

DEPURACIÓN DE ESCALAS EN MARKETING. UNA APLICACIÓN PRÁCTICA.

Sonia San Martín Gutiérrez
Universidad de Burgos

RESUMEN

El diseño de las escalas como parte integrante de un cuestionario es, a menudo, un problema al que se enfrentan los investigadores de diferentes áreas de conocimiento. Los conceptos de fiabilidad y validez adquieren gran importancia en cualquier estudio con escalas que aspire a ser mínimamente riguroso. Este trabajo trata de ilustrar los pasos a seguir en el diseño y depuración de las escalas con una aplicación práctica, extraída de una investigación más amplia realizada en el área de marketing. Para lograr este fin, se hará especial hincapié en la utilidad del programa LISREL para la validación de las escalas y de un modelo de ecuaciones estructurales.

Palabras clave: *fiabilidad, validez, depuración de escalas, LISREL, compromiso.*

Introducción

En las ciencias sociales, en ocasiones falta consenso entre los investigadores acerca de la escala más adecuada para medir ciertos conceptos, lo cual puede llevar a resultados muy diferentes y confusos en función de la escala elegida. Además, la fijación por los datos y los resultados así como la rapidez en la realización de encuestas se erigen como factores que actúan en detrimento de la calidad de las escalas utilizadas, pudiendo llegar a conclusiones que realmente no son útiles para la toma de decisiones ni generalizables a otros contextos.

Así, el *objetivo* de este trabajo es ilustrar los pasos a seguir en el diseño y depuración de las escalas con una aplicación práctica, extraída de una investigación más amplia realizada en el ámbito del marketing relacional.

En primer lugar, se comentarán las principales fases que son necesarias para la depuración de escalas con el programa LISREL para posteriormente, utilizarlas en el ejemplo práctico. Los programas informáticos que se han empleado para la aplicación práctica han sido el SPSS/PC+ en una primera fase y, principalmente, el LISREL para obtener las escalas definitivas.

La depuración de escalas con el programa Lisrel. Fases de análisis.

Una de las técnicas más apropiadas para validar escalas y modelos y que goza de gran aceptación en disciplinas como el marketing, es el enfoque de ecuaciones estructurales. En este enfoque, la medida es el proceso por el cual un concepto viene asociado a una o más variables latentes (constructos, variables teóricas, abstractas) y éstas a variables observables (medidas). Así, el modelo de medida especifica un modelo estructural relacionando variables latentes a una o más medidas o variables observables (Bollen, 1989). El proceso de medida se puede resumir en el esquema de la figura 1. Un problema que existe en algunas disciplinas, como el marketing, es la ausencia de consenso en torno a la definición y medida de determinadas variables (como el compromiso), lo que dificulta la comparación y generalización de los resultados de diferentes investigaciones sobre la(s) misma(s) variable(s). Un desarrollo de medidas inadecuadas para las variables (por omisión de ítems importantes o por la inclusión de ítems redundantes o que miden otra variable diferente) puede conllevar unos malos resultados finales o poco relacionados con el objetivo que se persigue en un determinado trabajo.

Una vez recogida la información en el trabajo de campo, las principales fases a seguir en la validación de un modelo mediante el enfoque de ecuaciones estructurales son las que se exponen a continuación. Para ello, podemos utilizar los programas SPSS y LISREL¹.

¹ Jöreskog y Sorböm (Universidad de Upsala) fueron los investigadores que crearon el programa informático LISREL basado en el análisis de las estructuras de covarianzas, como ayuda de la modelización en econometría y psicometría. No obstante, existen otros sistemas como el COSAN (*Covariance Structure Analysis*), el RAM (*Reticular Analysis Model*) y el EQS (*Equation Structu-*

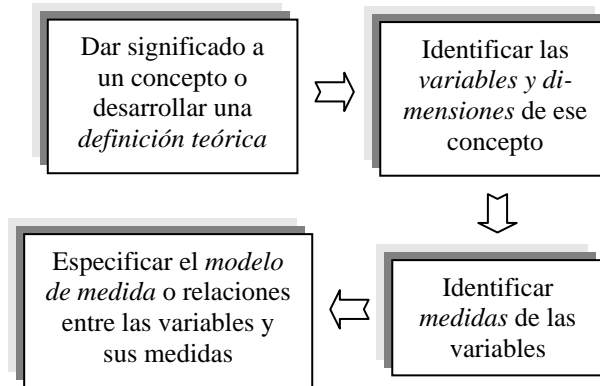


Figura 1: *fases del proceso de medida.*

Fuente: adaptado de Bollen (1989).

Análisis y depuración inicial de la base de datos

Tal como apuntan Bagozzi y Baumgartner (1994), es importante asegurarse de que no existen errores en la codificación de las variables y de que se tratan correctamente los valores ausentes. También recomiendan comprobar la normalidad de las variables mediante histogramas, gráficos *stem-and-leaf* y *Q* (los residuos se deben distribuir simétricamente alrededor del valor cero o a lo largo de una línea de 45 grados, respectivamente) y el cálculo de los coeficientes de asimetría y curtosis (que deben ser cero y tres, respectivamente, para hablar de normalidad). También se calcula en esta etapa la matriz de covarianzas o de correlaciones. Seguidamente, tiene lugar el análisis de la fiabilidad y validez de las escalas con la ayuda de los paquetes estadísticos SPSS, para el análisis de la fiabilidad, y LISREL, para el análisis de la validez. El método de estimación de parámetros que utilizamos en el caso práctico es el de *máxima verosimilitud*, que se basa en el supuesto de normalidad de las variables y que, como señala Bisquerra (1989), es “*lo suficientemente robusto como para no verse seriamente afectado por ligeras oscilaciones respecto a la distribución normal multivariable*”, pero siempre que el tamaño de la muestra sea mayor que 100 (Gerbing y Anderson, 1985).

Determinación de la fiabilidad de una escala

La *fiabilidad* de una medida o escala es la consistencia, exactitud y capacidad de predicción de los resultados que proporciona una escala o, dicho de otro modo, el grado en que el proceso de medición está libre de errores aleatorios (Grande y Abascal, 1994; Kinnear y Taylor, 1995). Algunos factores que influyen en la fiabilidad de una medida

ral System) que ofrecen otras posibilidades, pero que no van a ser comentados aquí por exceder el alcance del presente trabajo.

son el número de cuestiones realizado y su dificultad, el tipo de escala, el número de dimensiones, la denominación utilizada, etc (Peter y Churchill, 1986).

Cuadro 1: *principales ventajas e inconvenientes de diferentes técnicas de evaluación de la fiabilidad de una escala*

Técnica de evaluación de la fiabilidad	Ventajas	Desventajas
<i>Prueba-reprueba o test-retest</i> (mide la estabilidad temporal de la medida)	<ul style="list-style-type: none"> - Más exacta la estimación de fiabilidad cuantas más mediciones. - Simplicidad. 	<ul style="list-style-type: none"> - Son necesarios datos longitudinales. - Efecto memoria. Variaciones de los individuos muestrales, cansancio de éstos. - El mero hecho de repetir las cuestiones o la existencia de un evento entre t y t+1 puede sesgar las respuestas.
<i>Formas alternativas</i> (mide la consistencia interna de la medida)	<ul style="list-style-type: none"> - Reduce el efecto memoria. - Menor posibilidad de correlación de los errores de medida que en el anterior. 	<ul style="list-style-type: none"> - Son necesarios datos longitudinales; el fenómeno medido debe ser relativamente duradero. - Es difícil que las medidas sean equivalentes en contenido. - También aquí el proceso de medida puede cambiar las respuestas.
<i>Método de la fiabilidad bipartida, de división por la mitad o de las dos mitades</i> (mide la consistencia interna de la medida)	<ul style="list-style-type: none"> - Basta un solo momento de tiempo (corte transversal). - Menor esfuerzo y coste de realización. 	<ul style="list-style-type: none"> - Dificultad de agrupación de las cuestiones en dos mitades (porque el criterio de agrupación puede influir en el grado de correlación obtenido)².
<i>Coefficiente alpha de Cronbach</i> (la medida más popular en ciencias sociales y mide la consistencia interna de la medida)	<ul style="list-style-type: none"> - Basta una sola medición con la misma medida y en un solo momento de tiempo (corte transversal). - Facilidad de realización. 	<ul style="list-style-type: none"> - Aunque con un límite, este coeficiente aumenta según incorporamos items a la escala, incluso cuando éstos tienen unos coeficientes de regresión muy pequeños o no significativos. - No estima adecuadamente los errores debidos a factores externos al instrumento de medición (ej. factores personales).
<i>Correlación ítem-total</i> (entre un ítem y la escala global) y <i>correlación inter-items</i> (entre ítems)	<ul style="list-style-type: none"> - Sirve como complemento de las medidas de fiabilidad de las escalas globales. - Mide simultáneamente la validez y la fiabilidad. 	<ul style="list-style-type: none"> - Es una medida de fiabilidad individual. No valora la fiabilidad de la escala completa.
<i>Técnica de ecuaciones estructurales - AFC</i>	<ul style="list-style-type: none"> - Es aplicable cuando existe más de una variable latente. - Permite la correlación de los errores de medida. 	<ul style="list-style-type: none"> - Requiere una determinada distribución estadística (normalidad multivariable).

