

"Probamos por medio de la lógica,
pero descubrimos por medio de la intuición"

Henri Poincaré (1854-1912); matemático francés

Capítulo 05

Metodología utilizada para la realización del estudio de campo de la tesis doctoral

5.1. Introducción

El presente capítulo intenta exponer la metodología utilizada para evaluar y analizar el modelo presentado en el capítulo anterior. Para ello se ha dividido en cuatro apartados.

En el primer apartado se introduce el uso, cada vez más frecuente, de los métodos cuantitativos dentro de las ciencias sociales mediante un pequeño preámbulo de la investigación científica en las ciencias.

El siguiente apartado presenta la metodología cuantitativa utilizada en la presente tesis: las ecuaciones estructurales. Con este fin se introduce los conceptos básicos y las convenciones seguidas en secciones y capítulos posteriores para ayudar a la comprensión de la metodología. Además se plantea los pasos a seguir para un análisis multivariante con ecuaciones estructurales.

El modelo estructural y los modelos de medida de las variables exógenas y endógenas del objeto a investigar se desarrollan en el tercer apartado del presente capítulo, dando lugar al modelo de ecuaciones estructurales a analizar. Dicho modelo se expone a través de diversas representaciones como en un diagrama de caminos, a través ecuaciones y en forma matricial.

Una vez el modelo de ecuaciones estructurales ha sido definido, se procede a elaborar el cuestionario para la obtención de información cuantitativa para el análisis del modelo anterior. Un primer paso es la elección de tipo de entrevista a realizar, para continuar con la elaboración de cada uno de los elementos del cuestionario sobre las estructuras organizativas y sobre el constructo capacidad de absorción.

El último apartado expone los criterios seguidos para selección de la muestra utilizada en la investigación, así como el tamaño mínimo necesario para obtener una significatividad razonable.

5.2. Métodos cuantitativos en las ciencias sociales

El término investigación que, en general, significa indagar o buscar, cuando se aplica a las ciencias sociales, toma la connotación específica de crear conocimientos sobre la realidad social, es decir, sobre su estructura, las relaciones entre sus componentes, su funcionamiento, los cambios que experimenta el sistema en su totalidad o en esos componentes. Los conocimientos generados por una investigación en particular se unen a otros conocimientos ya existentes, acumulados durante mucho tiempo por otros investigadores, sea en la forma de un aporte original o como confirmación o refutación de hallazgos ya existentes. Cualquiera que sea la situación que se enfrente, la investigación es siempre la búsqueda de la solución a algún problema de conocimiento. Esa solución constituye un nuevo conocimiento que se mantiene mientras no haya otras propuestas mejor fundamentadas de acuerdo con criterios teóricos y metodológicos y sean aceptadas por la comunidad de científicos pertenecientes a la misma área de indagación.

La investigación social, como práctica científica que es, implica un camino largo y de mayor o menor dificultad según los problemas que se pretenden resolver. En todo caso, el punto de partida de tal camino comienza con la socialización del investigador en los primeros antecedentes de un determinado paradigma que luego se explicita en una o más teorías centrales, como también múltiples en un sistema metodológico de reglas y de técnicas de investigación que la persona que ha elegido el camino de la investigación debe conocer y saber en que momentos aplicarlas.

La investigación social cuantitativa está directamente basada en el paradigma explicativo. Este paradigma utiliza preferentemente información cuantitativa o cuantificable para describir o tratar de explicar los fenómenos que estudia, en las formas que es posible hacerlo en el nivel de estructuración lógica en el cual se encuentran las ciencias sociales actuales.

La reconstrucción de la práctica científica, que se hace habitualmente en los textos de metodología de la investigación, presenta ese proceso en una forma un tanto idealizada,

debido a que – por su naturaleza puede ser bastante complejo – como consisten en una serie de etapas y tareas que se darían en forma lineal, unas tras otras. Si bien ésta será la forma didáctica que utilizaremos en este texto, adhiriendo a Kaplan plenamente, reconocemos que “una lógica reconstruida no es una descripción (exacta), sino más bien una idealización de la práctica científica. Ni siquiera el más grande de los científicos tiene un estilo cognitivo que sea entera y perfectamente lógico...” (Kaplan, 1964).

Reconocido el carácter didáctico de la lógica reconstruida, según la expresión del epistemólogo norteamericano citado, podemos distinguir tres momentos principales en el proceso de investigación: 1) uno de acercamiento inicial al tema y a la problemática existente en una cierta área de interés; 2) otro segundo momento, representado por la preparación del proyecto de investigación; y 3) un tercer momento, en el cual se realiza o ejecuta la investigación, es decir, se pone en práctica el proyecto de investigación.

Como se ha recordado, toda investigación se propone crear conocimiento sobre un cierto aspecto objeto de la realidad social. Esa intención no es posible de realizar a partir de la “nada”. Muy por el contrario, el investigador debe conocer y tomar en cuenta la existencia de un conjunto de conocimientos ya acumulados, producto de un largo proceso de avances y retrocesos en el cual han intervenido otros investigadores dentro de una tradición científica en la cual se encuentran teorizaciones, propuestas metodológicas y resultados con mayores o menores niveles de confirmación.

Cuando un investigador decide hacer un nuevo estudio, no solo se basa en sus investigaciones anteriores, sino que debe conocer los trabajos de otros investigadores, lo cual lo obliga a estar al día en la literatura pertinente, sea para comprobar resultados presentados en ella o para proponerse otros problemas que se relacionan con su línea de trabajo y con la línea de trabajo de la comunidad científica con la cual se identifica.

El proyecto de investigación es el documento en el cual se sistematizan los conocimientos de la fase de acercamiento al tema y problemática de la investigación por realizar y se agregan otros menos o no tratados anteriormente. De manera convencional, el proyecto de investigación comprende los componentes y tareas que se indicarán en

cada uno de ellos: (1) Planteamiento del problema de investigación, (2) Marco conceptual del problema, (3) Objetivos de la investigación, (4) Finalidades, (5) Diseño metodológico, (6) Cronograma de trabajo.

Una vez que se ha terminado y revisado el proyecto de investigación, el investigador está en condiciones de iniciar la ejecución del proyecto. Esta ejecución implica cumplir las etapas y tareas allí especificadas como son la preparación del marco población con el cual se trabajará, la selección del tipo de muestra que se va a emplear y determinación de su tamaño, la elaboración de los instrumentos que se utilizarán en la recolección de información, la revisión de la información recogida, la codificación de la información recogida y revisada de acuerdo con un código previamente preparado, la preparación de un plan de análisis de acuerdo con los objetivos de la investigación, la preparación de la estructura del informe, análisis e interpretación de los resultados, y la redacción del informe definitivo.

5.3. Ecuaciones estructurales

En la actualidad los investigadores se ven obligados a recoger un gran número de medidas para poder captar de forma adecuada la complejidad de los fenómenos de las ciencias sociales. Para tratar con este conjunto de muestras se han implantado métodos multivariantes o multivariados, los cuáles permiten analizar simultáneamente conjuntos amplios de variables.

El principal impulso en la utilización de estos métodos y técnicas multivariantes ha sido el desarrollo de los ordenadores personales que han permitido a los investigadores tratar y analizar una gran cantidad de información simultáneamente. Su aparición ha colaborado al desarrollo de paquetes de software que permiten implementar con cierta facilidad los complejos y numerosos cálculos que demandan estas técnicas.

El número de métodos y técnicas multivariantes que se utilizan en las ciencias sociales es bastante amplio. Entre las más comunes destacan la regresión múltiple, el análisis factorial, el análisis multivariante de la varianza, y el análisis discriminante entre otras técnicas. Cada una de estas técnicas es una poderosa herramienta a la hora de tratar un amplio abanico de cuestiones prácticas y teóricas, sin embargo todas las mencionadas tienen una limitación en común: sólo pueden examinar una relación al mismo tiempo.

