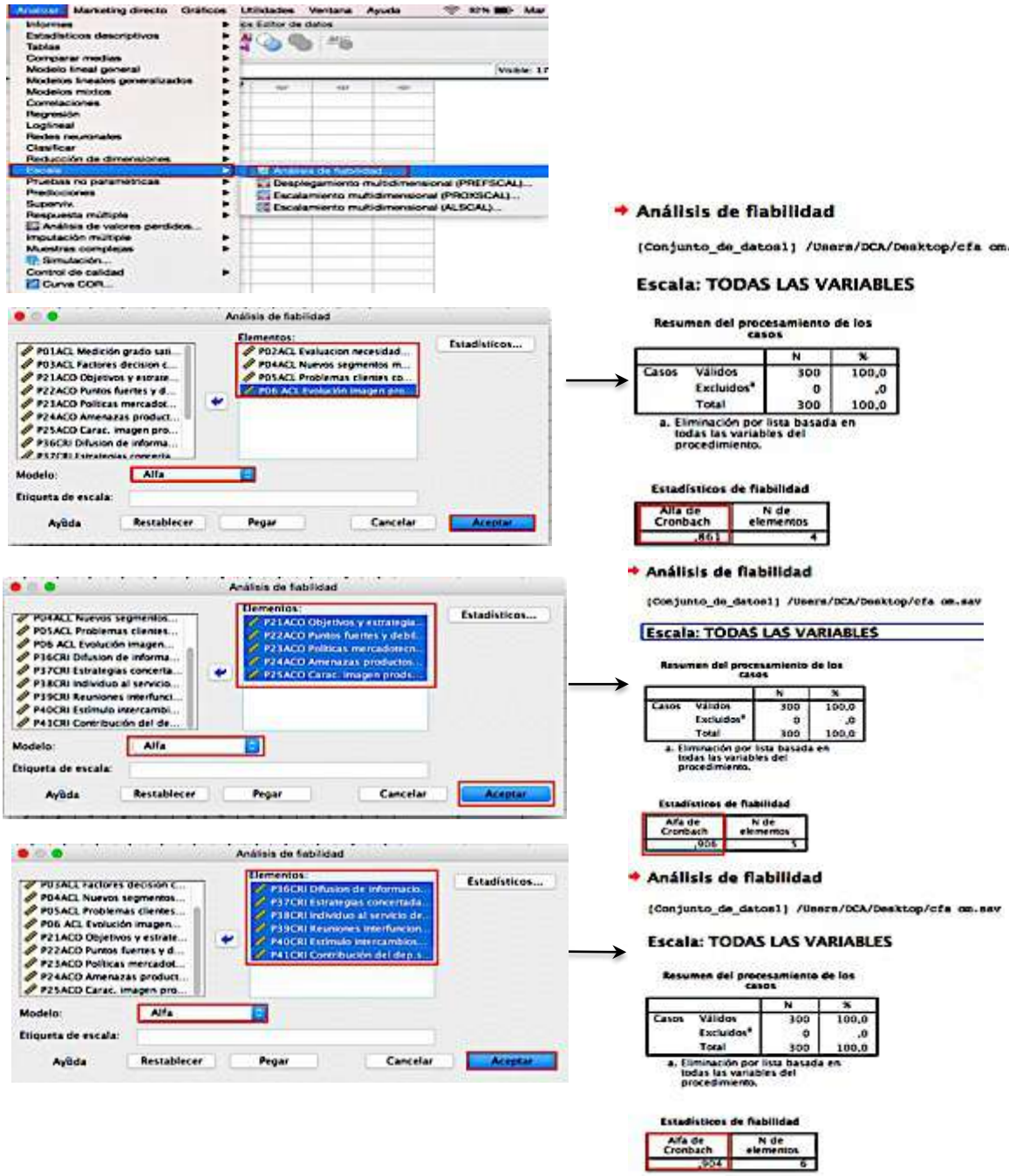



Fuente: SPSS 21 IBM, con datos y adaptación propia

Así, se procederá a realizar el cálculo de *alfa de Cronbach*, para cada uno de los ítems de los factores, con la siguiente secuencia inicial de comandos:

**Analizar->Escala->Análisis de fiabilidad->Modelo: Alfa->Elementos: V2; V4; V5; V6 (en el caso del archivo, los marcados como: P02ACL; P04ACL; P05ACL; P06ACL)->Aceptar. Ver Figura 7.79.**

**Figura 7.79. Secuencia de comandos cálculo alfa de Cronbach con SPSS, caso de estudio 2.**




Fuente: SPSS 21 IBM, con datos y adaptación propia

En consecuencia, el *alfa de Cronbach* por factor, es:

**F1** (V2; V4; V5; V6)= .861
**F2** (V7; V8; V9; V10; 11)= .906
**F3** (V12; V13; V14; V15; V16; V17)= .904

## *Cálculo de IFC, sin V1 y V3*

Normalmente, en la investigación científica los modelos teóricos desarrollados en la literatura no aparecen implicados un solo constructo, sino varios, por lo cual el *alfa de Cronbach* para cada factor por separado no considera la influencia sobre la fiabilidad del resto de constructos. Por esta razón, Fornell y Larcker (1981) proponen el cálculo del índice de la fiabilidad compuesta (**IFC**) para cada factor que, interpretándose exactamente igual que el *alfa de Cronbach,* Sí que tiene en cuenta las interrelaciones. Por lo tanto, para determinarlo partimos del reporte de solución estandarizada (*standardized solution*) de la **Figura 7.77.**

STANDARDIZED SOLUTION: R-SQUARED

| | | | | | |
|---|---|---|---|---|---|
| X2 | =V2 | = | .763 F1 | + .646 E2 | .583 |
| X4 | =V4 | = | .839*F1 | + .544 E4 | .704 |
| X5 | =V5 | = | .713*F1 | + .702 E5 | .508 |
| X6 | =V6 | = | .809*F1 | + .588 E6 | .654 |
| X7 | =V7 | = | .817 F2 | + .577 E7 | .667 |
| X8 | =V8 | = | .880*F2 | + .475 E8 | .774 |
| X9 | =V9 | = | .909*F2 | + .417 E9 | .826 |
| X10 | =V10 | = | .834*F2 | + .553 E10 | .695 |
| X11 | =V11 | = | .606*F2 | + .795 E11 | .368 |
| X12 | =V12 | = | .725 F3 | + .688 E12 | .526 |
| X13 | =V13 | = | .842*F3 | + .539 E13 | .710 |
| X14 | =V14 | = | .772*F3 | + .636 E14 | .595 |
| X15 | =V15 | = | .803*F3 | + .596 E15 | .644 |
| X16 | =V16 | = | .778*F3 | + .628 E16 | .605 |
| X17 | =V17 | = | .776*F3 | + .631 E17 | .602 |
| F1 | =F1 | = | .924*F4 | + .383 D1 | .854 |
| F2 | =F2 | = | .850*F4 | + .527 D2 | .723 |
| F3 | =F3 | = | .876*F4 | + .483 D3 | .767 |

Fuente: EQS 6, con datos y adaptación propia

Calculando el reporte **IFC**, como se observa en la **Figura 7.80**

**Figura 7.80. CFA, SEM y EQS. Recopilación de información que calcula el IFC del caso de estudio 2 sin V1 y V3**

| Constructo e indicadores | Carga Factorial Estandarizada $L_{ij}$ | Varianza del Término de Error $Var(E_{ij}) = 1 - (L_{ij})^2$ | Índice de Fiabilidad Compuesta IFC >.7 $\sum(L_{ij})^2/[\sum(L_{ij})^2 + \sum Var(E_{ij})]$ |
|---|---|---|---|
| **Orientación al Cliente ($F_1$)** | | | |
| $V_2$ | .763 | .418 | **.863** |
| $V_4$ | .839 | .296 | |
| $V_5$ | .713 | .492 | |
| $V_6$ | .809 | .346 | |
| Total | 3.124 | 1.552 | |
| **Orientación a la Competencia ($F_2$)** | | | |
| $V_7$ | .817 | .333 | **.907** |
| $V_8$ | .880 | .226 | |
| $V_9$ | .909 | .174 | |
| $V_{10}$ | .833 | .306 | |
| $V_{11}$ | .606 | .633 | |
| Total | 4.045 | 1.672 | |
| **Coordinación Interfuncional ($F_3$)** | | | |
| $V_{12}$ | .725 | .474 | **.905** |
| $V_{13}$ | .842 | .291 | |
| $V_{14}$ | .772 | .404 | |
| $V_{15}$ | .803 | .355 | |
| $V_{16}$ | .778 | .395 | |
| $V_{17}$ | .776 | .398 | |
| Total | 4.696 | 2.317 | |
| **Orientación al Mercado** | | | |
| $F_1$ | .924 | .146 | **.915** |
| $F_2$ | .850 | .277 | |
| $F_3$ | .876 | .233 | |
| Total | 2.650 | .656 | |

Fuente: EQS 6. con datos y adaptación propios.

## *Cálculo de IVE, sin V1 y V3*

Fornell y Larcker (1981) presentan el índice **IVE** como la relación entre la varianza que es capturada por un factor *i* en relación a la varianza total debida al error de medida de ese factor. Nótese que la única diferencia con el **IFC** es que cada carga factorial estandarizada es primero elevada al cuadrado antes de ser sumadas. Fornell y Larcker (1981) sugieren que es deseable que el constructo tenga valores de **IVE** >= .5, es decir, que sea superior la varianza capturada por el factor que la debida al error de medida. Ver **Figura 7.81.**

**Figura 7.81. CFA, SEM y EQS. Recopilación de información que calcula el IVE del caso de estudio 2 sin V1 y V3**

| Constructo e indicadores | Cuadrado de la Carga Factorial Estandarizada $(L_{ij})^2$ | Varianza del Término de Error $Var(E_{ij}) = 1 - (L_{ij})^2$ | Índice de Varianza Compuesta Extraída IVE>= .5 $\sum(L_{ij})^2/[\sum(L_{ij})^2 + \sum Var(E_{ij})]$ |
|---|---|---|---|
| **Orientación al Cliente ($F_1$)** | | | |
| $V_2$ | .582 | .418 | **.612** |
| $V_4$ | .704 | .296 | |
| $V_5$ | .508 | .492 | |
| $V_6$ | .654 | .346 | |
| Total | 2.448 | 1.552 | |
| **Orientación a la Competencia ($F_2$)** | | | |
| $V_7$ | .667 | .333 | **.666** |
| $V_8$ | .774 | .226 | |
| $V_9$ | .826 | .174 | |
| $V_{10}$ | .694 | .306 | |
| $V_{11}$ | .367 | .633 | |
| Total | 3.328 | 1.672 | |
| **Coordinación Interfuncional ($F_3$)** | | | |
| $V_{12}$ | .526 | .474 | **.614** |
| $V_{13}$ | .709 | .291 | |
| $V_{14}$ | .596 | .404 | |
| $V_{15}$ | .645 | .355 | |
| $V_{16}$ | .605 | .395 | |
| $V_{17}$ | .602 | .398 | |
| Total | 3.683 | 2.317 | |
| **Orientación al Mercado** | | | |
| $F_1$ | .854 | .146 | **.806** |
| $F_2$ | .723 | .277 | |
| $F_3$ | .767 | .233 | |
| Total | 2.334 | .565 | |

Fuente: EQS 6. con datos y adaptación propia

## *Análisis de la validez de la escala, sin V1 y V3*

Recordando a Bohrnstedt (1976): *validez es el grado en que un instrumento mide el concepto bajo estudio*. Desafortunadamente, el concepto tiene múltiples dimensiones que deben explicarse y analizarse por separado.

## Validez de contenido

La **Figura 7.77** muestra los estadísticos de bondad de ajuste proporcionados por **EQS**

```
GOODNESS OF FIT SUMMARY FOR METHOD = ROBUST

ROBUST INDEPENDENCE MODEL CHI-SQUARE =    2196.220 ON   105 DEGREES OF FREEDOM

INDEPENDENCE AIC =    1986.220   INDEPENDENCE CAIC =    1492.323
       MODEL AIC =      35.128          MODEL CAIC =    -369.398

SATORRA-BENTLER SCALED CHI-SQUARE =    207.1276 ON      86 DEGREES OF FREEDOM
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS      0.00000
```

Fuente: EQS 6. con datos y adaptación propia

De su análisis se desprende que el valor de la *Chi-Cuadrado* y el *número de grados de libertad* es superior a 2 (207.1276/86 = 2.408). Parece indicar que el modelo de *Orientación al Mercado no tiene un buen ajuste*, pero es posible seguir analizando otros estadísticos. Entre el conjunto de indicadores de ajuste que proporciona **EQS**, *se recomienda analizar aquellos no normalizados*, dado que tienen menor tendencia a ofrecer resultados sesgados en casos de pequeñas muestras (Bentler y Wu, 1993). Se observa que, tanto el **NNFI** (Bentler y Bonnet, 1980) como el **CFI** (Bentler y Wu, 1993), que cumplen este requisito, tienen valores > .9, indicando un ajuste razonable. **RMSEA**= .069 cumple con el requisito de ser <.08 De a misma **Figura 7.77,** ver la sección de reporte ajuste de índices (*fit indices*).

```
FIT INDICES
-----------
BENTLER-BONETT      NORMED FIT INDEX =      0.906
BENTLER-BONETT NON-NORMED FIT INDEX =      0.929
COMPARATIVE FIT INDEX (CFI)         =      0.942
BOLLEN'S          (IFI) FIT INDEX   =      0.943
MCDONALD'S        (MFI) FIT INDEX   =      0.817
ROOT MEAN-SQUARE ERROR OF APPROXIMATION (RMSEA)    =      0.069
90% CONFIDENCE INTERVAL OF RMSEA  (       0.057,          0.080)
```

Fuente: EQS 6, con datos y adaptación propia

Ya hemos estimado el *modelo de medida* y hemos comprobado que ofrece un buen ajuste como base de la *validez de contenido*. Pues bien, sobre este modelo de medida, ya es posible analizar la *validez convergente* y la *validez discriminante* de las escalas implicadas. Para analizar la *validez nomológica,* será necesario estimar el *modelo teórico*.

## Validez de constructo o de concepto

La validez de constructo o de concepto se divide en dos: *validez convergente y validez discriminante.* Una escala tiene *validez de constructo*, cuando tiene *validez convergente y validez discriminante.*

### 7.22.10.3. Validez convergente

La *validez convergente* existe cuando se emplean distintos instrumentos para medir un mismo constructo (distintos ítems para una misma variable latente) y estos instrumentos están fuertemente correlacionados. Si analizamos de la misma **Figura 7.77**, las secciones de reporte las ecuaciones de medida (*measurement equations*) se deberá observar que las cargas factoriales sean < 1 significativas al 5%

```
MEASUREMENT EQUATIONS WITH STANDARD ERRORS AND TEST STATISTICS
STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.
(ROBUST STATISTICS IN PARENTHESES)

X2     =V2  =          1.000 F1     +        1.000 E2

X4     =V4  =          1.164*F1     +        1.000 E4
                        .078
                      14.879@
               (        .081)
               (      14.284@

X5     =V5  =           .943*F1     +        1.000 E5
                        .076
                      12.410@
               (        .093)
               (      10.125@
```

Fuente: EQS 6, con datos y adaptación propia

La sección de reporte de soluciones estandarizadas (*standardized solutions*), así como para el modelo de medida observamos que *todas las cargas factoriales son significativas, como se desprende de sus respectivas t* para todas las variables. Asimismo, si calculáramos el promedio de los mismos veríamos que superan claramente .7, y de la misma **Figura 7.77** en su sección de prueba del multiplicador de Lagrange (*Lagrange multiplier test*) el ajuste del modelo mejoraría significativamente con la consideración, con suficiente marco teórico y **EFA** validado

de cambiar la **V10** a **F3** ambas como intento de mejora del modelo.

```
MULTIVARIATE LAGRANGE MULTIPLIER TEST BY SIMULTANEOUS PROCESS IN STAGE  1

PARAMETER SETS (SUBMATRICES) ACTIVE AT THIS STAGE ARE:

   PVV PFV PFF PDD GVV GVF GFV GFF BVF BFF
```

| | CUMULATIVE MULTIVARIATE STATISTICS | | | | UNIVARIATE INCREMENT | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| STEP | PARAMETER | CHI-SQUARE | D.F. | PROB. | CHI-SQUARE | PROB. | HANCOCK'S SEQUENTIAL D.F. | PROB. | ROBUST PREDICTED RMSEA | CFI |
| 1 | V10,F3 | 18.638 | 1 | 0.000 | 18.638 | 0.000 | 86 | 1.000 | 0.064 | 0.951 |
| 2 | V7,F3 | 29.246 | 2 | 0.000 | 10.607 | 0.001 | 85 | 1.000 | 0.061 | 0.955 |
| 3 | V6,F3 | 33.486 | 3 | 0.000 | 4.241 | 0.039 | 84 | 1.000 | 0.060 | 0.957 |

Fuente: EQS 6. con datos y adaptación propia

## Validez discriminante

Por otro lado, la *validez discriminante* se consigue cuando, si diversos instrumentos de medida están diseñados para medir distintas variables latentes, entonces las correlaciones entre dichos instrumentos son bajas. Una escala tiene, por lo tanto, *validez discriminante cuando no mide un constructo para el que no se diseñó.* En nuestro caso *NO es posible medir la validez discriminante porque estamos utilizando un factor de segundo orden (**F4**), que No tiene covarianza con otro factor de segundo orden (es el único)*

## Validez nomológica

El *modelo teórico* tendrá *validez nomológica si, efectivamente, No hay diferencias significativas entre los ajustes del modelo de medida y el teórico, dado que las escalas habrán sido capaces de establecer relaciones predictivas de otras variables tan sustantivas que, siendo menos, igualan la bondad del modelo*.
En este momento, es posible hacer un comparativo de las *Chi-Cuadrada del modelo teórico* **(Figura 7.67)**

```
SATORRA-BENTLER SCALED CHI-SQUARE =    369.7671 ON     116 DEGREES OF FREEDOM
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS      0.00000
```

***vs.*** *el de medida* de la **Figura 7.77**

```
SATORRA-BENTLER SCALED CHI-SQUARE =    207.1276 ON      86 DEGREES OF FREEDOM
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS      0.00000
```

Fuente: EQS 6. con datos y adaptación propia

Nos muestran los estadísticos de bondad de ajuste que son necesarios para determinar la validez nomológica. Así, es posible comprobar que las respectivas *Cuadrado del modelo teórico y de medida* (que obtuvimos al analizar la validez discriminante) son:

- *Modelo Teórico: 369.7671; grados de libertad, 116*

***vs***

- *Modelo de Medida: 207.1276; grados de libertad, 86*

La *validez nomológica* se establecería calculando un test de diferencias entre las *Chi-Cuadrado*, tal como hicimos para la validez discriminante:

- *Diferencias entre las Chi-Cuadrado: (369.7671- 207.1276)= 162.6395*
- *Diferencias en grados de libertad: (116-86)= 30*
- *Valor crítico p<0.01: 50.892* (Ver **Figura 7.82**)

**Figura 7.82. Área correspondiente al extremo derecho de una distribución Chi-Cuadrada**

| | | | Área en el extremo derecho | | Grados de |
|---|---|---|---|---|---|
| 0.20 | 0.10 | 0.05 | 0.25 | 0.01 | libertad |
| 1.642 | 2.706 | 3.841 | 5.024 | 6.635 | 1 |
| 3.219 | 4.605 | 5.991 | 7.378 | 9.210 | 2 |
| 4.642 | 6.251 | 7.815 | 9.348 | 11.345 | 3 |
| 5.989 | 7.779 | 9.488 | 11.143 | 13.277 | 4 |
| 7.289 | 9.236 | 11.070 | 12.833 | 15.086 | 5 |
| 8.558 | 10.645 | 12.592 | 14.449 | 16.812 | 6 |
| 9.803 | 12.017 | 14.067 | 16.013 | 18.475 | 7 |
| 11.030 | 13.362 | 15.507 | 17.535 | 20.090 | 8 |
| 12.242 | 14.684 | 16.919 | 19.023 | 21.666 | 9 |
| 13.442 | 15.987 | 18.307 | 20.483 | 23.209 | 10 |
| 14.631 | 17.275 | 19.675 | 21.920 | 24.725 | 11 |
| 15.812 | 18.549 | 21.026 | 23.337 | 26.217 | 12 |
| 16.985 | 19.812 | 22.362 | 24.736 | 27.688 | 13 |
| 18.151 | 21.064 | 23.685 | 26.119 | 29.141 | 14 |
| 19.311 | 22.307 | 24.996 | 27.488 | 30.578 | 15 |
| 20.465 | 23.542 | 26.296 | 28.845 | 32.000 | 16 |
| 21.615 | 24.769 | 27.587 | 30.191 | 33.409 | 17 |
| 22.760 | 25.989 | 28.869 | 31.526 | 34.805 | 18 |
| 23.900 | 27.204 | 30.144 | 32.852 | 36.191 | 19 |
| 25.038 | 28.412 | 31.410 | 34.170 | 37.566 | 20 |
| 26.171 | 29.615 | 32.671 | 35.479 | 38.932 | 21 |
| 27.301 | 30.813 | 33.924 | 36.781 | 40.289 | 22 |
| 28.429 | 32.007 | 35.172 | 38.076 | 41.638 | 23 |
| 29.553 | 33.196 | 36.415 | 39.364 | 42.980 | 24 |
| 30.675 | 34.382 | 37.652 | 40.647 | 44.314 | 25 |
| 31.795 | 35.563 | 38.885 | 41.923 | 45.642 | 26 |
| 32.912 | 36.741 | 40.113 | 43.194 | 46.963 | 27 |
| 34.027 | 37.916 | 41.337 | 44.461 | 48.278 | 28 |
| 35.139 | 39.087 | 42.557 | 45.722 | 49.588 | 29 |
| 36.250 | 40.256 | 43.773 | 46.979 | 50.892 | 30 |

Fuente: Levin y Rubin (2004)

En otras palabras, el *modelo teórico ofrece un ajuste significativamente mejor que el modelo de medida (diferencia inferior al valor crítico), lo que permitiría constatar la validez nomológica de las escalas.*

## *Reportando resultados: consistencia interna y validez convergente*

En este punto, es importante tener una referencia del acomodo de los datos, como se sugiere en el ejemplo de la **Figura 7.83.**

**Figura 7.83. Reporte tabla de consistencia interna y validez convergente del modelo teórico sin V1 y V3**

| Factor | Ítem V | Carga Factorial >0.6 (a) | Valor Robusto t | Factor Promedio de Carga Factorial | Fiabilidad de la Escala | | |
|---|---|---|---|---|---|---|---|
| | | | | | Alfa de Cronbach >=0.7 (b) | IFC >0.7 (b) | IVE >0.5 (c) |
| Orientación al Cliente ($F_1$) | $V_2$ | .763*** | 1.000a | 0.781 | .861 | **.863** | **.612** |
| | $V_4$ | .839*** | 14.284 | | | | |
| | $V_5$ | .713*** | 10.125 | | | | |
| | $V_6$ | .809*** | 13.394 | | | | |
| Orientación a la Competencia ($F_2$) | $V_7$ | .817*** | 1.000a | .809 | .906 | **.907** | **.666** |
| | $V_8$ | .880*** | 19.924 | | | | |
| | $V_9$ | .909*** | 17.996 | | | | |
| | $V_{10}$ | .833*** | 14.630 | | | | |
| | $V_{11}$ | .606*** | 9.886 | | | | |
| Coordinación Interfuncional ($F_3$) | $V_{12}$ | .725*** | 1.000a | .782 | .904 | **.905** | **.614** |
| | $V_{13}$ | .842*** | 15.555 | | | | |
| | $V_{14}$ | .772*** | 11.990 | | | | |
| | $V_{15}$ | .803*** | 13.575 | | | | |
| | $V_{16}$ | .778*** | 12.409 | | | | |
| | $V_{17}$ | .776*** | 11.769 | | | | |
| Orientación al Mercado | $F_1$ | .924*** | 13.244 | .883 | - | **.915** | **.806** |
| | $F_2$ | .850*** | 13.981 | | | | |
| | $F_3$ | .876*** | 11.605 | | | | |

***Resultados:***
**Chi-Cuadrada=** 207.1276; **gl=** 86; **p** < .01; **NFI** = .906; **NNFI** = .929; **CFI** = . 942; **RMSEA** = .069
***Conclusion:***
*Las relaciones entre las variables y los factores tienen un ajuste adecuado a los datos. Esto es, no hay error en el ajuste del modelo.Aunqe existe suficiente evidencia de validación convergente y fiabilidad , que justifica la fiabilidad interna de la escala aplicada (Nunnally y Bernstein, 1994; Hair et al., 1999), no se tiene validez discriminante porque sólo existe un Factor F4 (sólo una variable independiente). Por lo tanto, no es posible determinar un Reporte tabla de validez discriminante del modelo teórico.*

Notas:*** Parámetros constreñidos a ese valor en el proceso de identificación **= p < 0.001**; (a) De acuerdo a Bagozzi y Yi, 1988; (b) De acuerdo a Hair et al. 1999; (c) Índice de Varianza Extraída (**IVE**), de acuerdo a Fornell y Larcker, 1981.

Fuente: EQS 6. y SPSS IBM 21, con datos y adaptación propios.

Donde las columnas significan:

- *Factores,* listado de los involucrados en el estudio. En nuestro caso:

Orientación al Cliente **($F_1$)**; Orientación a la Competencia **($F_2$)**; Coordinación Interfuncional **($F_3$)** y Orientación al Mercado

- *Ítem,* consecutivo de las variables (**V**) no suprimidas mostradas, como : **V1; V2; V3,** etc.
- *Carga Factorial > .6 (a)* (Ver **Figura 7.77**) donde sólo se reportan los factores no suprimidos y con carga factorial *>.6***.** Es posible que el investigador utilice escalas previas que si a su juicio, por el marco teórico, deban asignarse más ítems, es válido realizarlo siempre y cuando, se mencione que la nueva escala es una *adaptación* de la previa más la añadida como propuesta del investigador dadas las condiciones de la realidad en la que se aplica contra las que originalmente la escala se creó.En esta etapa es donde se valida o no la pertinencia de los nuevos ítems basados en el marco teórico planteado.
- *Valor Robusto t* (ver **Figura 7.40** para mayor detalle**)**,

Con lo anterior, es posible afirmar que si los:

- *Valores de t (sección de reporte ecuaciones de medida (measurement equations) y ecuaciones de constructo (construct equations) para los factores, Figura 7.79) son mayores (>) 2.58 errores estándar de la media, entonces se anotarán tres asteriscos (***) por carga factorial, lo cual indica que p<.0001. Así se tienen una seguridad que las cargas factoriales aportan sus valores a un 99% de confiabilidad*
- *Alfa de Cronbach,* calculada y reportada de acuerdo a **Figura 7.81** mediante uso de base de datos **SPSS cfa om.sav**
- **IFC.** Se toman los datos de la **Figura 7.82** *y es muy similar al alfa de Cronbach.*
- **IVE.** Se toman los datos de la **Figura 7.83**

## Conclusión

Hasta este momento, tomando en cuenta la **Figura 7.85** afirmamos que para **maximizar** la orientación a mercado deberá ser atendido en mayor medida, cada una de las cargas factoriales con valor absoluto alto y hacer una explicación de la realidad, basada en el marco teórico del porqué debe ser así (o no), según los argumentos del investigador. Todo lo anterior basados en las cargas factoriales encontradas con una probabilidad de equivocación de *1/100 o 99%* de que lo afirmado ocurra (ver las ***)

Donde los renglones por debajo de las cargas factoriales, significan:

- ***S-Bχ2*** *Indice Satorra Bentler Chi-Cuadrada*, es el recomendado cuando se tienen más de dos factores para generar datos robustos. Así, se explica la lectura reportada en la base de la **Figura 7.39**:

- ***Resultados:***
- ***Chi-Cuadrada****= 207.1276;* ***gl****= 86; p < .01;* ***NFI*** *= .906;* ***NNFI*** *= .929;* ***CFI*** *= . 942;* ***RMSEA*** *= .069*
- ***Conclusión:***

*Las relaciones entre las variables y los factores tienen un ajuste adecuado a los datos. Esto es, no hay error en el ajuste del modelo.Aunqe existe suficiente evidencia de validación convergente y fiabilidad , que justifica la fiabilidad interna de la escala aplicada (Nunnally y Bernstein, 1994; Hair et al., 1999), no se tiene validez discriminante porque sólo existe un Factor* ***F4*** *(sólo una variable independiente). Por lo tanto no es posible determinar un Reporte tabla de validez discriminante del modelo teórico****.***

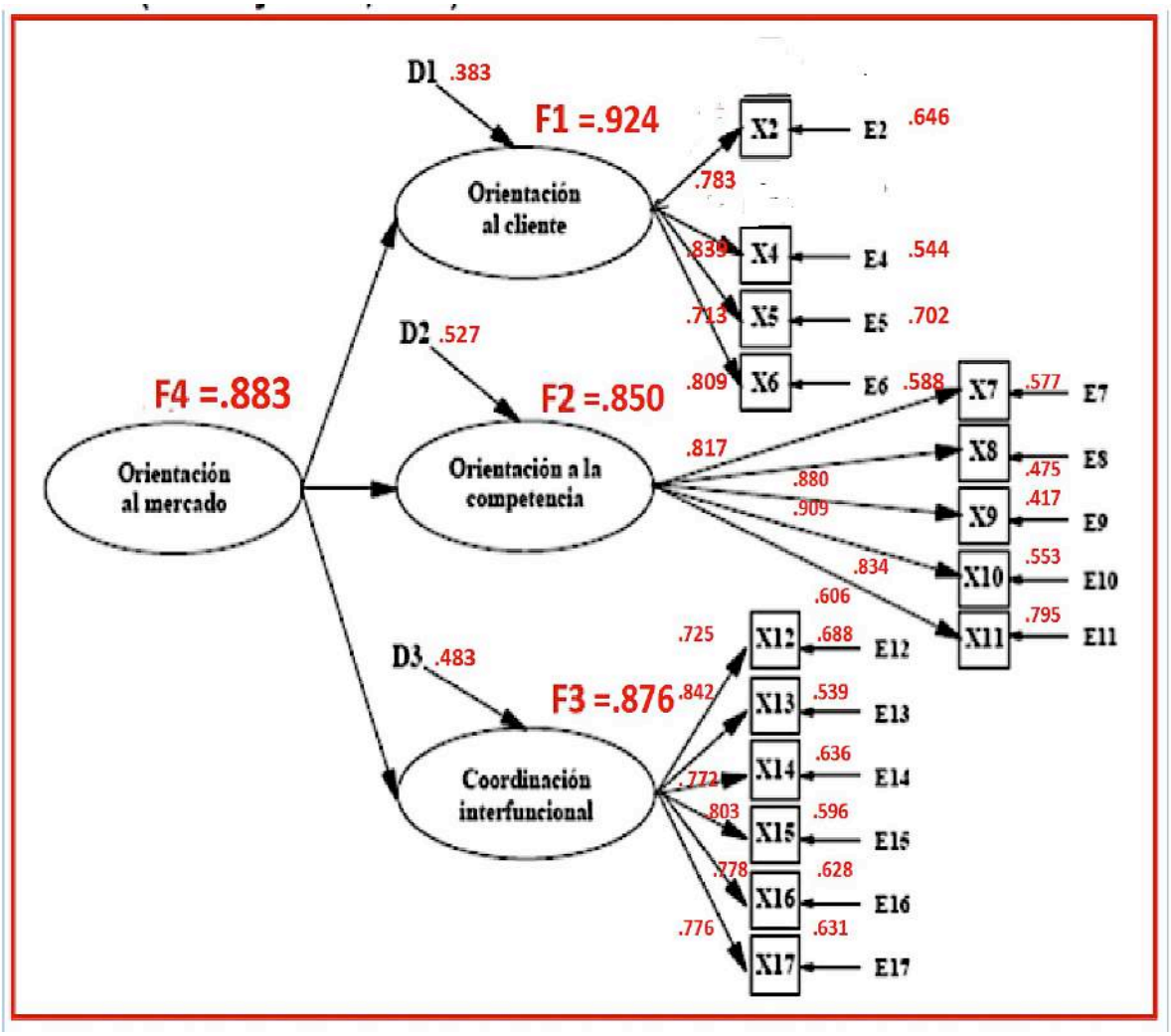
Así, una posible redacción de los resultados de la **Figura 7.85**, sería: *Por otro lado, para evaluar la fiabilidad y validez de las escalas de medida se realizó un Análisis Factorial Confirmatorio (****CFA****) utilizando el método de máxima verosimilitud con el software* ***EQS 6*** *(Bentler, 2005; Brown, 2006; Byrne, 2006). Asimismo, la fiabilidad de las escalas de medida se evaluó a partir del coeficiente alfa de Cronbach y el índice de fiabilidad compuesta* ***(IFC****, en (Bagozzi y Yi, 1988). Todos los valores de la escala excedieron el nivel recomendado de* ***.70*** *para el alfa de Cronbach y el* ***IFC*** *que proporciona una evidencia de fiabilidad y justifica la fiabilidad interna de las escalas (Nunally y Bernstein, 1994; Hair et al., 1995).*