En términos estadísticos y, más concretamente, en los modelos de ecuaciones estructurales, la fiabilidad viene dada por el cuadrado de la correlación que existe entre un

² Una solución es la que propone Peter (1979): tomar como coeficiente de fiabilidad la media de los coeficientes de fiabilidad obtenidos agrupando las cuestiones en dos bloques de todas las formas posibles.

constructo y sus medidas, es decir, por el grado de relación directa que existe entre la variables latente y la variable observable (Bollen, 1989).

En el cuadro 1 se exponen diferentes formas de medir la fiabilidad de una escala (para datos cuantitativos) junto con sus ventajas y desventajas principales (Nunnally, 1978; Churchill, 1979; Bollen, 1989; Steenkamp y van Trijp, 1991; Sáez y Sánchez, 1997).

Uno de los métodos expuestos con más aceptación en los últimos años es el análisis factorial confirmatorio (AFC) y los modelos de ecuaciones estructurales. Un primer paso anterior al AFC es el análisis factorial exploratorio (AFE), como veremos en la aplicación práctica. Mientras el AFE ayuda a generar hipótesis, modelos y estructuras, el AFC sirve para contrastar hipótesis, modelos o estructuras a partir de la información que, sobre la estructura de los datos, ofrece la teoría o la hipótesis (Jöreskog y Sörbom, 1993; Lévy, 1994). Además, en el AFE, frente al AFC, no se especifica previamente el número de variables latentes, de forma que todas ellas pueden influir sobre todas las observables, mientras que los errores de medida no pueden correlacionar (Bollen, 1989).

Unidimensionalidad

Cuando se mide un único constructo, se hace necesario comprobar que se cumpla el requisito de la *unidimensionalidad* (Bagozzi y Baumgartner, 1994), por la cual debe ser precisamente un único constructo el que subyace en un conjunto de ítems que constituyen una escala. Pero si hay algún constructo intrínsecamente multidimensional, como es el caso del ejemplo de este trabajo, la unidimensionalidad se ha de analizar para cada factor o dimensión individual. La unidimensionalidad de una variable se puede detectar mediante la realización de un *análisis factorial exploratorio* (Steenkamp y van Trijp, 1991). En nuestro ejemplo, lo haremos a través de un análisis de ejes principales con rotación *varimax* (Hair y otros, 1999), que permite elegir los ítems que tienen una carga factorial alta. Seguidamente, se depura el conjunto de ítems resultante mediante un *análisis factorial confirmatorio*, como luego veremos.

Conviene advertir que, dado que la fiabilidad de la escala depende especialmente de su longitud y del número de categorías, el coeficiente *alpha* es una medida de fiabilidad y no de unidimensionalidad porque no permite probar la estructura factorial del constructo (Steenkamp y van Trijp, 1991).

Concepto y tipos de validez de una escala

La *validez* de una escala hace referencia al grado en el cual el proceso de medida está libre tanto de errores sistemáticos como de errores aleatorios (Bisquerra, 1989; Kinnear y Taylor, 1995). Por tanto, la validez es un concepto más amplio que el de fiabilidad, de forma que una medida válida generalmente es fiable, pero una medida fiable no tiene por qué ser válida.

El cuadro 2 muestra algunas características e inconvenientes vinculados a diferentes tipos de validez (Churchill, 1979; Bollen, 1989; Steenkamp y van Trijp, 1991; Grande y Abascal, 1994 y Kinnear y Taylor, 1995).

Un problema de la medición de la validez mediante técnicas tradicionales es que únicamente utilizan variables observables, no incorporan variables latentes. De esta forma, se estaría suponiendo implícitamente que la correlación entre dos variables observables refleja exactamente la asociación entre variables latentes. Así, en la validez de criterio, la variable criterio es una *proxy* de la variable latente y en la validez de constructo, la correlación entre las medidas observables es una *proxy* de la correlación entre los constructos latentes (Bollen, 1989; Camisón y Bou, 2000).

Cuadro 2: principales características e inconvenientes de diferentes tipos de validez de una escala

Tipo de validez	Características	Inconvenientes
Validez de contenido (es cualitativa)	- Implicaría un consenso sobre el concepto entre los investigadores.	- Subjetividad (validez facial). - Depende en gran medida de la definición teórica subyacente (lo cual es un problema si falta consenso sobre el concepto en cuestión).
Validez de criterio (concurrente o predictiva)	- La validez <u>concurrente</u> es útil para comprobar la validez de una nueva forma o técnica de medición mediante su correlación con medidas ya existentes y admitidas. - La validez <u>predictiva</u> indicaría la estabilidad o correlación de una medida registrada en dos momentos de tiempo distintos.	- Es necesario una variable estándar con la que comparar la medida. - El grado de validez para la misma medida difiere según el criterio elegido.
Validez de constructo (convergente y discriminante)	- Es adecuada cuando no se puede medir la validez de contenido. - La validez <u>convergente</u> implicaría un grado suficiente de correlación entre las mediciones realizadas con técnicas independientes. - La validez <u>discriminante</u> muestra una falta decorrelación entre constructos diferentes, aunque conceptualmente similares.	- Es un proceso largo y complejo. No existe un test empírico para calcularla.
Validez nomológica	- Indica si una escala o medida demuestra las relaciones cuya existencia se deriva de la teoría o de la investigación previa.	- Es un proceso largo y complicado (ej. con modelos estructurales rivales).

Por este motivo, hay autores como Bollen (1989) que prefieren definir la validez en el contexto de las ecuaciones estructurales (complementaria de las formas tradicionales), de forma que la validez de una medida de la variable latente es la magnitud de la relación *directa* entre la variable latente y la variable observable o medida, relación ésta que debe ser estructural, causal. De este modo, se exige que exista una relación directa entre la variable y su medida, evitando aproximaciones. En esta línea, Steenkamp y van Trijp (1991) destacan la utilidad del programa LISREL frente a otros métodos tradicionales para apreciar la validez de constructo ya que proporciona un test de la estructura teórica del instrumento de medida (la relación entre el constructo y sus medidas) y permite contrastar así la relación entre diferentes constructos sin el sesgo que produce el error de medida.

Así, los *coeficientes lambda* o *cargas factoriales*, que miden la relación directa entre una variable latente y la variable observable o medida en una ecuación estructural, pueden interpretarse como coeficientes de validez de la segunda para medir la primera. Para que indiquen validez de la escala, se considera como requisitos que sean significativos y que su valor sea al menos 0,5 (Bollen, 1989; Bisquerra, 1989).

El problema es asignar una escala a la variable latente para comenzar el proceso de validación de un modelo. Una solución es igualar uno de los coeficientes *lambda* de una variable latente a 1 de forma que su escala sea igual que la de su variable observable y el resto de *lambdas* de esa variable latente quedan medidos de este modo en la escala de la variable observable. Para facilitar la comparación de los coeficientes en el caso de que cada variable observable venga medida en una escala muy diferente, se suelen calcular los coeficientes *lambda* estandarizados (similares a los coeficientes de regresión estandarizados).

En este trabajo nos vamos a centrar en la técnica de las ecuaciones estructurales y, por tanto, del AFC para depurar las escalas del ejemplo práctico. Por ello, en el siguiente apartado, se exponen las fases que se siguen en la realización de un AFC para validar escalas. Asimismo, se comentan las medidas que indican la bondad del ajuste.

El análisis factorial confirmatorio. Fases y medidas de bondad del ajuste

Las fases a seguir en la realización de un AFC son similares a las que se realizan para elaborar y validar un modelo. El objetivo es llegar, a través de sucesivos AFC, a un modelo satisfactorio tanto desde el punto de vista de bondad del ajuste como de significación de los resultados (Jöreskog, 1971). En la figura 2 aparecen las principales fases a seguir en la validación de un modelo.