La técnica del modelo de ecuaciones estructurales se considera una extensión de varias técnicas multivariantes como la regresión múltiple, el análisis factorial principalmente y el análisis de senderos. El modelo de ecuaciones estructurales abraza una familia entera de modelos conocidos por muchos nombres, entre ellos análisis de la estructura de covarianza, análisis de variable latente, análisis factorial confirmatorio e incluso análisis LISREL⁶³.

⁶³ LISREL es una de las aplicaciones más conocidas para el desarrollo de modelos de ecuaciones estructurales, y desarrollado por Jöreskog y Sörbom.

5.3.1. Características de los modelos de ecuaciones estructurales

Los modelos de ecuaciones estructurales se caracterizan por dos elementos:

1. Estimación de relaciones de dependencias múltiples y cruzadas, y
2. 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 observable entre los sistemas de ecuaciones estructurales y las otras técnicas de análisis multivariante es la capacidad de analizar relaciones para cada subconjunto de variables. Además de permitir interrelaciones entre algunas variables de estos subconjuntos.

Los sistemas de ecuaciones estructurales permiten, además de estimar relaciones múltiples, introducir conceptos o variables latentes⁶⁴ al análisis, mientras que las que estas interrelaciones establecidas fortalecen las correlaciones utilizadas. De esta manera se puede realizar estimaciones más precisas de los coeficientes estructurales.

5.3.2. Estrategias de modelización

La flexibilidad de los sistemas de ecuaciones estructurales permite a los investigadores utilizar la técnica de múltiples formas y en función de los objetivos a conseguir. Sin embargo el investigador debe seleccionar al inicio de la investigación que estrategia⁶⁵ quiere seguir para alcanzar su propósito.

Existen tres estrategias que un investigador puede adoptar en la utilización de sistemas de ecuaciones estructurales:

1. La estrategia de modelización confirmatoria,
2. La estrategia de modelos rivales, y
3. La estrategia de desarrollo del modelo.

⁶⁴ Una variable latente es un concepto supuesto y no observado que sólo puede ser aproximado mediante variables medibles u observables.

⁶⁵ El término estrategia hace referencia a un plan de acción orientado a la consecución de los resultados.

La aplicación más directa de los modelos de ecuaciones estructurales es una estrategia de modelización confirmatoria, donde el investigador especifica un modelo aislado y el modelo de ecuaciones estructurales se utiliza para evaluar su significación estadística.

La estrategia de modelos rivales permite evaluar el modelo estimado con modelos alternativos. En muchas ocasiones el investigador se encuentra con modelos alternativos para representar una misma situación sin conocer cuál proporciona un nivel de ajuste mejor. Esta estrategia proporciona las líneas básicas para realizar esta comparación y seleccionar el más idóneo.

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 modificaciones de los modelos de medida y/o estructurales. El investigador que sigue esta estrategia intenta reespecificar un modelo básico de teoría para encontrar uno nuevo.

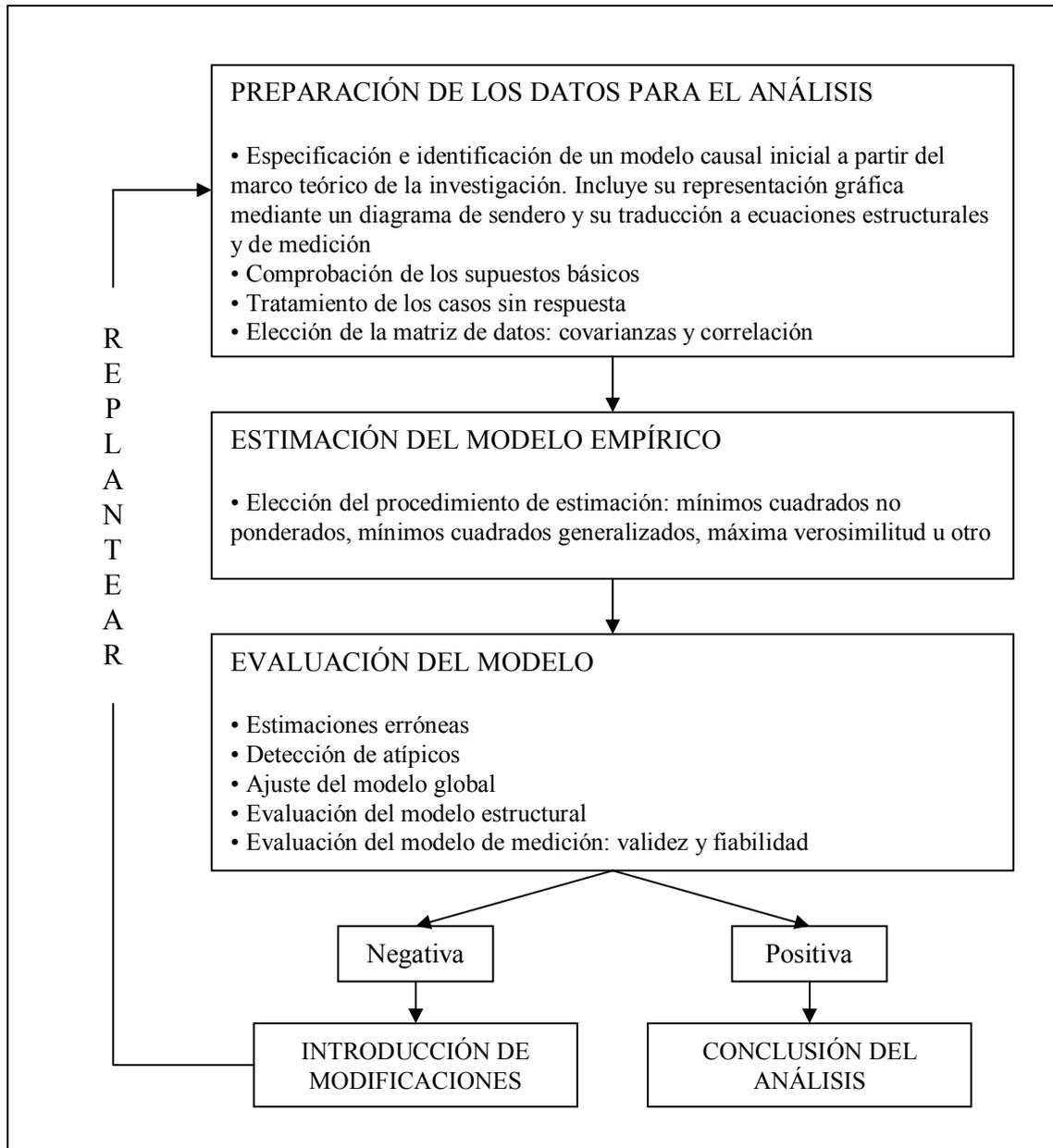
5.3.3. Fases de un modelo de ecuaciones estructurales

La Figura 5. 1 resume las distintas fases en la obtención de un modelo de ecuaciones estructurales. Comienza con la definición de un modelo teórico, el cuál se intentará contrastar estadísticamente, y finaliza con una verificación – parcial o global – del modelo.

La falta de confirmación de una o varias hipótesis establecidas en la elaboración del modelo de ecuación estructural conlleva a la modificación de este, manteniendo un significado lógico-sustantivo, además de una significatividad estadística. Toda modificación del modelo inicial lleva forzosamente a la repetición de todo el análisis estadístico.

El proceso de análisis acaba en el momento en que se consigue un modelo con un nivel de ajuste – global o parcial – igual o superior al establecido por el investigador.