*Los resultados de la aplicación del* ***CFA*** *se presentan en la Tabla 1 y sugiere que el modelo de medición proporciona un buen ajuste de los datos (****S-BX2*** *= XXX.YYY;* ***df*** *= XX;* ***p*** *= 0.000;* ***NFI*** *>. 9XXX ;* ***NNFI*** *>.9XXX;* ***CFI****>.9XXX; y* ***RMSEA*** *<.08XXX). Como evidencia de la validez convergente, los resultados del* ***CFA*** *indican que todos los ítems de los factores relacionados son significativos (p < 0.001) y el tamaño de todas las cargas factoriales estandarizadas son superiores a 0.60 (Bagozzi y Yi, 1988).*

*La Tabla xxx muestra una alta consistencia interna de los constructos, en cada caso, la α de Cronbach excede el valor de .70 recomendado por Nunnally y Bernstein (1994). El* ***IFC*** *representa la varianza extraída entre el grupo de variables observadas y el constructo fundamental (Fornell y Larcker, 1981). Generalmente un* ***IFC*** *superior a .60 es considerado como deseable (Bagozzi y Yi, 1988), en este estudio este valor es superior. El índice de la varianza extraída* ***(IVE****) fue calculado para cada par de constructos, resultando un* ***IVE*** *superior a .5 en todos los factores (Fornell y Larcker, 1981)…*

El *diagrama de taryectoria final*, es el mostrado en la **Figura (7.84)**

**Figura 7.84. Diagrama de trayectoria final caso de estudio 2**




Fuente: propia

# Modelo de segundo orden. Análisis de caso de estudio 3

Tomemos como ejemplo el concepto de *marketing mix*, propuesto por el Dr. Gonzalo Maldonado Guzmán (2016) el cual ha sido un concepto muy tratado y que a primera vista es interesante actualizar llevando a cabo una búsqueda a través de las múltiples herramientas de bases de datos que existen hoy en día, basadas por cierto en palabras clave o por fases completas mediante el empleo de operadores lógicos. Este fue aplicado en la Cd. de Aguascalientes, México, a empresarios PyME manufactureros, con resultados reflejados en la base de datos *marketing mix. Sav*. Adjunto mostramos la propuesta de escala de nuestro estudio de caso, como se aprecia en la **Figura 7.85**

**Figura 7.85. Marketing mix**

| Ítem | Var | Indicador | Total Desacuerdo | | | Total Acuerdo | |
|---|---|---|---|---|---|---|---|
| **Indique si su empresa a nivel de producto:** | | | **1** | **2** | **3** | **4** | **5** |
| **MPP1** | **V1** | Tiene una identidad o marca de sus productos | | | | | |
| **MPP2** | **V2** | Desarrolla constantemente nuevos productos | | | | | |
| **MPP3** | **V3** | Desarrolla nuevas líneas de productos y servicios | | | | | |
| **MPP4** | **V4** | Ha modificado líneas de productos o servicios por emergencia | | | | | |
| **MPP5** | **V5** | En comparación con la competencia, mi empresa es a menudo la primera en introducir nuevos productos o servicios | | | | | |
| **MPP6** | **V6** | Se distingue por la calidad de sus productos | | | | | |
| **MPP7** | **V7** | Se distingue por la especialización en sus productos | | | | | |
| **MPP8** | **V8** | Toma en cuenta las necesidades del mercado para producir sus productos | | | | | |
| **MPP9** | **V9** | Se centra en la maximización de las necesidades de sus clientes en cuento a los requerimientos de sus productos | | | | | |
| **MPP10** | **V10** | Invierte recursos necesarios en el desarrollo de nuevos productos o servicios | | | | | |
| **MPP11** | **V11** | Realiza estudios de mercado para desarrollar nuevos productos o servicios | | | | | |
| **MPP12** | **V12** | Es muy sensible a cómo evalúa el cliente sus productos y servicios, por lo que si se requieren modificaciones se realizan inmediatamente. | | | | | |
| **MPP13** | **V13** | Tiene un diseño de la marca, logotipo, símbolo, lema, el embalaje, etc. de sus productos y servicios para maximizar su imagen y comercialización. | | | | | |
| **Indique si su empresa a nivel de precio:** | | | **1** | **2** | **3** | **4** | **5** |
| **MPR1** | **V14** | Optimiza el precio, el costo y la calidad del producto / servicio que ofrecen para satisfacer o exceder las expectativas de los clientes. | | | | | |
| **MPR2** | **V15** | Los precios de nuestros productos son inferiores a los de la competencia | | | | | |
| **MPR3** | **V16** | Los precios de nuestros productos son los adecuados de acuerdo a los costos | | | | | |
| **MPR4** | **V17** | Los precios de nuestros productos varían en función de la cantidad de productos que nos compren. | | | | | |
| **MPR5** | **V18** | Aplicamos una política de descuento por pronto | | | | | |

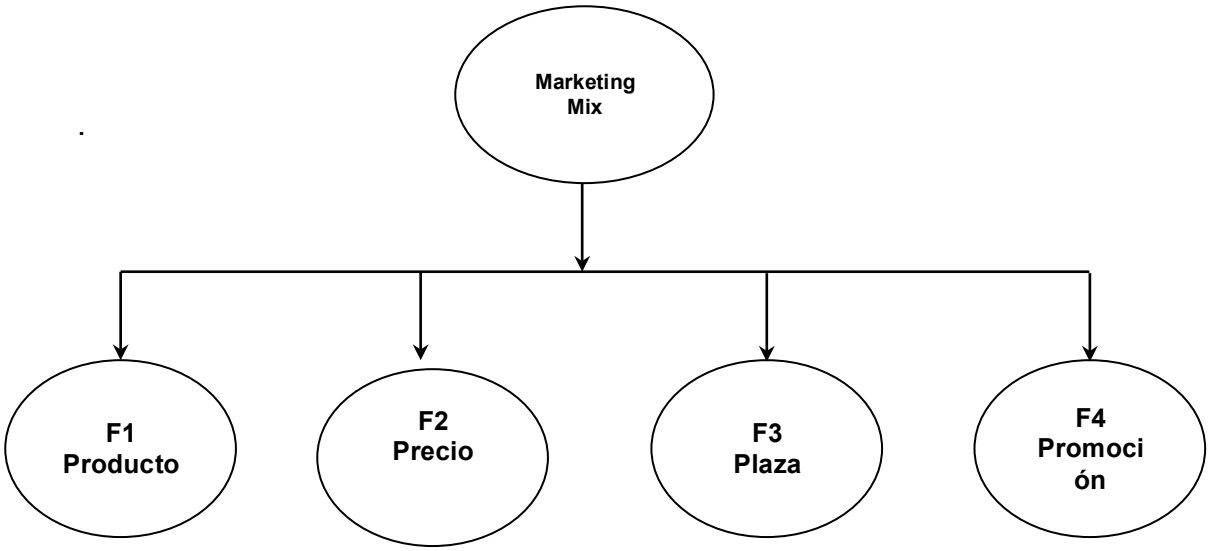
| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | pago | | | | | |
| **MPR6** | **V19** | Aplicamos una estrategia de precios | | | | | |
| **MPR7** | **V20** | Comúnmente negociamos el precio de nuestros productos con nuestros clientes | | | | | |
| **Indique si su empresa a nivel de plaza:** | | | **1** | **2** | **3** | **4** | **5** |
| **MPL1** | **V21** | Influye o controla los canales de distribución de sus productos | | | | | |
| **MPL2** | **V22** | Desarrolla o implementa técnicas innovadores de distribución | | | | | |
| **MPL3** | **V23** | Utiliza agentes de ventas altamente calificados y eficientes | | | | | |
| **MPL4** | **V24** | Cuenta con productos que son muy aceptados por los intermediarios del canal de distribución. | | | | | |
| **MPL5** | **V25** | Resuelve eficientemente sus problemas de logística | | | | | |
| **MPL6** | **V26** | Tiene una flexibilidad en sus procesos de logística | | | | | |
| **MPL7** | **V27** | Gestiona adecuadamente la cadena de suministro | | | | | |
| **MPL8** | **V28** | Permanentemente se mantiene en contacto con sus distribuidores | | | | | |
| **MPL9** | **V29** | Utiliza un software para controlas los pedidos y las entregas | | | | | |
| **MPL10** | **V30** | Tiene un sistema para controlar los la percepción de valor de marca de sus productos por parte de los intermediarios y distribuidores. | | | | | |
| **MPL11** | **V31** | Subcontrata frecuentemente las actividades de distribución y logística | | | | | |
| **Indique si su empresa a nivel de promoción:** | | | **1** | **2** | **3** | **4** | **5** |
| **MPO1** | **V32** | Aprovecha cada herramienta de comunicación para promocionar sus productos o servicios. | | | | | |
| **MPO2** | **V33** | Tiene una persona responsable capaz de monitorear la promoción de sus productos o servicios. | | | | | |
| **MPO3** | **V34** | La publicidad que realizamos es mejor que la que realiza la competencia | | | | | |
| **MPO4** | **V35** | Los medios de comunicación que utilizamos son los adecuados | | | | | |
| **MPO5** | **V36** | La inversión que realizamos en publicidad es la adecuada | | | | | |
| **MPO6** | **V37** | La publicidad que realizamos está dirigida a nuestro mercado meta | | | | | |
| **MPO7** | **V38** | Las ventas de la empresa se han incrementado gracias a la publicidad que realizamos | | | | | |
| **MPO8** | **V39** | Realizamos constantemente campañas promocionales de nuestros productos | | | | | |

Fuente: Maldonado-Guzmán (2016)

Cuyo diagrama de trayectoria se muestra en la **Figura 7.86**

**Figura 7.86. Diagrama de trayectoria modelo marketing mix**

.

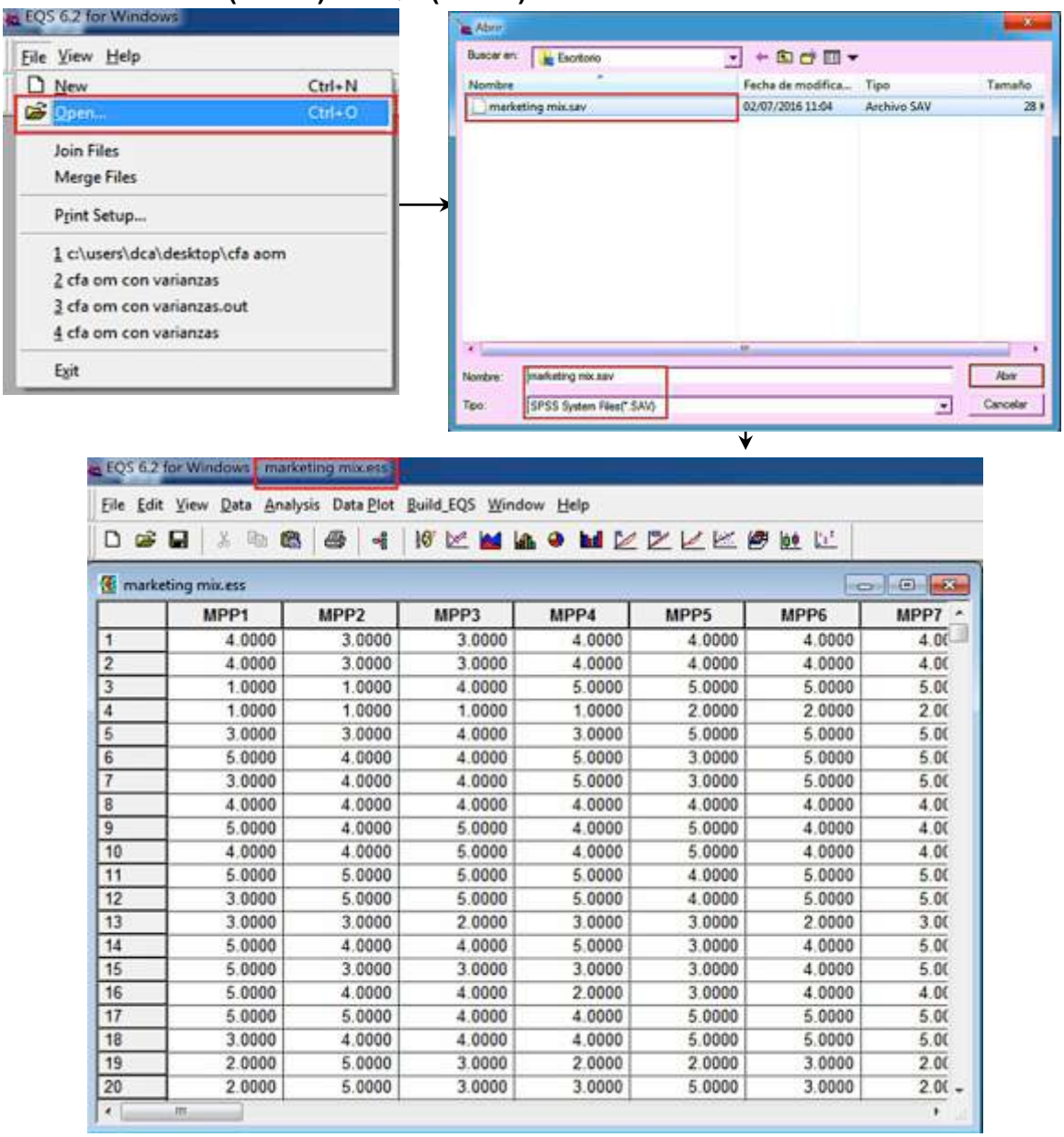



Fuente: Maldonado-Guzmán (2016)

## *EQS. Entrada de datos al programa*

Se sugiere seguir lo recomendado en las **Figuras 7.11,** accesando el archivo de entrada de programa (***.EQS**), llamado **marketing mix.sav (**formato **SPSS)**, el cual, al ser abierto, cambia su formato original ***.SAV** a ***.EES** (**marketing mix.ess**) para su despliegue y manipulación**.** Ver **Figura 7.87.**

**Figura 7.87. EQS. Selección de archivo externo para caso de estudio 2. Cambio de formato SPSS (*.SAV) a EQS (*.ESS)**




Fuente: SPSS 21 IBM con datos y adaptación propia.

Así también es posible verificar el archivo de programa de entrada **EQS**, del que resaltamos el nombre del archivo en el párrafo **/TITLE:** CFA y SEM de Marketing MIX Nombre: marketing mix. En el párrafo: **/SPECIFICATIONS,** la declarativa de analizar 39 variables con 400 casos.

**Nota: EQS**, para el caso de sistema operativo **WINDOWS** debe especificarse, en un archivo externo, la trayectoria del archivo de programa de entrada; por ejemplo: **DATAFILE='E:///cfa om.eqs'.** Si es **IOS**, sólo se menciona el nombre del archivo dentro de la computadora a nivel de escritorio con los archivos apareciendo en el emulador de Windows (***Parallels,*** en nuestro caso: **DATAFILE='marketing**

**mix.ess'**;**)**. Ver **Figura 7.88.**

**Figura 7.88. Vista de archivos desde emulador Windows (parallels) manejo de EQS con IOS**

Fuente: EQS 6. con datos y adaptación propia

Otra instrucción es **MATRIX=RAW**, que indica que la matriz de datos será generada por el mismo software. La instrucción **METHOD=ML** que indica que se aplicará máxima verosimilitud (**ML.*Maximum Likelihood***), la instrucción **ROBUST** en la que se da el indicativo que los datos tienen filtros para garantizar mínimos errores (no se recomienda trabajar con los datos originales) y finalmente, la instrucción **ANALYSIS= COV**, sobre la aplicación de covarianzas (es mejor a la matriz de correlaciones).
En el párrafo **/EQUATIONS**, se tienen las ecuaciones de todas y cada una de las variables con sus errores señalando con la supresión de *asterisco* **(*)** *al menos una variable que tenga varianza 1.* En el párrafo de **/VARIANCES**, observe en particular las expresiones que involucran a los constructos o factores que identifican al modelo de segundo orden: **F1=*; F2=*; F3=*; y F4=*** en las que al incluir al **F4** en las mismas, se da la idea de que *se correrá al mismo tiempo* ***CFA con SEM*****,** situación que es válida y que facilita el análisis, dado lo complejo del modelo. Se incluye el párrafo **/COVARIANCES,** ya que se consideran que corren como variables independientes, en todas las combinaciones cubiertas como:

**F1, F2 =*;**
**F1, F3 =*;**
**F1, F4 =*;**
**F2, F3 =*;**
**F2, F4 =*;**
**F3, F4 =*;**

Es importante hacer notar que tanto los nombres de la base de datos proveniente se **SPSS** (**.SAV**), tanto el nombre de variables y su aparición sea estrictamente como corresponde al modelo teórico**,** con lectura de izquierda a derecha, ya que esto es un

a condición necesaria para que el modelo **EQS**, pueda correr correctamente y/o subdividir las bases de datos. En otras palabras, debe haber correspondencia en nombre y orden de aparición de las variables de la base de datos, con el modelo teórico, con el fin de que **EQS** ejecute correctamente las instrucciones con las variables correspondientes y produzca los resultados esperados. **Ver Figura 7.89.**

**Figura 7.89. EQS Codificación de entrada de datos caso de estudio 2**

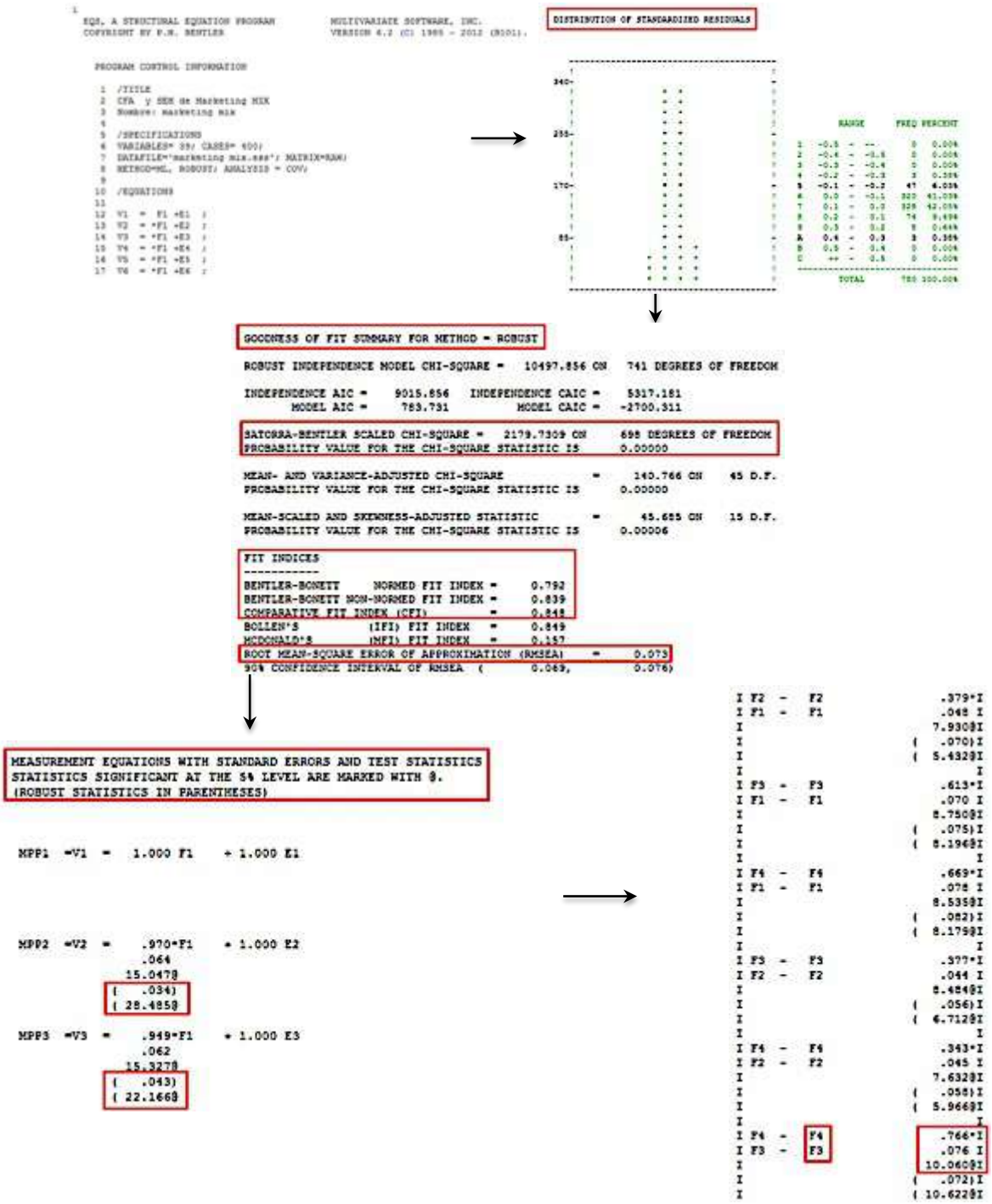
```
/TITLE
CFA  y SEM de Marketing MIX
Nombre: marketing mix

/SPECIFICATIONS
VARIABLES= 39; CASES= 400;
DATAFILE='marketing mix.ess'; MATRIX=RAW;
METHOD=ML, ROBUST; ANALYSIS = COV;

/EQUATIONS

V1  =  F1 +E1  ;
V2  = *F1 +E2  ;
V3  = *F1 +E3  ;
V4  = *F1 +E4  ;
V5  = *F1 +E5  ;
V6  = *F1 +E6  ;
V7  = *F1 +E7  ;
V8  = *F1 +E8  ;
V9  = *F1 +E9  ;
V10 = *F1 +E10 ;
V11 = *F1 +E11 ;
V12 = *F1 +E12 ;
V13 = *F1 +E13 ;

V14 =  F2 +E14 ;
V15 = *F2 +E15 ;
V16 = *F2 +E16 ;
V17 = *F2 +E17 ;
V18 = *F2 +E18 ;
V19 = *F2 +E19 ;
V20 = *F2 +E20 ;

V25 =  F3 +E25 ;
V26 = *F3 +E26 ;
V27 = *F3 +E27 ;
V28 = *F3 +E28 ;
V29 =  F3 +E29 ;
V30 = *F3 +E30 ;
V31 = *F3 +E31 ;

V32 =  F4 +E32 ;
V33 = *F4 +E33 ;
V34 = *F4 +E34 ;
V35 = *F4 +E35 ;
V36 = *F4 +E36 ;
V37 = *F4 +E37 ;
V38 = *F4 +E38 ;
V39 = *F4 +E39 ;

/VARIANCES

E1  =*;
E2  =*;
E3  =*;
E4  =*;
E5  =*;
E6  =*;
E7  =*;
E8  =*;
E9  =*;
E10 =*;
E11 =*;
E12 =*;
E13 =*;
E14 =*;
E15 =*;
E16 =*;
E33 =*;
E34 =*;
E35 =*;
E36 =*;
E37 =*;
E38 =*;
E39 =*;

F1 =*;
F2 =*;
F3 =*;
F4 =*;

/COVARIANCES

F1, F2 =*;
F1, F3 =*;
F1, F4 =*;
F2, F3 =*;
F2, F4 =*;
F3, F4 =*;

/LMTEST

/WTEST

/PRINT
FIT = ALL;

/END
```

Fuente: EQS 6. con datos y adaptación propia

Realizado lo anterior, ejecute el programa **EQS**, siguiendo lo mostrado en las instrucciones del apartado de *cómo ejecutar programa.*

Se procede a revisar rápidamente las secciones de reporte de resultados como se ve en la **Figura 7.90.**

**Figura 7.90. CFA, SEM y EQS. Proceso de revisión rápida de resultados caso estudio 3.**




Fuente: EQS 6. con datos y adaptación propia

```
STANDARDIZED SOLUTION:                                                    R-SQUARED

  MPP1  =V1  =   .762 F1    + .648 E1                                        .580
  MPP2  =V2  =   .720*F1    + .694 E2                                        .518
  MPP3  =V3  =   .731*F1    + .682 E3                                        .534
  MPP4  =V4  =   .618*F1    + .786 E4                                        .382
  MPP5  =V5  =   .735*F1    + .679 E5                                        .539
  MPP6  =V6  =   .723*F1    + .691 E6                                        .522
  MPP7  =V7  =   .733*F1    + .680 E7                                        .537
  MPP8  =V8  =   .803*F1    + .596 E8                                        .645
  MPP9  =V9  =   .820*F1    + .572 E9                                        .673
MPP10   =V10 =   .810*F1    + .587 E10                                       .656
MPP11   =V11 =   .729*F1    + .684 E11                                       .532
MPP12   =V12 =   .778*F1    + .629 E12                                       .605
MPP13   =V13 =   .695*F1    + .719 E13                                       .483
  MPR1  =V14 =   .592 F2    + .806 E14                                       .350
  MPR2  =V15 =   .595*F2    + .804 E15                                       .354
  MPR3  =V16 =   .601*F2    + .800 E16                                       .361
  MPR4  =V17 =   .773*F2    + .634 E17                                       .598
  MPR5  =V18 =   .768*F2    + .640 E18                                       .590
  MPR6  =V19 =   .819*F2    + .573 E19                                       .671
  MPR7  =V20 =   .699*F2    + .715 E20                                       .488
  MPL1  =V21 =   .725 F3    + .689 E21                                       .525
  MPL2  =V22 =   .792*F3    + .611 E22                                       .627
  MPL3  =V23 =   .765*F3    + .643 E23                                       .586
  MPL4  =V24 =   .806*F3    + .592 E24                                       .650
  MPL5  =V25 =   .815 F3    + .579 E25                                       .664
  MPL6  =V26 =   .815*F3    + .580 E26                                       .664
  MPL7  =V27 =   .805*F3    + .593 E27                                       .648
  MPL8  =V28 =   .703*F3    + .711 E28                                       .494
  MPL9  =V29 =   .713 F3    + .701 E29                                       .508
MPL10   =V30 =   .813*F3    + .582 E30                                       .661
MPL11   =V31 =   .728*F3    + .686 E31                                       .529
  PMO1  =V32 =   .776 F4    + .630 E32                                       .603
  MPO2  =V33 =   .765*F4    + .644 E33                                       .585
  MPO3  =V34 =   .884*F4    + .467 E34                                       .782
  MPO4  =V35 =   .862*F4    + .507 E35                                       .743
  MPO5  =V36 =   .893*F4    + .450 E36                                       .797
  MPO6  =V37 =   .835*F4    + .550 E37                                       .697
  MPO7  =V38 =   .871*F4    + .491 E38                                       .759

 STANDARDIZED SOLUTION:                                                   R-SQUARED

03-NOV-17       PAGE:  20  EQS      Licensee:
 TITLE:    CFA  y SEM de Marketing MIX

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

  MPO8  =V39 =   .729*F4    + .685 E39                                       .531

    MULTIVARIATE LAGRANGE MULTIPLIER TEST BY SIMULTANEOUS PROCESS IN STAGE  1

     PARAMETER SETS (SUBMATRICES) ACTIVE AT THIS STAGE ARE:

        PVV PFV PFF PDD GVV GVF GFV GFF BVF BFF

          CUMULATIVE MULTIVARIATE STATISTICS        UNIVARIATE INCREMENT
          ----------------------------------        --------------------------------
                                                                      HANCOCK'S     ROBUST
                                CHI-                    CHI-          SEQUENTIAL   PREDICTED
    STEP   PARAMETER            SQUARE  D.F. PROB.    SQUARE PROB.   D.F. PROB.   RMSEA   CFI
    ----   ---------            ------  ---- -----    ------ -----   ---- -----   -----------
       1     V14,F1              52.315     1 0.000   52.315 0.000    698 1.000   0.072 0.853
       2     V11,F4              87.292     2 0.000   34.976 0.000    697 1.000   0.071 0.857
       3     V32,F2             111.710     3 0.000   24.419 0.000    696 1.000   0.070 0.859
       4      V9,F3             135.570     4 0.000   23.859 0.000    695 1.000   0.070 0.862
       5     V24,F2             154.242     5 0.000   18.672 0.000    694 1.000   0.069 0.863
       6     V16,F1             169.282     6 0.000   15.040 0.000    693 1.000   0.069 0.865
```

Fuente: EQS 6 con datos y adaptación propi

## *Discusión rápida de resultados.*

De los reportes emitidos en la **Figura 7.90**, podemos observar y concluir de cada sección de reporte lo siguiente:

- Gráfico de distribución de residuos (*distribution of standardized residuals*). Tiene problemas de datos alejados de la media. *No Cumple.*
- Índices de ajuste (*fit indices*) robustos (*robust*),valores a analizar:

*Chi-Cuadrada de Satorra-Bentler*=2179.7309; gl= 698 (**Nota:** 2179.7309/698=3.122>2 Tendencia a que los índices de ajuste sean <.9); valor p<0.01**.** *No Cumple.*

**NFI=** .792; **NNFI=** .839 ; **CFI=** .848. *No Cumple* >.9 (para nueva escala). **RMSEA=.**073<.08. *Cumple.*

- Ecuaciones de medición (*measurment equations),*valores de t (@) =.05.*Cumple.*
- No genera Ecuaciones de constructo (*construct equations*)
- Covarianzas de las variables independientes (*covariances among independent variables*), valores de t (@)= .05, pero hay casos con correlación >.6 (nunca > .8) como **F3-F4= .766.** *No Cumple.*
- **Nota:** si cambia la matriz de **COV** a **COR** baja la correlación; de lo que se trata es generar la mayor correlación a partir de las covarianzas) y tiene potencial problema de *validez discriminante* la cual se consigue cuando, si diversos instrumentos de medida están diseñados para medir distintas variables latentes, entonces las correlaciones entre dichos instrumentos son bajas. Una escala tiene, pues, validez discriminante solamente cuando no mide un constructo para el que no se diseñó. Generalmente se realiza a través de tres Tests: 1. Test de las diferencias entre las Chi cuadrado 2. Test del intervalo de confianza 3. Test de la varianza extraída.

Se deben aplicar entre cada par de factores.