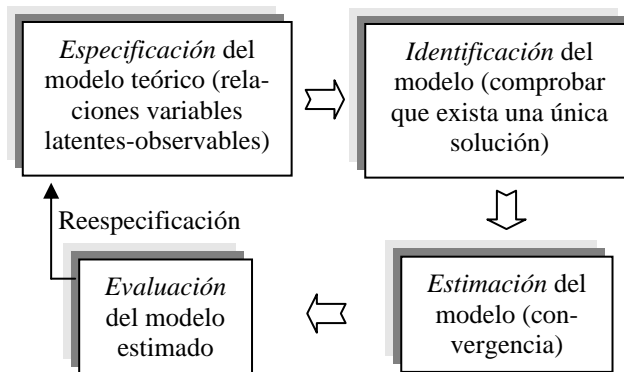


Figura 2: *fases del proceso de validación de un modelo.*

Fuente: adaptado de Bollen (1989).

A partir de las medidas de las variables, hay que *especificar* las relaciones entre las variables latentes y sus medidas. Al igual que lograr una correcta definición y medición de cada variable, es fundamental especificar correctamente y a la luz de la teoría, las

relaciones entre las variables. En nuestra opinión, es mucho más importante el marco teórico que subyace detrás de un determinado modelo que las diferentes relaciones que se pueden probar empíricamente. Por tanto, aunque en ocasiones se podrían establecer relaciones entre todas las variables, hay que realizar un esfuerzo para justificar cada relación e hipótesis que se va a contrastar empíricamente. Pese a que este razonamiento parece obvio, varios trabajos parecen guiarse primero por los resultados que por la teoría pertinente.

En segundo lugar, para *identificar* un modelo hay que introducir restricciones para los parámetros. Bollen (1989) analiza diferentes reglas y tests que ayudan en la identificación de un modelo. Sin embargo, hay programas como el LISREL y el EQS que calculan automáticamente un test empírico de la matriz de datos que también sugiere el grado de identificación del modelo.

Un problema que puede surgir en la validación de un modelo es la no convergencia del mismo cuando, tras repetidas iteraciones, los valores obtenidos no están suficientemente próximos como para llegar a una solución. La convergencia depende de (Bollen, 1989): a) el criterio que defina lo que son “valores insuficientemente próximos”; b) el número de iteraciones que se permitan antes de dejar de intentar la validación del modelo (250 iteraciones en el programa LISREL); c) los valores de partida que se otorguen a los parámetros desconocidos (LISREL los proporciona automáticamente; si son buenos, bastarán pocas iteraciones para la validación de un modelo); d) la correcta especificación del modelo; e) las fluctuaciones muestrales de las varianzas y covarianzas de las variables observables; f) el tamaño de la muestra y g) el número de indicadores por factor (es más probable la convergencia para muestras con tamaño mayor a 150 y con más de dos indicadores por factor).

Una vez contrastado el modelo, es necesario *evaluar* los resultados obtenidos con el fin de averiguar si son o no aceptables de acuerdo con lo previsto, si el ajuste del modelo es estadísticamente bueno y, si no, cuáles pueden ser las razones de desviación respecto a lo previsto. Para ello, aparte de observar la factibilidad de los parámetros obtenidos (signo del parámetro y su valor, etc), algunas de las medidas más utilizadas para conocer la bondad del ajuste de un modelo cuando se utiliza un enfoque de ecuaciones estructurales son las que aparecen a continuación (Bagozzi y Yi, 1988; Bisquerra, 1989; Bollen, 1989; Bentler, 1990; Jöreskog y Sörbom, 1993; Bagozzi y Baumgartner, 1994 y Baumgartner y Homburg, 1996). Como recomiendan diferentes autores, conviene utilizar varias medidas de bondad del ajuste en la interpretación de los resultados -índices absolutos e incrementales, de centralidad y de no-centralidad- con el fin de evitar los inconvenientes de cada una de ellas (Bagozzi y Yi, 1988; Baumgartner y Homburg, 1996; Mackenzie, 2001).

- El valor de la *Chi-cuadrado* indica si el modelo se ajusta a los datos. En concreto, mide la desviación que existe entre la matriz de covarianzas o correlaciones muestral y la ajustada; contrasta la hipótesis nula de que la matriz de varianzas-covarianzas estimada se desvía de la muestral sólo debido al error muestral. El ajuste será mejor cuanto menor sea el valor de la Chi-cuadrado o cuanto más exceda de 0,05 el valor de su probabilidad. Pero hay que tener presente que estamos ante un test relativamente sensible al número de parámetros del modelo: el valor de la Chi-cuadrado disminuye conforme se añaden parámetros al modelo. Y también es sensi-

ble al tamaño muestral. Si la disminución del valor de la Chi-cuadrado al intentar mejorar un modelo es grande respecto de la pérdida de grados de libertad, se habrá conseguido una mejora real del modelo (si las disminuciones son similares, el ajuste se ha conseguido mediante “prueba y error” y los parámetros obtenidos no tendrán un verdadero significado-Jöreskog, 1971). Por tanto, cuando alguno de los supuestos de la Chi-cuadrado no se cumple, pierde valor como índice de bondad del ajuste.

- El *índice de bondad del ajuste (GFI, Goodness of Fit Index)* y el *índice de bondad ajustado (AGF, Adjusted Goodness of Fit Index)* son dos medidas de bondad del ajuste absolutas y miden la variabilidad explicada por el modelo y pueden tomar valores entre 0 y 1. Cuanto más cerca se encuentren de 1, mejor es el ajuste. No depende del tamaño de la muestra y son, por tanto, menos sensibles que la Chi-cuadrado a las desviaciones de la normalidad. Sin embargo, aunque estas medidas no dependen explícitamente del tamaño muestral, su distribución muestral sí depende del tamaño (Jöreskog y Sörbom, 1993). Así, en ocasiones, los valores fluctúan en torno a 0,9 pese a ser modelos con errores de especificación.
- El *índice de bondad del ajuste comparativo (CFI, Comparative Fit Index)* es una medida de bondad del ajuste incremental. Indica en qué grado es mejor el ajuste del modelo en comparación con un modelo de referencia, normalmente el modelo de las variables independientes. Al igual que los índices GFI y AGFI, su valor puede oscilar entre 0 y 1, de modo que el ajuste será mejor cuanto más cerca de 1 sea el valor de este índice. Su principal inconveniente es que no refleja la parsimonia de los modelos alternativos.
- La *raíz media residual (RMR, Root Mean Residual)* es una medida resumen del tamaño medio de los residuos. Un valor cercano a cero indica que el modelo está ajustado, pero es difícil valorar en qué medida no está ajustado.

Otros medios que ayudan a detectar desajustes en el modelo son:

- Los *valores “t”* de cada parámetro estimado (deben ser mayores que 1.96 si se busca una significación al 95%);
- Los valores de los *residuos normalizados* (son residuos corregidos en cierta medida del efecto tamaño de la muestra y de diferencias de escala) y la *gráfica de residuales normalizados* (si la pendiente es mayor que 1 indica un buen ajuste; si los puntos no se disponen en una línea recta, pueden existir desviaciones de la normalidad o errores de especificación);
- La *correlación múltiple al cuadrado* (mide el grado en que una relación está bien especificada para una determinada variable del modelo y la fiabilidad de cada variable en los modelos con variables latentes);
- El *coeficiente de determinación total* (es igual a la correlación múltiple para todas las variables conjuntamente y señala el grado en que las relaciones están bien especificadas en el modelo de forma global).

Las medidas anteriores nos pueden indicar la necesidad de *reespecificar* el modelo. Esta necesidad surge cuando el ajuste es malo o no se logra el ajuste del modelo propuesto. Hay que prestar mayor cuidado en la reespecificación del modelo estructural que en el caso de la modificación del modelo de medición porque el modelo estructural no puede dejar de ser la representación fiel de las hipótesis objeto del contraste (Bagozzi y

Baumgartner, 1994). Aunque es el investigador el que debe saber cómo puede reespecificar y mejorar el modelo a partir de la teoría, para detectar problemas en la especificación del modelo, es especialmente útil analizar los residuos estandarizados (y los gráficos *stemleaf* y Q) y los denominados índices de modificación (Bisquerra, 1989; Jöreskog y Sörbom, 1993; Bagozzi y Baumgartner, 1994 y Camisón y Bou, 2000).