Figura 5. 1. Fases principales en la obtención de un modelo de ecuaciones estructurales



La modelización de ecuaciones estructurales se fundamenta en relaciones causales, en las que el cambio en una variable se supone que produce un cambio en otra variable. Se encuentra este tipo de enunciado cuando se define 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 físicos, tales como una reacción química, a las relaciones menos definidas encontradas

en la investigación del comportamiento. La fuerza y la convicción con que el investigador asume la causación entre dos variables no descansan en los métodos analíticos escogidos sino en la justificación teórica ofrecida por los análisis.

El problema de error de especificación, caracterizado por la omisión de una o más variables predictivas claves, es el más crítico en el desarrollo de modelos de fundamentación teórica. La consecuencia de omitir una variable significativa consiste en sesgar la evaluación de la importancia de otras variables.

El modelado de ecuación estructural incluye dos submodelos o componentes:

1. El modelo estructural (o de variables latentes), y
2. El modelo de medición.

5.3.4. El modelo estructural

El modelo estructural es aquel componente del modelo general que describe relaciones causales entre variables latentes. Las relaciones entre las variables observadas se incluyen cuando estas variables no actúan como indicadores de las variables latentes.

Las ecuaciones estructurales lineales representan las relaciones estructurales representando la línea de causalidad entre las variables latentes endógenas. Hay tantas ecuaciones como constructos endógenos que sean explicados por otras variables exógenas (latentes u observadas).

La estructura de las ecuaciones estructurales lineales es la siguiente:

$$\eta = B\eta + \Gamma\xi + \zeta$$

Donde:

“ η ” (eta) es un vector “ $p \times 1$ ” de variables endógenas latentes (las predictoras o variables no explicadas por otras incluidas en el modelo).

“ ξ ” (xi) es un vector “ $q \times 1$ ” de variables exógenas latentes (las predictoras o variables no explicadas por otras incluidas en el modelo).

“ Γ ” (gamma) es una matriz “ $p \times q$ ” de coeficientes (γ_{ij}) que relacionan las variables latentes exógenas (ξ) con las endógenas (η) o variables a explicar. Indican que una unidad de cambio en la variable exógena ξ_i resulta en un cambio en η_j γ_{ij} unidades, manteniendo todas las otras variables constantes. Este coeficiente estructural se interpreta, al igual que β_{ij} como efectos directos en las variables endógenas.

“ β ” (beta) es una matriz “ $q \times q$ ” de coeficientes que relacionan las variables latentes endógenas entre sí. Cada β_{ij} indica una unidad de cambio en la variable endógena η_i , manteniendo todas las demás variables constantes. Para cada efecto hipotetizado de una variable latente endógena en otra, de las mismas características, se tendrá un coeficiente estructural β_{ij} .

“ ζ ” (zeta) es un vector “ $q \times 1$ ” de errores o términos de perturbación. Indican que las variables endógenas no son perfectamente predichas por las ecuaciones estructurales. Se supone que no existe correlación entre los errores y las variables exógenas.

5.3.5. Representación gráfica de los modelos de ecuaciones estructurales

La representación mediante un gráfico de sendero es muy usual en el modelado de ecuaciones estructurales, sin embargo existen convenciones para su realización. Según la notación LISREL⁶⁶, las variables observadas se representan con un cuadro. Estas variables pueden ser indicadores de variables latentes en el modelo de medición. En el modelo estructural pueden igualmente ser variables independientes (exógenas) o dependientes (endógenas), cuando éstas son observadas y no latentes. En este caso, se representarían en un cuadrado en lugar de un círculo.

Las variables latentes son representadas a través de un círculo o de una elipse, aunque los errores de predicción – en el modelo estructural – y de medición – en el modelo de medición – pueden no aparecer en un círculo.

⁶⁶ La notación LISREL es hasta la fecha la representación más ampliamente aceptada (Bollen, 1998: 8).

Las flechas rectas y unidirecciones indican relaciones causales entre las variables exógenas (ξ) y las variables endógenas (η). Por este motivo es imposible encontrar que una flecha termine en una variable exógena.

Las correlaciones son representadas a través de flechas bidireccionales y curvas, mientras que las líneas discontinuas indican senderos no significativos. A cada flecha de la representación se le asocia un coeficiente: $\lambda_{ij}^{(X)}$, $\lambda_{ij}^{(Y)}$, β_{ij} , γ_{ij} , ϕ_{ij} , ψ_{ij} , $\theta_{ij}^{(\delta)}$, $\theta_{ij}^{(\epsilon)}$.

La falta de flecha entre os variables significa que dichas variables no están directamente relacionadas, aunque indirectamente puedan estarlo.

5.3.6. El modelo de medición

Para especificar el modelo de medición del sistema de ecuaciones estructurales, se hace una 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.

Por lo tanto, el modelo de medición representa las relaciones de las variables latentes (o constructor) con sus indicadores (o variables empíricas). Para cada constructo que aparezca en el modelo es necesario determinar cuáles serán sus indicadores. Estas variables latentes son variables no observadas, que resultan de las covarianzas entre dos o más indicadores. Son las que representan a los conceptos en los modelos de medición.

El objetivo fundamental del modelo de medición es corroborar la idoneidad de los indicadores seleccionados en la medición de los constructos de interés. La estructura general de las relaciones entre las variables latentes y sus indicadores se describen de la siguiente forma:

$$X = \Lambda_x \xi + \delta$$

Donde:

“X” es un vector “p X 1” de variables observadas independientes o exógenas.

Aquellas que no son explicadas por otras variables incluidas en el modelo.

“p” es el número de indicadores de “ξ”.

“ξ” (xi) es un vector “q X 1” de variables exógenas latentes.

“Λ_X” (lambda) es una matriz “p X m” de pesos factoriales (o “factores de carga”:

“λ_i”) que relacionan las variables manifiestas “X” con las latentes “ξ”. “m”

expresa el número de variables latente “ξ”. Los coeficientes lambda

representan las influencias lineales de las variables latentes en los

indicadores.

“δ” (delta) es un vector “p X 1” de errores de medición. Convencionalmente,

cada indicador se representa siendo también influido por un término de error.

El modelo de medición asume que los errores están incorrelacionados entre

sí y con respecto a “ξ”.

La ecuación del modelo de medición mide la relación entre variables exógenas latentes y manifiestas. Las interrelaciones entre las variables latentes se indican mediante sus covarianzas. Las relaciones de las variables latentes con sus indicadores se definen como “direccionales”, de cada variable latente a su manifiesta, y no a la inversa.

Para las variables endógenas existe un modelo de medición, cuando el modelo incluye variables latentes dependientes (o endógenas) medidas mediante uno o varios indicadores (Y_i). en este caso, las relaciones entre las variables endógenas latentes y las manifiestas se describen con la ecuación:

$$Y = \Lambda_Y \eta + \varepsilon$$

Donde:

“Y” es un vector “q X 1” de variables observadas dependientes.

“Λ_Y” es una matriz “q X n” de coeficientes factoriales de Y en las variables

endógenas latentes “η”. “n” identifica el número de “η”. Esta matriz muestra

las relaciones de “η” a “Y”.

“ε” (épsilon) es un vector “q X 1” de errores de medición (o “factores únicos”)

de los indicadores endógenos “Y_i”.

El modelo de ecuaciones estructurales es el resultado de la combinación del modelo estructural y del modelo de medición de cuyo resultado aparece un modelo comprensivo de relaciones entre variables endógenas y exógenas, latentes y manifiestas.

5.3.7. Elección de la matriz de entrada

El modelado de ecuaciones estructurales se diseñó para analizar matrices de varianzas-covarianza, sin embargo también es posible utilizar matrices de correlación⁶⁷ como ocurre en otras técnicas multivariantes como en el análisis de regresión o en el análisis factorial.

La utilización de matrices de correlación favorece la comparabilidad de los coeficientes estructurales y colabora a la interpretación de los resultados, en especial si la intención es comprender el modelo de relaciones entre los constructos. En cambio, si el objetivo de la investigación es comparar muestras o poblaciones diferentes en lugar de comparar distintas variables de una misma muestra es más adecuado utilizar las matrices de varianzas-covarianzas.