Se suele aplicar sólo al par de factores que más problema pueden causar, es decir, los que tienen una covarianza más elevada entre sí. En este caso **F3-F4**=.766, siendo el *Test de las diferencias entre las Chi-Cuadrado* (Baggozi y Phillips, 1982), de la siguiente forma:

- *Se estima el modelo de medida pero fijando a 1 las covarianzas entre los dos factores cuya validez discriminante suscita dudas* **(F3 y F4**).
- *Se restan las Chi cuadrado de ambos modelos y los grados de libertad.*
- *El valor del estadístico (la resta) debe ser superior al valor crítico de la Chi-Cuadrado para ese nivel de grados de libertad.*
- *Esto querrá decir que el modelo en el que los factores están relacionados es mejor que aquel en que esos factores son el mismo (ausencia de validez discriminante).*

Esto se verá más adelante al correr nuevamente el programa sin las variables problema

- Solución estandarizada (*standardized solutions*), *¿todas las cargas factoriales son >.6?.* Bagozzi y Yi (1988) que cada carga debe ser factorial debe ser >.6. *No Cumple en* **V14** (.592) ; **V15** (.595); **V16** (.601)
- Prueba del multiplicador de Lagrange multivariado *(Lagrange multiplier test*), se observa que **V14** y **V16** *tienen problemas.*

## *Discusión rápida de resultados sin V14, V15 y V16*

Por lo anterior, se sugiere eliminar las variables **V14,V15** junto a **V16** (recuerde colocar en el programa de entrada **marketing mix.eqs** el símbolo **!** a las ecuaciones que se encuentran en párrafo **/EQUATIONS** para **V14,V15,V16** así como en el párrafo **VARIANCES** los errores **E14; E15 y E16** y hacer revisión rápida. Ver **Figura 7.91.**

**Figura 7.91. CFA y SEM caso de estudio 3 sin V14, V15 y V16**

```
/TITLE
CFA  y SEM de Marketing MIX
Nombre: marketing mix sin V14,V15 Y V16

/SPECIFICATIONS
VARIABLES= 39; CASES= 400;
DATAFILE='marketing mix.ess'; MATRIX=RAW;
METHOD=ML, ROBUST; ANALYSIS = COV;

/EQUATIONS

V1  =  F1 +E1  ;
V2  = *F1 +E2  ;
V3  = *F1 +E3  ;
V4  = *F1 +E4  ;
V5  = *F1 +E5  ;
V6  = *F1 +E6  ;
V7  = *F1 +E7  ;
V8  = *F1 +E8  ;
V9  = *F1 +E9  ;
V10 = *F1 +E10 ;
V11 = *F1 +E11 ;
V12 = *F1 +E12 ;
V13 = *F1 +E13 ;

!V14 =  F2 +E14 ;
!V15 = *F2 +E15 ;
!V16 = *F2 +E16 ;
V17 = F2 +E17 ;
V18 = *F2 +E18 ;
V19 = *F2 +E19 ;
V20 = *F2 +E20 ;

V21 =  F3 +E21 ;
V22 = *F3 +E22 ;
```

```
E10 =*;
E11 =*;
E12 =*;
E13 =*;
!E14 =*;
!E15 =*;
!E16 =*;
E17 =*;
E18 =*;
E19 =*;
E20 =*;
E21 =*;
```

```
E33 =*;
E34 =*;
E35 =*;
E36 =*;
E37 =*;
E38 =*;
E39 =*;

F1 =*;
F2 =*;
F3 =*;
F4 =*;

/COVARIANCES

F1, F2 =*;
F1, F3 =*;
F1, F4 =*;
F2, F3 =*;
F2, F4 =*;
F3, F4 =1;

/LMTEST

/WTEST

/PRINT
FIT = ALL;

/END
```

```
PROGRAM CONTROL INFORMATION

 1  /TITLE
 2  CFA  y SEM de Marketing MIX
 3  Nombre: marketing mix sin V14,V15 Y V16
 4
 5  /SPECIFICATIONS
 6  VARIABLES= 39; CASES= 400;
 7  DATAFILE='marketing mix.ess'; MATRIX=RAW;
 8  METHOD=ML, ROBUST; ANALYSIS = COV;
 9
10  /EQUATIONS
11
12  V1  =  F1 +E1  ;
13  V2  = *F1 +E2  ;
14  V3  = *F1 +E3  ;
15  V4  = *F1 +E4  ;
16  V5  = *F1 +E5  ;
17  V6  = *F1 +E6  ;
18  V7  = *F1 +E7  ;
19  V8  = *F1 +E8  ;
20  V9  = *F1 +E9  ;
21  V10 = *F1 +E10 ;
22  V11 = *F1 +E11 ;
23  V12 = *F1 +E12 ;
24  V13 = *F1 +E13 ;
25
26  !V14 =  F2 +E14 ;
27  !V15 = *F2 +E15 ;
28  !V16 = *F2 +E16 ;
29  V17 = F2 +E17 ;
30  V18 = *F2 +E18 ;
31  V19 = *F2 +E19 ;
```

```
DISTRIBUTION OF STANDARDIZED RESIDUALS

 300-                 *
                      *
                      *
                      *
                      *
 225-            *    *
                 *    *
                 *    *
                 *    *
                 *    *
 150-            *    *
                 *    *
                 *    *
                 *    *    *
                 *    *    *
  75-            *    *    *
                 *    *    *
                 *    *    *
                 *    *    *
              *  *    *    *  *
     1  2  3  4  5  6  7  8  9  A  B  C

                 RANGE      FREQ PERCENT
  1  -0.5  -  --            0   0.00%
  2  -0.4  -  -0.5          0   0.00%
  3  -0.3  -  -0.4          2   0.30%
  4  -0.2  -  -0.3         12   1.80%
  5  -0.1  -  -0.2        230  34.53%
  6   0.0  -  -0.1        303  45.50%
  7   0.1  -   0.0        175  26.28%
  8   0.2  -   0.1         10   1.50%
  9   0.3  -   0.2          1   0.15%
  A   0.4  -   0.3          0   0.00%
  B   0.5  -   0.4          0   0.00%
  C   --   -   0.5          0   0.00%
                 TOTAL      666 100.00%

EACH "*" REPRESENTS   15 RESIDUALS
```

Fuente: EQS 6. con datos y adaptación propia

```
GOODNESS OF FIT SUMMARY FOR METHOD = ROBUST

ROBUST INDEPENDENCE MODEL CHI-SQUARE =     9877.374 ON    630 DEGREES OF FREEDOM

INDEPENDENCE AIC =     8617.374    INDEPENDENCE CAIC =     5472.751
       MODEL AIC =      669.030           MODEL CAIC =    -2289.926

SATORRA-BENTLER SCALED CHI-SQUARE =    1842.0295 ON     591 DEGREES OF FREEDOM
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS       0.00000

MEAN- AND VARIANCE-ADJUSTED CHI-SQUARE              =     134.114 ON     43 D.F.
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS       0.00000

MEAN-SCALED AND SKEWNESS-ADJUSTED STATISTIC         =      44.076 ON     14 D.F.
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS       0.00006

FIT INDICES
-----------
BENTLER-BONETT      NORMED FIT INDEX =      0.814
BENTLER-BONETT NON-NORMED FIT INDEX =      0.856
COMPARATIVE FIT INDEX (CFI)          =      0.865
BOLLEN'S          (IFI) FIT INDEX    =      0.865
MCDONALD'S        (MFI) FIT INDEX    =      0.209
ROOT MEAN-SQUARE ERROR OF APPROXIMATION (RMSEA)     =      0.073
90% CONFIDENCE INTERVAL OF RMSEA  (        0.069,          0.077)
```

```
MEASUREMENT EQUATIONS WITH STANDARD ERRORS AND TEST STATISTICS
STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.
(ROBUST STATISTICS IN PARENTHESES)

MPF1  =V1  =   1.000 F1    + 1.000 E1

MPF2  =V2  =    .971*F1    + 1.000 E2
                 .062
               15.7038
             (   .032)
             ( 30.5438)

MPF3  =V3  =    .950*F1    + 1.000 E3
                 .059
               15.9848
             (   .040)
             ( 23.8598)
```

```
MPF1  =V1  =   [illegible]  + .633 E1                    .339
MPF2  =V2  =   .734*F1      + .679 E2                    .539
MPF3  =V3  =   .745*F1      + .667 E3                    .555
MPF4  =V4  =   .631*F1      + .776 E4                    .398
MPF5  =V5  =   .749*F1      + .663 E5                    .560
MPF6  =V6  =   .731*F1      + .682 E6                    .535
MPF7  =V7  =   .742*F1      + .670 E7                    .551
MPF8  =V8  =   .812*F1      + .583 E8                    .660
MPF9  =V9  =   .829*F1      + .559 E9                    .687
MPF10 =V10 =   .820*F1      + .572 E10                   .673
MPF11 =V11 =   .743*F1      + .669 E11                   .552
MPF12 =V12 =   .759*F1      + .615 E12                   .622
MPF13 =V13 =   .706*F1      + .708 E13                   .499
MPF4  =V17 =   .785 F2      + .619 E17                   .614
MPF5  =V18 =   .841*F2      + .541 E18                   .708
MPF6  =V19 =   .835*F2      + .550 E19                   .698
MPF7  =V20 =   .715*F2      + .699 E20                   .512
MPL1  =V21 =   .754 F3      + .657 E21                   .568
MPL2  =V22 =   .813*F3      + .582 E22                   .661
MPL3  =V23 =   .790*F3      + .613 E23                   .624
MPL4  =V24 =   .826*F3      + .563 E24                   .683
MPL5  =V25 =   .839 F3      + .545 E25                   .703
MPL6  =V26 =   .833*F3      + .553 E26                   .694
MPL7  =V27 =   .823*F3      + .568 E27                   .678
MPL8  =V28 =   .727*F3      + .687 E28                   .528
MPL9  =V29 =   .744 F3      + .669 E29                   .553
MPL10 =V30 =   .832*F3      + .554 E30                   .693
MPL11 =V31 =   .753*F3      + .658 E31                   .567
PMO1  =V32 =   .811 F4      + .585 E32                   .658
MPO2  =V33 =   .785*F4      + .615 E33                   .621
MPO3  =V34 =   .897*F4      + .441 E34                   .805
MPO4  =V35 =   .877*F4      + .480 E35                   .770
MPO5  =V36 =   .905*F4      + .425 E36                   .819
MPO6  =V37 =   .852*F4      + .524 E37                   .725
MPO7  =V38 =   .885*F4      + .465 E38                   .784
MPO8  =V39 =   .754*F4      + .657 E39                   .568
```

```
V                                   F
---                                 ---
                I F2  -   F2                .704*I
                I F1  -   F1                .074 I
                I                         9.528@I
                I                     (   .075)I
                I                     (  9.377@I
                I                             I
                I F3  -   F3                .742*I
                I F1  -   F1                .062 I
                I                        11.887@I
                I                     (   .064)I
                I                     ( 11.654@I
                I                             I
                I F4  -   F4                .830*I
                I F1  -   F1                .067 I
                I                        12.366@I
                I                     (   .063)I
                I                     ( 13.273@I
                I                             I
                I F3  -   F3                .805*I
                I F2  -   F2                .063 I
                I                        12.778@I
                I                     (   .069)I
                I                     ( 11.687@I
                I                             I
                I F4  -   F4                .764*I
                I F2  -   F2                .059 I
                I                        12.880@I
                I                     (   .056)I
                I                     ( 13.711@I
                I                             I
                I F4  -   F4                1.000 I
                I F3  -   F3                      I
```

```
MULTIVARIATE LAGRANGE MULTIPLIER TEST BY SIMULTANEOUS PROCESS IN STAGE  1

PARAMETER SETS (SUBMATRICES) ACTIVE AT THIS STAGE ARE:

   PVV PFV PFF PDD GVV GVF GFV GFF BVF BFF

   CUMULATIVE MULTIVARIATE STATISTICS      UNIVARIATE INCREMENT
   ----------------------------------      --------------------------------
                                                          MANCOCK'S    ROBUST
                 CHI-                          CHI-     SEQUENTIAL   PREDICTED
STEP  PARAMETER  SQUARE  D.F. PROB.          SQUARE PROB.  D.F. PROB.  RMSEA   CFI
----  ---------  ------  ---- -----          ------ -----  ---- -----  -----  -----
  1     V11,F4    32.474   1 0.000          32.474 0.000   591 1.000  0.072  0.865
  2      V9,F3    54.696   2 0.000          22.221 0.000   590 1.000  0.071  0.870
  3     V24,F2    74.309   3 0.000          19.613 0.000   589 1.000  0.071  0.872
  4     V10,F2    91.238   4 0.000          16.929 0.000   588 1.000  0.070  0.874
  5     V36,F2   106.802   5 0.000          15.564 0.000   587 1.000  0.070  0.876
  6     V26,F1   121.372   6 0.000          14.569 0.000   586 1.000  0.070  0.877
  7      V5,F4   135.437   7 0.000          14.066 0.000   585 1.000  0.069  0.879
  8     V39,F1   147.940   8 0.000          12.503 0.000   584 1.000  0.069  0.880
  9     V37,F3   160.791   9 0.000          12.851 0.000   583 1.000  0.069  0.881
 10     V34,F3   178.233  10 0.000          17.442 0.000   582 1.000  0.068  0.883
 11      V6,F3   190.444  11 0.000          12.211 0.000   581 1.000  0.068  0.884
```

```
MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

CORRELATIONS AMONG INDEPENDENT VARIABLES
----------------------------------------

V                                   F
---                                 ---
                I F2  -   F2                .627*I
                I F1  -   F1                      I
                I                                 I
                I F3  -   F3                .622*I
                I F1  -   F1                      I
                I                                 I
                I F4  -   F4                .634*I
                I F1  -   F1                      I
                I                                 I
                I F3  -   F3                .726*I
                I F2  -   F2                      I
                I                                 I
                I F4  -   F4                .628*I
                I F2  -   F2                      I
                I                                 I
                I F4  -   F4                .773 I
                I F3  -   F3                      I
                I                                 I
```

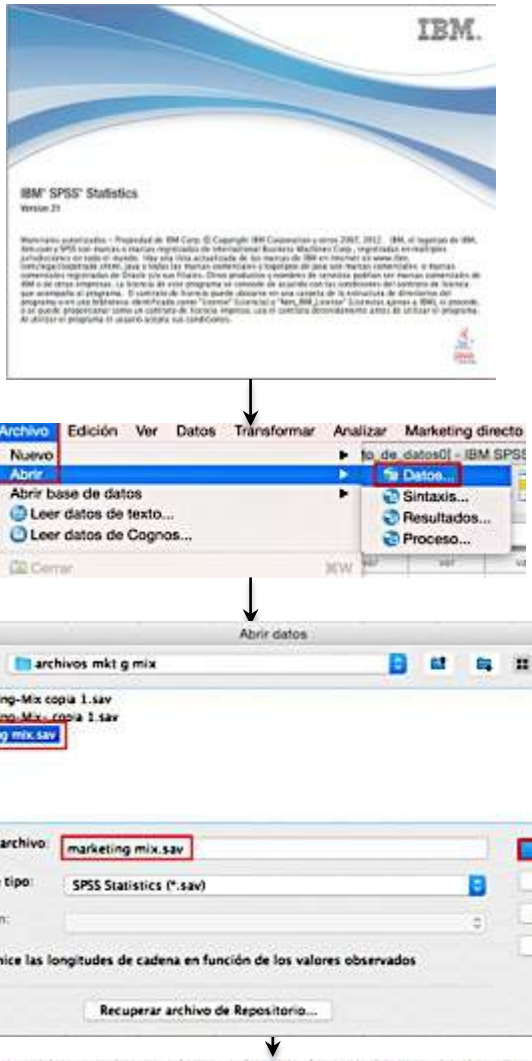
Fuente: EQS 6. con datos y adaptación propia

Nuevamente, de los reportes emitidos podemos observar y concluir de cada sección de reporte lo siguiente:
- Gráfico de distribución de residuos (*distribution of standardized residuals*). Mejor que el anterior aunque aún tiene problemas de datos alejados de la media. *No Cumple.*
- *Índices de ajuste* (*fit indices*) *robustos* (*robust*),valores a analizar:
*Chi-Cuadrada de Satorra-Bentler*= 1842.0295; **gl**= 591
**Nota:** 1842.0295/591=3.116>2 Tendencia a que los índices de ajuste sean <.9); valor p<0.01**.** *No Cumple.*
**NFI=** .814; **NNFI=** .856 ; **CFI=** .865. *No Cumple >.9* (para nueva escala). **RMSEA=**.073<.08. *Cumple.*
- Ecuaciones de medición (*measurment equations),*valores de t (@) =.05.*Cumple*
- No genera Ecuaciones de constructo (*construct equations*).
- Covarianzas de las variables independientes (*covariances among independent variables*), valores de t (@)= .05, y dado que condicionamos **F3-F4=1** el resto de los casos incrementan sus correlaciones aunque <1 (normalmente nunca > .8). Cabe destacar que sería un máximo de 3.9 variables=4 (10%) del total que podrían ser eliminadas. Se anima al lector a decidir cuál otra sería la candidata a partir de observar el multiplicador de Lagrange multivariado. Por lo pronto **V14, V15, V16** hablan de precios; se tiene que revisar de la teoría el porque estos tres indicadores aportan muy poco a la escala; aparentemente generan confusión en los encuestados , por lo que también se debe revisar su redacción y/o presentación. Para fines didácticos, se acuerda que el modelo en este momento está ajustado.
- Solución estandarizada (*standardized solutions*), *¿todas las cargas factoriales son >.6?.* Bagozzi y Yi (1988) que cada carga debe ser factorial debe ser >.6. *Cumple.*

## *Cálculo del alfa de Cronbach, sin V14,V15 y V16*

Dado que se requiere el *alfa de Cronbach*, por constructo (*No por modelo, que es lo que se reporta en **EQS***), requerimos calcularlo.Este cálculo debará proceder de la base de datos **SPSS**, en nuestro caso **marketing mix.sav** y proceder únicamente con las variables que quedaron del modelos estructura, o sea, *No incluir las variables* **V14, V15 ni V16.** Es importante considerar que si está utilizando **SPSS** soportado con **IOS** o **Windows (***con un emulador como **Parallels***),** ubique su archivo **.SAV**, de acuerdo al sistema operativo con el cual va a trabajar. En nuestro caso es deforma directa en sistema operativo **IOS.**
Así, la secuencia de comandos para ubicar y abrir el archivo **marketing mix.sav**, es:
**Teclear en ícono SPSS desde aplicaciones o barra de herramientas>Archivo->Abrir (ir a subdirectorio correspondiente)->Nombre de archivo: marketing mix.sav->Abrir. Ver Figura 7.92.**

**Figura 7.92. Secuencia de comandos apertura de archivo base de datos SPSS, caso de estudio 3.**



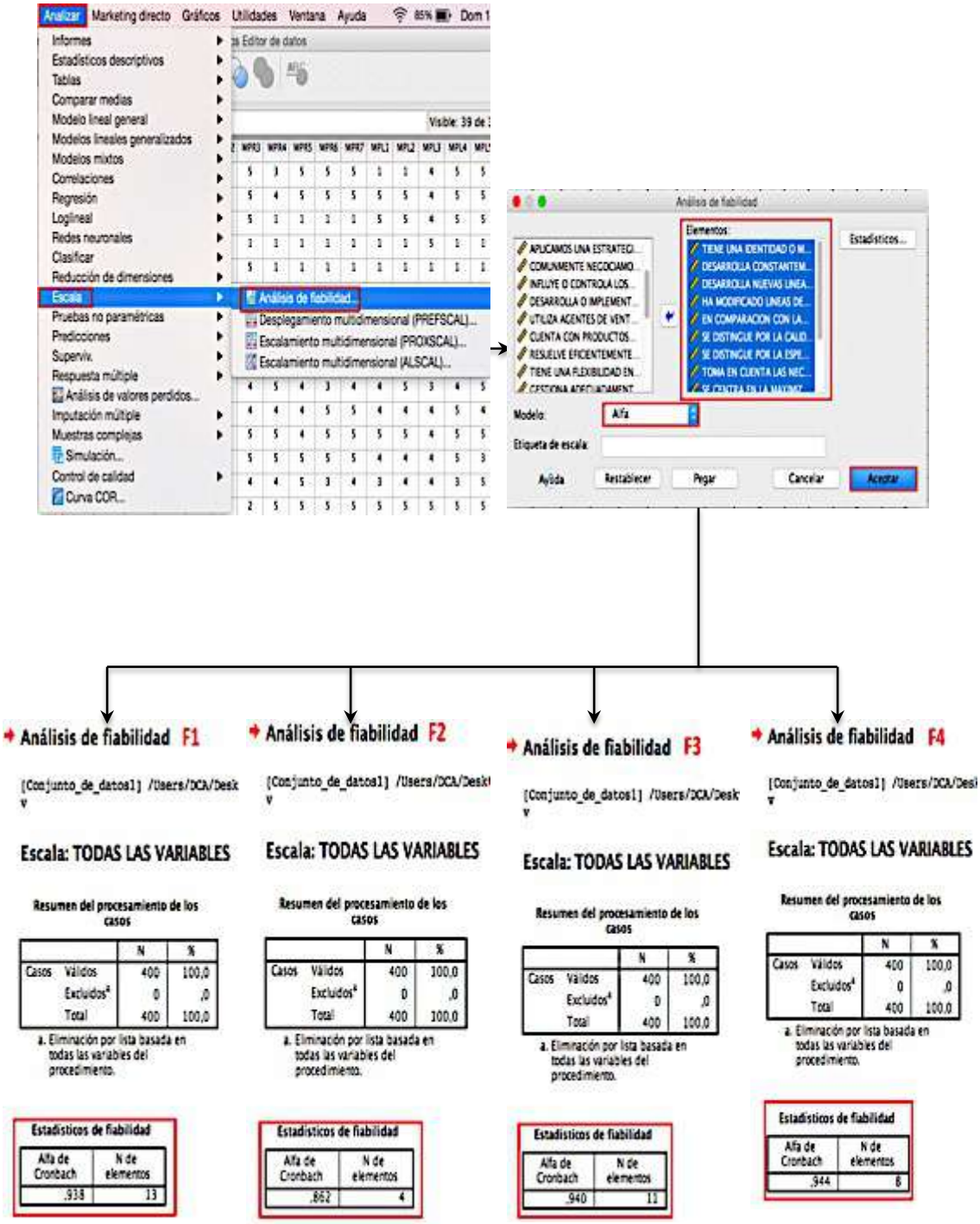

| | MPP1 | MPP2 | MPP3 | MPP4 | MPP5 | MPP6 | MPP7 | MPP8 | MPP9 | MPP10 | MPP11 | MPP12 | MPP13 | MPR1 | MPR2 | MPR3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 4 | 3 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 3 | 4 | 4 | 5 | 4 | 5 | 5 |
| 2 | 4 | 3 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 3 | 4 | 4 | 5 | 5 | 5 | 5 |
| 3 | 1 | 1 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 4 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 4 | 3 | 1 | 5 | 2 | 1 | 1 |
| 5 | 3 | 3 | 4 | 3 | 5 | 5 | 5 | 5 | 5 | 5 | 2 | 5 | 5 | 5 | 1 | 5 |
| 6 | 5 | 4 | 4 | 5 | 3 | 5 | 5 | 5 | 5 | 4 | 2 | 4 | 5 | 4 | 1 | 4 |
| 7 | 3 | 4 | 4 | 5 | 3 | 5 | 5 | 5 | 5 | 4 | 2 | 4 | 5 | 4 | 3 | 5 |
| 8 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 3 | 5 |
| 9 | 5 | 4 | 5 | 4 | 5 | 4 | 4 | 4 | 5 | 4 | 4 | 4 | 4 | 4 | 3 | 3 |
| 10 | 4 | 4 | 5 | 4 | 5 | 4 | 4 | 5 | 4 | 5 | 4 | 4 | 4 | 4 | 3 | 4 |

Fuente: SPSS 21 IBM. con datos y adaptación propia

Así, se procederá a realizar el cálculo de *alfa de Cronbach*, para cada uno de los ítems de los factores, con la siguiente secuencia inicial de comandos:

**Analizar->Escala->Análisis de fiabilidad->Modelo: Alfa->Elementos: (F1= V1 hasta V13; F2=V17 hasta V20; F3=V21 hasta V31 y F4=V32 hasta V39, según el caso) >Aceptar. Ver Figura 7.93.**

**Figura 7.93. Secuencia de comandos cálculo alfa de Cronbach con SPSS, caso de estudio 3.**




Fuente: SPSS 21 IBM. con datos y adaptación propia

En consecuencia, el *alfa de Cronbach* por factor, es:
**F1** (V2; V3; V4; V5; V6; V7; V8; V9; V10; V11; V12; V13)= .938
**F2** (V17; V18; V19; V20)= .862
**F3** (V21; V22; V23; V24; V25; V26; V27; V28; V29; V30; V31)= .940
**F4** (V32; V33; V34; V35; V36; V37; V38; V39)= .944

## *Cálculo de IFC, sin V14,V15 y V16*

Normalmente, en la investigación científica los modelos teóricos desarrollados en la literatura no aparecen implicados un solo constructo, sino varios, por lo cual el *alfa de Cronbach* para cada factor por separado no considera la influencia sobre la fiabilidad del resto de constructos. Por esta razón, Fornell y Larcker (1981) proponen el cálculo del índice de la fiabilidad compuesta (**IFC**) para cada factor que, interpretándose exactamente igual que el *alfa de Cronbach,* Sí que tiene en cuenta las interrelaciones. Por lo tanto, para determinarlo partimos del reporte de solución estandarizada (*standardized solution*) de la **Figura 7.91.**

| | | | | | |
|---|---|---|---|---|---|
| MPP1 | =V1 | = | .774 F1 | + .633 E1 | .599 |
| MPP2 | =V2 | = | .734*F1 | + .679 E2 | .539 |
| MPP3 | =V3 | = | .745*F1 | + .667 E3 | .555 |
| MPP4 | =V4 | = | .631*F1 | + .776 E4 | .398 |
| MPP5 | =V5 | = | .749*F1 | + .663 E5 | .560 |
| MPP6 | =V6 | = | .731*F1 | + .682 E6 | .535 |
| MPP7 | =V7 | = | .742*F1 | + .670 E7 | .551 |
| MPP8 | =V8 | = | .812*F1 | + .583 E8 | .660 |
| MPP9 | =V9 | = | .829*F1 | + .559 E9 | .687 |
| MPP10 | =V10 | = | .820*F1 | + .572 E10 | .673 |
| MPP11 | =V11 | = | .743*F1 | + .669 E11 | .552 |
| MPP12 | =V12 | = | .789*F1 | + .615 E12 | .622 |
| MPP13 | =V13 | = | .706*F1 | + .708 E13 | .499 |
| MPR4 | =V17 | = | .785 F2 | + .619 E17 | .616 |
| MPR5 | =V18 | = | .841*F2 | + .541 E18 | .708 |
| MPR6 | =V19 | = | .835*F2 | + .550 E19 | .698 |
| MPR7 | =V20 | = | .715*F2 | + .699 E20 | .512 |
| MPL1 | =V21 | = | .754 F3 | + .657 E21 | .568 |
| MPL2 | =V22 | = | .813*F3 | + .582 E22 | .661 |
| MPL3 | =V23 | = | .790*F3 | + .613 E23 | .624 |
| MPL4 | =V24 | = | .826*F3 | + .563 E24 | .683 |
| MPL5 | =V25 | = | .839 F3 | + .545 E25 | .703 |
| MPL6 | =V26 | = | .833*F3 | + .553 E26 | .694 |
| MPL7 | =V27 | = | .823*F3 | + .568 E27 | .678 |
| MPL8 | =V28 | = | .727*F3 | + .687 E28 | .528 |
| MPL9 | =V29 | = | .744 F3 | + .669 E29 | .553 |
| MPL10 | =V30 | = | .832*F3 | + .554 E30 | .693 |
| MPL11 | =V31 | = | .753*F3 | + .658 E31 | .567 |
| PMO1 | =V32 | = | .811 F4 | + .585 E32 | .658 |
| MPO2 | =V33 | = | .788*F4 | + .615 E33 | .621 |
| MPO3 | =V34 | = | .897*F4 | + .441 E34 | .805 |
| MPO4 | =V35 | = | .877*F4 | + .480 E35 | .770 |
| MPO5 | =V36 | = | .905*F4 | + .425 E36 | .819 |
| MPO6 | =V37 | = | .852*F4 | + .524 E37 | .725 |
| MPO7 | =V38 | = | .885*F4 | + .465 E38 | .784 |
| MPO8 | =V39 | = | .754*F4 | + .657 E39 | .568 |

Calculando el reporte **IFC**, como se observa en la **Figura 7.94**

**Figura 7.94. CFA, SEM y EQS. Recopilación de información que calcula el IFC del caso de estudio 3 sin V14,V15 y V16**

| Constructo e indicadores | Carga Factorial Estandarizada $L_{ij}$ | Varianza del Término de Error $Var(E_{ij}) = 1 - (L_{ij})^2$ | Indice de Fiabilidad Compuesta IFC >.7 $\sum(L_{ij})^2/[\sum(L_{ij})^2 + \sum Var(E_{ij})]$ |
|---|---|---|---|
| **Producto ($F_1$)** | | | |
| $V_1$ | .774 | 0.401 | **.945** |
| $V_2$ | .734 | 0.461 | |
| $V_3$ | .745 | 0.445 | |
| $V_4$ | .631 | 0.602 | |
| $V_5$ | .749 | 0.439 | |
| $V_6$ | .731 | 0.466 | |
| $V_7$ | .742 | 0.449 | |
| $V_8$ | .812 | 0.341 | |
| $V_9$ | .829 | 0.313 | |
| $V_{10}$ | .820 | 0.328 | |
| $V_{11}$ | .743 | 0.448 | |
| $V_{12}$ | .789 | 0.377 | |
| $V_{13}$ | .706 | 0.502 | |
| Total | 9.805 | 5.571 | |
| **Precio ($F_2$)** | | | |
| $V_{17}$ | .785 | 0.384 | **.873** |
| $V_{18}$ | .841 | 0.293 | |
| $V_{19}$ | .835 | 0.303 | |
| $V_{20}$ | .715 | 0.489 | |
| Total | 3.176 | 1.468 | |
| **Plaza ($F_3$)** | | | |
| $V_{21}$ | .754 | 0.431 | **.950** |
| $V_{22}$ | .813 | 0.339 | |
| $V_{23}$ | .790 | 0.376 | |
| $V_{24}$ | .826 | 0.318 | |
| $V_{25}$ | .839 | 0.296 | |
| $V_{26}$ | .833 | 0.306 | |
| $V_{27}$ | .823 | 0.323 | |
| $V_{28}$ | .727 | 0.471 | |
| $V_{29}$ | .744 | 0.446 | |
| $V_{30}$ | .832 | 0.308 | |
| $V_{31}$ | .753 | 0.433 | |
| Total | 8.734 | 4.048 | |
| **Promoción (F4)** | | | |
| $V_{32}$ | .811 | 0.342 | **.953** |
| $V_{33}$ | .788 | 0.379 | |
| $V_{34}$ | .897 | 0.195 | |

| | | | |
|---|---|---|---|
| $V_{35}$ | .877 | 0.231 | |
| $V_{36}$ | .905 | 0.181 | |
| $V_{37}$ | .852 | 0.274 | |
| $V_{38}$ | .885 | 0.217 | |
| $V_{39}$ | .754 | 0.431 | |
| Total | 6.769 | 2.251 | |

Fuente: EQS 6. con datos y adaptación propios.