Por lo que respecta a los *residuos estandarizados*, un ítem que está en el factor incorrecto presentará grandes residuos negativos con otros ítems de ese factor y positivos con los ítems del factor “correcto” (Steenkamp y van Trijp, 1991) o puede que un indicador cargue en varios factores, en cuyo caso se le puede relacionar con esos factores o dejar correlacionar sus errores (Bagozzi, y Baumgartner, 1994). Esta última solución sólo se debe utilizar como excepción y siempre que haya una justificación teórica para ello, ya que, en caso contrario, se podría sacrificar la unidimensionalidad de la medida (Bagozzi, y Baumgartner, 1994) y “la principal asunción en los modelos de ecuaciones estructurales es que el término de error en cada relación está incorrelacionado con todas las variables independientes” (Jöreskog y Sörbom, 1993).

Por su parte, los *índices de modificación* nos permiten analizar en qué medida se puede mejorar el ajuste si se relajan algunas de las restricciones que impone el modelo. Hay un índice de modificación para cada parámetro fijo del modelo (especificado como 0) igual al decremento esperado en la Chi-cuadrado en caso de que sólo fuese libre ese parámetro y el resto se mantuvieran con sus valores estimados. Se procede buscando el índice de modificación más alto y dejándole libre siempre que haya explicación para ello. Seguidamente, tratamos de validar el modelo resultante (Bisquerra, 1989; Jöreskog y Sörbom, 1993 y Bagozzi y Baumgartner, 1994). Obviamente, todos los cambios que se realicen deben estar justificados por la teoría.

Una aplicación práctica con el programa *lisrel*

Planteamiento del problema:

análisis del compromiso del consumidor en sus relaciones comerciales

Como ejemplo práctico, se expone una parte de un estudio realizado desde el enfoque de marketing relacional para las relaciones que entabla el consumidor con la empresa; nos centramos en esta aplicación práctica en el estudio de la variable compromiso (San Martín y otros, 2000). En primer lugar y haciendo referencia al proceso de medida señalado previamente en la figura 1, podemos comenzar definiendo el compromiso. Tras revisar las definiciones de la literatura (Anderson y Weitz, 1992; Gundlach y otros, 1995; Morgan y Hunt, 1994; Rylander y otros, 1997), podemos entender por compromiso *un deseo de desarrollar y mantener a largo plazo relaciones de intercambio, deseo que se materializa en la realización de unas promesas implícitas y explícitas y de unos sacrificios en pro del beneficio económico y social de todas las partes que tienen algún interés en la relación* (San Martín y otros, 2000).

A continuación, hay que identificar las posibles dimensiones de esa variable. En este sentido, cabe decir que el análisis del compromiso en el marketing relacional se ha centrado generalmente en un único aspecto del mismo. De hecho, la dificultad de materializar la naturaleza abstracta del compromiso en una acepción general ha llevado a la

discusión acerca de su consideración como un constructo unidimensional o multidimensional. Este dilema se ha resuelto habitualmente mediante el tratamiento del compromiso de una forma global (Dwyer y otros, 1987; Garbarino y Johnson, 1999; Grayson y Ambler, 1999; Morgan y Hunt, 1994), sin hacer referencia explícita a las posibles dimensiones del mismo. Además, con frecuencia se ha medido el compromiso a través de escalas globales que no captan todos los aspectos que encierra el compromiso (Rylander y otros, 1997). En esa elección por una medición unidimensional no se ha tenido en cuenta que los modelos que contemplan el compromiso de forma multidimensional tienen mayor probabilidad de alcanzar poder explicativo e, incluso, predictivo que los modelos que tratan el compromiso de forma global, cuya validez metodológica es cuestionable (Jaros y otros, 1993; Pritchard y otros, 1999).

Geyskens y otros (1996) advierten que la definición tradicionalmente otorgada al compromiso —intención de un miembro del canal de continuar la relación— no comprende las diferentes razones que llevan a esa intención y que dan lugar a diversos componentes del compromiso. Así, es posible que los factores señalados como determinantes del compromiso lo sean tan sólo de alguna de sus dimensiones básicas (Meyer y Allen, 1991; Gundlach y otros, 1995 y Kim y Frazier, 1997). Del repaso de la investigación sobre el término se desprende que el compromiso relacional implica aspectos de *tiempo* —intención de mantenimiento duradero de la relación—, de *comportamiento* —inversión en la relación (por ejemplo, inversiones en activos específicos) — y de *afecto* —sentimientos de las partes—, dimensión esta última más implícita en la relación y, por tanto, más difícil de medir (Rylander y otros, 1997).

En consecuencia, en este caso práctico diferenciamos tres componentes del compromiso: uno referido a deseos y sentimientos —*dimensión afectiva*—, otro que contempla el comportamiento efectivo actual —*dimensión comportamental*— y un tercero que indica la intención futura de compromiso —*dimensión temporal*. El siguiente paso en el proceso de medición, la búsqueda de medidas para cada dimensión, se comenta en el apartado 3.2.

El último paso en el proceso de medición era especificar las relaciones entre las variables (figura 3). La literatura sobre el compromiso sugiere un ciclo de reforzamiento del mismo (Anderson y Weitz, 1992; Shemwell y otros, 1994; Gundlach y otros, 1995; Rylander y otros, 1997). Por ejemplo, Gundlach y otros (1995) consideran el compromiso como un proceso que comprende un *input* (credibilidad de la otra parte y recursos comprometidos por ambas partes), un componente afectivo (atadura psicológica, identificación y congruencia de valores) y una dimensión temporal (consistencia de inputs y actitudes en el tiempo). Por tanto, estamos ante un ciclo en el que todas las dimensiones están interrelacionadas y se refuerzan mutuamente. En concreto, los tres componentes que nosotros hemos diferenciado están relacionados del modo siguiente (San Martín y otros, 2000; San Martín, 2003): el compromiso afectivo es capaz de generar unos vínculos de tipo emocional que pueden desembocar en un compromiso de comportamiento y en un compromiso temporal. El compromiso comportamental también puede conducir al temporal, pero es de esperar que no sea tan duradero y sólido como el que proviene de los lazos emocionales del compromiso afectivo.

Debido a que el objetivo de este epígrafe es la aplicación del programa LISREL para la depuración de escalas, seguidamente serán comentadas las fases de análisis utiliza-

das en la obtención de las escalas definitivas para cada dimensión del compromiso. La justificación teórica y el contraste de las relaciones entre todas las variables del modelo global se puede encontrar en San Martín (2003).

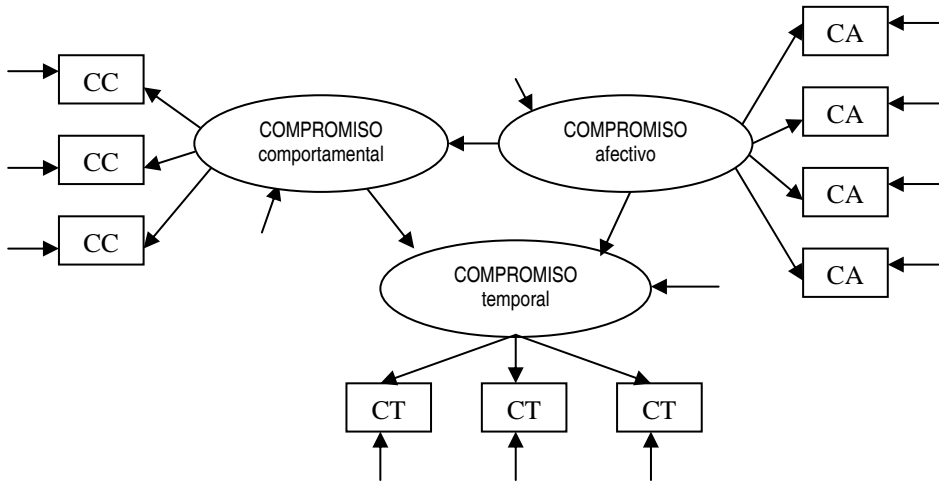


Figura 3: *diagrama del modelo de ecuaciones estructurales que refleja las relaciones entre las dimensiones del compromiso.*

Recogida de información. Cuestionario, variables y medidas

El caso objeto de estudio exigía la obtención de información primaria mediante la realización de una encuesta. El universo del estudio lo constituyeron los usuarios de agencias de viajes de Burgos y Valladolid capitales. La información fue recogida mediante cuestionarios autoadministrados en marzo, abril y mayo de 2001 con una tasa de respuesta del 35.26% (171 encuestas de un total de 485 entregadas). La caracterización de la muestra aparece en el cuadro 3. A los individuos muestrales se les pedía que seleccionaran la agencia de viaje con la que tenían más relación y que refirieran sus respuestas a ella.