El uso de matrices de correlación comportan tres posibles consecuencias negativas a la utilización de la técnica: (1) la modificación del modelo que se analiza, (2) la generación de valores de χ^2 y de otras medidas de bondad de ajustes incorrectos, y (3) la obtención de errores típicos incorrectos.

5.3.8. La estimación del modelo

Una vez los modelos estructurales y de medida han sido definidos y formulados, es necesario elegir cómo se estimará el modelo. En el caso de los modelos de ecuaciones estructurales hay varias técnicas y procesos a elegir.

⁶⁷ La matriz de correlaciones es la matriz de varianzas-covarianzas estandarizada.

Los primeros modelos de ecuaciones estructurales utilizaban la regresión de los mínimos cuadrados ordinarios (OLS) para su estimación. Sin embargo, rápidamente aparecieron nuevas estimaciones que la superaban como la de máximo verosímil (MLE), que es eficiente y no sesgada cuando se cumplen los supuestos de normalidad multivariante. La sensibilidad de la estimación de máximo verosímil a la normalidad condujo a nuevos métodos como los mínimos cuadrados ponderados (WLS), los mínimos cuadrados no ponderados (ULS), los mínimos cuadrados generalizados (GLS) y el asintóticamente libre de distribución (AGL).

Además de la técnica de estimación empleada, es necesario elegir entre varios procesos de estimación. Estos procesos van desde la estimación directa del modelo, que es similar a la utilizada en la mayoría de técnicas multivariantes, a métodos que generan miles de estimaciones del modelo para las cuales se obtienen resultados finales del modelo. Algunos de estos procesos son: la estimación directa, el bootstrapping, la simulación y el análisis de jackknife.

La estimación directa es el proceso de estimación más habitual, en el cual se estima directamente un modelo con un procedimiento de estimación elegido. Los otros tres procesos de estimación – el bootstrapping, la simulación y el análisis de jackknife – no descansan en un único modelo de estimación sino en estimaciones del parámetro calculado y sus intervalos de confianza basados en estimaciones múltiples.

5.3.9. La evaluación del modelo

Esta sección trata de mostrar distintas técnicas de analizar el ajuste global del modelo de ecuaciones estructurales a través de índices. Para facilitar la comprobación de la equivalencia entre las matrices de varianzas-covarianzas observada y predicha, o el ajuste global, existe una amplia variedad de estadísticos.

Existen tres grupos genéricos de índices globales: los índices de ajuste absoluto, los índices de ajuste incremental y los índices de ajuste parsimonia.

Los índices de ajuste absoluto miden el ajuste global del modelo, considerando las diferencias entre la matriz de varianzas y co-varianzas observada y predicha (matriz residual). De los índices incluidos en este grupo genérico, el de mayor aplicación es el índice de razón de verosimilitud χ^2 , el único que proporciona una prueba de significatividad estadística.

Otros índices de ajuste global son el de bondad de ajuste – similar al índice de razón de verosimilitud χ^2 pero sin la dependencia respecto al tamaño de la muestra observada –, la raíz cuadrada de la media de residuos cuadrados o el índice de centralidad.

Los índice de ajuste incremental no comparan las matrices de varianza-covarianza (o de correlación) observada o predicha, sino que comparan el ajuste del modelo propuesto con un modelo base especificado a priori por el investigador. El modelo base suele referirse como modelo “nulo” o modelo de “independencia”. Entre los índices más utilizados de ajuste incremental está el índice de Tucker-Lewis, el índice de ajuste relativo, el índice de ajuste normado, el índice de ajuste incremental, el índice de ajuste comparativo o el índice de centralidad relativa.

Los índices de ajuste de parsimonia relacionan la bondad de ajuste del modelo con el número de coeficientes estimados. El propósito es equilibrar la bondad de ajuste con la “parsimonia” o simplicidad: incluir los menos parámetros posibles. Un modelo es de elevada parsimonia cuando tiene relativamente pocos parámetros y, en cambio, muchos grados de libertad. Para ello su tamaño muestral ha de ser elevado. Entre los índices de ajuste de parsimonia más utilizados están el índice de ajuste parsimonioso, el χ^2 cuadrado, el índice de bondad de ajuste ajustado, el índice de bondad de ajuste de parsimonia, el error de la raíz cuadrada media de aproximación, el criterio de información de Akaike, y el estadístico N crítico.

5.4. Modelo estructural del objeto de investigación

En el siguiente apartado se presentará el modelo estructural y el modelo de medición a del sistema de ecuaciones estructurales través de distintas representaciones, como gráficos de sendero, ecuaciones o formas matriciales.

5.4.1. El modelo estructural

Tal y como se ha definido en el cuarto capítulo de la presente tesis doctoral, los dos objetos de investigación – la estructura organizativa y el constructo capacidad de absorción – pueden ser analizados a través de sus componentes o dimensiones.

El objetivo de la investigación de la tesis es averiguar cómo la estructura organizativa afecta a cada una de las fases del constructo capacidad de absorción. La Figura 5. 2 incluye la ilustración del modelo estructural mediante un diagrama de sendero de tres variables latentes exógenas y cuatro endógenas.

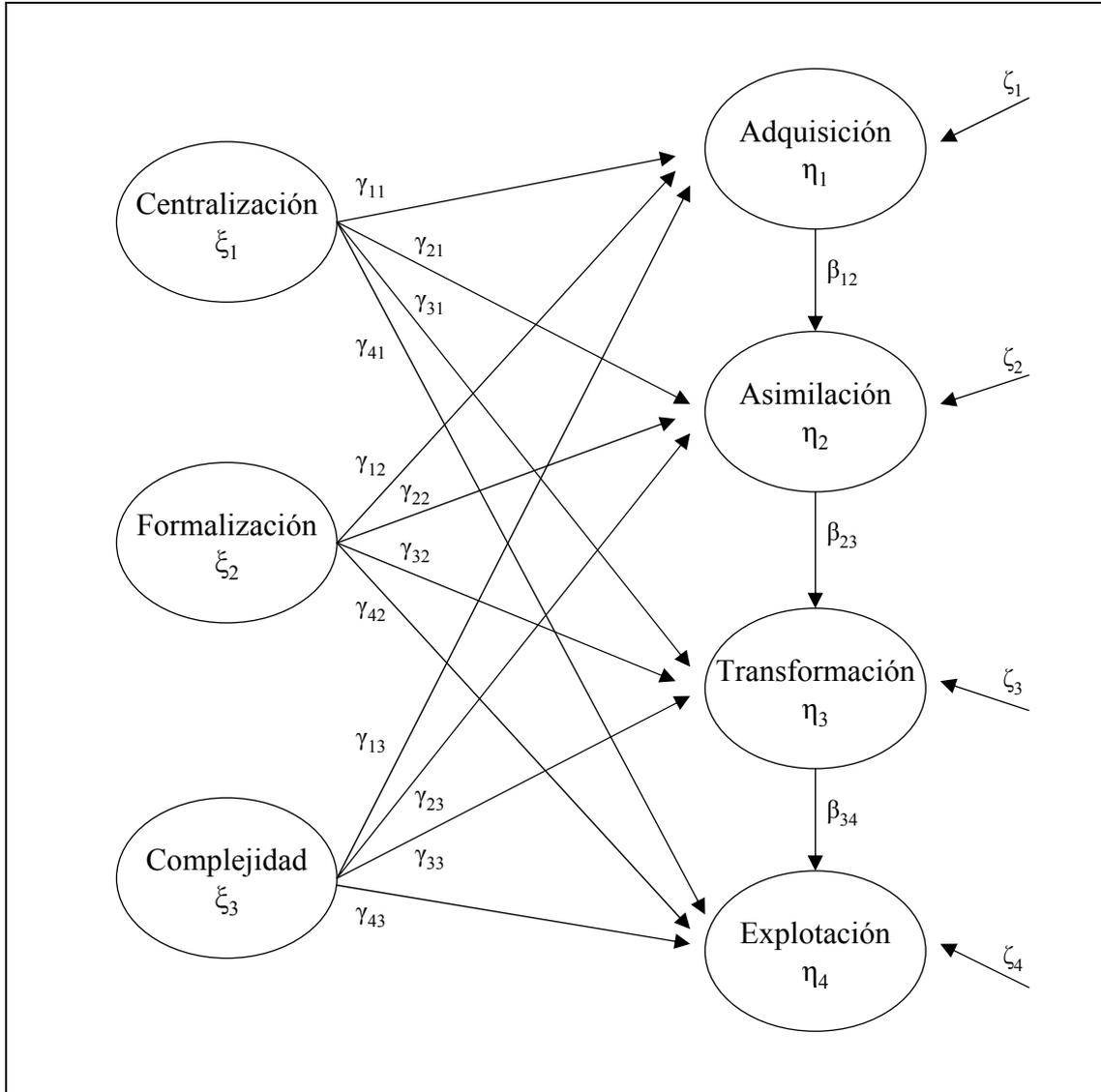
El modelo de la Figura 5. 2 establece que la variable latente endógena Adquisición (η_1) está causalmente relacionada con las tres variables latentes exógenas Centralización (ξ_1), Formalización (ξ_2) y Complejidad (ξ_3). Estas tres variables latentes no logran explicar perfectamente la Adquisición, por lo que se incluye el término de perturbación ζ_1 .