## *Cálculo de IVE, sin V14,V15 y V16*

Fornell y Larcker (1981) presentan el índice **IVE** como la relación entre la varianza que es capturada por un factor ***i*** en relación a la varianza total debida al error de medida de ese factor. Nótese que la única diferencia con el **IFC** es que cada carga factorial estandarizada es primero elevada al cuadrado antes de ser sumadas. Fornell y Larcker (1981) sugieren que es deseable que el constructo tenga valores de **IVE** >= .5, es decir, que sea superior la varianza capturada por el factor que la debida al error de medida. Ver **Figura 7.95.**

**Figura 7.95. CFA, SEM y EQS. Recopilación de información que calcula el IVE del caso de estudio 3 sin V14,V15 y V16**

| Constructo e indicadores | Cuadrado de la Carga Factorial Estandarizada $(L_{ij})^2$ | Varianza del Término de Error $Var(E_{ij}) = 1 - (L_{ij})^2$ | **Índice de Varianza Compuesta Extraída IVE>= .5** $\sum(L_{ij})^2/[\sum(L_{ij})^2 + \sum Var(E_{ij})]$ |
|---|---|---|---|
| **Producto ($F_1$)** | | | |
| $V_1$ | 0.599 | 0.401 | **.571** |
| $V_2$ | 0.539 | 0.461 | |
| $V_3$ | 0.555 | 0.445 | |
| $V_4$ | 0.398 | 0.602 | |
| $V_5$ | 0.561 | 0.439 | |
| $V_6$ | 0.534 | 0.466 | |
| $V_7$ | 0.551 | 0.449 | |
| $V_8$ | 0.659 | 0.341 | |
| $V_9$ | 0.687 | 0.313 | |
| $V_{10}$ | 0.672 | 0.328 | |
| $V_{11}$ | 0.552 | 0.448 | |
| $V_{12}$ | 0.623 | 0.377 | |
| $V_{13}$ | 0.498 | 0.502 | |
| Total | 7.429 | 5.571 | |
| **Precio (F2)** | | | |
| $V_{17}$ | 0.616 | 0.384 | **.633** |
| $V_{18}$ | 0.707 | 0.293 | |
| $V_{19}$ | 0.697 | 0.303 | |
| $V_{20}$ | 0.511 | 0.489 | |

| | | | |
|---|---|---|---|
| Total | 2.532 | 1.468 | |
| **Plaza (F3)** | | | |
| $V_{21}$ | 0.569 | 0.431 | **.632** |
| $V_{22}$ | 0.661 | 0.339 | |
| $V_{23}$ | 0.624 | 0.376 | |
| $V_{24}$ | 0.682 | 0.318 | |
| $V_{25}$ | 0.704 | 0.296 | |
| $V_{26}$ | 0.694 | 0.306 | |
| $V_{27}$ | 0.677 | 0.323 | |
| $V_{28}$ | 0.529 | 0.471 | |
| $V_{29}$ | 0.554 | 0.446 | |
| $V_{30}$ | 0.692 | 0.308 | |
| $V_{31}$ | 0.567 | 0.433 | |
| Total | 6.952 | 4.048 | |
| **Promoción (F4)** | | | |
| $V_{32}$ | 0.658 | 0.342 | **.718** |
| $V_{33}$ | 0.621 | 0.379 | |
| $V_{34}$ | 0.805 | 0.195 | |
| $V_{35}$ | 0.769 | 0.231 | |
| $V_{36}$ | 0.819 | 0.181 | |
| $V_{37}$ | 0.726 | 0.274 | |
| $V_{38}$ | 0.783 | 0.217 | |
| $V_{39}$ | 0.569 | 0.431 | |
| Total | 5.749 | 2.251 | |

Fuente: EQS 6. con datos y adaptación propios.

## *Análisis de la validez de la escala, sin V14,V15 y V16*

Recordando a Bohrnstedt (1976): *validez es el grado en que un instrumento mide el concepto bajo estudio*. Desafortunadamente, el concepto tiene múltiples dimensiones que deben explicarse y analizarse por separado.

## *Validez de contenido*

La **Figura 7.91** contiene la sección de reporte índices de ajuste (*fit indices*) proporcionados por **EQS.** De su análisis se desprende que el valor de la *Chi-Cuadrado* y el *número de grados de libertad* es superior a 2 (1842.0295/591= 3.11)**.**Esto nos indica que el modelo de Marketing Mix *no tiene un buen ajuste*, en sus índices ya que varios serán <.9, sin embargo, esto no es limitante para analizar el resto de satos estadísticos. Entre el conjunto de indicadores de ajuste que proporciona **EQS***, se recomienda analizar aquellos no normalizados*, dado que tienen menor tendencia a ofrecer resultados sesgados en casos de pequeñas muestras (Bentler y Wu, 1993).

Se observa que, tanto e **NNFI** (Bentler y Bonnet, 1980) como el **CFI** (Bentler y Wu, 1993), *SI cumplen este requisito*, tienen valores > .8 pero <.9, indicando un ajuste razonable. Se debe recordar que estos ajustes, (*>.8 y <.9 escalas probadas; >.9*

*escalas por crear*) aunque No cumplen el nivel deseado de ser >.9, es posible continuar con el análisis de los reportes estadísticos. Por ejemplo **RSMEA**= 0.073, ya que debe ser <.08.De a misma **Figura 7.91,** ver la sección de reporte ajuste de índices (*fit indices*)

```
GOODNESS OF FIT SUMMARY FOR METHOD = ROBUST

ROBUST INDEPENDENCE MODEL CHI-SQUARE =    9877.374 ON   630 DEGREES OF FREEDOM

INDEPENDENCE AIC =    8617.374   INDEPENDENCE CAIC =    5472.751
       MODEL AIC =     660.030          MODEL CAIC =   -2289.926

SATORRA-BENTLER SCALED CHI-SQUARE =   1842.0295 ON     591 DEGREES OF FREEDOM
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS      0.00000

MEAN- AND VARIANCE-ADJUSTED CHI-SQUARE              =     134.114 ON    43 D.F.
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS      0.00000

MEAN-SCALED AND SKEWNESS-ADJUSTED STATISTIC         =      44.076 ON    14 D.F.
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS      0.00006

FIT INDICES
-----------
BENTLER-BONETT      NORMED FIT INDEX =     0.814
BENTLER-BONETT NON-NORMED FIT INDEX =     0.856
COMPARATIVE FIT INDEX (CFI)          =     0.865
BOLLEN'S            (IFI) FIT INDEX  =     0.865
MCDONALD'S          (MFI) FIT INDEX  =     0.209
ROOT MEAN-SQUARE ERROR OF APPROXIMATION (RMSEA)     =     0.073
90% CONFIDENCE INTERVAL OF RMSEA  (       0.069,          0.077)
```

Fuente: EQS 6, con datos y adaptación propios.

Hasta este momento, se ha estimado el *modelo de medida* y se ha comprobado que ofrece un buen ajuste, para dar base de *validez de contenido.* Pues bien, sobre este *modelo de medida,* ya es posible analizar la *validez convergente* y la *validez discriminante* de las escalas implicadas. Para analizar la *validez nomológica,* será necesario estimar el *modelo teórico*.

## Validez de constructo o de concepto

La validez de constructo o de concepto se divide en dos: *validez convergente y validez discriminante. Una escala tiene validez de constructo, cuando tiene validez convergente y validez discriminante.*

## Validez convergente

La *validez convergente* existe cuando se emplean distintos instrumentos para medir un mismo constructo (distintos ítems para una misma variable latente) y estos instrumentos están fuertemente correlacionados. Si analizamos de la misma **Figura 7.91**, las secciones de reporte las ecuaciones de medida (*measurement equations*), se deberá observar que las cargas factoriales sean < 1 significativas al 5%.

```
MEASUREMENT EQUATIONS WITH STANDARD ERRORS AND TEST STATISTICS
STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.
(ROBUST STATISTICS IN PARENTHESES)
```

```
MPP1  =V1  =   1.000 F1    + 1.000 E1


MPP2  =V2  =    .971*F1    + 1.000 E2
                .062
              15.703@
             (   .032)
             ( 30.543@

MPP3  =V3  =    .950*F1    + 1.000 E3
                .059
              15.984@
             (   .040)
             ( 23.859@
```

Fuente: EQS 6, con datos y adaptación propia

De la **Figura 7.91,** la sección de reporte de soluciones estandarizadas (*standardized solutions*), así como para el modelo de medida observamos que *todas las cargas factoriales son significativas, como se desprende de sus respectivas t* para todas las variables. Asimismo, si calculáramos el promedio de los mismos veríamos que superan claramente .7.

| | | | | | |
|---|---|---|---|---|---|
| MPP1 | =V1 | = | .774 F1 | + .633 E1 | .599 |
| MPP2 | =V2 | = | .734*F1 | + .679 E2 | .539 |
| MPP3 | =V3 | = | .745*F1 | + .667 E3 | .555 |
| MPP4 | =V4 | = | .631*F1 | + .776 E4 | .398 |
| MPP5 | =V5 | = | .749*F1 | + .663 E5 | .560 |
| MPP6 | =V6 | = | .731*F1 | + .682 E6 | .535 |
| MPP7 | =V7 | = | .742*F1 | + .670 E7 | .551 |
| MPP8 | =V8 | = | .812*F1 | + .583 E8 | .660 |
| MPP9 | =V9 | = | .829*F1 | + .559 E9 | .687 |
| MPP10 | =V10 | = | .820*F1 | + .572 E10 | .673 |
| MPP11 | =V11 | = | .743*F1 | + .669 E11 | .552 |
| MPP12 | =V12 | = | .789*F1 | + .615 E12 | .622 |
| MPP13 | =V13 | = | .706*F1 | + .708 E13 | .499 |
| MPR4 | =V17 | = | .785 F2 | + .619 E17 | .616 |
| MPR5 | =V18 | = | .841*F2 | + .541 E18 | .708 |
| MPR6 | =V19 | = | .835*F2 | + .550 E19 | .698 |
| MPR7 | =V20 | = | .715*F2 | + .699 E20 | .512 |
| MPL1 | =V21 | = | .754 F3 | + .657 E21 | .568 |
| MPL2 | =V22 | = | .813*F3 | + .582 E22 | .661 |
| MPL3 | =V23 | = | .790*F3 | + .613 E23 | .624 |
| MPL4 | =V24 | = | .826*F3 | + .563 E24 | .683 |
| MPL5 | =V25 | = | .839 F3 | + .545 E25 | .703 |
| MPL6 | =V26 | = | .833*F3 | + .553 E26 | .694 |
| MPL7 | =V27 | = | .823*F3 | + .568 E27 | .678 |
| MPL8 | =V28 | = | .727*F3 | + .687 E28 | .528 |
| MPL9 | =V29 | = | .744 F3 | + .669 E29 | .553 |
| MPL10 | =V30 | = | .832*F3 | + .554 E30 | .693 |
| MPL11 | =V31 | = | .753*F3 | + .658 E31 | .567 |
| PMO1 | =V32 | = | .811 F4 | + .585 E32 | .658 |
| MPO2 | =V33 | = | .788*F4 | + .615 E33 | .621 |
| MPO3 | =V34 | = | .897*F4 | + .441 E34 | .805 |
| MPO4 | =V35 | = | .877*F4 | + .480 E35 | .770 |
| MPO5 | =V36 | = | .905*F4 | + .425 E36 | .819 |
| MPO6 | =V37 | = | .852*F4 | + .524 E37 | .725 |
| MPO7 | =V38 | = | .885*F4 | + .465 E38 | .784 |
| MPO8 | =V39 | = | .754*F4 | + .657 E39 | .568 |

Fuente: EQS 6. con datos y adaptación propia

y de la misma **Figura 7.91** en su sección de prueba del multiplicador de Lagrange (*Lagrange multiplier test*) el ajuste del modelo mejoraría significativamente con la consideración, de tener suficiente marco teórico y **EFA** validado para considerar cambiar, por ejemplo la **V11** a **F4** ambas como intento de mejora del modelo.

```
MULTIVARIATE LAGRANGE MULTIPLIER TEST BY SIMULTANEOUS PROCESS IN STAGE  1

PARAMETER SETS (SUBMATRICES) ACTIVE AT THIS STAGE ARE:

   PVV PFV PFF PDD GVV GVF GFV GFF BVF BFF
```

| | CUMULATIVE MULTIVARIATE STATISTICS | | | | UNIVARIATE INCREMENT | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| STEP | PARAMETER | CHI-SQUARE | D.F. | PROB. | CHI-SQUARE | PROB. | HANCOCK'S SEQUENTIAL D.F. | HANCOCK'S SEQUENTIAL PROB. | ROBUST PREDICTED RMSEA | ROBUST PREDICTED CFI |
| 1 | V11,F4 | 32.474 | 1 | 0.000 | 32.474 | 0.000 | 591 | 1.000 | 0.072 | 0.868 |
| 2 | V9,F3 | 54.696 | 2 | 0.000 | 22.221 | 0.000 | 590 | 1.000 | 0.071 | 0.870 |
| 3 | V24,F2 | 74.309 | 3 | 0.000 | 19.613 | 0.000 | 589 | 1.000 | 0.071 | 0.872 |
| 4 | V10,F2 | 91.238 | 4 | 0.000 | 16.929 | 0.000 | 588 | 1.000 | 0.070 | 0.874 |
| 5 | V36,F2 | 106.802 | 5 | 0.000 | 15.564 | 0.000 | 587 | 1.000 | 0.070 | 0.876 |
| 6 | V26,F1 | 121.372 | 6 | 0.000 | 14.569 | 0.000 | 586 | 1.000 | 0.070 | 0.877 |
| 7 | V8,F4 | 135.437 | 7 | 0.000 | 14.066 | 0.000 | 585 | 1.000 | 0.069 | 0.879 |
| 8 | V39,F1 | 147.940 | 8 | 0.000 | 12.503 | 0.000 | 584 | 1.000 | 0.069 | 0.880 |
| 9 | V37,F3 | 160.791 | 9 | 0.000 | 12.851 | 0.000 | 583 | 1.000 | 0.069 | 0.881 |
| 10 | V34,F3 | 178.233 | 10 | 0.000 | 17.442 | 0.000 | 582 | 1.000 | 0.068 | 0.883 |
| 11 | V6,F3 | 190.444 | 11 | 0.000 | 12.211 | 0.000 | 581 | 1.000 | 0.068 | 0.884 |

Fuente: EQS 6. con datos y adaptación propia

## Cálculo de validez discriminante

Es aquí donde continuamos el estudio de la validez discriminante a partir del método

1. ***Test de las diferencias entre las Chi cuadrado*,** que como se había indicado, se realiza de la siguiente forma:

• *Se estima el modelo de medida pero fijando a 1 las covarianzas entre los dos factores cuya validez discriminante suscita dudas* **(F3 y F4**).

• *Se restan las Chi cuadrado de ambos modelos y los grados de libertad*.

***Modelo de medida***

*Chi-Cuadrada de Satorra-Bentler= 2179.7309; gl= 698*

***Modelo de medida con COV F3-F4=1***

*Chi-Cuadrada de Satorra-Bentler= 1842.0295; gl= 591*

• *El valor del estadístico (la resta) debe ser superior al valor crítico de la Chi-Cuadrado para ese nivel de grados de libertad. Esto querrá decir que el modelo en el que los factores están relacionados es mejor que aquel en que esos factores son el mismo (ausencia de validez discriminante).*

*La diferencia de las Chi-Cuadrada= 337.704; gl=17. Ver* ***Figura 7.96.*** *De la que se desprende el razonamiento siguiente:*

*Valor crítico con 17 gl:*

*p< .005= 35.7184*

*p< .0025= 37.9462*

*p< .001= 40.7911 por lo tanto , a 337.704<.001. Se constata así que se tiene validez discriminante a más del 99%.*

**Figura 7.96. Tabla de distribución de Chi-Cuadrada**

**TABLA 3-Distribución Chi Cuadrado $\chi^2$**

P = Probabilidad de encontrar un valor mayor o igual que el chi cuadrado tabulado, v = Grados de Libertad

| v/p | 0,001 | 0,0025 | 0,005 | 0,01 | 0,025 | 0,05 | 0,1 | 0,15 | 0,2 | 0,25 | 0,3 | 0,35 | 0,4 | 0,45 | 0,5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 10,8274 | 9,1404 | 7,8794 | 6,6349 | 5,0239 | 3,8415 | 2,7055 | 2,0722 | 1,6424 | 1,3233 | 1,0742 | 0,8735 | 0,7083 | 0,5707 | 0,4549 |
| 2 | 13,8150 | 11,9827 | 10,5965 | 9,2104 | 7,3778 | 5,9915 | 4,6052 | 3,7942 | 3,2189 | 2,7726 | 2,4079 | 2,0996 | 1,8326 | 1,5970 | 1,3863 |
| 3 | 16,2660 | 14,3202 | 12,8381 | 11,3449 | 9,3484 | 7,8147 | 6,2514 | 5,3170 | 4,6416 | 4,1083 | 3,6649 | 3,2831 | 2,9462 | 2,6430 | 2,3660 |
| 4 | 18,4662 | 16,4238 | 14,8602 | 13,2767 | 11,1433 | 9,4877 | 7,7794 | 6,7449 | 5,9886 | 5,3853 | 4,8784 | 4,4377 | 4,0446 | 3,6871 | 3,3567 |
| 5 | 20,5147 | 18,3854 | 16,7496 | 15,0863 | 12,8325 | 11,0705 | 9,2363 | 8,1152 | 7,2893 | 6,6257 | 6,0644 | 5,5731 | 5,1319 | 4,7278 | 4,3515 |
| 6 | 22,4575 | 20,2491 | 18,5475 | 16,8119 | 14,4494 | 12,5916 | 10,6446 | 9,4461 | 8,5581 | 7,8408 | 7,2311 | 6,6948 | 6,2108 | 5,7652 | 5,3481 |
| 7 | 24,3213 | 22,0402 | 20,2777 | 18,4753 | 16,0128 | 14,0671 | 12,0170 | 10,7479 | 9,8032 | 9,0371 | 8,3834 | 7,8061 | 7,2832 | 6,8000 | 6,3458 |
| 8 | 26,1239 | 23,7742 | 21,9549 | 20,0902 | 17,5345 | 15,5073 | 13,3616 | 12,0271 | 11,0301 | 10,2189 | 9,5245 | 8,9094 | 8,3505 | 7,8325 | 7,3441 |
| 9 | 27,8767 | 25,4625 | 23,5893 | 21,6660 | 19,0228 | 16,9190 | 14,6837 | 13,2880 | 12,2421 | 11,3887 | 10,6564 | 10,0060 | 9,4136 | 8,8632 | 8,3428 |
| 10 | 29,5879 | 27,1119 | 25,1881 | 23,2093 | 20,4832 | 18,3070 | 15,9872 | 14,5339 | 13,4420 | 12,5489 | 11,7807 | 11,0971 | 10,4732 | 9,8922 | 9,3418 |
| 11 | 31,2635 | 28,7291 | 26,7569 | 24,7250 | 21,9200 | 19,6752 | 17,2750 | 15,7671 | 14,6314 | 13,7007 | 12,8987 | 12,1836 | 11,5298 | 10,9199 | 10,3410 |
| 12 | 32,9092 | 30,3182 | 28,2997 | 26,2170 | 23,3367 | 21,0261 | 18,5493 | 16,9893 | 15,8120 | 14,8454 | 14,0111 | 13,2661 | 12,5838 | 11,9463 | 11,3403 |
| 13 | 34,5274 | 31,8830 | 29,8193 | 27,6882 | 24,7356 | 22,3620 | 19,8119 | 18,2020 | 16,9848 | 15,9839 | 15,1187 | 14,3451 | 13,6356 | 12,9717 | 12,3398 |
| 14 | 36,1239 | 33,4262 | 31,3194 | 29,1412 | 26,1189 | 23,6848 | 21,0641 | 19,4062 | 18,1508 | 17,1169 | 16,2221 | 15,4209 | 14,6853 | 13,9961 | 13,3393 |
| 15 | 37,6978 | 34,9494 | 32,8015 | 30,5780 | 27,4884 | 24,9958 | 22,3071 | 20,6030 | 19,3107 | 18,2451 | 17,3217 | 16,4940 | 15,7332 | 15,0197 | 14,3389 |
| 16 | 39,2518 | 36,4555 | 34,2671 | 31,9999 | 28,8453 | 26,2962 | 23,5418 | 21,7931 | 20,4651 | 19,3689 | 18,4179 | 17,5646 | 16,7795 | 16,0425 | 15,3385 |
| 17 | 40,7911 | 37,9462 | 35,7184 | 33,4087 | 30,1910 | 27,5871 | 24,7690 | 22,9770 | 21,6146 | 20,4887 | 19,5110 | 18,6330 | 17,8244 | 17,0646 | 16,3382 |
| 18 | 42,3119 | 39,4220 | 37,1564 | 34,8052 | 31,5264 | 28,8693 | 25,9894 | 24,1555 | 22,7595 | 21,6049 | 20,6014 | 19,6993 | 18,8679 | 18,0860 | 17,3379 |

Fuente: Levin y Rubin (2004) con datos y adaptación propia

**2.Test del intervalo de confianza (Anderson y Gerbing, 1988)**

Se calcula un intervalo de confianza para la covarianza de ± dos errores estándar alrededor de la estimación resultante del **CFA**.
Si incluye el 1 no se puede confirmar la validez discriminante, en caso contrario sí.

En nuestro caso, de la **Figura 7.91,** de la sección de reporte covarianzas entre las variables independientes (*covariances among independent variances*), de la relación más alta, se tiene:

```
I F2  -   F2                .379*I
I F1  -   F1                .048 I
I                          7.930@I
I                       (   .070)I
I                       (  5.432@I
I                                I
I F3  -   F3                .613*I
I F1  -   F1                .070 I
I                          8.750@I
I                       (   .075)I
I                       (  8.196@I
I                                I
I F4  -   F4                .669*I
I F1  -   F1                .078 I
I                          8.535@I
I                       (   .082)I
I                       (  8.179@I
I                                I
I F3  -   F3                .377*I
I F2  -   F2                .044 I
I                          8.484@I
I                       (   .056)I
I                       (  6.712@I
I                                I
I F4  -   F4                .343*I
I F2  -   F2                .045 I
I                          7.632@I
I                       (   .058)I
I                       (  5.966@I
I                                I
I F4  -   F4                .766*I
I F3  -   F3                .076 I
I                         10.060@I
I                       (   .072)I
I                       ( 10.622@I
```

Fuente: EQS 6. con datos y adaptación propia

Así, el razonamiento de cálculo por intervalo, es:

**F3-F4**

Covarianza al cuadrado $(.766)^2$= .586

Inferior = .766 – 2 x .076= .614

Superior= .766 + 2 x .076= .918

**F2-F4**

Covarianza al cuadrado $(.343)^2$=.118

Inferior = .343 -2 x .045= .253

Superior = .343+2 x .045= .433

**F2-F3**

Covarianza al cuadrado $(.377)^2$=.142

Inferior = .377 – 2 x .044= .289

Superior = .377 + 2 x .044= .465

**F1-F4**

Covarianza al cuadrado $(.669)^2$=.447

Inferior = .669 – 2 x .078= .513

Superior = .669 + 2 x .078= .825

**F1-F3**

Covarianza al cuadrado $(.613)^2$=.376

Inferior = .613 – 2 x .070= .473

Superior = .613 + 2 x .070= .753

**F1-F2**

Covarianza al cuadrado $(.379)^2$=.144

Inferior = .379 – 2 x .048= .283

Superior = .379 + 2 x .048= .475

*Se comprueba, que al ser la correlación más alta entre factores es <1, por lo tanto, existe validez discriminante*

**3.Test de la varianza extraída (Fornell y Larcker, 1981)**

Se procede de forma sistemática, como sigue:

- Se calcula el **IVE** para cada uno de los factores implicados como se explicó al analizar la confiabilidad.
- Se compara el **IVE** con el cuadrado de las covarianzas entre los dos factores.

Hay validez discriminante si los **IVE** de los dos factores superan el cuadrado de la covarianza.

- Cuadrado de la covarianza: $0.766^2$ = 0.587 .

Ver **Figura 7.97**

**Figura 7.97. Reporte tabla de validez discriminante del modelo teórico**

| Factor | Producto (F1) | Precio (F2) | Plaza (F3) | Promoción (F4) |
|---|---|---|---|---|
| **Producto (F1)** | .571 | .118 | .376 | .447 |
| **Precio (F2)** | .283-.475 | .633 | .142 | .118 |
| **Plaza (F3)** | .473-.753 | .289-.465 | .632 | .586 |
| **Promoción (F4)** | .513-.825 | .253-.433 | .614-.918 | .718 |

**Nota:** La diagonal representa el índice de varianza extraída (**IVE**), mientras que arriba de la diagonal se presenta la varianza (la correlación al cuadrado), abajo de la diagonal, se muestra la estimación de la correlación de los factores con un índice de fiabilidad del 95%.

Fuente: EQS 6. con datos y adaptación propios.

## *Cálculo de validez nomológica.*

Se debe recordar que, el modelo teórico que se estaba contrastando planteaba relaciones teóricas entre variables que podrían constatar la validez nomológica de ser ciertas. La mejor forma de hacerlo es estimar el modelo, no el de medida, como hemos hecho con el **CFA,** sino el teórico. Si con las modificaciones pertinentes para lograr un buen ajuste, los constructos cuya validez se está evaluando mantienen relaciones significativas previstas por la teoría, habremos establecido la validez nomológica (concurrente). El modelo teórico ha de tener un ajuste significativamente mejor que el de medida (lo comprobaremos con las diferencias entre las Chi cuadradas). De nos ser así, no podremos afirmar la validez nomológica. Así, como se había visto el cálculo de validez discriminante por el test de diferencias de Chi-Cuadradas, se tiene:

- *Se restan las Chi cuadrado de ambos modelos y los grados de libertad.*

***Modelo de medida***

*Chi-Cuadrada de Satorra-Bentler= 2179.7309; gl= 698*

***Modelo de medida con COV F3-F4=1***

*Chi-Cuadrada de Satorra-Bentler= 1842.0295; gl= 591*

- *El valor del estadístico (la resta) debe ser superior al valor crítico de la Chi cuadrado para ese nivel de grados de libertad. Esto querrá decir que el modelo en el que los factores están relacionados es mejor que aquel en que esos factores son el mismo (ausencia de validez discriminante).*
- *La diferencia de las Chi-Cuadrada= 337.704; gl=17. Ver* ***Figura 7.100.*** *De la que se desprende el razonamiento siguiente:*

  *Valor crítico con 17 gl:*

  *p< .005= 35.7184*

  *p< .0025= 37.9462*

  *p< .001= 40.7911 por lo tanto , a 337.704<.001. Se constata así que se tiene validez discriminante a más del 99%, y que también, el modelo teórico ofrece un ajuste significativamente mejor que el modelo de medida, lo cual permite constatar la validez nomológica de las escalas.*

## *Reportando resultados: consistencia interna y validez convergente*

En este punto, es importante tener una referencia del acomodo de los datos, como se sugiere en el ejemplo de la **Figura 7.98.**

**Figura 7.98. Reporte tabla de consistencia interna y validez convergente del modelo teórico sin V1 y V3**

| Factor | Ítem V | Carga Factorial >0.6 (a) | Valor Robusto t | Factor Promedio de Carga Factorial | Fiabilidad de la Escala | | |
|---|---|---|---|---|---|---|---|
| | | | | | Alfa de Cronbach >=0.7 (b) | IFC >0.7 (b) | IVE >0.5 (c) |
| **Producto ($F_1$)** | $V_1$ | .774 | 1.000a | .754 | .938 | **.945** | **.571** |
| | $V_2$ | .734 | 30.543 | | | | |
| | $V_3$ | .745 | 23.859 | | | | |
| | $V_4$ | .631 | 18.145 | | | | |
| | $V_5$ | .749 | 20.440 | | | | |
| | $V_6$ | .731 | 12.083 | | | | |
| | $V_7$ | .742 | 12.857 | | | | |
| | $V_8$ | .812 | 17.794 | | | | |
| | $V_9$ | .829 | 20.182 | | | | |
| | $V_{10}$ | .820 | 20.196 | | | | |
| | $V_{11}$ | .743 | 22.787 | | | | |
| | $V_{12}$ | .789 | 20.807 | | | | |
| | $V_{13}$ | .706 | 12.382 | | | | |
| **Precio ($F_2$)** | $V_{17}$ | .785 | 1.000a | .794 | .862 | **.873** | **.633** |
| | $V_{18}$ | .841 | 19.147 | | | | |
| | $V_{19}$ | .835 | 15.118 | | | | |
| | $V_{20}$ | .715 | 12.427 | | | | |
| **Plaza ($F_3$)** | $V_{21}$ | .754 | 1.000a | .794 | .940 | **.950** | **.632** |
| | $V_{22}$ | .813 | 26.640 | | | | |
| | $V_{23}$ | .790 | 21.947 | | | | |
| | $V_{24}$ | .826 | 27.235 | | | | |
| | $V_{25}$ | .839 | 25.555 | | | | |
| | $V_{26}$ | .833 | 24.449 | | | | |
| | $V_{27}$ | .823 | 26.245 | | | | |
| | $V_{28}$ | .727 | 19.193 | | | | |
| | $V_{29}$ | .744 | 22.333 | | | | |
| | $V_{30}$ | .832 | 30.486 | | | | |
| | $V_{31}$ | .753 | 28.996 | | | | |
| **Promoción ($F_4$)** | $V_{32}$ | .811 | 1.000a | .846 | .944 | **.953** | **.718** |
| | $V_{33}$ | .788 | 24.686 | | | | |
| | $V_{34}$ | .897 | 25.868 | | | | |
| | $V_{35}$ | .877 | 25.295 | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | $V_{36}$ | .905 | 25.817 | | | | |
| | $V_{37}$ | .852 | 21.998 | | | | |
| | $V_{38}$ | .885 | 25.344 | | | | |
| | $V_{39}$ | .754 | 22.254 | | | | |

•**Resultados:**
**Chi-Cuadrada= 1842.0295; gl= 591 p < .01; NFI = .814; NNFI = .856; CFI = .865; RMSEA = .073**
•**Conclusion:** las relaciones entre las variables y los factores tienen un ajuste adecuado a los datos. Esto es, no hay error en el ajuste del modelo.Por lo tanto, existe suficiente evidencia de validación convergente y confiabilidad , que justifica la consistencia interna de la escala aplicada (Nunnally y Bernstein, 1994; Hair et al., 1999)

Notas:
*** Parámetros constreñidos a ese valor en el proceso de identificación = **p** < 0.001
(a) De acuerdo a Bagozzi y Yi, 1988.; (b) De acuerdo a Hair et al. 1999.; (c) Índice de Varianza Extraída (**IVE**), de acuerdo a Fornell y Larcker, 1981.
Fuente: EQS 6. y SPSS 21 IBM con datos y adaptación propios.