A los efectos de elaborar el cuestionario soporte de la encuesta, realizamos previamente una revisión bibliográfica exhaustiva de los trabajos que proporcionan escalas de medición de las variables contempladas en nuestro estudio. Así, se tiene una buena base para cumplir con los requisitos de fiabilidad y validez de las escalas. No obstante, los ítems obtenidos tuvieron que ser corregidos para conseguir su adaptación a nuestro contexto particular. Además, generamos otros nuevos³. Los ítems que finalmente utilizamos

³ En lo que al número de ítems por escala se refiere, intentamos no restringirlo en exceso, ya que, como puntualizan Nunnally y Berstein (1994), la utilización de escalas con uno o dos ítems reduce su capacidad para capturar completamente la naturaleza de las variables en estudio.

aparecen medidos con escalas de tipo Likert⁴ con valores de 1 a 5 (desde “en total desacuerdo” hasta “totalmente de acuerdo” con la proposición formulada). El cuestionario inicial fue depurado mediante la realización de un pretest a 15 individuos. Las variables del cuestionario referentes al compromiso aparecen codificadas en el ANEXO junto con la escala de referencia, en su caso.

Cuadro 3: descripción de la muestra

Sexo	Porcentaje	Renta	Porcentaje
<i>Hombres</i>	50.9	<i>Menos de 450.76 euros</i>	3.9
<i>Mujeres</i>	49.1	<i>450.76-601.01 euros</i>	2.6
		<i>601.92-901.52 euros</i>	7.1
		<i>901.53-1202.02 euros</i>	14.7
		<i>1202.03-1803.04 euros</i>	29.7
		<i>1803.05-240405 euros</i>	18.7
		<i>2404.06-3005.06 euros</i>	12.3
		<i>Más de 3005.06 euros</i>	11.0
Agencia de viajes*	Porcentaje	Renta	Porcentaje
Halcón	25.7	<i>Menos de 450.76 euros</i>	3.9
Barceló	12.3	<i>450.76-601.01 euros</i>	2.6
El Corte Inglés	7.6	<i>601.92-901.52 euros</i>	7.1
ODA	7.0	<i>901.53-1202.02 euros</i>	14.7
Indra	5.8	<i>1202.03-1803.04 euros</i>	29.7
Solinieve	5.3	<i>1803.05-240405 euros</i>	18.7
Marsans	4.7	<i>2404.06-3005.06 euros</i>	12.3
Todomundo	4.1	<i>Más de 3005.06 euros</i>	11.0
Oriente	3.5		
Ecuador	2.9		

Fases de análisis

En este apartado, implementamos las fases expuestas en el epígrafe 2 para el caso práctico que nos ocupa y que se resumen en la figura 4.

Fase I: Análisis univariante y bivariante

En nuestro ejemplo, no detectamos ningún dato extraño. Si analizamos los valores de los coeficientes de curtosis y asimetría (o apuntamiento) para ver si las variables siguen una distribución normal, observamos que, en general, se encuentran comprendidos en el intervalo que se considera aceptable para hablar de variables distribuidas nor-

⁴ Pese a que las escalas Likert son escalas ordinales y no cumplen la característica de proporcionalidad entre las distancias (debido a que no existe un origen que indique ausencia de característica), fueron consideradas, como es frecuente en el área de marketing, escalas de intervalos a la hora de analizar los resultados. Ello permite aplicarles el tratamiento estadístico propio de las variables cuantitativas. Grande y Abascal (1994) se pronuncian en este sentido y sostienen que estamos ante un tipo de escala que admite tanto un análisis de componentes principales como un análisis factorial de correspondencias múltiples. Asimismo, Bentler y Chou (1987) indican que se pueden utilizar métodos continuos cuando una variable tiene cuatro o más niveles.

malmente. En esta fase, procedimos también al cálculo de la matriz de correlaciones entre las variables (tabla 1).

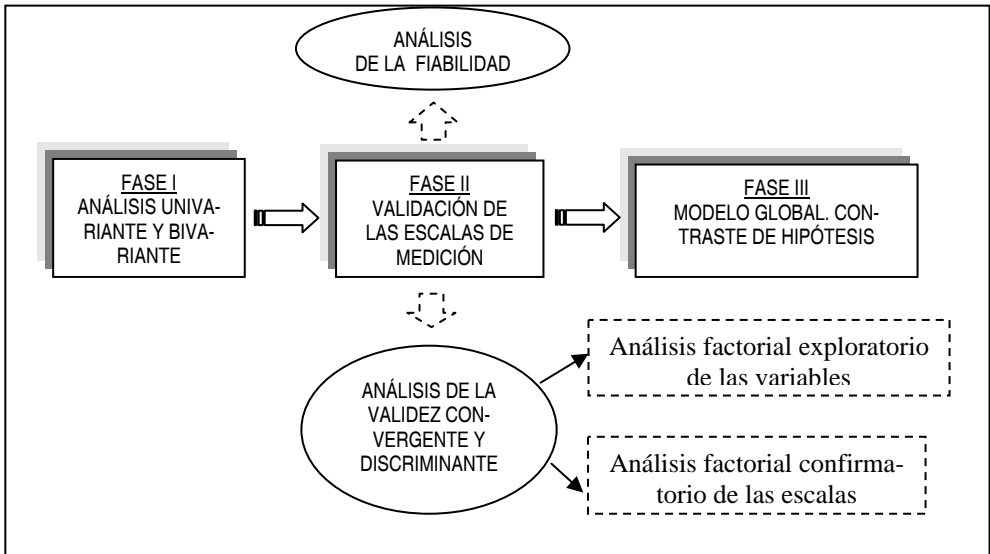


Figura 4: fases metodológicas para el tratamiento de la información.

Tabla 1: matriz de correlaciones entre las variables que miden el compromiso.

	V1	V2	V3	V4	V5
V1 (% de veces)	1,000	0,937	0,388	0,386	0,138
V2 (% de gasto)	0,937	1,000	0,400	0,380	0,111
V3 (frecuencia)	0,388	0,400	1,000	0,491	0,385
V4 (sentido lealtad)	0,386	0,380	0,491	1,000	0,436
V5 (valor positivo)	0,138	0,111	0,385	0,436	1,000
V6 (recomendación)	0,190	0,224	0,427	0,442	0,571
V7 (cordialidad)	0,335	0,317	0,422	0,490	0,491
V8 (afecto)	0,304	0,285	0,452	0,589	0,491
V9 (intenc. continuar)	0,392	0,337	0,408	0,638	0,529
V10 (deseo continúe)	0,329	0,273	0,400	0,569	0,474
V11 (deseo intensa)	0,269	0,249	0,364	0,523	0,379

Tabla 1 (continuación)

	V6	V7	V8	V9	V10	V11
V1	0,190	0,335	0,304	0,392	0,329	0,269
V2	0,224	0,317	0,285	0,337	0,273	0,249
V3	0,427	0,422	0,452	0,408	0,400	0,364
V4	0,442	0,490	0,589	0,638	0,569	0,523
V5	0,571	0,491	0,491	0,529	0,474	0,379
V6	1,000	0,454	0,586	0,483	0,436	0,475
V7	0,454	1,000	0,610	0,634	0,567	0,555
V8	0,586	0,610	1,000	0,639	0,701	0,691
V9	0,483	0,634	0,639	1,000	0,753	0,658
V10	0,436	0,567	0,701	0,753	1,000	0,804
V11	0,475	0,555	0,691	0,658	0,804	1,000

Fase II: Validación de las escalas de medición

→ Coeficiente alpha de Cronbach

Como primer paso para el análisis de la fiabilidad de las escalas de medición propuestas para cada constructo se calculó el correspondiente coeficiente *alpha de Cronbach*. Un valor mínimo bastante aceptado para poder decir que la escala multi-ítem en cuestión es fiable y goza de consistencia interna es 0,70, aunque depende del tipo de investigación (Sáez y Sánchez, 1997)⁵. Para las escalas de nuestro caso práctico, los valores que arroja dicho coeficiente son de 0.9673, 0.8508 y 0.8945 para el compromiso comportamental, el compromiso afectivo y el compromiso temporal respectivamente. Asimismo, cabe señalar que los valores de las correlaciones ítem-total corregidas también superan el valor aceptable de 0.50 (compromiso comportamental: 0.9366 para V1 y V2; compromiso afectivo: 0.5581 para V3, 0.6408 para V4, 0.6148 para V5, 0.6426 para V6, 0.6435 para V7 y 0.7278 para V8; compromiso temporal: 0.7433 para V9, 0.8550 para V10 y 0.7810 para V11). Por tanto, a juzgar por estos valores, las tres escalas son fiables.