Por su parte, la variable endógena latente Asimilación (η_2) se encuentra relacionada causalmente con otras cuatro variables latentes. Tres de ellas exógenas y una endógena. La variable latente endógena es la Adquisición (η_1), y está relacionada a través del coeficiente β_{12} , mientras que las variables latentes exógenas son la Centralización (ξ_1), la Formalización (ξ_2) y la Complejidad (ξ_3). También se incluye el término de perturbación ζ_2 .

Paralelamente, las variables endógenas Transformación (η_3) y Explotación (η_4) están causalmente relacionadas con tres variables latentes exógenas Centralización (ξ_1),

Formalización (ξ_2) y Complejidad (ξ_3), y una variable latente endógena Asimilación (η_2) y Transformación (η_3) respectivamente.

Figura 5. 2. Modelo estructural del objeto de investigación



Fuente propia

A continuación se presenta las ecuaciones estructurales del modelo anterior, siguiendo la anotación de la Figura 5. 2.

$$\eta_1 = \gamma_{11}\xi_1 + \gamma_{12}\xi_2 + \gamma_{13}\xi_3 + \zeta_1$$

$$\eta_2 = \beta_{12}\eta_1 + \gamma_{21}\xi_1 + \gamma_{22}\xi_2 + \gamma_{23}\xi_3 + \zeta_2$$

$$\eta_3 = \beta_{23}\eta_2 + \gamma_{31}\xi_1 + \gamma_{32}\xi_2 + \gamma_{33}\xi_3 + \zeta_3$$

$$\eta_4 = \beta_{34}\eta_3 + \gamma_{41}\xi_1 + \gamma_{42}\xi_2 + \gamma_{43}\xi_3 + \zeta_4$$

5.4.2. El modelo de medición de las variables latentes exógenas

El modelo de medición de las variables exógenas incluye las relaciones entre estas variables con sus indicadores o variables empíricas. Para representar las tres dimensiones o variables latentes exógenas se han utilizado seis variables empíricas, que se caracterizan por haber sido utilizadas con asiduidad y aceptadas por la comunidad científica en varias ocasiones.

La Figura 5. 3 muestra la representación gráfica del modelo de medición de las variables exógenas, según la convención introducida en las secciones anteriores del presente capítulo.

A continuación se introduce cada una de las variables empíricas de dicho modelo:

La dimensión formalización es medida usando las escalas diseñadas por el grupo Aston para evaluar las variables *especialización* (SPEC) y *formalización* (FORMAL) (Inkson, Pugh, y Hickson, 1970; Pugh y Hickson, 1976), además de la variable *control* (CONTROL) desarrollada por Khandwalla (1974).

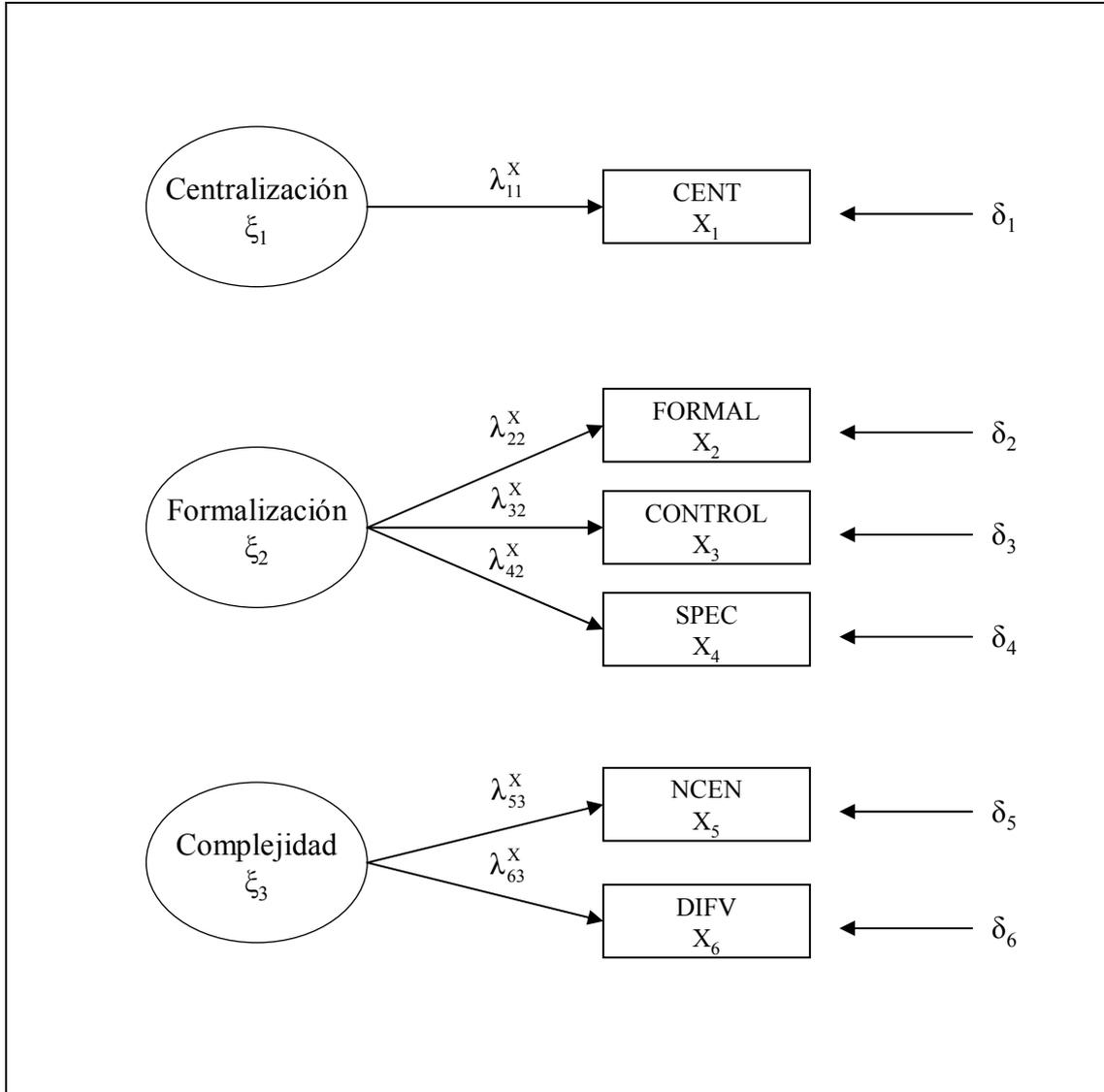
La dimensión complejidad se mide a través de dos variables empíricas: la *diferenciación vertical o niveles de autoridad* (VERT SPAN) y el *número de centros de operaciones*. Ambas variables también fueron creadas y escaladas por el grupo Aston.