Donde las columnas significan:
•*Factores,* listado de los involucrados en el estudio. En nuestro caso:
Producto **($F_1$);** Precio **($F_2$);** Plaza **($F_3$)** y Promoción **($F_4$)**
•*Ítem,* consecutivo de las variables (**V**) no suprimidas mostradas, como : **V1; V2; V3,** etc.
•*Carga Factorial > .6 (a)* (Ver **Figura 7.91**) donde sólo se reportan los factores no suprimidos y con carga factorial >.6. Es posible que el investigador utilice escalas previas que si a su juicio, por el marco teórico, deban asignarse más ítems, es válido realizarlo siempre y cuando, se mencione que la nueva escala es una **adaptación** de la previa más la añadida como propuesta del investigador dadas las condiciones de la realidad en la que se aplica contra las que originalmente la escala se creó.En esta etapa es donde se valida o no la pertinencia de los nuevos ítems basados en el marco teórico planteado.
•*Valor Robusto t* (Ver **Figura 7.40** para mayor detalle**)**,
Con lo anterior, es posible afirmar que si los valores de t (sección de reporte ecuaciones de medida *measurement equations y ecuaciones de constructo construct equations para los factores* , ***Figura 7.91***) son mayores (>) 2.58 errores estándar de la media, entonces se anotarán tres asteriscos (***) por carga factorial, lo cual indica que p<.0001. Así se tienen una seguridad que las cargas factoriales aportan sus valores a un 99% de confiabilidad
•*Alfa de Cronbach***,** cañculada y reportada de acuerdo al apartado *de cálculo del alfa de Cronbach, sin V14,V15 y V16* mediante uso de base de datos **SPSS marketing mix.sav**
•**IFC.** Se toman los datos de la **Figura 7.94** y es muy similar al *alfa de Cronbach.*
•**IVE.** Se toman los datos de la **Figura 7.95.**

## Conclusión

Así, afirmamos que para *maximizar la mercadotecnia mix* deberá ser atendido en mayor medida, por cada una de las cargas factoriales con valor absoluto alto y hacer una explicación de la realidad, basada en el marco teórico del porqué debe ser así (o no), según los argumentos del investigador. Todo lo anterior basados en las cargas factoriales encontradas con una probabilidad de equivocación de *1/100 o 99%* de que lo afirmado ocurra (ver las ***)

Donde los renglones por debajo de las cargas factoriales, significan:

- ***S-Bχ2*** *Indice Satorra Bentler Chi-Cuadrada,* es el recomendado cuando se tienen más de dos factores para generar datos robustos. Así, se explica la lectura reportada en la base de la **Figura 7.39**:
- ***Resultados:***

***Chi-Cuadrada=*** *1842.0295****; gl****= 591;* ***p <*** *.01****; NFI*** *= .814****; NNFI*** *= .856****; CFI*** *= .865****; RMSEA*** *= .073*

*Las relaciones entre las variables y los factores tienen un ajuste adecuado a los datos. Esto es, no hay error en el ajuste del modelo.Aunqe existe suficiente evidencia de validación convergente y fiabilidad , que justifica la fiabilidad interna de la escala aplicada (Nunnally y Bernstein, 1994; Hair et al., 1999), Sí se tiene validez discriminante), por lo que es posible determinar un Reporte tabla de validez discriminante del modelo teórico.*

Así, una posible redacción de los resultados, sería:

*Por otro lado, para evaluar la fiabilidad y validez de las escalas de medida se realizó un Análisis Factorial Confirmatorio (***CFA***) utilizando el método de máxima verosimilitud con el software* ***EQS 6*** *(Bentler, 2005; Brown, 2006; Byrne, 2006). Asimismo, la fiabilidad de las escalas de medida se evaluó a partir del coeficiente alfa de Cronbach y el índice de fiabilidad compuesta* ***(IFC****) (Bagozzi y Yi, 1988). Todos los valores de la escala excedieron el nivel recomendado de* ***.70*** *para el alfa de Cronbach y el* ***IFC*** *que proporciona una evidencia de fiabilidad y justifica la fiabilidad interna de las escalas (Nunally y Bernstein, 1994; Hair et al., 1995).*

*Los resultados de la aplicación del* ***AFC*** *se presentan en la Tabla 1 y sugiere que el modelo de medición proporciona un buen ajuste de los datos (****S-BX2 =*** *XXX.YYY****; df =*** *XX****; p =*** *0.000****; NFI >****. 9XXX* ***; NNFI*** *>.9XXX****; CFI****>.9XXX****; y RMSEA*** *<.08XXX). Como evidencia de la validez convergente, los resultados del AFC indican que todos los ítems de los factores relacionados son significativos (p < 0.001) y el tamaño de todas las cargas factoriales estandarizadas son superiores a 0.60* ***(Bagozzi y Yi, 1988).***

*La Tabla xxx muestra una alta consistencia interna de los constructos, en cada caso, la α de Cronbach excede el valor de .70 recomendado por Nunnally y Bernstein (1994). El* ***IFC*** *representa la varianza extraída entre el grupo de variables observadas y el constructo fundamental (Fornell y Larcker, 1981). Generalmente un IFC superior a .60 es considerado como deseable (****Bagozzi y Yi, 1988****), en este estudio este valor es superior. El índice de la varianza extraída* ***(IVE****) fue calculado para cada par de constructos, resultando un* ***IVE*** *superior a .5 en todos los factores (Fornell y Larcker, 1981)…*

Con estos resultados, *ya estamos, en posibilidades de aprobar o desaprobar, hipótesis* en función de las cargas factoriales encontradas y realizar la discusión del porqué está sucediendo (o no). Con base a filtros que aplique a su base de datos de **marketing. sav**, ; algunas de ellas podrían ser, como ejemplo:

**H1:** *Ha modificado líneas de productos o servicios por emergencia (**$V_4$**), es el indicador más representativo del factor Producto (**$F_2$**). Se desaprueba,* ya que es incluso menos representativa (.631). La más representativa es **$V_9$**: *Se centra en la maximización de las necesidades de sus clientes en cuento a los requerimientos de sus productos.*

**H2:** *Los factores más importantes que considera la PyME manufacturera de Aguascalientes, en orden decreciente son: Precio, Promoción, Producto, Plaza en sus planes de administración de mercadotecnia***.** *Se desaprueba*, ya que son: Promoción (.846), Precio (.794) , Plaza (.794) y finalmente, Producto (.754).

**H3.** *Con base al modelo de marketing mix, aplicado en las PyMES manufactureras de Aguascalientes, el cual muestran que los esfuerzos en Precio (**$F_2$**) y Plaza (**$F_3$**), son muy similares y que lo más representativo es el esfuerzo en Promoción (**$F_4$**) y el menos, Producto (**$F_1$**). Por lo tanto, se esperan los mismos resultados en un rango de cargas factoriales de +- 5% al aplicarlo al mismo sujeto de estudio en la Cd. de Guadalajara.* El resultado sería a revisar una vez obtenidos los datos del caso Guadalajara.

De ser posible, el resultado final de las cargas factoriales (**Figura 7.98**), la puede complementar con un diagrama de trayectorias que muestre los datos en el mismo.

# Conclusión final SEM y CFA

Se anexa en la **Figura 7.99**, una tabla de datos técnicos, como referencia para documentar proyectos **SEM-CFA**.

**Figura 7.99. Tabla de datos técnicos, pruebas y valores utilizados en la investigación**

| | DATOS TÉCNICOS DE LA INVESTIGACIÓN | | |
|---|---|---|---|
| | CARACTERÍSTICAS | ESTUDIO | |
| | Universo | **400** especialistas PyMES manufactureras de Aguascalientes | |
| | Alcance | Zona Metropolitana Ciudad de Aguascalientes, Aguascalientes, México | |
| | Método de recolección de datos | e-Mail y encuestas directas | |
| | Escala | Likert 5 | |
| | Pruebas de validez y confiabilidad | 4 profesores (*visión académica*); 1 CEO pyme; 1 gerente back office; 1 gerente front office; 1 diseñador de software; 1 consultor de tecnologías de información (*visión de expertos*) | |
| | Fecha de levantamiento de datos | Enero 2017- Abril-2017 | |
| | Total de entrevistas efectivas | **400** que incluyen: **100** empresarios Pyme; **100** administradores de proceso interno (*back office*) Pyme/**100** admnistradores de servicio Pyme; **100** consultores que conforman la *visión de expertos* | |
| | Pruebas aplicadas en este estudio | **Descripción/ Valor** | **Autor** |
| **1** | Radio NC/VoQ= Número de casos(NC) / Variables del | **NC= 400 (>=100 y <=1000)** especialistas PyME **. VoQ = 39 Radio NC/VoQ= 400/39=10.25>5** ( >=5 recomendación de Hair, 2014) | Hair et al.(2014) |

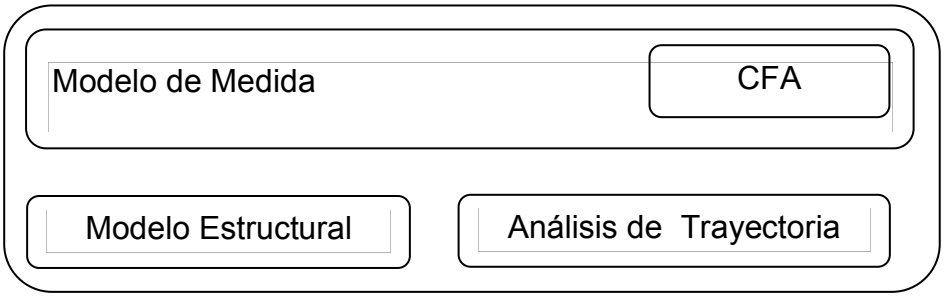
| | | | |
|---|---|---|---|
| | Cuestionario (VoQ) | | |
| **2** | Análisis Factorial Confirmatorio (CFA. *Confirmatory Factor Analysis*) Por el método de Máxima Versosimilitud y Análisis de Covarianza con uso del software EQS 6 | Para verificar la Fiabilidad y Validez de las escalas de medición | Bentler, (2006); Brown, (2006); Byrne, (2006) |
| **3** | **Alfa de Cronbach** (**CHA**) e Indice de Fiabilidad Compuesta (**IFC**) | **CHA (Por Factor, Via SPSS) y IFC>=0.7** para la fiabilidad de la escala de medición | Bagozzi y Yi, (1988); Nunnally y Bernestain, (1994); Hair et al., (2014) |
| **4** | Estimado de Mardia Normalizado (**M**) | **M>5.00.** Distribuído como una unidad de variación normal tal que los valores grandes reflejan una curtosis positiva significativa y un gran número de valores negativos reflejan una curtosis negativa significativa. Bentler (2006) sugiere en la práctica, valores >5.00 que son indicativos de datos, que no son distribuídos normalmente | Mardia (1970) Mardia (1974) Bentler (2006); Byrne, (2006) |
| **5** | Estadística Satorra–Bentler (**S-Bχ2**) | **SBχ2.-** Al especificarse ME=ML, ROBUST, la salida genera una estadística de chi cuadrada robusta, llamada (**χ2**). Esto permite minimizar los valores atípicos y lograr la bondad de ajuste | Satorra y Bentler, (1988) |
| **6** | Indice de ajuste normalizado (*Normed-Fit Index*) **(NFI)** | **NFI>=0.8 y <=.89.** El índice ha sido utilizado por más de 2 décadas por Bentler y Bonett's (1980), como criterio práctico de elección, como lo demuestra en gran parte el estatus *clásico* actual del trabajo original de Bentler (1992), y Bentler y Bonett (1987), (citado por Byrne, 2006). Sin embargo, el NFI ha mostrado una tendencia a subestimar el ajuste en muestras pequeñas, | Bentler y Bonnet,(1980); Byrne (2006) |
| **7** | Indice de ajuste comparativo (Comaprative-Fit Index) (CFI) | **CFI>=0.8 y <=.89.** Bentler (1990, citado por Byrne, 2006) revisó la **NFI** para considerar el tamaño de la muestra y el **CFI** propuesto. Los valores para el **NFI** y el **CFI** oscilan entre **cero y 1.00** y se derivan de la comparación entre los modelos de hipótesis e independencia. Como tal, cada uno proporciona una medida de la covariación completa en los datos. Aunque se consideró inicialmente que un valor> 0.90 era representativo de un modelo bien ajustado (véase Bentler, 1992, citado por Byrne, 2006), se ha recomendado un valor de corte revisado cercano a 0.95 (Hu y Bentler, 1999, citado por Byrne, 2006). Aunque ambos índices de ajuste se informan en la producción de EQS, Bentler (1990, citado por Byrne, 2006) sugirió que el **CFI** debería ser el índice de elección | |
| **8** | Indice No-Normalizado de ajuste (*Non Normed-Fit Index*) (NNFI) | **NNFI>=0.8 y <=.89.** Es una variante del **NFI** que tiene en cuenta la complejidad del modelo. Los valores para el **NNFI** pueden exceder los reportados para el **NFI** y también pueden caer fuera del rango cero a 1.00 (Byrne, 2006) | |
| **9** | Raíz cuadrática media de error por aproximación (*Root Mean Square Error*) (RMSEA) | **RMSEA>=0.05 y <=0.08.** El **RMSEA** considera el error de aproximación en la población haciendo la pregunta: *¿Qué tan bueno es el modelo, con valores de parámetros desconocidos pero óptimamente elegidos, de tal forma que se ajuste en la matriz de covarianza de población de estar disponible?* (Browne y Cudeck, 1993, pp. 137-8, citado por Byrne, 2006). Esta discrepancia, medida por la **RMSEA**, se expresa con el grado de libertad, por lo que | Hair et al, (2014); Byrne, (2006); Chau, (1997); Heck, (1998) |

| | | | |
|---|---|---|---|
| | | es sensible al número de parámetros estimados en el modelo (es decir, la complejidad del modelo). Los valores inferiores a 0.05 indican un buen ajuste, y valores tan altos como 0.008 representan errores razonables de aproximación en la población (Brown y Cudeck, 1993, citado por Byrne, 2006). Al dirigirse a Steiger (1990) se pide el uso de intervalos de confianza para evaluar la precisión de las estimaciones de **RMSEA**, **EQS** informa un intervalo del **90%** alrededor del valor **RMSEA**. A diferencia de las estimaciones puntuales del ajuste del modelo (que no reflejan la imprecisión de la estimación), los intervalos de confianza pueden proporcionar esta información, proporcionando así al investigador más ayuda en la evaluación del ajuste del modelo. | |
| **10** | Validez convergente (*Convergent Validity*) (CV) | Todos los ítems de los factores relacionados son significativos (**p < 0.001**), eltamaño de todas las cargas factoriales son mayores a **0.60** (Bagozzi y Yi, 1988) . Es el grado en el cual, las diferentes mpetodos de evaluación concurren en sus mediciones del mismo constructo. Idealmente, estos valores deben de ser moderadamente altos (Byrne, 2006) | Bagozzi y Yi, (1988); Byrne, (2006) |
| **11** | Índice de varianza compuesta extraída (AVE.*Average Variance Extracted*) (IVE) | **IVE > 0.50** En todos los factores pareados como constructos a través de una matriz de relaciones, donde la diagonal representa al **IVE**, mientras arriba de la diagonal se representa la varianza ( la correlación al cuadrado); debajo de la diagonal, se encuentra el estimado de la correlación de los factores con un intervalo de confianza del 95%. Vea la Tabla de *validez discriminante del modelo teórico*. | Fornell y Larcker, (1981) |
| **12** | Validez discriminante (*Discriminant Validity*) (DV) | DV, es la medida en que los métodos de evaluación independientes divergen en la medición con diferentes rasgos, idealmente estos valores deberían demostrar una convergencia mínima (Byrne, 2006). DV se proporciona en dos formas: En primer lugar, con un intervalo de 95% de fiabilidad, ninguno de los elementos individuales de la matriz de correlación de factores latentes contiene 1.0 (Anderson y Gerbing, 1988). En segundo lugar, IVE entre cada par de factores es mayor que su IVE correspondiente (Fornell y Larcker, 1981). Por lo tanto, sobre la base de estos criterios, las diferentes mediciones efectuadas en la escala muestran suficiente evidencia de fiabilidad, CV y DV | Byrne, 2006; Anderson y Gerbing,(1988); Fornell y Larcker,(1981) |
| **13** | Validez nomológica (*Nomologic Validity*) (NV) | Se prueba con el chi cuadrado, a través del cual se comparó el modelo teórico con el modelo ajustado. Los resultados indican que si no hay diferencias significativas, son buenos modelos teórico en la explicación de las relaciones observadas entre las construcciones latentes | Anderson y Gerbing,(1988); Hatcher, (1994) |

Fuente: recopilación propia

La Tabla **7.100** presenta un resumen de lo visto hasta ahora, planteado a manera de preguntas y respuestas para su mayor comprensión

## Tabla 7.100. Lo que es SEM y CFA

| Pregunta | Respuesta |
|---|---|
| ¿Qué es **SEM y CFA**? | **SEM** (*Structural Equation Modeling*), o modelaje de ecuación estructural es una metodología estadística que realiza una aproximación de cálculo confirmatoria (por ejemplo, en prueba de hipótesis) al análisis de la teoría estructural que soporta a un fenómeno. Típicamente, esta teoría representa el *proceso de causa* que genera observacionbes en múltiples variabñes (Bentler, 1988, en Byrne, 2010, p.3)<br>**CFA** (*Confirmatory Factor Analysis*), o análisis factorial confirmatorio. En contratste con EFA (*Exploratory Factor Analysis*), el cfa es apropiado usarlo cuando el investigador tiene algún conocimiento de la estructura de las variables latentes subyacentes. Basado en la teoría del conocimiento, la investigación empírica o ambas, el investigador postula las relaciones entre las mediciones observadas y los factores subyacentes *a priori* y entonces prueba esta estructura hipotética, estadísticamente. (Byrne, 2010, p.6)<br>Existen tres tipos de **SEM**: *Modelo de medida; Modelo Estructural* siendo éstas dos, las mas populares enntre las ciencias económico-administrativas y las ciencias sociales en general (Byrne, 2010) y por último, el *Análisis de Trayectoria* (Hair, et a.,2010; Awang, 2015). *El Modelo de Medida*, implica al **CFA.**El Modelo Estructural implica al *análisis causal* (análisis de regresión).<br>  <br><br>**SEM** es considerada como un análisis tipo Covarianza (**CB-SEM.***Covariance-Type Analysis*), la cual se encuentra entre un *análisis de varianza y un análisis de regresión*.<br>Por lo tanto, **SEM** utiliza los supuestos del análisis de datos paramétrico (*Parametric Data Analysis*), especialmente los de análisis de regresión, por lo que éstas características hacen al **SEM**, más riguroso en términos del análisis de datos comparados con técnicas como el el **PLS-SEM** (Partial Least Square-Structural Equation Modeling,el cual es considerado com plementario a **CB-SEM** más que como contraparte). |

| Aspectos | CB-SEM | PLS-SEM |
|---|---|---|
| **Teoría** | **Robusta** | **Flexible** |
| Supuestos de distribución | Normalidad Multivariada | No-Paramétrica |
| Tipo de análisis | Basado en Análisis de Covarianza | Basado en Análisis de Varianza |
| Tamaño de muestra | Grande (mínimo 100) | Pequeño (mínimo de 30-100) |
| Enfoque de análisis | Confirma teóricamente las relaciones supuestas | Predicción y/o identificación de relaciones entre constructos |
| Tipo de medición | Intervalo a la muestra (*nominal* para análisis moderador y *ordinal* para variables latentes observadas-con técnica *Certain Model of Fitness*-) | Categórica a la muestra |
| Complejidad de modelo | Grandes modelos (>100 indicadores) puede ser problemática | Puede tratar con modelos grandes |

**¿Cuántas muestras requiere SEM?**

| Tamaño de la muestra | Número de constructos en el modelo | Número de ítems por constructo | Ítems en comunalidad |
|---|---|---|---|
| Desde n=100 | <5 | >3 variables observadas | Alto (=>0.6) |
| Desde n=150 | <7 | >3 variables observadas | Modesto (0.5 sin ítem / constructo bajo identificación) |
| Desde n=300 | <7 | >3 variables observadas | Bajo (<0.45 y/o con ítem / constructo por identificar menos de 3) |
| Desde n=500 | >7 | <3 | Y/O algunos con baja comunalidad |

*El tamaño de la muestra, al igual que con cualquier otra inferencia estadística, debe ser adecuado para representar a la población de interés, y esta puede ser a menudo la principal preocupación del investigador. (Hair et al.,1999).*
*El promedio de error de varianza (Average Error Variance) de los indicadores: investigaciones recientes indican que el concepto de comunalidad es la forma más relevante de cálculo que el asunto del tamaño de la muestra. Las comunalidades representan la cantidad promedio de variación entre las variables medidas / indicadoras explicadas por el modelo de medición.*
*La comunalidad de un ítem puede ser directamente calculada como el cuadrado del ítem estandarizado/ cargas de los constructos.*
*Los estudios muestran que se requieren tamaños de muestra más grandes a medida que las comunidades se vuelven más pequeñas* (es decir, los constructos no observados no explican tanta varianza en los elementos medidos).
Los modelos que contienen *constructos múltiples* con valores compartidos menores que .5 (es decir, estimaciones de carga estandarizadas menores que .7) también requieren tamaños más grandes para la convergencia y la estabilidad del modelo (Enders y Bandalos, 2001).
El problema se exagera cuando los modelos tienen constructos con solo uno o dos elementos (Hair et al. 1999).

**¿De dónde viene la teoría?**

| **Revisión de la Literatura** a través de un análisis documental (aproximación cualitativa) | **Una postulación teórica** a través de estudios previos | **Técnica por adoptar y/o adaptar** de estudios previos | **Una postulación teórica** a través de una aproximación (aproximación cualitativa) | **Revisión de la literatura y EFA** |
|---|---|---|---|---|

**¿Cuáles son los supuestos del Análisis de Datos Pramétrico y Multivariante?**

Existen cuatro principales supuestos de Análisis de Datos Pramétrico: *Datos Normalmente Distribuídos; Independencia de Datos; Homogeneidad de la Varianza; Datos de Intervalo y/o Ratio.*
Como supuestos del Análisis Multivariante, se tienen:.

- Tamaño de la muestra; al menos 15 participantes por predictor (Stevens, 1996)
- Multicolinealidad y Singularidad; Correlaciones entre variables independientes menores a 0.9 (Pallant, 2013). Sin embargo, para el caso CFA, es <0.85 (Hair et al. 1999, Awang,2015).
- Datos atípicos (outliers), deben eliminarse.
- Normalidad, Linealidad, Homoscedasticidad e independencia de residuos (Tabchnick y Fidell,2007)

**¿Puede la escala de Likert ser usda en SEM-CFA?**

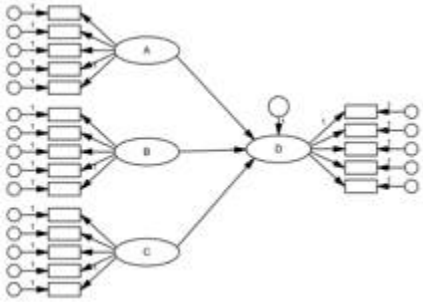
Una vez entendidas las suposiciones de los datos paramétricos y el análisis multivariable, al parecer la escala de Likert pudiera no ser adecuada para SEM-CFA, ya que es una escala de medida de tipo ordinal.
Sin embargo, de acuerdo con Kline (2005) y Hooper, Coughlan y Mullen (2008), el uso de la escala de tipo Likert como variable observada para la variable latente es aceptable si sólo el modelo estándar de raíz cuadrática media residual (SRMR.Standardized Root Mean Square Residual) como modelo de índice de ajuste se implementa en el SEM-CFA.
La escala tipo Likert, es aconsejable no elegirla como su unidad de medida para su

instrumento. Hair et al. (1999) sugirió una escala de intervalo de al menos 4 puntos, y Awang (2015) sugirió una escala de intervalo de 10 puntos.

| Escala de Likert de 5 puntos (L-S) | Escala de 5 puntos de intervalo (I-S) |
|---|---|
| 1-En fuerte desacuerdo<br>2-Desacuerdo<br>3-No seguro<br>4-Acuerdo<br>5-En fuerte acuerdo<br>-L-S debe tener un punto medio<br>-Los encuestados deben sólo escoger el número entero de rango y etiquetas dado (es decir, 1, 2,3,4 y 5)<br>-El intervalo entre los dos números no es cuantificado | 1-En fuerte desacuerdo<br>2-<br>3-<br>4-<br>5-En fuerte acuerdo<br>ó<br>1-En fuerte desacuerdo<br>.<br>.<br>.<br>.<br>5-En fuerte acuerdo<br>-I-S- no tiene un rango designado<br>-Solamente el valor más bajo y/o más alto es etiquetado<br>-Los encuestados pueden escoger ya sea un número entero o no entero (es decir, 2.5,3.8 y 4.9)<br>-El intervalo entre dos números es cuantificado. |

**¿Qué son los Modelos**

El modelo en **SEM-CFA** es un gráfico que representa una teoría o un concepto



**SEM** depende principalmente de varios índices de ajuste (*Model of Fit-Indices*) para determinar su aptitud y viabilidad, de los cuales, existen tres: Índices de ajuste absoluto (*Absolut Fit Indices*); Índices de ajuste incremental (*Incremental Fit Indices*); Índices de ajuste de parsimonia (*Parsimony Fit Indices*)(Hooper, Coughlan y Mullen, 2008; Hair et al., 1999; Awang, 2015)

**¿Qué son los Modelos de Indice de Ajuste (Model-of-fit Indices)?**

- Índices de ajuste absoluto (*Absolut Fit Indices*);

Determinan qué tan bien encaja un modelo a priori con los datos de la muestra (McDonald y Ho, 2002) y demuestra qué modelo propuesto tiene el ajuste más adecuado.
Los índices son; Modelo *Chi-Cuadrado (X2), Relativo / Chi-Cuadrado Normalizada (X2/df),* **RMSEA, GFI, AGFI, RMR y SRMR** *(Hooper, Coughlan y Mullen, 2008).*

- Índices de ajuste incremental (*Incremental Fit Indices*)

También conocidos como comparativos (*comparatives*) (Miles y Shevlin, 2007) o *índices de ajuste relativo* (McDonald y Ho, 2002).
Los índices son; **NFI, NNFI y CF**I (Hooper, Coughlan y Mullen, 2008)

- Indices de ajuste por Parsimonia (*Parsimony Fit Indices*)

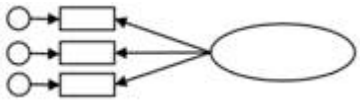
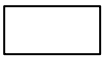
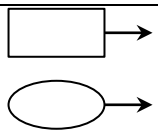
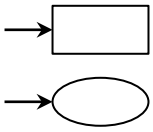
Tener un modelo complejo casi saturado, significa que el proceso de estimación depende de los datos de la muestra. Esto da como resultado un modelo teórico menos riguroso que, paradójicamente, *produce índices de mejor ajuste* (Mulaik et al., 1989; Crowley y Fan, 1997). Los índices son: **PGFI** y **PNFI** (Hooper, Coughlan y Mullen, 2008).

**¿Cuáles son los Modelos de índice de Ajuste?**

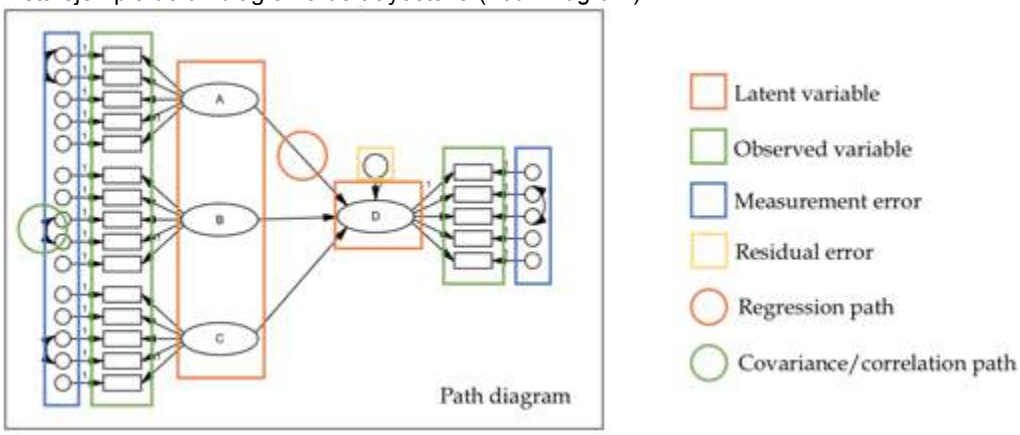
| Modelo de Índice de Ajuste | Nombre Completo | Valores de Corte |
|---|---|---|
| Índices de ajuste absoluto (Absolut Fit Indices) | | |
| Chi-Cuadrado ($X^2$) | Chi-Cuadrado. Uso sólo para muestras n<200 o p>.05 (Bentler y Nonnet,1980; Jöreskog y Sörbom,1993). Sin embargo, debe ser reportado en la descripción | Resultado insignificante (p> .05) (Barret, 2007). |
| $X^2$/*df* | Chi-Cuadrada relativa normalizada | Valores:<2.0,Tabachnick y |

| | | | |
|---|---|---|---|
| | | (*Relative/Normed Chi-Square*, Wheaton et al., 1977). Uso solamente para muestras n>200 o si p<.05) (Bentler y Bonnet, 1980; Jöreskog y Sörbom, 1993). | Fidell, 2007; <2.0, Hu y Bentler, 1999 y <5.0, Wheaton et al, 1977). |
| | **RMSEA** | Raíz cuadrática media de error por aproximación. (*Root Mean Square Error*) (Steiger y Lind, 1980) | Valor entre .08 a .10 (ajuste mediocre); <.08 (ajuste adecuado) (McCallum et al., 1996). Valor cercano a .60 (Hu y Bentler, 1999); Valor <.07 (Steiger, 2007) |
| | **GFI** | Índice de Bondad de ajuste estadístico (*Goodness-of-Fit Index*) (Jöreskog and Sörbom, 1993). Sin embargo, se recomienda no usarla debido a la tendencia a sesgarse debido al tamaño de la muestra ((Bollen, 1990; Miles y Shevlin, 1998). | Valor >.9 o >.95 (.95 si la carga factorial y el número de muestras es bajo) (Miles y Shevlin, 1998) |
| | **AGFI** | Índice de bondad ajustada *ajustado* (*Adjusted Goodness-of-Fit Index*) (Tabachnick and Fidell, 2007) | Valor >.8 (Hu y Bentler, 1999) y >=.9 (Hooper, Coughan y Mullen, 2008) |
| | **RMR** | Raíz cuadrática media residual (*Root Mean Square Residual*) | No Aplica |
| | **SRMR** | Raíz cuadrática media residual Estandarizada (Root Mean Square Residual) | <.05 (Byrne, 1998; Diamantopoulos y Siguaw, 2000) <.09 (Hu y Bentler, 1999). |
| | Índices de ajuste incremental (*Incremental Fit Indices*) | | |
| | **NFI** | Índice de ajuste normalizado (*Normed-Fit Index*) (Bentler y Bonnet, 1980). Sensitivo a tamaño de la muestra n <200 (Mulaik et al, 1989; Bentler, 1990), por lo tanto, debe acompañarse de otros índices (Kline, 2005). | Valor >.9 (Bentler y Bonnet, 1980) y >.95 (Hu y Bentler, 1999) |
| | **NNFI** | Índice de ajuste no normalizado (*Non-Normed-Fit Index*) (Ticker y Lewis, 1973) | Valor >.8 (Bentler y Bonnet, 1980) y >.95 (Hu y Bentler, 1999) |
| | **CFI** | Índice de ajuste comparativo (*Comparative-Fit Index*) (Bentler, 1990). Versión revisada por Byrne, (1998), y menos afectada por el tamaño de la muestra (Fan et al., 1999). | Valor >.9 (Bentler y Bonnet, 1980) y >.95 (Hu y Bentler, 1999) |
| | Indices de ajuste por Parsimonia (*Parsimonous Normed Fit Index*) | | |
| | **PNFI** | Índice de ajuste normalizado por Parsimonia (*Parsimonous Normed-Fit Index*) (Mulaik et al., 1989) | Valor >.9 (Mulaik et al., 1989). Sin embargo, no tiene un valor de corte recomendado |
| | **PGFI** | Índice de bondad de ajuste por Parsimonia (Parsimony Goodness-of-Fit-Index) (Mulaik et al., 1989) | |
| ¿Cuántos Modelos de índice de Ajuste deben ser reportados en una investigación ? | De acuerdo a Hair, et al. (1999), *al menos un índice por cada tipo.* Sin embargo, debe ser consciente de la limitación de cada modelo de índice de ajuste ya que varios de estos, se reafirman con otros. | | |