→ Análisis factorial exploratorio de las variables (AFE)

Con el objetivo de obtener las mejores escalas posibles mediante un proceso de depuración, se realizó, en segundo lugar, un AFE para el compromiso (en concreto, realizamos un análisis factorial de ejes principales), cuyos resultados aparecen en la tabla 2. Con este análisis, además, conseguimos una primera prueba de que los constructos cum-

⁵ Sáez y Sánchez (1997) apuntan cuáles son los valores aceptables del coeficiente *alpha* de Cronbach en función del tipo de investigación en cuestión y según la opinión de diferentes autores.

plen el requisito de *unidimensionalidad* (Steenkamp y van Trijp, 1991; Bagozzi y Baumgartner, 1994). En este caso, el compromiso es un constructo multidimensional, por lo que la unidimensionalidad se comprobó para cada dimensión del compromiso. Como podemos observar, el análisis agrupa los indicadores del compromiso en tres factores que coinciden con las tres dimensiones propuestas: dimensión de comportamiento, afectiva y temporal. No obstante, el factor de mayor peso es el temporal seguido del afectivo, mientras que el compromiso comportamental explica únicamente un 8% del compromiso.

Tabla 2: resultados del análisis factorial exploratorio para el compromiso.

Factor	VARIABLES QUE RECOGE EL FACTOR		Peso de cada variable
<i>F1</i> (Compromiso temporal)	V9	Intención futura de continuación	0.659
	V10	Deseo de continuidad relación	0.868
	V11	Deseo de intensidad relación	0.811
<i>F2</i> (Compromiso afectivo)	V3	Frecuencia de la relación	0.480
	V4	Sentido de lealtad	0.474
	V5	Valoración positiva relación	0.694
	V6	Recomendación de X	0.690
	V7	Grado de cordialidad	0.493
<i>F3</i> (Compromiso comportamental)	V8	Sentimiento de afecto	0.526
	V1	% de veces que acude a X/total	0.932
	V2	% de gasto en X/total	0.960

Factor	VARIABLES	Porcentaje de información explicada	Porcentaje de explicación acumulado
<i>F1</i> (Compromiso temporal)	V9	52.044	52.044
	V10		
	V11		
<i>F2</i> (Compromiso afectivo)	V3	14.655	66.699
	V4		
	V5		
	V6		
	V7		
<i>F3</i> (Compromiso comportamental)	V8	8.016	74.715
	V1		
	V2		

→ Análisis factorial confirmatorio (AFC)

En este apartado, describimos los resultados alcanzados tras efectuar el AFC para los distintos factores obtenidos en el AFE del compromiso⁶. Así, hemos planteado un análisis factorial confirmatorio para el conjunto de las tres dimensiones del compromiso con el fin de obtener modelos congenéricos⁷ y, al mismo tiempo, probar la validez discriminante de las dimensiones estudiadas (tabla 3). Este modelo resultante para el constructo compromiso no ajusta bien, a juzgar por los valores de la Chi-cuadrado ($p < 0,05$), por lo que ha sido necesario volver a especificarle mediante la eliminación de las variables V12, V13, V15 y V18 al encontrar correlaciones entre sus errores de medida y por las sugerencias de los índices de modificación. Finalmente, la salida para el modelo de medida del compromiso reespecificado y ajustado se recoge en la tabla 4. Como vemos, el modelo ajusta bien a juzgar por los valores de las medidas de bondad del ajuste, porque la probabilidad correspondiente a la Chi-Cuadrado es $> 0,05$ y los índices GFI, AGFI y CFI son claramente superiores a 0.9.

No obstante lo anterior y como se ha mencionado en el apartado 2.5, el hecho de que algunas veces la Chi-cuadrado nos indique que el ajuste no es correcto puede ser debido a su dependencia y sensibilidad respecto al tamaño de la muestra, de forma que el rechazo de un modelo conforme aumenta el tamaño muestral no es extraño puesto que un mayor tamaño facilita la detección de discrepancias entre la matriz observada y la implícita (Bagozzi y Yi, 1988; Bentler, 1990). Siguiendo a Bagozzi y Yi (1988), si la Chi-cuadrado y otras medidas de bondad del ajuste (incrementales y/o el AGFI) indican un modelo insatisfactorio, no se debe apoyar ese modelo independientemente del resto de medidas; si la Chi-cuadrado o/y otras medidas de bondad del ajuste (incrementales o el AGFI) indican que el modelo es satisfactorio, se deberían verificar otros criterios internos, globales y de validez y, en su caso, considerar que los criterios tienden a apoyar el modelo. Por ello, siempre que los indicadores GFI, AGFI y CFI arrojen valores adecuados (al menos de 0.9), se puede aceptar el modelo de medida y la validez convergente de la escala.

La matriz de correlaciones entre las dimensiones del compromiso (tabla 5) nos indica que las tres dimensiones están correlacionadas, resultado que no nos sorprende en absoluto y que se corresponde con las hipótesis propuestas. No obstante, para comprobar la discriminación entre los tres constructos, se estimó un modelo confirmatorio en el que se restringieron a 1 dichas correlaciones y el resultado, en cuanto a las medidas de bon-

⁶ Para la realización de los análisis confirmatorios de este apartado, hemos fijado la unidad de medida de la variable latente (la hemos fijado en 1) respecto a las variables observadas, para que aquella esté medida siempre en la misma escala, como se comentó previamente. Además, los coeficientes *lambda* que proporcionamos están completamente estandarizados, es decir, son los que resultan de utilizar variables latentes y observadas ya estandarizadas.

⁷ Cuando realizamos el análisis confirmatorio para una única dimensión del compromiso, que viene medida a través de menos de 4 ítems (variables de medida), puede que no existan suficientes grados de libertad para estimar el modelo (lo cual lleva a que el modelo no sea identificado o sea perfectamente identificado). Una solución es trabajar con modelos congenéricos de dos o más factores cuando hay al menos 2 medidas para cada factor (Bagozzi, 1994; Jöreskog, 1971).

dad del ajuste, fue a todas luces peor que en el primer modelo. Disponemos, así, de alguna garantía respecto de la validez discriminante de las dimensiones del compromiso. Es interesante notar que la correlación entre el compromiso afectivo y el temporal, aunque no llega a ser 1, es alta (en torno a 0.9), por lo que cabe esperar que el componente afectivo refleje la mayor parte del efecto determinante de la dimensión temporal del compromiso.

Tabla 3: *resultados del análisis factorial confirmatorio para el constructo compromiso.*

Variable latente	Variables de medida		Coeficientes Lambda	
			Lambda	t
Compromiso comportamental	V1	% de veces que acude a X	0.996	--
	V2	% de gasto en X	0.941	16.184
	V3	Frecuencia relación	0.638	--
	V4	Sentido de lealtad	0.735	11.748
Compromiso afectivo	V5	Valorac.posit. relación	0.642	9.567
	V6	Recomendación de X	0.675	10.300
	V7	Grado de cordialidad	0.743	11.984
	V8	Sentimiento de afecto	0.853	15.458
Compromiso temporal	V9	Intención fut.continuar	0.844	--
	V10	Deseo continuar relac.	0.918	19.387
	V11	Deseo intensidad relac	0.859	16.560

Variable latente	Varianza de los errores de medida		R ²	Bondad del ajuste
Compromiso comportamental	V1	0.009	0.991	$\chi^2 = 118.061$ (p=0.00) GFI=0.887 AGFI=0.830 CFI=0.943 RMR=0.0793
	V2	0.115	0.885	
Compromiso afectivo	V3	0.593	0.407	
	V4	0.460	0.540	
	V5	0.588	0.412	
	V6	0.544	0.456	
Compromiso temporal	V7	0.447	0.553	
	V8	0.272	0.728	
	V9	0.288	0.712	
	V10	0.156	0.844	
	V11	0.262	0.738	

Como conclusión de este análisis global del compromiso, recogemos en el cuadro 4 el listado de ítems que definitivamente forman las escalas de medida de cada una de las dimensiones del compromiso.

Una vez que hemos obtenido un modelo de medida satisfactorio (fase II), el siguiente paso consistiría en buscar que también lo sea el modelo estructural o latente (fase III) (ver en San Martín, 2003).