Finalmente, la dimensión centralización es medida a través de una variable observable (CENT) definida también por el grupo Aston.

La anterior definición de las dimensiones y de las variables estructurales del modelo de medición se basa en la realizada por Miller y Dröge (1986) en su investigación de las

estructuras organizativas basándose en la personalidad del director general de una empresa.

Figura 5. 3. Modelo de medición de las variables latentes exógenas



Fuente propia

A continuación se presenta las ecuaciones estructurales del modelo de medición anterior, siguiendo la anotación de la Figura 5. 3:

$$X_1 = \lambda_{11}^x \xi_1 + \delta_1$$

$$X_2 = \lambda_{22}^X \xi_2 + \delta_2$$

$$X_3 = \lambda_{32}^X \xi_2 + \delta_3$$

$$X_4 = \lambda_{42}^X \xi_2 + \delta_4$$

$$X_5 = \lambda_{53}^X \xi_3 + \delta_5$$

$$X_6 = \lambda_{63}^X \xi_3 + \delta_6$$

Debido a la gran cantidad de variables y ecuaciones que acostumbran a surgir, es común representar las ecuaciones del modelo de medición en forma matricial, como se presenta a continuación:

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \\ X_6 \end{pmatrix} = \begin{pmatrix} \lambda_{11}^X & 0 & 0 \\ 0 & \lambda_{22}^X & 0 \\ 0 & \lambda_{32}^X & 0 \\ 0 & \lambda_{42}^X & 0 \\ 0 & 0 & \lambda_{53}^X \\ 0 & 0 & \lambda_{63}^X \end{pmatrix} \cdot \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \end{pmatrix}$$

5.4.3. Modelo de medición de las variables endógenas

El modelo de medición de las variables endógenas incluye las relaciones entre estas variables con sus indicadores o variables empíricas. Para representar las cuatro dimensiones o variables latentes endógenas se han utilizado seis variables empíricas.

Para evaluar cada una de las dimensiones del modelo desarrollado por Zahra y George (2002) se ha utilizado distintas escalas debido a la falta de un consenso en la medición del constructo.

Para evaluar la dimensión adquisición de conocimiento se ha utilizado el porcentaje destinado a la investigación y al desarrollo de nuevos productos o servicios sobre las ventas totales (GASTO), de forma similar a como la realizó Cohen y Levinthal (1990) o

como lo propone Zahra y George (2002). Además se complementa con una segunda variable como es el *número de proyectos de investigación en marcha* (PROY).

La asimilación es la dimensión más compleja de medir debido a su naturaleza tácita. Sin embargo, se han seguido las indicaciones de Cockburn y Henderson (1998), o Zahra y George (2002) de cuantificarla a través del número de publicaciones internas y externas realizadas por una organización (PUBLIC).

La transformación es medida a través del número de nuevas ideas o proyectos iniciados recientemente (NPROY) dentro de la organización, tal y como Leonard-Barton (1995) propuso y llevó a cabo en sus investigaciones.

Por último, la dimensión explotación es representada a través del número de nuevos productos o servicios que una empresa llega a comercializar en un año (NPROD). Además se incluye el número de patentes, licencias y modelos de utilidad (PAT) que han sido registrados, y por lo que una organización puede obtener nuevos beneficios.

La Figura 5.4 muestra la representación gráfica del modelo de medición de las variables endógenas, según la convención introducida en las secciones anteriores del presente capítulo.

A continuación se presenta las ecuaciones estructurales del modelo de medición anterior, siguiendo la anotación de la Figura 5. 4:

$$Y_1 = \lambda_{11}^Y \eta_1 + \varepsilon_1$$

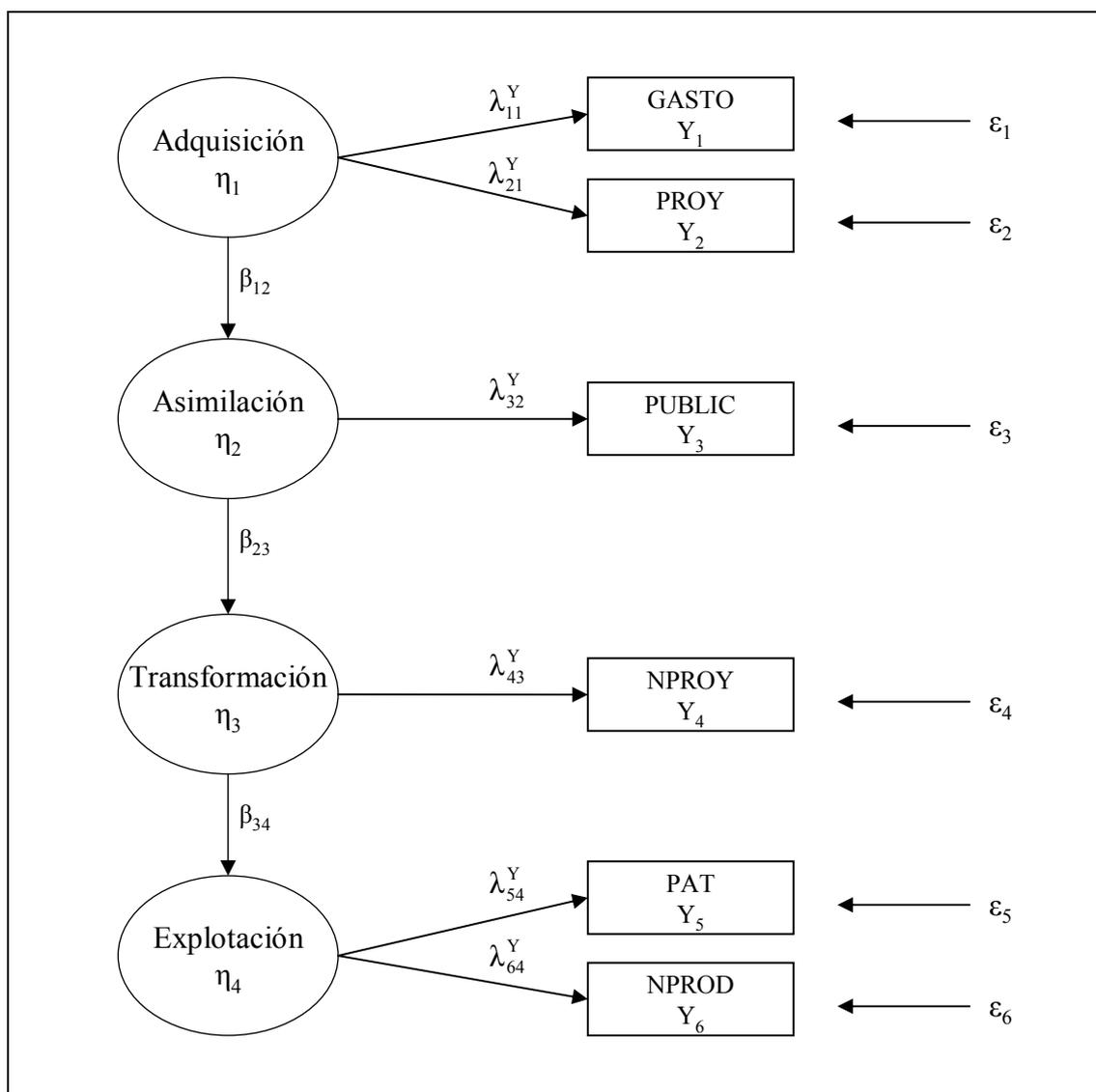
$$Y_2 = \lambda_{21}^Y \eta_1 + \varepsilon_2$$

$$Y_3 = \lambda_{32}^Y \eta_2 + \varepsilon_3$$

$$Y_4 = \lambda_{43}^Y \eta_3 + \varepsilon_4$$

$$Y_5 = \lambda_{54}^Y \eta_4 + \varepsilon_5$$

$$Y_6 = \lambda_{64}^Y \eta_4 + \varepsilon_6$$

Figura 5. 4. Modelo de medición de las variables latentes endógenas*Fuente propia*