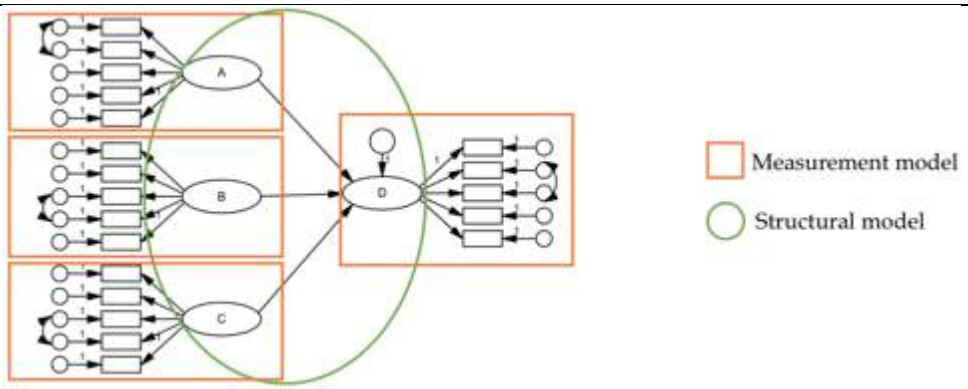
| | |
|---|---|
| ¿Cuáles son los escenarios de prueba del **SEM**? | Jöreskog (1993) distingue tres escenarios, llamados:<br>Estrictamente confirmatorios (**SC.***Strictly Confirmatory*); Modelos alternativos (**AM.***Alternative Models*; Modelo General (**MG.***Model Generating*) |
| ¿Cómo se ve un modelo de ecuación estructural general? | Algunos **términos usados en los modelos SEM, son**: |

| Término | Definición | Símbolos en AMOS/EQS |
|---|---|---|
| Constructo Latente (*Latent Construct*) | Constructo No directamente observado |  |
| Variable Formativa (*Formative Variable*) | Variable directamente observada |  |
| Variable Reflectiva (*Reflective Variable*) | Variable no directamente observada | |
| Variable exógena (*Exogenous Variable*) | Variable independiente |  o |
| Variable endógena (*Endogenous Variable*) | Variable dependiente |  o |

Vista ejemplo de un diagrama de trayectoria (*Path Diagram*)




Ejemplo de Modelo de Medida (*Measurement Model*) y Modelo Estructural (*Structural Model*)

| | |
|---|---|
| |  |
| Además de los modelos de índices de ajuste ¿hay otras medidas que **SEM-CFA** pueda proporcionar? | Dado que **SEM-CFA** se ocupa del *modelo de medida (Measurment Model)* o, en otras palabras, del proceso de validación del instrumento, también proporciona otras medidas.<br>• Validez de constructo (*Construct Validaty*)<br>En **SEM, CFA** se usa comúnmente para evaluar la validez de constructo (Jöreskog, 1967).<br>La validez de constructo se puede evaluar en función de los índices de ajuste del modelo. Si todos los índices de ajuste del modelo son adecuados, por lo tanto, se ha logrado la validez de constructo en general.<br>Sin embargo, el ajuste general de un modelo no es suficiente para garantizar la capacidad de un instrumento para medir un determinado fenómeno.<br>Bajo la validez de constructo, hay dos términos prominentes usados en **CFA**; validez convergente (*Convergent Validity*) y validez discriminante (*Discriminant Validity*) (Campbell y Fiske, 1959).<br>• Validez convergente (*Convergent Validity*); criterio de Fornell-Larcker (1981)<br>Esta se refiere al grado de confianza (*Degree of Confidence*) que tenemos de un rasgo (*constructo*) es bien medida por sus indicadores (Campbell y Fiske, 1959). En otras palabras, si se infringe, muestra que su ítem no se correlaciona bien entre sí dentro de su factor principal; es decir, el factor latente no está bien explicado por sus variables observadas. La validez convergente se evalúa en función del Índice de Varianza Extraída (**IVE** / **AVE**. *Average Variance Extracted*).<br>• Validez discriminante (*Discriminant Validity*)<br>La validez discriminante es el grado en que las medidas de diferentes rasgos (constructos) no están relacionadas (Campbell y Fiske, 1959).<br>En otras palabras, si se viola, muestra que su constructo se correlaciona más con el constructo fuera de su factor origen que con las variables dentro de dicho factor; es decir, *el factor latente se explica mejor por algunas otras variables (de un factor diferente), que por sus propias variables observadas.*<br>La validez discriminante se evalúa en función de la raíz cuadrada de **IVE** y las correlaciones entre constructos latentes. En primer lugar, las correlaciones entre constructo latente deben estar por debajo de .85, y los niveles de raíz cuadrada del **IVE** para cada constructo deberían ser mayores que la correlación que involucra los constructos.<br>Existen enfoques para evaluar la validez discriminante a través de la relación *Heterotrait-Monotrait* (**HTMT**) de Henseler, Ringle y Sarstedt (2015), en el que **HTMT** es el promedio de las correlaciones *heterotrait-heteromethod* en relación con el promedio de las correlaciones *monotrait-heteromethod*. El **HTMT** se deriva de la matriz clásica *multitrait-multimethod* (**MTMM**) (Campbell & Fiske, 1959). • El valor de corte para **HTMT** es; <.85 (Kline, 2011) y <.90 (Gold et al., 2001).<br>y<br>• Confiabilidad (*Reliability*)<br>Representado por el *alfa de Cronbach*, el cual, se usa comúnmente para evaluar la confiabilidad de un instrumento. Aunque no es una práctica común informar sus valores en **CFA**, no hay literatura que prohíba su uso.<br>Si un investigador desea informar el valor de *alfa de Cronbach* para cada constructo latente y el modelo de medición general, es solo con el propósito de confirmar el valor de confiabilidad compuesta (**CR.***Composite Reliability*) para cada constructo latente. El *alfa de Cronbach* se |

calcula utilizando el **SPSS**.
**CR** se usa comúnmente para informar la fiabilidad de un constructo latente, así como un instrumento en su conjunto. La razón por la que se elige comúnmente **CR** sobre *Alfa de Cronbach* en **CFA** *se debe a su estimación menos sesgada en comparación con Alfa de Cronbach* (Fornell-Larcker, 1981).

| Aspecto | Si hay violación (no cumple) | Medida y valor de corte |
|---|---|---|
| Validez de constructo (Construct Validity) | Usted construye, pero en general no está en forma, por lo tanto, no pudo confirmar el modelo de medición hipotético. | Seleccione un tipo de modelo de índice de ajuste y reintente |
| Validez convergente (*Convergent Validity*) | Su ítem no se correlaciona bien entre sí dentro de su factor principal; es decir, el factor latente no está bien explicado por sus variables observadas. | Varianza media extraída * (***AVE.*** *Average Variance Extracted*). El valor para cada construcción debe ser> .05 (Hair et al., 1999). |
| Validez discriminante (*Discriminant Validity*) | Su constructo se correlaciona más con conceptos externos que el factor propio; es decir, el factor latente se explica mejor por algunas otras variables (*de un factor diferente*), que por sus propias variables observadas. | La correlación entre constructos debe ser <.85 (Hair et al., 1999; Awang, 2015) y la raíz cuadrada de **IVE (AVE.** *Average Variance Extracted***)** es mayor que las correlaciones entre constructos (Fornell y Lacker, 1981) |
| Confiabilidad (*Reliability*) | Su constructo, en general, no es confiable para medir un cierto fenómeno. | El valor de confiabilidad compuesta (**CR.***Composite Reliability*) para cada constructo debe ser ≥.60 (Awang, 2015), o> .70 (Hair et al., 1999). |

- Hay otro aspecto que debe examinarse durante el **CFA**, que es la unidimensionalidad.

-La unidimensionalidad se logra cuando todos los elementos de medición tienen cargas de factores aceptables para el respectivo constructo latente. Para asegurar la unidimensionalidad de un modelo de medición, cualquier ítem con baja carga de factor debería ser eliminado. (Awang, 2015, p.54)

- Awang (2015) sugiere *además que cualquier carga factorial< .50 debe eliminarse si los artículos se desarrollan por primera vez*, y las cargas factoriales por debajo de .60 si los ítems están bien establecidos.

-Además, Hair et al. (1999) sugirieron, aunque el proceso de eliminación es necesario, uno debe tener en cuenta; *cada constructo debe tener al menos tres elementos, cuatro es mejor.*

¿Cuáles son las pasos iniciales para realizar **SEM-CFA**?

- Como ya sabemos que **SEM-CFA** debe cumplir con los *supuestos de los datos paramétricos* y el *análisis multivariante*, la detección de datos debe implementarse primero en nuestros datos. -Primero, debemos detectar a los *encuestados no comprometidos*. Estos son el tipo de que *prefieren calificar todos los ítems con el mismo valor*. Podríamos detectarlos calculando la desviación estándar de respuestas para cada encuestado.
- En segundo lugar, debemos *detectar cualquier valor perdido en nuestro conjunto de datos* y deshacernos de ese encuestado en particular.
- En tercer lugar, *también debemos detectar valores atípicos y extremos en nuestro conjunto de datos.* Necesitamos llevar a cabo el análisis de datos exploratorios (**EDA.** *Exploratory Data Analysis*), particularmente el diagrama de caja (*Box-plot*). Esta técnica es aplicable solo para datos de intervalo.

¿Qué hacer para asegurar un modelo de ajuste adecuado?

Una vez que se realiza la estimación, verifique los índices de ajuste de acuerdo a su modelo. Si todos los índices cumplen con los requisitos mínimos, entonces su modelo está en forma y el trabajo ya está resuelto. Pero, la mayoría de las veces no es así. Hay formas de modificar su modelo para cumplir con los requisitos mínimos de los índices de ajuste según el modelo seleccionado.

| | | | |
|---|---|---|---|
| Detecte la carga factorial por ítem <.50 y elimínelos uno por uno después de cada estimación. *Recuerde que la eliminación de ítems no debe ser mayor al 20% (3 elementos) por cada 15 o la cantidad total de ítems.* | Después de que se eliminen todos los elementos de carga factorial y si el modelo aún no se ajusta, consulte el índice de modificación (**MI**. *Modification Index*), principalmente al usar **AMOS**. | Detecte cualquier índice de modificación (**MI**. *Modification Index,* **AMOS**), de error superior a 15. Cree un estimado libre (*free estimate*) o elimina el ítem que corresponde al error | Detecte cualquier ítem con valores bajos de $\mathbf{R^2}$ y elimínelos |

El ajuste del modelo es evaluado con base al modelo de índice de ajuste (*Model Fit Indices*), como se muestra abajo para el caso de comparación de dichos modelos para *n=653* de acuerdo a **APA**.

| Model | No. of Item | Chi-square ($X^2$) | df | CMIN/DF ($X^2$/df) | RMSEA | AGFI | SRMR | CFI | PNFI |
|---|---|---|---|---|---|---|---|---|---|
| Six Factor** | 36 | 2133.14* | 579 | 3.68 | .06 | .82 | .05 | .80 | .69 |
| Six Factor** | 34 | 1932.14* | 519 | 3.77 | .07 | .82 | .05 | .81 | .70 |
| Six Factor** | 31 | 1612.95* | 419 | 3.85 | .07 | .84 | .05 | .83 | .71 |
| Six Factor** | 29 | 1204.24* | 357 | 3.37 | .06 | .86 | .05 | .87 | .73 |

Nota: *Significant at $p$<.001, therefore, $X^2/df$ value is referred (Wheaton *et al.*, 1997; Tabachnick & Fidell, 2007). $X^2/df$ value is also referred since $n$>200 (Marsh & Hocevar, 1985; Bentler, 1990). SRMR index is used since Likert 5-point scale (Kline, 2005) is implemented in the instrument. AGFI index is used since it is less sensitive than the GFI value (Sharma *et al.*, 2005; Hooper, Coughlan & Mullen, 2008). **Factor refers to latent constructs in the study.

Para lograr la unidimensionalidad, elimine los ítems problemáticos / redundantes.
Así, se deben eliminar los ítems *con carga factorial* < *.50* (ítems recientemente desarrollados) o *.60* (ítems ya establecidos) (Awang, 2015).
Sin embargo, en algunos casos, para garantizar que *NO VIOLEN* la validez discriminante, se *deben eliminar los elementos con factor de carga inferior a .70*.
Si se eliminaron todos los elementos que lo merecían y el modelo sigue siendo inadecuado, se deben convertir los valores de índice de modificación (**MI.** *Modification index)* y $\mathbf{R^2}$. **MI**> 15 debe ser inspeccionado y los valores de $\mathbf{R^2}$ < *.40* también deben ser inspeccionados.

Fuente: recopilación propia

# Apéndice 1

## Análisis de validez y confiabilidad: ¿Qué es?

**SPSS**, tiene importantes herramientas que permiten analizar los cuestionarios diseñados por los investigadores. Para saber mpas, consulte: IBM, 2011a; IBM, 2011b; IBM,2011c.

Así, es posible por ejemplo:

- Partir del **cruce-tabular** que muestre las relaciones entre los diferentes indicadores o preguntas, por ejemplo, si Usted requiere conocer de los gerentes de la industria de las telecomunicaciones su género como primer pregunta y grado académico máximo como la segunda pregunta, lo podrá visualizar mejor a través de aplicar el **cruce-tabular**
- Ahora bien, pruebas como un ***chi-cuadrado***, también se pueden aplicar para comparar los patrones de estos datos.
- Diferentes cuestionarios conducirán al uso de diferentes pruebas como las **pruebas *t*, o ANOVA.** Incluso la pertinencia de las variables a utilizar a través del análisis factorial, suele emplearse para asegurar la **validez** del instrumento o grado en el que un instrumento en verdad mide la variable que se busca medir. (por ejemplo, si está diseñado para medir el instrumento la innovación, no debe medir competitividad). Al respecto de la **validez**, la tenemos clasificada en :
  - **Validez de Contenido:** grado en que un instrumento refleja un dominio específico de contenido de lo que se mide. Por ejemplo: una prueba de operaciones aritméticas no tendrá validez de contenido si incluye sólo problemas de adición y excluye problemas de sustracción, multiplicación y división (**Validez de juicio de experto**).
  - **Validez de Criterio:** se establece al validar un instrumento de medición al compararlo con algún criterio externo que pretende medir lo mismo.
  - **Validez Concurrente y la Validez Predictiva**. En las campañas electorales, los sondeos se comparan con los resultados finales de las elecciones. Por ejemplo: Coeficiente de Contingencias, ***Spearman – Brow, Pearson,* Alfa de *Cronbach*** y la Técnica ***Aiken***.
  - **Validez de Constructo:** debe explicar el modelo teórico empírico que subyace a la variable de interés. Por ejemplo El Análisis de Factorial y Análisis de Cofactores, el Análisis de Covarianza.

  **Validez Total= (Validez de Contenido + Validez de Criterio + Validez de Constructo)/3**

Sin embargo, ninguno de estos diferentes análisis será significativo a menos que nuestro cuestionario sea **confiable.** La **confiabilidad es la capacidad del cuestionario para medir de forma consistente el tema bajo estudio en diferentes momentos y a través de diferentes poblaciones**. Imagine tener una cinta métrica que mide a una persona como 1 metro 70 centímetros en un día y 1 metro 50 centímetros en otro día. Esta sería una cinta métrica de muy baja

confiabilidad. Existen diferentes formas de evaluar la fiabilidad como el método de las dos mitades.

Existen diversas formas de evaluar la confiabilidad, tales como:

- **Medida de estabilidad**: Un mismo instrumento de medición se aplica dos más veces a un mismo grupo de personas, después de cierto periodo. Se le conoce también como **Confiabilidad por test-retest, "*r*" de *Pearson.***
- **Método de formas alternativas o paralelas**. Aquí no se administra el mismo instrumento de medición, sino dos o más versiones equivalentes de este. Se le conoce también como **Coeficiente de correlación producto-momento de *Pearson*.**
- **Método de mitades partidas**: Se necesita solo una aplicación, el total de los ítems se divide en dos partes y se comparan los resultados. (***Pearson y Spearman-Brown***).
- **Medidas de consistencia interna: Requiere sólo una administración.** Confiabilidad del test según el método de división de las mitades por ***Rulon y Guttman***, ***Fórmula 20 de Kuder-Richardson*** y Coeficiente del **Alfa de *Cronbach***.

Por su amplia aplicabilidad, el presente documento se referirá al Alfa de ***Cronbach***

## *El Alfa de Cronbach*

Cuando llevamos a cabo un cuestionario que a menudo lo estamos utilizando para medir un determinado constructo, tales como la **innovación**. En el cuestionario se diseñan una serie de preguntas o indicadores acerca de este constructo, por lo que para la **"*innovación*"** se puede plantear: ¿se realiza en la organización?, ¿es a nivel de producto y/o servicio?, ¿es a nivel de mercadotecnia?, ¿se aplica a en los procesos o en el modelo de negocios? De tal manera que un gerente de una empresa de alta tecnología, con alto sentido de la estrategia que implica la innovación al responder el cuestionario, deberá medir con **validez** el constructo. Debe ser capaz de medir a otras empresas dedicadas a la innovación apropiadamente, reportando correctamente su nivel de innovación. **El cuestionario deberá No sólo ser sólo válido sino también confiable**. Con esto, deberíamos esperar que si aplicamos el cuestionario al gerente de la empresa de alta tecnología nuevamente, **entonces mostraría el mismo resultado**. La **confiabilidad** puede evaluarse de diferentes maneras. Para lograrlo, una alternativa es la de **dividir nuestro cuestionario en dos y ver si la primera mitad de las preguntas se resuelven con el mismo resultado que la segunda mitad (confiabilidad de dos mitades).** Podemos llevar el cuestionario a niveles de negación y/o afirmación de las preguntas a fin de lograr las mismas respuestas y asegurarnos de que el instrumento es totalmente confiable, de lo contrario se deberán descartar y/o reemplazar aquellas preguntas que no logren dicho objetivo. La prueba de **Alfa de *Cronbach*** es el método más popular de examinar la **confiabilidad** y se basa en el número de **ítems, preguntas o indicadores** de un cuestionario así como el promedio de la correlación entre **ítems**. Se deberá asumir que toda respuesta a las preguntas del cuestionario por cada individuo existirá una desviación de la respuesta esperada vs. La obtenida la cual se denomina **error aleatorio**. Así:

•**Una alta correlación entre los diferentes ítems** indicará que se **está midiendo lo mismo** y por lo tanto existirán valores pequeños de **error**.

•**Una baja correlación entre los diferentes ítems** indicará que hay una **gran cantidad de errores y los elementos no son confiable medir la misma cuestión.**

**Alfa de *Cronbach*** oscila entre **0** para una prueba completamente confiable (aunque técnicamente puede bajar por debajo de **0**) a **1** para una prueba completamente fiable. ¿Qué valor de **Alfa de *Cronbach*** se debe obtener para que un cuestionario o una medida sean fiables? Hay un cierto debate alrededor de esto, con algunos estadísticos sugiriendo **0.7** o más, mientras que otros recomiendan **0.8**. Esto dependerá en cierta medida del número de elementos de la prueba y del número de participantes, pero **0.75** es un valor de compromiso razonable que tomar como punto de referencia, siendo el referente sugerido:

- **0.90 y más, excelente confiabilidad**
- **0.70 a 0.90, alta confiabilidad**
- **0.50 a 0.70 moderada confiabilidad**
- **0.50 y menor, baja confiabilidad**

**Alfa de *Cronbach*** se basa en los mismos supuestos que hemos considerado en nuestros capítulos sobre la correlación lineal simple y múltiple, ya que emplea un análisis correlacional. También se emplea un modelo en el que se supone que las puntuaciones observadas son las **verdadero puntaje más error**, y así, como se ha explicado, para que este modelo sea apropiado **los errores deben ser aleatorios** (y por lo tanto imparciales). **Si los supuestos NO se cumplen, el valor de Alfa de *Cronbach* puede ser una subestimación o una sobreestimación del valor correcto.**

## *Ejemplo*

**Paso 1: Objetivos**

•**Problema 1:** La empresa **MKT Digital** ha construido un cuestionario para examinar su modelo de negocios, y desea probar su confiabilidad con una muestra de **200** participantes y utilizando una escala de **10** puntos siendo 1 totalmente en desacuerdo y **10**, totalmente de acuerdo. Ver **Figura Ap.1 y Ap.2**

## Figura Ap1. Visor de Variables de la base de datos BM_MKT_Digital.sav

BM_MKT_Digital_OK.sav [Conjunto_de_datos1] - IBM SPSS Statistics Editor de datos

Archivo Edición Ver Datos Transformar Analizar Marketing directo Gráficos Utilidades Ventana Ayuda

| | Nombre | Tipo | Anchura | Decimales | Etiqueta | Valores | Perdidos | Columnas | Alineación | Me |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | id | Numérico | 3 | 0 | id.-Identidad | Ninguna | Ninguna | 4 | Centrado | Nom |
| 2 | X1 | Numérico | 2 | 0 | X1 - Antiguedad del consumidor | {1, < a 1 afi... | Ninguna | 11 | Centrado | Nom |
| 3 | X2 | Numérico | 2 | 0 | X2 - Tipo de industria | {0, Software... | Ninguna | 15 | Centrado | Nom |
| 4 | X3 | Numérico | 2 | 0 | X3 - Tamaño de la empresa | {0, PyME (0... | Ninguna | 17 | Centrado | Nom |
| 5 | X4 | Numérico | 2 | 0 | X4 - País | {0, MEX/Nor... | Ninguna | 18 | Centrado | Nom |
| 6 | X5 | Numérico | 2 | 0 | X5 - Sistema de distribución | {0, Indirecto... | Ninguna | 4 | Centrado | Nom |
| 7 | X6 | Numérico | 5 | 1 | X6 - Calidad del servicio | {.0, Mala}... | Ninguna | 4 | Centrado | Esc |
| 8 | X7 | Numérico | 5 | 1 | X7 - Comercio electrónico (e-Commerce) | {.0, Mala}... | Ninguna | 4 | Centrado | Esc |
| 9 | X8 | Numérico | 5 | 1 | X8 - Soporte técnico | {.0, Mala}... | Ninguna | 4 | Centrado | Esc |
| 10 | X9 | Numérico | 5 | 1 | X9 - Respueta a quejas | {.0, Mala}... | Ninguna | 4 | Centrado | Esc |
| 11 | X10 | Numérico | 5 | 1 | X10 - Publicidad | {.0, Mala}... | Ninguna | 4 | Centrado | Esc |
| 12 | X11 | Numérico | 5 | 1 | X11 - Línea de servicios | {.0, Mala}... | Ninguna | 4 | Centrado | Esc |
| 13 | X12 | Numérico | 5 | 1 | X12 - Imagen de la fuerza de ventas | {.0, Mala}... | Ninguna | 4 | Centrado | Esc |
| 14 | X13 | Numérico | 5 | 1 | X13 - Precio competitivo | {.0, Mala}... | Ninguna | 4 | Centrado | Esc |
| 15 | X14 | Numérico | 5 | 1 | X14 - Garantías | {.0, Mala}... | Ninguna | 4 | Centrado | Esc |
| 16 | X15 | Numérico | 5 | 1 | X15 - Nuevos productos y servicios | {.0, Mala}... | Ninguna | 4 | Centrado | Esc |
| 17 | X16 | Numérico | 5 | 1 | X16 - Ordenes y facturación | {.0, Mala}... | Ninguna | 4 | Centrado | Esc |
| 18 | X17 | Numérico | 5 | 1 | X17 - Flexibilidad de precios | {.0, Mala}... | Ninguna | 4 | Centrado | Esc |
| 19 | X18 | Numérico | 5 | 1 | X18 - Velocidad de entrega | {.0, Mala}... | Ninguna | 4 | Centrado | Esc |

Fuente: SPSS 20 IBM

## Figura Ap2.Visor de Datos de la base de datos BM_MKT_Digital.sav

BM_MKT_Digital_OK.sav [Conjunto_de_datos1] - IBM SPSS Statistics Editor de datos

Archivo Edición Ver Datos Transformar Analizar Marketing directo Gráficos Utilidades Ventana Ayuda

1 : id 1 Visible: 24 de 24 variab

| | id | X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 | X10 | X11 | X12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 a 5 años | Software empresarial | Grande (500+) | Fuera de MEX/Norteamérica | Dire... | 8.5 | 3.9 | 2.5 | 5.9 | 4.8 | 4.9 | 6.0 |
| 2 | 2 | Más de 5 años | Software para juegos | PyME (0 to 499) | MEX/Norteamérica | Indir... | 8.2 | 2.7 | 5.1 | 7.2 | 3.4 | 7.9 | 3.1 |
| 3 | 3 | Más de 5 años | Software empresarial | Grande (500+) | Fuera de MEX/Norteamérica | Dire... | 9.2 | 3.4 | 5.6 | 5.6 | 5.4 | 7.4 | 5.8 |
| 4 | 4 | < a 1 año | Software para juegos | Grande (500+) | Fuera de MEX/Norteamérica | Indir... | 6.4 | 3.3 | 7.0 | 3.7 | 4.7 | 4.7 | 4.5 |
| 5 | 5 | 1 a 5 años | Software empresarial | Grande (500+) | MEX/Norteamérica | Dire... | 9.0 | 3.4 | 5.2 | 4.6 | 2.2 | 6.0 | 4.5 |
| 6 | 6 | < a 1 año | Software para juegos | PyME (0 to 499) | Fuera de MEX/Norteamérica | Indir... | 6.5 | 2.8 | 3.1 | 4.1 | 4.0 | 4.3 | 3.7 |
| 7 | 7 | < a 1 año | Software para juegos | Grande (500+) | Fuera de MEX/Norteamérica | Indir... | 6.9 | 3.7 | 5.0 | 2.6 | 2.1 | 2.3 | 5.4 |
| 8 | 8 | 1 a 5 años | Software empresarial | Grande (500+) | Fuera de MEX/Norteamérica | Indir... | 6.2 | 3.3 | 3.9 | 4.8 | 4.6 | 3.6 | 5.1 |
| 9 | 9 | 1 a 5 años | Software para juegos | Grande (500+) | Fuera de MEX/Norteamérica | Indir... | 5.8 | 3.6 | 5.1 | 6.7 | 3.7 | 5.9 | 5.8 |
| 10 | 10 | < a 1 año | Software empresarial | Grande (500+) | Fuera de MEX/Norteamérica | Indir... | 6.4 | 4.5 | 5.1 | 6.1 | 4.7 | 5.7 | 5.7 |
| 11 | 11 | Más de 5 años | Software empresarial | Grande (500+) | MEX/Norteamérica | Dire... | 8.7 | 3.2 | 4.6 | 4.8 | 2.7 | 6.8 | 4.6 |
| 12 | 12 | < a 1 año | Software empresarial | Grande (500+) | Fuera de MEX/Norteamérica | Indir... | 6.1 | 4.9 | 6.3 | 3.9 | 4.4 | 3.9 | 6.4 |
| 13 | 13 | < a 1 año | Software para juegos | PyME (0 to 499) | MEX/Norteamérica | Dire... | 9.5 | 5.6 | 4.6 | 6.9 | 5.0 | 6.9 | 6.6 |
| 14 | 14 | Más de 5 años | Software para juegos | PyME (0 to 499) | MEX/Norteamérica | Dire... | 9.2 | 3.9 | 5.7 | 5.5 | 2.4 | 8.4 | 4.8 |
| 15 | 15 | 1 a 5 años | Software empresarial | Grande (500+) | Fuera de MEX/Norteamérica | Dire... | 6.3 | 4.5 | 4.7 | 6.9 | 4.5 | 6.8 | 5.9 |
| 16 | 16 | Más de 5 años | Software empresarial | PyME (0 to 499) | MEX/Norteamérica | Indir... | 8.7 | 3.2 | 4.0 | 6.8 | 3.2 | 7.8 | 3.8 |
| 17 | 17 | 1 a 5 años | Software para juegos | PyME (0 to 499) | Fuera de MEX/Norteamérica | Dire... | 5.7 | 4.0 | 6.7 | 6.0 | 3.3 | 5.5 | 5.1 |
| 18 | 18 | 1 a 5 años | Software empresarial | Grande (500+) | Fuera de MEX/Norteamérica | Indir... | 5.9 | 4.1 | 5.5 | 7.2 | 3.5 | 6.4 | 5.5 |

Fuente: SPSS 20 IBM

## Paso 2: Diseño y Paso 3: Condiciones de aplicabilidad

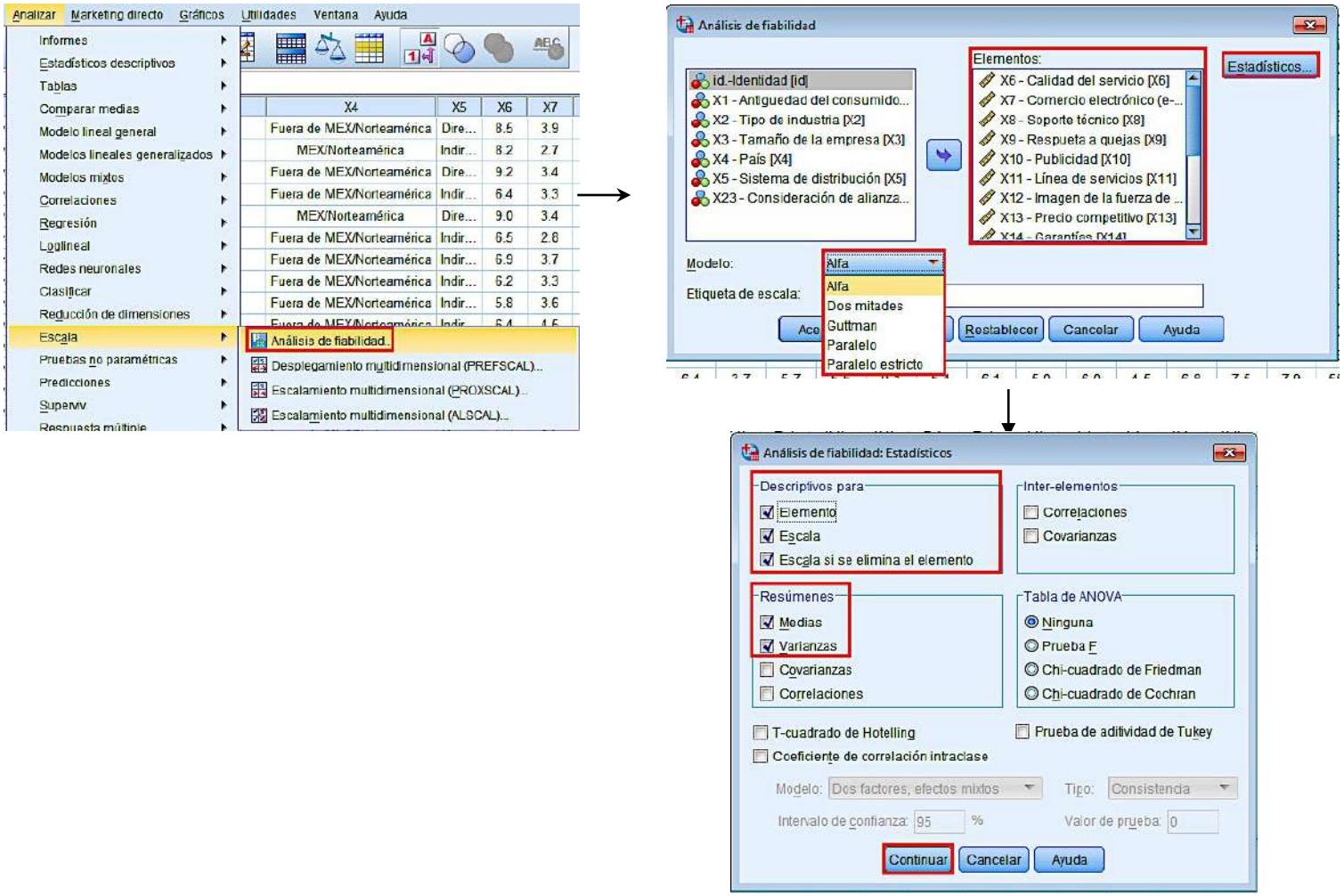
- Ampliamente vistos en Capítulo 6 del libro de Mejía-Trejo, J. (2017c).