Tabla 4: *resultados del análisis factorial confirmatorio ajustado para el constructo compromiso.*

Variable latente	Variables de medida	Coeficientes Lambda	
		Lambda	t
Compromiso comportamental	V1 % de veces que acude a X/total	0.975	--
	V2 % de gasto en X/total	0.962	35.257
	V3 Frecuencia relación	0.623	--
Compromiso afectivo	V4 Sentido de lealtad	0.758	12.085
	V5 Valorac.posit. relación		
	V6 Recomendación de X		
	V7 Grado de cordialidad	0.721	11.179
Compromiso temporal	V8 Sentimiento de afecto		
	V9 Intención fut.continuar	0.912	--
	V10 Deseo continuar relac.	0.830	14.383
	V11 Deseo intensidad relac		

Variable latente	Varianza de los errores de medida	R ²	Bondad del ajuste
Compromiso comportamental	V1	0.049	0.951
	V2	0.075	0.925
Compromiso afectivo	V3	0.612	0.388
	V4	0.426	0.574
	V5	Eliminada	
	V6	Eliminada	
	V7	0.480	0.520
Compromiso temporal	V8	Eliminada	
	V9	0.168	0.832
	V10	0.311	0.689
	V11	Eliminada	

$\chi^2 = 24.238$
 (p=0.0612)
 GFI=0.961
 AGFI=0.927
 CFI=0.988
 RMR=0.049

Tabla 5: *matriz de correlaciones entre las dimensiones del compromiso.*

	<i>Dimensión com- portamental</i>	<i>Dimensión afectiva</i>	<i>Dimensión temporal</i>
<i>Dimensión comportamental</i>	1		
<i>Dimensión afectiva</i>	0.55	1	
<i>Dimensión temporal</i>	0.42	0.91	1

Cuadro 4: *escala de items resultante para cada dimensión del compromiso.***COMPROMISO COMPORTAMENTAL**

- ◆ Porcentaje de veces que acudió a la agencia de viajes X el año pasado respecto al total (V1)
- ◆ Porcentaje de gasto en la agencia de viajes X el año pasado respecto al total (V2)

COMPROMISO AFECTIVO

- ◆ Intensidad de la relación con la agencia de viajes X (V3)
- ◆ Sentido de lealtad a la agencia de viajes X (V4)
- ◆ Grado de cordialidad en la relación con la agencia de viajes X (V5)

COMPROMISO TEMPORAL

- ◆ Intención futura de continuación de la relación con la agencia de viajes X (V9)
- ◆ Deseo de continuidad de la relación (V10)

Reflexión final

La clave en la obtención de unas buenas escalas y resultados en la validación de cualquier modelo radica en la existencia de una teoría bien fundamentada previamente al análisis empírico del modelo. La teoría debe ser la base sólida del desarrollo de un modelo y ayudar en la interpretación de los resultados. Por ejemplo, un valor de un coeficiente de fiabilidad o de validez inadecuado en un determinado estudio puede ser aceptable en otro a la luz del marco teórico que justifica dicho modelo.

Entre todas las técnicas que existen en la actualidad para la validación de modelos de ciencias sociales con datos cuantitativos, numerosos autores se decantan por la aplicación del enfoque de ecuaciones estructurales, dadas las ventajas que presenta sobre

otros (Bagozzi, 1994; Diamantopoulos, 1994; Lévy, 1994; Baumgartner y Homburg, 1996; Sáez y Sánchez, 1997; Hair y otros, 1999; Camisón y Bou, 2000; Mackenzie, 2001). Podemos resumir estas ventajas de la modelización mediante ecuaciones estructurales en que contemplan varias relaciones entre variables dependientes e independientes (frente a una sola relación en la correlación canónica o en el análisis multivariante de la varianza, por ejemplo), permiten controlar el error de medida, probar estructuras teóricas complejas, proporcionan medios para depurar las escalas de medida hasta lograr su fiabilidad y validez convergente y discriminante, son un buen método para tratar conceptos multidimensionales, valoran el grado de ajuste del modelo de medida y del modelo estructural propuestos a los datos empíricos obtenidos y, en definitiva, favorecen el desarrollo y prueba de la teoría. Y no debemos olvidar que según se avanza en las ciencias sociales, se observa una tendencia hacia el desarrollo de conceptos multidimensionales como parte del refinamiento de ideas anteriores de la investigación. De esta forma, pasamos de escalas globales para las variables a escalas para cada dimensión de la variable.

En definitiva, el enfoque de ecuaciones estructurales constituye una técnica de estimación robusta y versátil, adecuada para el tratamiento de un conjunto de ecuaciones simultáneas con regresiones múltiples (Hair y otros, 1999). Con esta guía, hemos presentado un caso en el que, con la ayuda del programa LISREL, se han logrado depurar las escalas de las dimensiones del compromiso que desarrolla el consumidor en su relación con la empresa. Así, de acuerdo con nuestra propuesta, el *compromiso relacional* del consumidor se revela como un concepto de naturaleza compleja y de estructura multidimensional, un constructo que engloba tres aspectos claramente diferenciados: un *comportamiento* observable (repetición de compras y fidelidad formal), un *afecto* (implicación amistosa o identificación con la empresa o vendedor) y una *intención de futuro* (deseo e intención de prolongar y estrechar aún más la relación). En consecuencia, podemos desglosar el compromiso en su dimensión comportamental, su dimensión afectiva y su dimensión temporal.

Referencias

- Anderson, E. y B. Weitz (1992): "The use of pledges to build and sustain commitment in distribution channels", *Journal of Marketing Research*, XXIX, 18-34.
- Bagozzi, R.P. (1994): Structural equations models in marketing research: basic principles, en *Principles of Marketing Research*, 317-385. Ed. Brasil Blackwell Ltd.
- Bagozzi, R.P. y Y.Yi (1988): "On the evaluation of structural equation models", *Journal of the Academy of Marketing Science*, 16 (1), 74-94.
- Bagozzi, R.P. y H. Baumgartner (1994): "The evaluation of structural equation models and hypothesis testing", en *Principles of Marketing Research*, cap. 10, 386-419. Ed. Brasil Blackwell Ltd.
- Baumgartner, H. y C. Homburg (1996): "Applications of structural equation modeling in marketing and consumer research: a review", *International Journal of Research in Marketing*, 13,139-161.
- Bentler, P.M. (1990): "Comparative fit indexes in structural models", *Psychological Bulletin*, 107 (2), 238-246.

- Bentler, P.M. y C.P. Chou (1987): "Practical issues in structural modeling", *Sociological Methods and Research*, 16, 78-117.
- Bisquerra, R. (1989): Introducción conceptual al análisis multivariante: un enfoque informático con los paquetes SPSS X, BMDP, LISREL y SPAD, cap. 15. Promociones y Publicaciones Universitarias, I y II.
- Bollen, K.A. (1989): *Structural equations with latent variables*. Nueva York: John Wiley and Sons.
- Camarero, M.C. y J. Gutiérrez (1999): "Dimensiones estructurales y procesales de las relaciones entre empresas. Un estudio empírico", en la obra colectiva *XI Encuentro de Profesores Universitarios de Marketing*, ed. ESIC, Madrid, 195-210.
- Camisón, C. y J.C. Bou (2000): "Calidad percibida de la empresa: desarrollo y validación de un instrumento de medida", *Revista Europea de Dirección y Economía de la Empresa*, 9 (1), 9-24.
- Churchill, G.A. Jr. (1979): "A paradigm for developing better measures of marketing constructs", *Journal of Marketing Research*, XVI, febrero, 64-73.
- Diamantopoulos, A. (1994): "Modeling with LISREL: a guide for the uninitiated", *Journal of Marketing Management*, 10, 105-136.
- Doney, P.M. y J.P. Cannon (1997): "An examination of the nature of trust in buyer-seller relationships", *Journal of Marketing*, 61, abril, 35-51.
- Dwyer, F.R., P.H. Schurr y S. Oh (1987): "Developing buyer-seller relationships", *Journal of Marketing*, 51, abril, 11-27.
- Ganesan, S. (1994): "Determinants of long-term orientation in buyer-seller relationships", *Journal of Marketing*, 58, abril, 1-19.
- Garbarino, E. y M.S. Johnson (1999): "The different roles of satisfaction, trust, and commitment in customer relationships", *Journal of Marketing*, 63, abril, 70-87.
- Gerbing, D.W. y J.C. Anderson (1985): "The effects of sampling error and model characteristics on parameter estimation for maximum likelihood confirmatory factor analysis", *Multivariate Behavioral Research*, 20, 255-271.
- Geyskens, I., J-B.E.M. Steenkamp, L.K. Scheer y N. Kumar (1996): "The effects of trust and interdependence on relationship commitment: a trans-atlantic study", *International Journal of Research in Marketing*, 13, 303-317.
- Gordon, M.E., K. McKeage y M.A. Fox (1998): "Relationship marketing effectiveness: the role of involvement", *Psychology and Marketing*, 15 (5), agosto, 443-459.
- Grande, I y E. Abascal (1994): *Fundamentos y técnicas de investigación comercial*. Ed. ESIC, Madrid.
- Grayson, K. y T. Ambler (1999): "The dark side of long-term relationships in marketing services" *Journal of Marketing Research*, XXXVI, febrero, 132-141.
- Gundlach, G.T., R. Achrol y J. Mentzer (1995): "The structure of commitment in exchange", *Journal of Marketing*, 59, enero, 78-92.
- Hair, J.F.; R.E. Anderson, R.L. Tatham y W.C. Black (1999): *Análisis multivariante de datos*. Ed. Prentice Hall, Madrid.
- Iacobucci, D. y A. Ostrom (1996): "Commercial and interpersonal relationships; using the structure of interpersonal relationships to understand individual-to-individual, individual-to-firm, and firm-to-firm relationships in commerce", *International Journal of Research in Marketing*, 13, 53-72.