Debido a la gran cantidad de variables y ecuaciones que acostumbran a surgir, es común representar las ecuaciones del modelo de medición en forma matricial, como se presenta a continuación:

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \\ Y_6 \end{pmatrix} = \begin{pmatrix} \lambda_{11}^Y & 0 & 0 & 0 \\ \lambda_{21}^Y & 0 & 0 & 0 \\ 0 & \lambda_{32}^Y & 0 & 0 \\ 0 & 0 & \lambda_{43}^Y & 0 \\ 0 & 0 & 0 & \lambda_{54}^Y \\ 0 & 0 & 0 & \lambda_{64}^Y \end{pmatrix} \cdot \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \end{pmatrix}$$

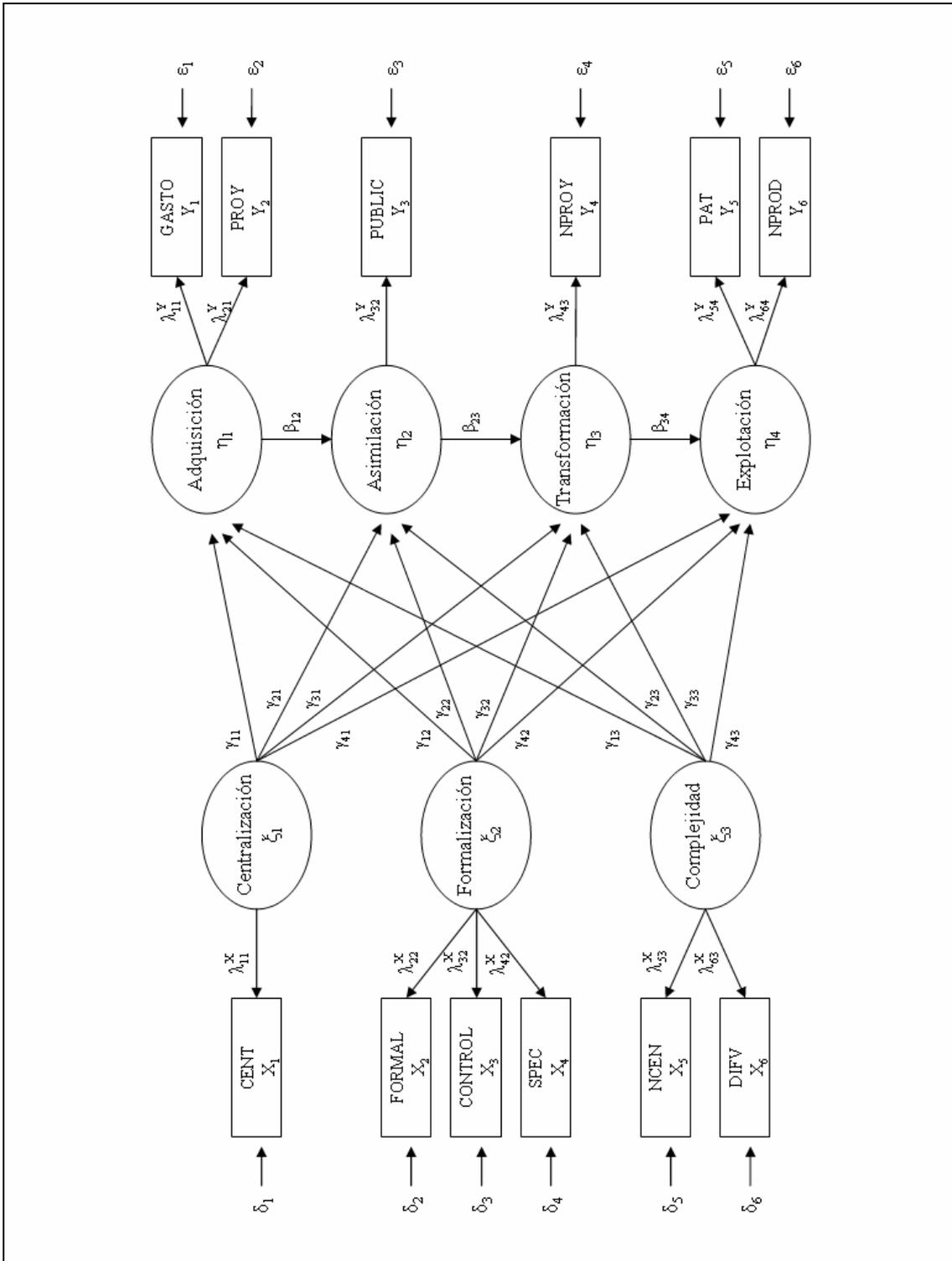
5.4.4. Modelo de ecuaciones estructurales final del objeto de investigación

De la combinación de los tres modelos anteriores – el modelo estructural, el modelo de medición de las variables latentes exógenas y el modelo de medición de las variables latentes endógenas – resulta un modelo comprehensivo de relaciones entre variables endógenas y exógenas, latentes y manifiestas: el modelo de ecuaciones estructurales.

La Figura 5. 5 incluye el modelo de ecuaciones estructurales que resulta de la combinación de los modelos de estructura (Figura 5. 2) y de medición para las variables exógenas (Figura 5. 3) y variables endógenas (Figura 5. 4) latentes.

El modelo ilustrado en la Figura 5.5 es un modelo “no recursivo” porque no incluye relaciones causales recíprocas. Además, los modelos “no recursivos” se caracterizan por permitir la correlación entre los términos de perturbación.

Figura 5. 5. Modelo de ecuaciones estructurales



Fuente propia

5.5. Metodología en el diseño del cuestionario para la captación de datos

Tras desarrollar el modelo de ecuaciones estructurales en la sección anterior, se debe proceder a la obtención de información primaria cuantitativa para el análisis del modelo. Con este objetivo se procede al desarrollo de un cuestionario para la captación de dicha información.

5.5.1. Selección del tipo de entrevista

Antes de desarrollar el cuestionario es necesario seleccionar el modelo de entrevista que se va a realizar: personal, telefónica, por correo, virtual, etc. La selección de una u otra condicionará la extensión, profundidad, así como el tipo de preguntas que se podrán plantear en el cuestionario.

Por el tipo de información que se necesita, se ha seleccionada la encuesta por correo electrónico, y la encuesta virtual como una alternativa. Las características principales de este tipo de entrevistas son un reducido coste, un alto grado de comodidad y accesibilidad, así como el nulo número de errores de transcripción.

Por el contrario, se tiene que vigilar con la duración de la encuesta ya que la falta de un entrevistador y una baja implicación por parte del entrevistado pueden llevar a que no se realice. Según Francisco José Sarabia (1999), en encuestas virtuales o por correo electrónico el tiempo máximo aconsejable para la realización son diez minutos. Además hay que constatar que existe un bajo control de las respuestas en este tipo de entrevistas.

5.5.2. Estructura del cuestionario

De la misma manera que el modelo estructural expuesto en el presente capítulo, el cuestionario está estructurado en dos partes: la relacionada con la estructura organizacional, y la relacionada con el constructo capacidad de absorción.