## Paso 4: Estimación y ajuste

- **Teclear: Analizar->Escala->Análisis de fiabilidad->Modelo: Alfa->Elementos: todas las variables métricas ($X_6$ a $X_{22}$ en nuestro ejemplo) ->Estadísticos;**

**Descriptivos para: Elementos; Escala; Escala si se elimina el elemento->Resúmenes: Medias; Varianzas->Continuar->Aceptar. Ver Figura Ap.3**

**Figura Ap3. Proceso para calcular Alfa de Cronbach**




Fuente: SPSS 20 IBM

- Modelos alternativos de análisis de confiabilidad se pueden encontrar en el menú desplegable **Modelo**. Sin embargo, se recomienda el uso del **Alfa de *Cronbach*** debido a su flexibilidad para aplicarlo tanto a las **respuestas dicotómicas/ binarias** como a los datos medidos en una **escala** más grande. **Alpha de *Cronbach*** nos dará un cálculo de confiabilidad basado en la **totalidad del cuestionario.**
- Otro tipo de confiabilidad, que puede que desee evaluar, es la confiabilidad de **Dos mitades (ampliamente visto en los capítulos anteriores)**. Aquí las preguntas se dividen en dos grupos. Si la escala es confiable, las dos mitades estarán positivamente correlacionadas. Si este no es el caso, entonces puede inferir que nuestros participantes están respondiendo de manera diferente en las dos partes de

la cuestionario. Tenga en cuenta que la opción por defecto para la confiabilidad de **Dos mitades** en **SPS**S es agrupar por la primera mitad las primeras preguntas y luego realizar lo mismo por las preguntas finales.

**Paso 5: Interpretación**

•La primera tabla generada por **SPSS**, es la de **Resumen del procesamiento de los casos,** la cual confirma la cantidad de elementos que son incluidos en el análisis así como los excluidos. Ver **Figura Ap.4**

**Figura Ap.4. Tabla Resumen del procesamiento de los casos**

**Resumen del procesamiento de los casos**

| | | N | % |
|---|---|---|---|
| Casos | Válidos | 200 | 100.0 |
| | Excluidos[a] | 0 | .0 |
| | Total | 200 | 100.0 |

a. Eliminación por lista basada en todas las variables del procedimiento.

Fuente: SPSS 20 IBM

•La segunda tabla generada por **SPSS**, es la **Estadísticos de fiabilidad**. En nuestro caso el coeficiente de fiabilidad para los **17 elementos** se muestra como un **Alfa de *Cronbach***, **basada en los elementos tipificados** (estandarizada); ambos valores suelen ser muy similares **y suele el elegirse éste último (el estandarizado).**

•Una puntuación Alfa por encima de **0.75** se toma generalmente para indicar una escala de alta de confiabilidad, **generalmente se acepta que 0.5 a 0.75** que indica una escala **moderadamente confiable**, mientras que una cifra inferior indica generalmente una escala de **baja fiabilidad. En nuestro caso 0.842 es considerado de alta confiabilidad.**

•Normalmente se elige el **alfa estandarizado**. De hecho, **si bien es preferible que nuestras escalas sean similares** (por ejemplo, todas las preguntas se midan en una escala de 10 puntos), puede también que lo sea el **Alfa de *Cronbach*** simple**.** Ver **Figura Ap5.**

**Figura Ap.5. Estadísticos de fiabilidad**

**Estadísticos de fiabilidad**

| Alfa de Cronbach | Alfa de Cronbach basada en los elementos tipificados | N de elementos |
|---|---|---|
| .669 | .842 | 17 |

Fuente: SPSS 20 IBM

- La siguiente tabla generada por **SPSS**, es la de **Estadísticos de los elementos.**
- La primera parte reporta un resumen de las respuestas de los participantes a las preguntas individuales, y proporciona información sobre la **Media** y la **Desviación típica** (desviación standard) para las respuestas a cada pregunta, y un informe que indica cuántos (**N**) participantes completó la pregunta.
- De nuestro ejemplo podemos ver que todos los **200** de participantes respondieron a todas nuestras preguntas. También podemos ver qué preguntas suscitan una amplia variedad de respuestas, a través de las desviaciones estándar más grandes.
- En el ejemplo, las respuestas se midieron en una escala de **1 a 10**. Altas puntuaciones de la **Media**, por lo tanto, indican preguntas donde los participantes estuvieron finalmente de acuerdo en la escala de puntuación. Ver **Figura Ap.6**

### Figura Ap.6. Tabla Estadísticos de los elementos

**Estadísticos de los elementos**

| | Media | Desviación típica | N |
|---|---|---|---|
| X6 - Calidad del servicio | 7.894 | 1.3830 | 200 |
| X7 - Comercio electrónico (e-Commerce) | 3.765 | .7689 | 200 |
| X8 - Soporte técnico | 5.243 | 1.6552 | 200 |
| X9 - Respueta a quejas | 5.368 | 1.2100 | 200 |
| X10 - Publicidad | 4.061 | 1.1471 | 200 |
| X11 - Línea de servicios | 5.815 | 1.3174 | 200 |
| X12 - Imagen de la fuerza de ventas | 5.248 | 1.1286 | 200 |
| X13 - Precio competitivo | 6.971 | 1.5813 | 200 |
| X14 - Garantías | 6.048 | .8753 | 200 |
| X15 - Nuevos productos y servicios | 5.211 | 1.4960 | 200 |
| X16 - Ordenes y facturación | 4.242 | .9119 | 200 |
| X17 - Flexibilidad de precios | 4.464 | 1.1927 | 200 |
| X18 - Velocidad de entrega | 3.816 | .7494 | 200 |
| X19 - Satisfacción | 6.952 | 1.2411 | 200 |
| X20 - Probabilidad de recomendación | 6.952 | 1.0829 | 200 |
| X21 - Probabilidad de compra | 7.665 | .8932 | 200 |
| X22 - Nivel de compra | 58.200 | 8.9662 | 200 |

Fuente: SPSS 20 IBM

•Otra tabla generada por **SPSS**, es la de **Estadísticos de resumen de los elementos.**

•La fila **Medias de los elementos** detalla las estadísticas descriptivas de una respuesta en las preguntas individuales. Como podemos ver en el ejemplo anterior, la puntuación **Media** de los ítems es **8.701**, tal como se esperaría al promediar las puntuaciones en una gama de elementos empleando una escala de **1-10.**

•Los valores **Mínimo** y **Máximo** son las dos puntuaciones más extremas seleccionadas Por los participantes. En nuestro caso se trata de **3.765** y **58.2**, lo que indica que no hay encuestados más extremos de la escala.

•La columna **Varianza de los elementos** muestra la varianza en puntaje cuando se observan las puntuaciones en el ítem individual. Ver **Figura Ap.7**

**Figura Ap.7. Estadísticos de resumen de los elementos**

**Estadísticos de resumen de los elementos**

| | Media | Mínimo | Máximo | Rango | Máximo/mínimo | Varianza | N de elementos |
|---|---|---|---|---|---|---|---|
| Medias de los elementos | 8.701 | 3.765 | 58.200 | 54.435 | 15.458 | 164.445 | 17 |
| Varianzas de los elementos | 6.074 | .562 | 80.392 | 79.830 | 143.160 | 367.187 | 17 |

Fuente: SPSS 20 IBM

•La siguiente tabla producida por **SPSS** permite examinar la confiabilidad de cada pregunta, y el efecto en el cuestionario general, **si se eliminara dicha pregunta individual**

•El resultado de **SPSS** muestra las conclusiones del análisis para cada ítem del cuestionario.

•En nuestro ejemplo, las preguntas que son potencialmente preocupantes podrían señalarse

•La columna **Correlación elemento-total corregida** muestra la relación entre las respuestas sobre preguntas individuales y la puntuación total en el cuestionario. Nosotros esperamos que una pregunta confiable tuviera una relación positiva con el total general, **idealmente por encima de 0.3.** Un elemento que muestre **una relación positiva débil o una relación negativa con el total e indica una pregunta que puede ser pobre en confiabilidad y por lo tanto afecta a las Conclusiones de toda la escala.** En el ejemplo anterior, podemos ver que la pregunta $\mathbf{X_{15}}$ es la más débil, ya que su correlación con el total global es de sólo **0.142**.

•Los efectos que pueden tener las preguntas individuales sobre la confiabilidad general del cuestionario se destacan por la **relación inversa** entre la **Correlación elemento-total corregida vs Alfa de *Cronbach* si se elimina el elemento.** La importancia de la relación débil entre la pregunta **15** y la puntuación global total en el cuestionario se refleja en el aumento de la puntuación alfa para el cuestionario **SI este elemento se omite.** Un examen de la tabla más abajo de la salida nos da un Valor alfa de **0.669 (Ver Figura 10.5)**. Si bien esta cifra es alta, la eliminación de la

pregunta 1 de la El cuestionario final vería esta cifra subir a **0.649** (como puede verse en la tabla). Ver **Figura Ap.8**

**Figura Ap.8. Estadísticos de resumen de los elementos**

| | Media de la escala si se elimina el elemento | Varianza de la escala si se elimina el elemento | Correlación elemento-total corregida | Correlación múltiple al cuadrado | Alfa de Cronbach si se elimina el elemento |
|---|---|---|---|---|---|
| X6 - Calidad del servicio | 140.022 | 259.017 | .395 | .757 | .649 |
| X7 - Comercio electrónico (e-Commerce) | 144.151 | 268.707 | .365 | .660 | .659 |
| X8 - Soporte técnico | 142.673 | 265.796 | .185 | .740 | .663 |
| X9 - Respueta a quejas | 142.548 | 250.078 | .705 | .795 | .632 |
| X10 - Publicidad | 143.854 | 264.557 | .339 | .464 | .656 |
| X11 - Línea de servicios | 142.101 | 251.000 | .617 | .974 | .635 |
| X12 - Imagen de la fuerza de ventas | 142.668 | 263.051 | .388 | .815 | .653 |
| X13 - Precio competitivo | 140.945 | 286.436 | -.195 | .416 | .691 |
| X14 - Garantías | 141.867 | 269.580 | .284 | .748 | .661 |
| X15 - Nuevos productos y servicios | 142.705 | 269.278 | .142 | .100 | .667 |
| X16 - Ordenes y facturación | 143.674 | 259.279 | .627 | .648 | .645 |
| X17 - Flexibilidad de precios | 143.452 | 271.275 | .148 | .970 | .666 |
| X18 - Velocidad de entrega | 144.100 | 259.004 | .785 | .980 | .644 |
| X19 - Satisfacción | 140.964 | 245.320 | .814 | .856 | .624 |
| X20 - Probabilidad de recomendación | 140.963 | 254.845 | .651 | .639 | .639 |
| X21 - Probabilidad de compra | 140.251 | 260.510 | .597 | .581 | .647 |
| X22 - Nivel de compra | 89.715 | 83.189 | .703 | .780 | .774 |

Fuente: SPSS 20 IBM

- La última tabla que el SPSS genera es la de **Estadísticos de la escala**
- **N of elementos** es el número de ítems en nuestro cuestionario En nuestro caso **17** indicadores,
- La estadística en los renglones nos reporta la parte de la estadística descriptiva para la escala como un todo. En el ejemplo, cuando el total del puntaje del cuestionario es analizado, la puntuación **Media** de los participantes es de **147.916** con una **varianza** de **278.511**, y una Desviación típica, de **16.6886** Una desviación standard pequeña nos indica que no hay amplias variaciones en los puntajes de los participantes para todo el modelo global de puntaciones de nuestro cuestionario. Ver **Figura Ap.9**

**Figura Ap.9. Estadísticos de la escala**

**Estadísticos de la escala**

| Media | Varianza | Desviación típica | N de elementos |
|---|---|---|---|
| 147.916 | 278.511 | 16.6886 | 17 |

Fuente: SPSS 20 IBM

# Apéndice 2

## Antecedentes

De acuerdo a Aguilar-Barojas (2005), cuando se desea realizar un estudio, lo que se desea compartir con la comunidad médica son:

- Observaciones realizadas en la práctica cotidiana,
- Datos específicos del expediente clínico en grupos de pacientes con alguna característica en común
- Diferencias que se consideran importantes al comparar grupos de pacientes.

En cualquiera de estos tres casos, el cálculo del tamaño de la muestra apoya al objetivo de cualquier investigación, el cual puede ser de dos tipos:

- Estimar un parámetro, o
- Contrastar una hipótesis. Entonces, el cálculo permite responder a cualquiera de las siguientes preguntas: ¿cuántos individuos se deben considerar para estimar un parámetro con un grado de confianza determinado? o ¿cuántos individuos se deben estudiar para detectar en los resultados de los dos grupos, una diferencia que sea estadísticamente significativa?

El resultado del cálculo de una muestra debe considerarse como orientativo, ya que se fundamenta en supuestos que pueden ser incorrectos y que en el momento de introducirlos numéricamente en las fórmulas, afectan la viabilidad del estudio, el costo y hasta los aspectos éticos. Por otro lado, un estudio con una muestra insuficiente, puede afectar la precisión y la sensibilidad para detectar diferencias entre los grupos y conducir a conclusiones falsas.

Existen, por lo menos cuatro razones para estudiar muestras, en vez de poblaciones:

- Las muestras pueden estudiarse con más rapidez que las poblaciones.
- Es menos costoso estudiar una muestra que una población.
- En la mayor parte de las situaciones el estudio de la población entera es imposible.
- Los resultados de una muestra son más precisos que los derivados de poblaciones. Esto es por la calidad de los datos, la capacitación de quien recoge los datos; la estimación del error en los parámetros resultantes y la homogeneidad de las muestras.

En una investigación, incluir más sujetos de estudio, no es sinónimo de mejor estudio. Por esta razón, se debe de planear el tamaño apropiado de la muestra antes de iniciarlo. Este proceso se conoce como la determinación de la fuerza o potencia del estudio (1-β) cuya definición es la capacidad de un estudio para identificar una diferencia de un tamaño dado, si en realidad la diferencia existe.

El tamaño de la muestra dará la fuerza o potencia del estudio (1-β), sobre todo en el caso de que se vaya a contrastar una hipótesis y se pueda cometer cualquiera de los dos tipos de errores, el tipo I, que consiste en rechazar una hipótesis nula, cuando es cierta y el II en aceptarla cuando es falsa. Un buen cálculo de muestra nos protege de ignorar una diferencia significativa entre los dos grupos y, por lo tanto, de cometer el error tipo II. Es importante recordar que la Hipótesis Nula (Ho) es un enunciado que afirma que no hay diferencia entre el valor XA del grupo A y el valor XB del grupo B. Mientras que la Hipótesis alternativa (H1) es un enunciado en desacuerdo con la hipótesis nula, que enuncia: si hay diferencia entre el valor XA del grupo A y el valor

XB del grupo B.

## Consideraciones previas

Aún cuando varios autores mencionan que hay que tomar en cuenta los objetivos de la investigación, para no confundir al lector es mejor precisar que es con respecto al objetivo general, sin tomar en cuenta los objetivos específicos. Ese objetivo general, que se desprendió de una pregunta de investigación, lleva insertada la variable de interés. Hay que determinar si es una variable cualitativa, con escala nominal u ordinal, o cuantitativa con escala discreta o continua.
En las variables cualitativas hay dos tipos de posibilidades:
**a) Nominales,** que pueden ser dicotómicas o multicotómicas y
**b) Ordinales**. En ambos casos los resultados se expresan como porcentajes o tasas, datos que se introducen en la fórmula para el cálculo de la muestra.
En variables cuantitativas también hay dos tipos de posibilidades: a) discretas, cuando los resultados se expresan numéricamente, pero no aceptan fracciones, por ejemplo, número de dientes, número de hijos, etc. y b) continuas, cuando los resultados numéricos sí aceptan fracciones, como peso, estatura, todos los parámetros de química sanguínea, etc. En ambos casos los resultados se resumen en forma de medias y varianzas, mismas que se introducen en la fórmula del cálculo de la muestra.
**Es importante destacar, que las variables cuantitativas son más poderosas para el análisis estadístico y no hay que convertirlas en cualitativas**

## Cálculo de la muestra

Deberá conocer:
La variabilidad del parámetro que se desea estudiar (**p, q, S**). Aunque esto parezca paradójico, ya que de eso se trata precisamente la investigación, hay que conocer, cómo se presenta en la población la variable que se quiere estudiar. La primera reflexión que asalta al investigador es: si supiera cómo se presenta, no estaría diseñando una investigación para conocerla, ¿cómo me piden eso...?
Si se trata de variables cualitativas, se necesita conocer el porcentaje con que se presenta la variable en la población. Por ejemplo, si deseamos conocer la presencia de hábito tabáquico en los médicos, debemos saber qué porcentaje de médicos fuman (**p**) y qué porcentaje no fuman (**q**) justamente en la población que deseamos estudiar. Si es una variable cuantitativa, cuál es la desviación estándar (**S**) con que se presenta en la población. Por ejemplo, si queremos realizar una investigación para establecer los picogramos séricos de nicotina en los fumadores, debemos precisar la media que esperamos encontrar y su desviación estándar. Como se puede apreciar, el que nos pidan los datos de algo que aún no estudiamos, resulta sumamente confuso. Sin embargo, sí hay solución a este problema. Estos datos se pueden obtener de tres formas:

- De estudios reportados en revistas, con una población parecida o similar a la que queremos estudiar;
- Con un *estudio piloto de 25 sujetos*, o si no se pude lo anterior, se asigna la máxima probabilidad con que se puede presentar la variable en cuestión, que es, en

variables cualitativas del 50% y en el caso de variables cuantitativas se determina la diferencia entre el máximo y el mínimo valor esperable, se divide entre cuatro y, por lo tanto, se tiene una cierta aproximación al valor de la desviación estándar.

**El nivel de confianza deseado (Z).** Indica el grado de confianza que se tendrá de que el valor verdadero del parámetro en la población se encuentre en la muestra calculada. Cuanta más confianza se desee, será más elevado el número de sujetos necesarios. Se fija en función del interés del investigador. Los valores más comunes son 99% 95% o 90%. Hay que precisar que los valores que se introducen en la fórmula son del cálculo del área de la curva normal para esos porcentajes señalados:

| %Error | Nivel de confianza | Vlorde Z calculado en tablas |
|---|---|---|
| 1 | 99% | 2.58 |
| 5 | 95% | 1.96 |
| 10 | 90% | 1.645 |

A veces se afirma erróneamente, por ejemplo, que un nivel de confianza al 95%, con relación a la media muestral, da la amplitud dentro de la cual se encontrará el 95% de futuras medias muestrales. Esto es incorrecto, ya que la distribución de medias muestrales se centra en la media poblacional y no en la media muestral particular. La mayoría de este tipo de afirmaciones son desorientadoras e incorrectas. Lo que sí se puede afirmar es que se tiene el 95% de probabilidad de que el valor verdadero de lo que se esté estudiando en la población, se encuentre en la muestra calculada.

**La precisión absoluta (d) .** Es la amplitud deseada del intervalo de confianza a ambos lados del valor real de la diferencia entre las dos proporciones (en puntos porcentuales). Su uso es para dejar margen al error aleatorio exigido en el modelo lineal aditivo, el cual expresa que la observación i-ésima es una observación de la media μ, pero está sujeta a un error de muestreo $\varepsilon_i$ (épsilon sub i). El modelo lineal aditivo más simple es:

$$Y_i = \mu + \varepsilon_i$$

Cuanta más precisión se desee, más estrecho será este intervalo y más sujetos serán estudiados. Se fija previamente tomando en cuenta la finalidad de la investigación. En algunos casos puede requerirse una gran precisión, mientras que en otros, solo se desea conocer aproximadamente entre qué valores se encuentra el parámetro, se requiere menor precisión y, por lo tanto, menos sujetos de estudio. Las precisiones absolutas comúnmente utilizadas son: la mayor de 0.1; una media 0.05 y la más pequeña de 0.01. El tamaño de la muestra es especialmente sensible a la precisión que se elija.

| % | Valor d |
|---|---|
| 90 | 0.1 |
| 95 | 0.05 |
| 99 | 0.001 |

La precisión también puede expresarse de forma relativa **(ε)** con respecto a la estimación. La diferencia con la absoluta es que esta última se expresa en puntos porcentuales, mientras que la relativa es en porcentaje y los más comunes son: 10%, 20%, 25% y el 50%.

## *Fórmulas para calcular la muestra en estudios descriptivos*
## *Para estudios cuya variable principal es de tipo cuantitativo*

• **Para una población infinita** (cuando se desconoce el total de unidades de observación que la integran o la población es mayor a 10,000):

$$n = \frac{Z^2 \; S^2}{d^2}$$

• **Para una población finita** (cuando se conoce el total de unidades de observación que la integran):

$$n = \frac{N \, Z^2 S^2}{d^2 (N-1) + Z^2 S^2}$$

Donde:
**n =** tamaño de la muestra
**N =** tamaño de la población
**Z =** valor de Z crítico, calculado en las tablas del área de la curva normal. Llamado también nivel de confianza.
**$S^2$ =** varianza de la población en estudio (que es el cuadrado de la desviación estándar y puede obtenerse de estudios similares o pruebas piloto)
**d =** nivel de precisión absoluta. Referido a la amplitud del intervalo de confianza deseado en la determinación del valor promedio de la variable en estudio.

## *Para estudios cuya variable principal es de tipo cualitativo:*

En las investigaciones donde la variable principal es de tipo cualitativo, que se reporta mediante la proporción del fenómeno en estudio en la población de referencia, la muestra se calcula a través de las fórmulas:

• **Para población infinita** (cuando se desconoce el total de unidades de observación que la integran o la población es mayor a 10,000):

$$n = \frac{Z^2 \quad pq}{d^2}$$

• **Para población finita** (cuando se conoce el total de unidades de observación que la integran):

$$n = \frac{N\,Z^2 \quad pq}{d^2 \;\; (N - 1) + Z^2 \quad pq}$$

Donde:
**p =** proporción aproximada del fenómeno en estudio en la población de referencia
**q =** proporción de la población de referencia que no presenta el fenómeno en estudio (1 -p).
**La suma de la p y la q siempre debe dar** 1. Por ejemplo, si p= 0.8 q= 0.2 Z, N y d se explicaron en el párrafo anterior.

## *Fórmulas para calcular la muestra en estudios explicativos*

Los conceptos presentados al inicio de esta revisión con respecto a los errores tipo I y tipo II y el poder de la prueba, son esenciales para una clara comprensión del problema del tamaño de la muestra. Al escoger un tamaño de muestra para detectar una diferencia en la variable que se quiere estudiar al comparar dos grupos, se debe admitir la posibilidad de un error tipo I o un error de tipo II y calcular el tamaño de muestra en consecuencia.
Además de los elementos para calcular la muestra en los estudios descriptivos ya mencionados, se deben conocer otros datos que intervienen en el cálculo del número de sujetos necesarios para contrastar una hipótesis, como son:

- Si la hipótesis que se quiere contrastar es uni o bilateral.
- Cuál es el riesgo que el investigador acepta de cometer error tipo Alfa
- El riesgo aceptado de cometer un error β (**potencia 1-β**)
- Cuál es la magnitud mínima de la diferencia o asociación que se considera importante detectar.

Cuál es la variabilidad de respuesta en el grupo de referencia.

- Otros factores:

Prueba estadística que se usará en el análisis
Tipos especiales de diseño del estudio
Comparación de dos o más grupos
Grupos de tamaño diferente
Equivalencia de las intervenciones
Estratificación por las características basales
Múltiples variables de respuesta
Porcentaje de pérdidas y abandonos esperados

## Cálculo de muestra para prueba de hipótesis en una proporción de población

El objetivo de estos estudios es someter a pruebas de hipótesis que la proporción de sujetos de una población que posee cierta característica es igual a un valor determinado y se utiliza la prueba de Chi cuadrada ($\chi^2$)
Información necesaria y notación:
Valor de prueba de la proporción en caso de hipótesis nula P0
Valor previsto de la proporción de la población Pa
Nivel de significación Alfa 100%
Potencia de la prueba (**1-β**) 100%
Hipótesis alternativa: Prueba unilateral: Pa>Po o Pa<Po
o bien: Prueba bilateral: Pa ? Po
Fórmula para prueba unilateral

$$n= \left[ \frac{Z_\alpha \sqrt{\pi_1(1-\pi_1)} + Z_\beta \sqrt{\pi_0(1-\pi_0)}}{\pi_1 - \pi_0} \right]^2$$

## Cálculo de muestra para probar hipótesis en estudios que comparan la media de una sola muestra contra otra media ya establecida

El objetivo de estos estudios es comparar una media de una población contra otra media estandarizada, como puede ser el de otra población, sin que el investigador haya calculado esta última. La contrastación de la hipótesis puede realizarse mediante una **t** de student
Información necesaria y notación:

- Valor de la diferencia entre medias que tiene significado (**μ1 – μ0** precisión relativa) **ε**
- Valor previsto de la desviación estándar de la población Alfa.
- Nivel de significación Alfa 100%
- Potencia de la prueba (**1-β**) 100%
- Hipótesis alternativa Prueba bilateral: Pa? Po

$$n= \left[ \frac{(Z_\alpha + Z_\beta)\sigma}{\varepsilon} \right]^2$$

## Cálculo de muestra para probar hipótesis en estudios que comparan dos medias

Una fórmula semejante puede usarse para calcular el tamaño de muestras para comparar las medias de dos grupos independientes. Se deben cumplir dos premisas: las desviaciones estándar de las dos poblaciones son iguales y los tamaños de muestras son iguales en los dos grupos. La contrastación de las hipótesis también se puede realizar con una t de student para dos medias.

Información necesaria y notación:

- Valor de la diferencia entre medias que tiene significado ($\mu_1$ - $\mu_o$ precisión relativa) **ε**
- Valor porevisto de las desviaciones estándar de las dos poblaciones Alfa
- Nivel de significación Alfa 100%
- Hipótesis alternativa
- Prueba bilateral Pa ? Po

$$n = 2\left[\frac{(Z_\alpha + Z_\beta)\sigma}{\varepsilon}\right]^2$$

## Cálculo de muestra para prueba de hipótesis en estudios de casos y controles

Este tipo de muestras permite que se calcule la probabilidad de la presencia del fenómeno mediante el odds ratio, razón de momios o razón de productos cruzados.

Información necesaria y notación:

- Hay que conocer dos de los siguientes parámetros:
- Probabilidad prevista de “exposición” para las personas con la enfermedad [a/(a+b)] $P_1$*
- Probabilidad prevista de “exposición” para las personas sin la enfermedad [c/(c+d)] $P_2$*
- Razón de probabilidad prevista OR
- Nivel de confianza (1-α)100%
- Precisión relativa **ε**

Fórmula:

$$n = Z^2_{1-a/2} \frac{\left[\frac{1}{\left[P_1^*(1-P_1^*)\right]} + \frac{1}{\left[P_2^*(1-P_2^*)\right]}\right]}{[\log_e(1-\varepsilon)]^2}$$

## Cálculo de muestra para probar hipotesis en estudios de cohortes

Este cálculo de muestra permite realizar el cálculo de un riesgo relativo con precisión relativa específica.

Información necesaria y notación:

Deben conocerse dos de los siguientes parámetros:

- Probabilidad prevista de enfermedad en las personas expuestas al factor investigado P1
- Probabilidad prevista de enfermedad en las personas no expuestas al factor investigado P2
- Riesgo relativo previsto RR
- Nivel de confianza (1-alfa) 100%
- Precisión relativa **ε**

Fórmula:

$$Z^2_{1-a/2} \frac{\left[\frac{(1-p_1)}{p_1} + \frac{(1-p_2)}{p_2}\right]}{\log_e(1-\varepsilon)]^2}$$

## Cálculo de muestra para asociar dos variables cuantitativas utilizando el coeficiente de correlación de Pearson

La asociación entre dos variables cuantitativas requiere habitualmente la utilización del coeficiente de correlación r de Pearson.

Información necesaria y notación:

- Valor del coeficiente de correlación que se supone existe en la población r
- Valor previsto de la desviación estándar de la población Alfa
- Nivel de significación Alfa 100%
- Potencia de la prueba (1-β)100%
- Logaritmo natural o neperiano ln

Fórmula:

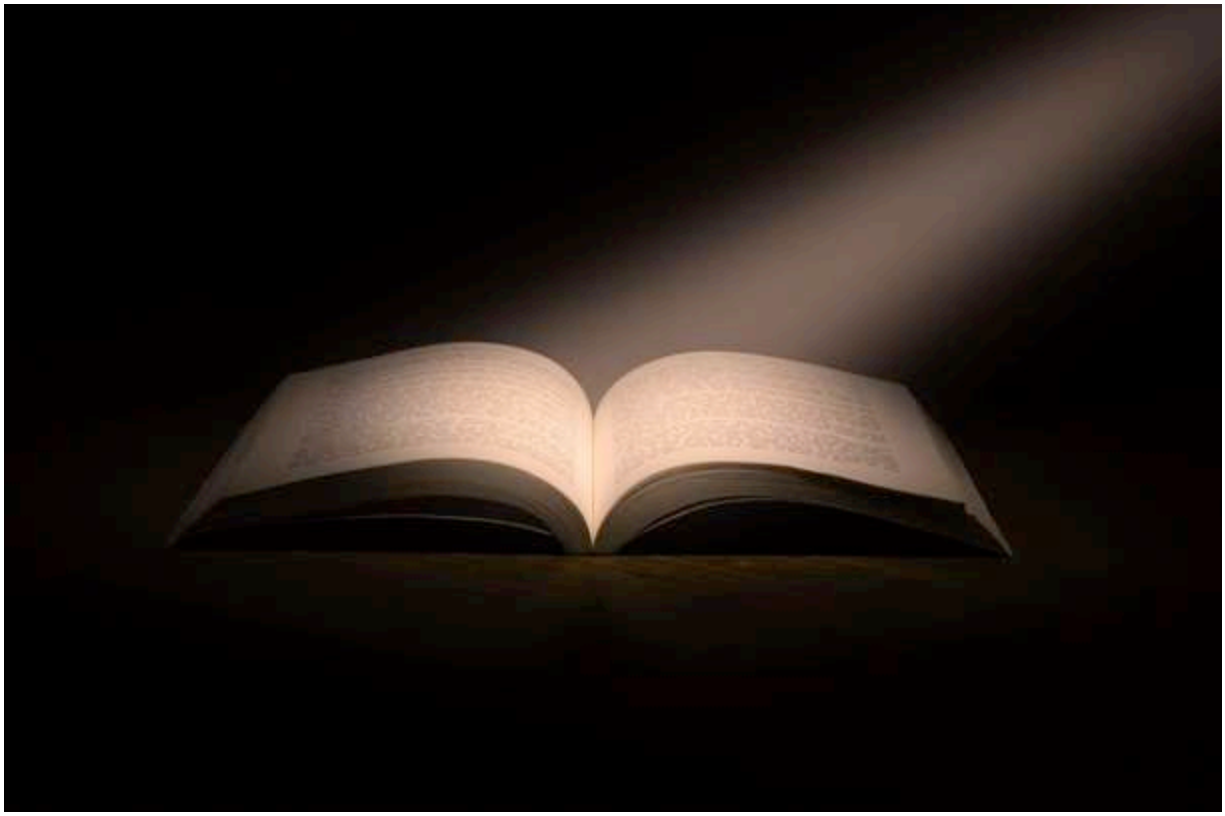
$$N= \left[\frac{(Z\alpha + Z\beta)}{0.5\ln\left[(1+r)/(1-r)\right]}\right]^2 + 3$$

En la planificación de un estudio de investigación, es fundamental decidir sobre el tamaño de la muestra de la población objeto de estudio. La respuesta a la pregunta ¿Qué tamaño de muestra necesito? dependerá básicamente del alcance del estudio, las variables y el plan de análisis. Si no se toman en cuenta estos aspectos, el investigador no podrá decidir la fórmula que utilizará y los datos que deberá incluir en ella. Además, deberá estar lo suficientemente familiarizado con el comportamiento de la variable en la población en estudio para que pueda establecer su variabilidad, decidir el nivel de confianza con el que puede trabajar y el grado de precisión más apropiado.