- Jaros, S.J.; J.M. Jermier; J.W. Koehler y T. Sincich (1993): "Effects of continuance, affective, and moral commitment on the withdrawal process: an evaluation of eight structural equation models", *Academy of Management Journal*, 36 (5), 951-995.
- Jöreskog, K.G. (1971): "Statistical analysis of sets of congeneric tests", *Psychometrika*, 36 (2), junio, 109-133.
- Jöreskog, K.G. y D. Sörbom (1993): *LISREL VIII: Structural equation modeling with the SIMPLIS command language*. Scientific Software International, INC. Chicago.
- Kim, K. y G.L. Frazier (1997): "On distributor commitment in industrial channels of distribution: a multicomponent approach", *Psychology and Marketing*, 14, 847-877.
- Kinney, J.C. y J.R. Taylor (1995): *Investigación de mercados: un enfoque aplicado*. Ed. McGraw Hill. Madrid.
- Lévy, J-P (1994): "El análisis factorial confirmatorio y las estructuras de covarianza, estudio diferenciado con el análisis factorial exploratorio", *ESIC MARKET*, octubre-diciembre, 121-130.
- Macintosh, G. y L.S. Lockshin (1997): "Retail relationships and store loyalty: a multi-level perspective", *International Journal of Research in Marketing*, 14, 487-497.
- Mackenzie, S.B. (2001): "Opportunities for improving consumer research through latent variable structural equation modeling", *Journal of Consumer Research*, 28, junio, 159-166.
- Meyer, J.P. y J.N. Allen (1991): "A three-component conceptualization of organizational commitment", *Human Resource Management Review*, 1 (1), 61-98.
- Morgan, R. y S. Hunt (1994): "The commitment-trust theory of relationship marketing", *Journal of Marketing*, 58, julio, 20-38.
- Nunnally, J.C. (1978): *Psychometric theory*, 2ª ed. New York: McGraw-Hill.
- Nunnally, J.C. y I.H. Bernstein (1994): *Psychometric theory*. 3ª ed. McGraw Hill, New York.
- Peter, J.P. (1979): "Reliability: a review of psychometrics basics and recent marketing practices", *Journal of Marketing Research*, XVI, febrero, 6-17.
- Peter, J.P. y G.A. Churchill (1986): "Relationships among research design choices and psychometric properties of rating scales: a meta-analysis", *Journal of Marketing Research*, XXIII (febrero), 1-10.
- Pritchard, M.P., M.E. Havitz y D.R. Howard (1999): "Analyzing the commitment-loyalty link in service contexts", *Journal of the Academy of Marketing Science*, 27 (3), 333-348.
- Ross, W.T. Jr., E. Anderson y B. Weitz (1997): "Performance in principal-agent dyads: the causes and consequences of perceived asymmetry of commitment to the relationship", *Management Science*, 43 (5), mayo, 680-704.
- Ruyter, K. de y M. Wetzels (1999): "Commitment in auditor-client relationships: antecedents and consequences", *Accounting, Organizations and Society*, 24, 57-75.
- Rylander, D., D. Strutton y L.E. Pelton (1997): "Toward a synthesized framework of relational commitment: implications for marketing channel theory and practice", *Journal of Marketing Theory and Practice*, primavera, 58-71.
- Sáez, E. y M. Sánchez (1997): "Análisis de los métodos de evaluación de la fiabilidad en marketing", en la obra colectiva *XI Encuentro de Profesores Universitarios de Marketing*, ed. ESIC, Madrid, 407-425.

- San Martín, S. (2003): *La relación del consumidor con las agencias de viajes*. Servicio de Publicaciones de la Universidad de Burgos.
- San Martín, S.; J. Gutiérrez y M.C. Izquierdo (2000): “La confianza como principal determinante del compromiso relacional”, *XI Encuentro de Profesores Universitarios de Marketing* a celebrar en septiembre en Santiago de Compostela.
- Shemwell, D.J., J.J. Cronin y W.R. Bullard (1994): “Relational exchange in services: an empirical investigation of ongoing customer-service provider relationships”, *International Journal of Service Industry Management*, 5 (3), 57-68.
- Sirohi, N., E.W. McLaughlin y D.R. Wittink (1998): “A model of consumer perceptions and store loyalty intentions for a supermarket retailer”, *Journal of Retailing*, 74 (2), 223-245.
- Steenkamp, J-B.E.M. y H.C.M. van Trijp (1991): “The use of LISREL in validating marketing constructs”, *International Journal of Research in Marketing*, 8, 283-299.
- Visauta, B. (1997): *Análisis estadístico con SPSS para Windows*. Ed. McGraw Hill, Madrid.
- Wetzels, M., K. Ruyter de y M.V. Birgelen (1998): “Marketing service relationships: the role of commitment”, *Journal of Business and Industrial Marketing*, 13 (4/5), 406-423.
- Wulf, K. De, G. Odekerken-Schröder y D. Iacobucci (2001): “Investments in consumer relationships: a cross-country and cross-industry exploration”, *Journal of Marketing*, 65, octubre, 33-50.

Anexo

*Codificación, descripción y referencia (si procede)
de las variables e ítems de la encuesta realizada*

Variable y su dimensión (si procede)	Codificación del ítem	Descripción del ítem
Compromiso comportamental	V1	CC1. Porcentaje de veces que acude a la agencia de viajes X respecto al total
	V2	CC2. Porcentaje de gasto en la agencia de viajes X el año pasado respecto al total
Compromiso afectivo o actitudinal	V3	CA1. Frecuencia de la relación con la agencia de viajes X
	V4	CA2. Sentido de lealtad a la agencia de viajes X
	V5	CA3. Valoración positiva de la relación con la agencia de viajes X
	V6	CA4. Recomendación de la agencia de viajes X
	V7	CA5. Grado de cordialidad en la relación con la agencia de viajes X
	V8	CA6. Sentimiento de afecto por la agencia de viajes X
Compromiso temporal	V9	CT1. Intención futura de continuación con la agencia de viajes X
	V10	CT2. Deseo de continuidad de la relación con la agencia de viajes X
	V11	CT3. Deseo de intensidad de la relación con la agencia de viajes en el futuro

Variable y su dimensión (si procede)	Codificación del ítem	Trabajos con escalas de referencia (adaptadas para el presente trabajo o que han sugerido algún ítem del mismo)
Compromiso comportamental	V1	Wulf y otros (2001)
	V2	Ganesan (1994); Macintosh y Lockshin (1997); Wulf y otros (2001)
Compromiso afectivo o actitudinal	V3	Iacobucci y Ostrom (1996)
	V4	Anderson y Weitz (1992); Wulf y otros (2001); Ross y otros (1997)
	V5	Elaboración propia
	V6	Sirohi, McLaughlin y Wittink (1998)
	V7	Shemwell y otros (1994); Doney y Cannon (1997); Gordon y otros (1998)
	V8	Elaboración propia
Compromiso temporal	V9	Kim y Frazier (1997); Shemwell y otros (1994)
	V10	Ruyter y Wetzels (1999); Ross y otros (1997); Wetzels y otros (1998)
	V11	Elaboración propia