5.5.3. Cuestionario sobre la estructura organizativa

Por la naturaleza de las variables observables o empíricas del modelo de medición de las variables exógenas se necesita tanto escalas simples o naturales como escalas compuestas.

Para la formación de escalas compuestas existe gran cantidad de técnicas como la diferencial semántico, la aditiva de Likert, el formato Stapel, la suma constante o la técnica de comparación-par Thurstone. Las dos primeras técnicas se caracterizan por facilitar la administración de las respuestas y por un bajo tiempo necesario para responderlas. Por estos motivos, se han elegido la técnica aditiva de Likert y la diferencial semántico.

La utilización de la técnica aditiva de Likert presupone que la variable o indicador que se está midiendo viene dada por la suma de los valores obtenidos en cada proposición o ítem elegido. Esta técnica utiliza una categorización del continuo actitudinal, graduada según la intensidad, que debe ser indicada por el sujeto cuestionado.

En cambio, la técnica diferencial semántica se utiliza cuando el objetivo es medir conceptos o fenómenos, bajo el supuesto de que no son independientes de las reacciones de los sujetos. Este método, desarrollado por Osgood y Tannenbaum, suele ofrecer altas cotas de fiabilidad y validez cuando está bien construido. Su aplicación necesita de la existencia de adjetivos que califiquen al objeto, de forma que cada adjetivo muestra una dimensión.

Para medir la variable observable CENT se utiliza once cuestiones que conciernen a distintas situaciones dentro de una organización en donde se tiene que decidir qué persona o grupo de personas tienen la autoridad necesaria para tomar una decisión sobre ellas. La categorización de la escala utilizada oscila entre el consejo de administración hasta los operarios, pasando por los distintos niveles jerárquicos de la empresa.

La variable observable CONTROL es medida a través de cuestiones sobre la frecuencia de uso de cinco dispositivos de control diferentes. En este caso se ha utilizado una escala basada en la técnica diferencial semántico.

La variable observable FORMAL representa el nivel de formalización de la empresa y es medida mediante once cuestiones que se caracterizan por tener respuestas binarias (Sí – No). De la misma forma, la variable ESP se utiliza para medir el nivel de especialización de la organización a través de dieciséis cuestiones con respuestas binarias.

Las dos últimas variables NCEN y DIFV representan el número de centros de operaciones y el número de niveles que hay en una organización respectivamente. Ambas variables son escaladas a través de una escala simple.

La elaboración de presente cuestionario sobre estructuras organizativas está basada en la realizada por Miller y Drögue en 1986. Ambos autores realizaron dicho cuestionario como mecanismo para evaluar las estructuras organizativas de una gran cantidad de empresas a finales de los años ochenta. Sus publicaciones basadas en este cuestionario han recibido más de seiscientas citas desde su publicación según el Social Science Citation Index.

5.5.4. Cuestionario sobre la capacidad de absorción

En contraposición al cuestionario sobre estructuras organizativas, no existe ningún cuestionario para evaluar el constructo capacidad de absorción. Hasta la actualidad, cada investigador ha cuantificado el constructo según sus necesidades.

Para captar la información necesaria para cuantificar el constructo capacidad de absorción, se ha utilizado escalas simples para cada una de las variables observables o empíricas definidas previamente.

5.5.5. Presentación del cuestionario

Como se ha especificado en las secciones anteriores, se ha realizado el cuestionario sobre dos formatos diferentes:

- A través de una página Web, y
- A través de un formulario realizado con MS Word (ver Anexo 1).

El cuestionario realizado sobre página Web ha sido diseñado y desarrollado en HTML, y proporciona la opción de enviar la información directamente a una base de datos para procesarla posteriormente.

En cambio, el cuestionario en formato MS Word no está enlazado directamente a la base de datos del investigador, por lo que el encuestado debe reenviar el archivo resultante por e-mail al investigador. Posteriormente se tendrá que introducir los resultados de la encuesta a la base de datos uno por uno.

Este segundo método tiene el inconveniente de introducir la posibilidad de errores de transcripción del formulario a la base de datos, pero debido a la infraestructura de ciertas organizaciones se ha creído conveniente realizarlo en paralelo para poder acceder a un mayor número de empresas.

5.6. Selección de la muestra

La finalidad de cualquier análisis estadístico no se limita a la descripción de los casos de los que se ha recogido información. Éstos constituyen la muestra de la investigación. Ante todo, se quiere describir, a partir de las características observadas en la muestra, al conjunto de la población a la que ésta pertenece.

Respecto al tamaño de la muestra, su cuantía incide directamente en la reducción del error de estimación y en la consiguiente significatividad de los resultados del análisis. Por este motivo, se precisa que el tamaño de la muestra sea lo más elevado posible, para favorecer la significatividad de los estadísticos muestrales.

En los modelos de ecuaciones estructurales se barajan distintos ratios mínimos de observaciones precisas por cada variable independiente introducida al análisis. Afifi y Clark (1990) establecieron que el ratio mínimo debía estar entre 5 y 10 veces más casos que variables predictoras.

Según el tipo de estimación utilizado en el análisis del modelo, el mínimo número de muestras necesarias para obtener una significatividad razonable varía. El procedimiento de estimación de parámetros de máxima verosimilitud precisa de un tamaño muestral mínimo de 100 unidades. Con otros puede aumentar hasta 200 unidades e incluso superior el millar de muestras.

Bibliografía

- Afifi, A. A. y Clark, V. (1990). *Computer-aided multivariate analysis*. Chapman & Hall: Nueva York.
- Batista, J. M. y Coenders, G. (2000). *Modelos de ecuaciones estructurales*. La Muralla: Madrid.
- Cea, M. A. (2002). *Análisis multivariable. Teoría y práctica en la investigación social*. Editorial Síntesis: Madrid.
- Cockburn, I., & Henderson, R. (1998). 'Absorptive capacity, coauthoring behaviour, and the organization research in drug discovery', *Journal of Industrial Economics*, 46, pp. 157-183.
- Cohen, W. M., y Levinthal, D. A. (1990). 'Absorptive capacity: A new perspective on learning and innovation', *Administrative Science Quarterly*, 35, pp. 128-153.
- Hoel, P. G. y Jessen, R. J. (1993). *Estadística básica para negocios y economía*. CECSA: Azcapotzalco.
- Inkson, J.H.K, Pugh, D.S. y Hickson D.J. (1970). Organization context and structure - abbreviated replication. *Administrative science quarterly*, 15, pp. 318-329.
- Leonard-Barton, D. (1995). *Wellsprings of knowledge*. Boston: Harvard Business School Press.
- Jöreskog, K. G. (1993). *Modelado de ecuaciones estructurales*. Euskadiko: Guipúzcoa.
- Kaplan, A. (1964). *The conduct of inquiry*. Chandler: San Francisco.
- Khandwalla P. N. (1974). 'Mass output orientation of operations technology and organizational structure', *Administrative Science Quarterly*, 19, pp. 74-79.
- Miller, D., & Dröge, C. (1986). 'Psychological and traditional determinants of structure', *Administrative Science Quarterly*, 31, pp. 539-560.
- Miller, D. (1987). 'Strategic making and structure: analysis and implications for performance', *Academy of Management Journal*, 30, pp. 7-32.
- Miller, D., Dröge, C., & Toulouse, J. M. (1988). 'Strategic process and content as mediators between organizational context and structure', *Academy of Management Journal*, 31, pp. 544-569.
- Pugh D.S. y Hickson D.J. (1976). *Organizational structure in its context*. Westmead, Farnborough: Saxon House.

- Sarabia, F. J. (1999). *Metodología para la investigación en marketing y dirección de empresas*. Ediciones Pirámide: Madrid.
- Zahra, S. A., & George, G. (2002). 'Absorptive capacity: a review, reconceptualization, and extension', *Academy of Management Review*, 27, pp. 185-203.