Todas las fórmulas presentadas, se pueden establecer con ayuda de una calculadora manual, computadora, tablas ya calculadas o con programas de estadística como **SPSS o Minitab**.

Finalmente, una vez que se ha obtenido el tamaño de la muestra, es necesario analizar las dificultades operativas, la disponibilidad de tiempo y recursos para que verdaderamente sea viable aplicarla, de lo contrario, por muy bien calculada que esté la muestra, será imposible llevarla a la realidad, lo que podría invalidar las conclusiones del estudio.

# BIBLIOGRAFÍA



Aguilar-Barojas, S. (2005).Fórmulas para el cálculo de la muestra en investigaciones de salud *Salud en Tabasco. 11*(1-2), 333-338. Recuperado el 20180710 de: http://www.redalyc.org/pdf/487/48711206.pdf

Aldás-Manzano y Maldonado-Guzmán (2016). Ecuaciones Estructurales. Universidad de Valencia.Resumen de los capítulos XIII y XIV del libro de Joaquín Aldás y Ezequiel Uriel *Análisis Multivariante Aplicado*, que es publicado por la editorial Thompson – Paraninfo.

Aldás-Manzano y Maldonado-Guzmán (2015). Desarrollo y Validación de Escalas de Medida. Universidad de Valencia.

Allen, S. J., y Hubbard, R. (1986). Regression equations of the latent roots of random data correlation matrices with unities on the diagonal. *Multivariate Behavioral Research, 21,* 393-398.

American Psychological Association (1985), *Standards for Educational and Psychological Tests*. Washington, D.C.: APA.

Anastasi, A., y Urbina, S. (1998). *Psychological testing*. Englewood Cliffs, NJ: Prentice-Hall.

Anderson, J. C., Gerbing, D. W. y Hunter J. E. (1987), On the Assessment of Unidimensional Measurement: Internal and External Consistency and Overall Consistency Criteria. *Journal of Marketing Research* 24 (November): 432-37.

Anderson James C., Gerbing David W. (1988). Structural equation modeling in practice: a review and recommended two-step approach. Psychological Bulletin.Vol. 1 (3): 411-423.

Arbuckle, J. L. (2003). Amos 5.0 [Computer software]. Chicago, IL: SPSS.

Awang, Z. (2015). *SEM made simple: a gentle approach to learning Structural Equation Modelling*. Bandar Baru Bangi: MPWS Rich Publication.

Bagozzi, R. P., y Heatherton, T. F. (1994). A general approach to representing multi-faceted personality constructs: Application to state self-esteem. *Structural Equation Modeling, 1*(1), 35-67.

Bagozzi, R. P., Yi, Y., y Phillips, L. W. (1991). Assessing construct validity in organizational research. *Administrative Science Quarterly, 36*(3), 421-458.

Bagozzi Richard P., Yi, Youjae.(1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*.Spring, Vol.16 (1): 74-94

Barrett, P. (2007). Structural equation modelling: adjudging model fit. *Personality and Individual Differences, 42*(5), 815-24.

Bearden, W. O., Hardesty, D., y Rose, R. (2001). Consumer self-confidence: Refinements in conceptualization and measurement. *Journal of Consumer Research, 28*(June), 121-134.

Bearden, W. O., y Netemeyer, R. G. (1998). *Handbook of marketing scales: Multi-ítem measures for marketing and consumer behavior research.* Thousand Oaks, CA: Sage.

Bearden, W. O., Netemeyer, R. G. y Mobic M. ( 1993), *Handbook of Marketing Sea!es: Multi-Item Measures for Marketing and Consumer Behavior*. Newbury Park, Calif.: Sage.

Bearden, W. O., Netemeyer, R. G., y Teel, J. E. (1989). Measurement of consumer susceptibility to interpersonal influence. *Journal of Consumer Research, 15*(March), 473-481.

Bearden, W. O., Sharma, S., y Teel, J. E. (1982). Sample size effects on chi-square and other statistics used in evaluating causal models. *Journal of Marketing Research, 19*(November), 425-430.

Bentler , P.M. (2006). EQS 6.1 Structural Equations Program Manual. CA: Multivariate Software Inc.

Bentler, P. M. (1992), *EQS Structural Equations Program Manual.* Los Angeles: BMDP Statistical Software.

Bentler, P. M. (1990). Comparative fit indices in structural equation models. *Psychological Bulletin, 107,* 238-246.

Bentler, P. M., y Chou, C. (1987). Practical issues in structural modeling. *Sociological Methods y Research, 16*(1), 78-117.

Bentler Peter M, Bonnet Douglas G.(1980) .Significance tests and goodness of fit in analysis of covariance structures. *Psychological Bulletin*, Sep-Dec. Vol 88 (3): 588-606.

BMDP Statistical Software, Inc. ( 1992), *BMDP Statistical Software Manual,* Release 7, vols. 1 and 2,Los Angeles: BMDP Statistical Software.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: John Wiley y Sons.

Bollen, K. A., y Lennox, R. (1991). Conventional wisdom on measurement: A

structural equations perspective. *Psychological Bulletin, 110*, 305-314.
Borgatta, E. F., Kercher, K. y Stull D. E. ( 1986), A Cautionary Note on the Use of Principal Components Analysis. *Sociological Methods and Research* 15:160-68.
Boyle, G. J. (1991). Does ítem homogeneity indicate internal consistency or ítem redundancy in psychometric scales? *Personality and Individual Differences, 3, :* 291-294.
Brown Timothy A. (2006).Confirmatory Factor Analysis for Applied Research: New York: The Guilford Press.
Browne, M. W., y Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen y J. S. Long (Eds.), *Testing structural equation models* (p. 136-162). Newbury Park, CA: Sage.
Bruner, G., y Hensel, P. (1997). *Marketing scales handbook: A compilation of multi-ítem measures* (2nd ed.). Chicago: American Marketing Association.
Bruner, G. C., y Hensel P. J. (1993). *Marketing Scales Handbook, A Compilation of Multi-Item Measures.* Chicago: American Marketing Association.
Byrne, B. M. (2010). Structural equation modelling with AMOS: basic concepts, applications, and programming (2nd ed.). New York, NY: Routledge.
Byrne, B. M. (2006).Structural Equation Modeling With EQS, basic concepts, applications, and programming.London: LEA Publishers.
Byrne, B. M. (2005a). Factor analytic models: Viewing the structure of an assessment instrument from three perspectives. *Journal of Personality Assessment*, 85: 17–30.
Byrne, B. M. (2005b). Factor analysis: Confirmatory. In B. S. Everitt D. C. Howell (Eds.), *Encyclopedia of Statistics in Behavioural Science*. 599–606. London, UK: Wiley.
Byrne, B. M. (2001). Structural equation modeling with AMOS: Basic concepts, applications, and programming. Mahwah, NJ: Lawrence Erlbaum Associates.
Calder, B. J., Phillips, L. W., y Tybout, A. M. (1982). The concept of external validity. *Journal of Consumer Research, 9*(December) . 240-244.
Campbell, D. T. (1960). Recommendations for APA test standards regarding construct, trait, or discriminant validity. *American Psychologist, 15,* 546-553.
Campbell, D. T., y Fiske, D. W. (1959). Convergent and discriminant validity by the multitrait-multimethod matrix. *Psychological Bulletin, 56 :*81-105.
Carver, C. S. (1989). How should multi-faceted personality constructs be tested? Issues illustrated by self-monitoring, attributional style, and hardiness. *Journal of Personality and Social Psychology, 56*(4), 577-585.
Cattell, R. B., Balear, K. R. Horn, J. L. y Nesselroade, J. R. (1969), Factor Matching Procedures: An Improvement ofthe s index; with tables. *Educational and Psychological Measurement* 29: 781-92
Cattell, R. B. (1966). The meaning and strategic use of factor analysis. In R. B. Cattell (Ed.), *Handbook of multivariate experimental psychology* : 174-243. Chicago: Rand McNally.
Cattell, R. B. (1966), The Scree Test for the Number of Factors. *Multivariate Behavioral Research* 1 (April): 245-76.
Chatterjee, S., Jamieson, L. y Wiseman F. (1991), Identifying Most Influential

Observations in Factor Analysis. *Marketing Science* 10 (Spring): 145-60.
Chau, Patrick Y.K. (1997). Reexamining a model for evaluating information center success using a structural equation modeling approach. *Decision Sciences.*Apr. Vol. 28 (2):309-334.
Chou, C. -P., Bentler, P. M., Satorra, A. (1991). Scaled test statistics and robust standard errors for nonnormal data in covariance structure analysis*. British Journal of Mathematical and Statistical Psychology*, 44: 347–357.
Churchill, G. A., y Iacobucci, D. (2002). *Marketing research methodological foundations* (8th ed.). Fort Worth, TX: Harcourt College Publishers.
Churchill, G.A. (1991). *Marketing Research. Methodological Foundations*. USA: Dryden Press
Churchill, G. A., y Peter, J. P. (1984). Research design effects on the reliability of rating scales: A meta-analysis. *Journal of Marketing Research, 21*(November), 360-375.
Churchill, G. A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research, 16*(February), 64-73.
Clark, L. A., y Watson, D. (1995). Constructing validity: Basic issues in scale development. *Psychological Assessment, 7*(3), 309-319.
Cliff, N. (1988). The eigenvalue-greater-than-one rules and reliability of components. *Psychological Bulletin, 103*(2), 276-279.
Cliff, N., y Hamburger, C. D. ( 1967), The Study of Sampling Errors in Factor Analysis by Means of Artificial Experiments. *Psychological Bulletin* 68: 430-45.
Comrey, A. L. (1988). Factor-analytic methods of scale development in personality and clinical psychology. *Journal of Consulting and Clinical Psychology, 56,* 754-761.
Cook, T. D., y Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Boston: Houghton Mifflin.
Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and application. *Journal of Applied Psychology, 78,* 98-104.
Creswell, J. (2007).*Qualitative Inquiry & Research Design.Choosing Among Five Approaches.*USA: SAGE Publications
Crocker, L., y Algina, J. (1986). *Introduction to classical and modern test theory.* Orlando, FL: Holt, Rinehart, y Winston.
Cronbach, L. J., y Meehl, P. E. (1955). Construct validity in psychological tests.*Psychological Bulletin, 52,* 281-302.
Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 31,* 93-96.
Crowley, S. L. y Fan, X. (1997). Structural equation modelling: basic concepts and applications in personality assessment research. *Journal of Personality Assessment, 68*(3), 508-31
DeVellis, R. F. (1991). *Scale development: Theory and applications.* Newbury Park, CA: Sage.
Diamantopoulos, A., y Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research, 36,* 269-277.
Diamantopoulos, A. y Siguaw, J.A. (2000). *Introducing LISREL*. London: Sage

Publications.
Dillon, W. R. y Goldstein M. (1984), *Multivariate Analysis: Methods and Applications.* New York: Wiley.
Dillon, W. R., Mulani, N. y Frederick, D. G. (1989), The Use of Component Scores in the Presence of Group Structure. *Journal of Consumer Research* 16:106-12.
DRAE (2017) Diccionario de la Real Academia Española, en: http://dle.rae.es/?id=DgIqVCc.
Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research, 20*(September), 303-315.
Floyd, F. J., y Widaman, K. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7*(3), 286-299.
Fornell, C., y Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(February), 39-50.
Friestad, M., y Wright, P. (2001). *Pre-adult education on marketplace persuasion tactics: Integrating marketplace literacy and media literacy*. Unpublished faculty working paper, University of Oregon, Eugene.
Ganster, D. C., Hennessey, H. W., y Luthans, F. (1983). Social desirability response effects: Three alternative models. *Academy of Management Journal, 26*(June), 321-331.
Gerbing, D. W., y Anderson, J. C. (1988). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research, 25*(May), 186-192.
Gerbing, D. W., y Anderson, J. C. (1984). On the meaning of within-factor correlated measurement errors. *Journal of Consumer Research, 11*(June), 572-580.
Gorsuch, R. L. ( 1983), *Factor Analysis*. Hillsdale,N.J.: Lawrence Erlbaum Associates.
Gorsuch, R. L. ( 1990), Common Factor Analysis versus Cornponent Analysis: Sorne Well and Little Known Facts.*Multivariate Behavioral Research* 25: 33-39.
Grande-Esteban, I y Abascal-Fernández, e (2017*). Fundamentos y Técnicas de Investigación Comercial.*13ª.Edición. Madrid: Libros profesionales de empresa ESIC. Business-Marketing-School
Green, D. P., Goldman, S. L., y Salovey, P. (1993). Measurement error masks bipolarity in affect ratings. *Journal of Personality and Social Psychology, 64,* 1029-1041.
Green, p. (1988). *Research for Marketing Decisions.* USA: Prentice Hall.
Grinell, R. (1997). *Social work research & evaluation: Quantitative and qualitative* approaches. E.E. Peacock Publishers, 5.ed. Illinois.
Hair, J. J., Anderson, R., Taham, R., y Black, W. (1999). *Análisis Multivariante.* Madrid: Prentice Hall Iberia.
Hattie, J. (1985). Methodology review: Assessing unidimensionality of tests and í tems. *Applied Psychological Measurement, 9*(June), 139-164.
Hayduk, L. A. (1996). *LISREL: Issues, debates, and strategies.* Baltimore, MD: Johns Hopkins University Press.
Heck Ronald H. (1998). Factor analysis: exploratory and confirmatory approaches in Marcoulides, G.A. Mahwah, *Modern Methods for Business Research*. NJ:

|Lawrence Erlbaum Associates Inc.

Haynes, S., Nelson, N. K., y Blaine, D. (1999). Psychometric issues in assessment research. In P. C. Kendall, J. N. Butcher, y G. Holmbeck (Eds.), *Handbook of research methods in clinical psychology* (pp. 125-154). New York: John Wiley y Sons.

Haynes, S., Richard, D. C., y Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7,* 238-247.

Herche, J., y Engellend, B. (1996). Reversed polarity ítems and scale dimensionality. *Journal of the Academy of Marketing Science, 24*(4), 366-374.

Hernández-Sampieri, R.; Fernández-Collado, C.; Baptista-Lucio M.P. (2010). Metodología de la investigación. 5ª,Ed.USA: McGrawHill

Hernández Sampieri, R., & Mendoza, C. P. (2008) El matrimonio cuantitativo cualitativo: el paradigma mixto. In J. L. Álvarez Gayou (Presidente), 6º Congreso de Investigación en Sexología. Congreso efectuado por el Instituto Mexicano de Sexología, A. C. y la Universidad Juárez Autónoma de Tabasco, Villahermosa, Tabasco, México.

Hinton, P. R; Bfownlow, C., McMurray, I., y Cozens, B. (2004) *spss explanmed*. Routledge Taylor and Francis group. London, New York.

Hooper, D., Coughlan, J. y Mullen, M. (2008). Structural equation modelling: guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 6(1), 53-60.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30,* 179-186.

Hoyle, R. (1995). *Structural equation modeling: Issues and applications.* Newbury Park, CA: Sage.

Hu, L., y Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.

Hu, L. -T., Bentler, P. M., Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin*, 112, 351–362.

Hull, J. G., Lehn, D. A., y Tedlie, J. (1991). A general approach to testing multifacted personality constructs. *Journal of Personality and Social Psychology, 61*(6), 932-945.

IBM (2011a). *IBM SPSS Statistics Base 20.* EUA. .Industrial Business Machines. Recuperado el 20161201 de: ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/20.0/es/client/Manuals/IBM_SPSS_Statistics_Base.pdf

IBM (2011b).*Guía breve de IBM SPSS Statistics 20.* EUA.Industrial Business Machines. Recuperado el 20161201 de: ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/20.0/es/client/Manuals/IBM_SPSS_Statistics_Brief_Guide.pdf

IBM (2011c). IBM SPSS Missing Values 20. EUA. .Industrial Business Machines. Recuperado el 20161201 de: ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/20.0/es/client/Manuals/IBM_SPSS_Missing_Values.pdf

Iacobucci, D., Ostrom, A., y Grayson, K. (1995). Distinguishing service quality and customer satisfaction: The voice of the consumer. *Journal of Consumer Psychology, 4*(3), 277-303.

Jarvis, W. B. G., y Petty, R. E. (1996). The need to evaluate. *Journal of Personality and Social Psychology, 70*(1), 172-194.

Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen J. S. Long (Eds.), *Testing structural equation models* (pp. 294–316). Newbury Park, CA: Sage

Jöreskog, K. G., y Sorbo, D. (1993), *LISREL 8*. Mooresville, Ind.: Scientific Software International.

Jöreskog, K., y Sörbom, D. (1989). *LISREL7: A guide to the program and applications* (2nd ed.). Chicago: SPSS.

Kaiser, H. F. (1970), A Second-Generation Little Jifty. *Psychometrika* 35: 401-15.

Kaiser, H. F. (1974), Little Jifty, Mark IV. *Educational and Psychology Measurement* 34: 111-17.

Kaplan, R. M., y Saccuzzo, D. P. (1997). *Psychological testing: Principles, applications, and issues* (4th ed.). Pacific Grove, CA: Brooks/Cole.

Kelley, J. R., y McGrath, J. E. (1988). *On time and method.* Beverly Hills, CA: Sage.

Kenny, D. A., y Kashy, D. A. (1992). Analysis of the multitrait-multimethod matrix by confirmatory factor analysis. *Psychological Bulletin, 122,* 165-172.

Kline, R. B. (2005). Principles and practice of structural equation modelling (2nd ed.). New York: The Guilford Press.

Kumar, A., y Dillon, W. R. (1987). Some further remarks on measurement-structure interaction and the unidimensionality of constructs. *Journal of Marketing Research, 24*(November), 438-444.

Lastovicka, J. L., Bettencourt, L. A., Hughner, R. S., y Kuntze, R. J. (1999). Lifestyle of the tight and frugal: Theory and measurement. *Journal of Consumer Research, 26*(June), 85-98.

Levin, R.I,Rubin, D.S. (2004) *Estadística para Administración y Economia,*7ª.Ed. México: Pearson Prentice Hall

Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports, 3,* 635-694.

MacCallum, R. C., Browne, M. W. y Sugawara, H., M. (1996). Power analysis and determination of sample size for covariance structure modelling. *Psychological Methods, 1*(2), 130-49.

McDonald, R. P. y Ho, M. H. R. (2002). Principles and practice in reporting statistical equation analyses. *Psychological Methods, 7*(1), 64-82.

McDonald, R. P. (1981), The Dirnensionality of Tests and Iterns. *British Journal of Mathematical and Social Psychology* 34: 100-117.

Mardia, K. V. (1974). Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. Sankhya, B36, 115–128.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. Biometrika, 57, 519–530.

Muthén, B. O (1997). Latent variable modeling of longitudinal and multilevel data. In A. E. Raftery (Ed.), *Sociological Methodology* 1997 (pp. 453–481). Washington, DC: American Sociological Association.

MacCallum, R. C., y Browne, M. W. (1993). The use of causal indicators in covariance structure models: Some practical issues. *Psychological Bulletin, 1 14,* 533-541.

McDonald, R. P., y Marsh, H. W. (1990). Choosing a multivariate model: Noncentrality and goodness of fit. *Psychological Bulletin, 105,* 430-445.

McDonald, R. P. (1981). The dimensionality of tests and ítems. *British Journal of Mathematical and Statistical Psychology, 34,* 100-117.

Maldonado-Guzmán, G. (2016). Apuntes de Creación de Escalas de Medida, presentados en asingatura del Doctrorado en Ciencias de la Administración, de CUCEA, Universidad de Guadalajara.

Mejía-Trejo, J. (2017a). *Banda ancha digital e Innovación abierta: Primeras Impresiones en el Sector de Tecnologías de la Información*. De Revista del Centro de Investigación Universidad Lasalle: http://ojs.dpi.ulsa.mx/index.php/rci/article/view/1146

Mejía-Trejo, J. (2017b). *The determinant factors of open business model*. De Nova Scientia:http://novascientia.delasalle.edu.mx/ojs/index.php/Nova/article/view/710/334

Mejía-Trejo, J. (2017c). *Las Ciencias de la Administracion y el Análisis Multivariante. Proyectos de investigación, análisis y discusión de resultados. Enfoque : Técnicas Interdependientes* (Vol. 2). Universidad de Guadalajara

Mejía-Trejo, J. (2017d). *Las Ciencias de la Administración y el Análisis Multivariante. Proyectos de investigación, análisis y discusión de resultados- Enfoque: Técnicas Dependientes* (Vol. 1). Guadalajara, Jalisco, México: Universidad de Guadalajara.

Mejía-Trejo, j., Sánchez-Gutiérrez, J., y Maldonado-Guzmán, G. (2016). The customer knowledge management and innovation. *Contaduría y Administración* (61), 456-477.

Mejía-Trejo, J., J., S.-G., y Vázquez-Ávila, G. (2015). *Empirical Model for Mobile Learning and their Factors. Case Study: Universities located at Metropolitan Zone of Guadalajara, México.* De Apertura. Revista de innovación educativa:http://www.udgvirtual.udg.mx/apertura/index.php/apertura/article/view/725

Mejía-Trejo, J., Sánchez-Gutiérrez, J., y Vázquez-Ávila, G. (2014). Culture of Competitiveness in the New Global Context. In J. Sánchez-Gutiérrez, J. Vargas-Barraza, y J. Mejía-Trejo, *Modeling MLQ5X for Innovation and Value Creation* (Vol. 1, pp. 1-18). Guadalajara, Jalisco: Universidad de Guadalajara.

Messick, S. (1993). Validity. In R. L. Linn (Ed.), *Educational measurement* (2nd ed., pp. 105-146). Phoenix, AZ: American Council on Education and the Oryx Press.

Mick, D. G. (1996). Are studies of dark side variables confounded by socially desirable responding? The case of materialism. *Journal of Consumer Research, 23*(September), 106-119.

Miles, J. y Shevlin, M. (1998). Effects of sample size, model specification and factor loadings on the GFI in confirmatory factor analysis. *Personality and Individual Differences, 25*, 85-90.

Mulaik, S. A. (1990), Blurring the Distinction Between Component Analysis and

Cornrnon Factor Analysis. *Multivariate Behavioral Research* 25: 53-59.
Mulaik, S. A., James, L. R., Van Alstine, J., Bennet, N., Lind, S. y Stilwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin, 105*(3), 430-45.
Mulaik, S. A., y McDonald R. P. (1978), The Effect of Additional Variables on Factor Indeterminancy in Models with a Single Cornrnon Factor. *Psychometrika* 43: 177-92.
Nederhof, A. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology, 15*(July-September), 263-280.
Netemeyer, R.G., Bearden, W.O, Sharma, S. (2003) *Scaling Procedures.Issues and Applications.*USA: Sage Publications
Netemeyer, R. G., Pullig, C., y Bearden, W. O. (2002). Observations on some key psychometric properties of paper-and-pencil measures. In A. G. Woodside y E. M. Moore (Eds.), *Essays by distinguished marketing scholars of the Society for Marketing Advances* (pp. 115-138). Amsterdam: JAI.
Netemeyer, R. G., Boles, J. S., y McMurrian, R. C. (1996). Development and validation of Work-Family Conflict and Family-Work Conflict scales. *Journal of Applied Psychology, 81*(4), 400-410.
Netemeyer, R. G., Durvasula, S., y Lichtenstein, D. R. (1991). A cross-national assessment of the reliability and validity of the CETSCALE. *Journal of Marketing Research, 28*(August), 320-327.
Neuberg, S. L., West, S. G., Thompson, M. M., y Judice, T. N. (1997). On dimensionality, discriminant validity, and the role of psychometric analyses in personality theory and measurement: Reply to Kruglanski et al.'s (1997) defense of the Need for Closure Scale. *Journal of Personality and Social Psychology, 73,* pp.1017-1029.
Nevo, B. (1985). Face validity revisited. *Journal of Educational Measurement, 22*, 287-293.
Nunnally, J., y Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
Nunnally, J. (1979), *Psychometric Theory.* New York: McGraw-Hill
Nunnally, J. L. (1978), *Psychometric Theory*. 2d ed. New York: McGraw-Hill..
OECD. (2003). Knowledge Management.Measuring Knowledge Management in the Business Sector. First Steps. In O. f.-o. Development (Ed.).
OECD. (2005). Oslo Manual. Guidelines for Collecting and Interpreting Innovation Data. In O. f.-o. Development (Ed.).
O'Rourke Norm y Hatcher Larry (2013).Step by Step Approach to Using the SAS System for Factor Analysis and Structural Equation Modeling. 2nd Ed.Cary,NC, USA: SAS Institute Inc.
Osgood, C. E., y Tannenbaum, P. H. (1955). The principle of congruence in the prediction of attitude change. *Psychological Bulletin, 62,* 42-55.
Park, C. W., Mothersbaugh, D. L., y Feick, L. (1994). Consumer knowledge assessment. *Journal of Consumer Research, 21*(June), 71-82.
Paulhus, D. L. (1993). *The Balanced Inventory of Desirable Responding: Reference manual for BIDR Version 6.* Toronto: Multi-Health Systems.
Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson,

P. R. Shaver, y L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes,* 17-59. New York: Academic Press.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46*(3), 598-609.

Perreault, W. D., Jr., y Leigh, L. E. (1989). Reliability of nominal data based on qualitative judgments. *Journal of Marketing Research, 26*(May), 135-148.

Peter, J. P. (1981). Construct validity: A review of basic issues and marketing practices. *Journal of Marketing Research, 18*(May), 133-145.

Peter, J. P. (1979), Reliability: A Review of Psychornetric Basics and Recent Marketing Practices. *Journal of Marketing Research* 16 (February): 6-17.

Peterson, R. A. (1994). A meta-analysis of Cronbach's coefficient alpha. *Journal of Consumer Research, 21*(September), 381-391.

Pitionyak, D. (2010). *From The importance of belonging*, en: https://www.hsri.org/files/NCRI/Belonging.pdf.

PNC. (2017). *Premio Nacional de Calidad México* en: http://www.pnc.org.mx/e-book-modelo-nacional-para-la-competitividad/

Price, J. P., y Mueller, C. W. (1986). Handbook of organizational measurement. Marshfield, MA: Pittman.

Richins, M. L., y Dawson, S. (1992). A consumer values orientation for materialism and its measurement: Scale development and validation. *Journal of Consumer Research, 19*(December), 303-316.

Richins, M. L. (1983). An analysis of consumer interaction styles in the marketplace. *Journal of Consumer Research, 10*(June), 73-82.

Robinson, J. P., Shaver, P. R., y Wrightsman, L. S. (1991). Criteria for scale selection and evaluation. In J. P. Robinson, P. R. Shaver, y L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes,*1-15. San Diego: Academic Press.

Robinson, J. P., y Shaver P. R. (1973), *Measures of Psychological Attitudes*. Ann Arbor, MI: Survey Research Center Institute for Social Research, University of Michigan.

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.

Rossiter, J. R. (2001). *The procedure for scale development in marketing*. Faculty working paper, University of Wollongong, Australia.

Rurnrnel, R. J. (1970). *Applied Factor Analysis*.Evanston, Ill.: Northwestern University Press.

SAS Institute, Inc. (1990). *SAS User 's Guide: Statistics, Version 6.* Cary, N.C.: SAS Institute.

Satorra, A., Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. Psychometrika, 66, 507–514.

Satorra, A., Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye C. C. Clogg (Eds.). Latent variables analysis: Applications for developmental research, 399–419. Thousand Oaks, CA: Sage.

Satorra, A. (1989). Alternative test criteria in covariance structure analysis: A unified approach. Psychometrika, 54, 131–151.

Satorra, A., Bentler, P. M. (1988). Scaling corrections for chi square statistics in covariance structure analysis. American Statistical

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. Multivariate Behavioral Research, 25, 173–180.

Saxe, R., y Weitz, B. A. (1982). The SOCO scale: A measure of the customer orientation of salespeople. *Journal of Marketing Research, 19*(August), 343-351.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment, 8,* 350-353.

Schumacker, R. E., y Lomax, R. G. (1996). *A beginner's guide to structural equation modeling.* Mahwah, NJ: Lawrence Erlbaum Associates.

Sharma, S. (1996). *Applied multivariate techniques.* New York: John Wiley y Sons.

Smith, G. T., y McCarthy, D. M. (1995). Methodological considerations in the refinement of clinical assessment instruments. *Psychological Assessment, 7,* 300-308.

Srnith, Scott M. (1989), *PC-MDS: A Multidimensional Statistics Package.* Provo, Utah: Brigharn Young University Press.

Snook, S. C., y Gorsuch, R. L. (1989), Principal Cornponent Analysis versus Cornrnon Factor Analysis: A Monte Cario Study. Psychological Bulletin 106:148-54.

Spector, P. E. (1992). *Summated rating scale construction: An introduction.* Newbury Park, CA: Sage.

Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural equation modelling. *Personality and Individual Differences, 42*(5), 893- 98

Steiger, J. H., y Lind, J. C. (1980, May). *Statistically-based tests for the number of common factors*. Paper presented at the meeting of the Psychometric Society, Iowa City, IA. Sumner, G. A. (1906). *Folkways.* New York: Ginn Custom Publishing.

Stewart, D. W. (1981), The Application and Misapplication of Factor Analysis in Marketing Research. *Journal of Marketing Research* 18 (February): 51-62.

Tabachnick, B.G. y Fidell, L.S. (2007). Using Multivariate Statistics (5th ed.). New York: Allyn and Bacon.

Tian, K. T., Bearden, W. O., y Manning, K. C. (2002). *Agents' socially desirable responding: Scale development and validation*. Faculty working paper, University of South Carolina, Columbia.

Tian, K. T., Bearden, W. O., y Hunter, G. L. (2001). Consumers' Need for Uniqueness: Scale development and validation. *Journal of Consumer Research, 28*(June), 50-66.

Trochim, A. (2002). *Construct validity*. Retrieved November 19, 2002, en: http:// trochim.human.cornell.edu/kb/constval.htm

Velicer, W. F., y Jackson, D. N. (1990), Cornponent Analysis versus Cornrnon Factor Analysis: Sorne Issues in Selecting an Appropriate Procedure. *Multivariate Behavioral Research* 25: 1-28

Wheaton, B., Muthen, B., Alwin, D., F. y Summers, G. (1977). Assessing Reliability and Stability in Panel Models. *Sociological Methodology, 8*(1), 84-136.

Widaman, K. F. (1985). Hierarchically nested covariance structure models for

multitrait-multimethod data. *Applied Psychological Measurement, 9,* 1-26.
Wright, P. L. (1975). Factors affecting cognitive resistance to advertising. *Journal of Consumer Research, 2*(June), 1-9.