

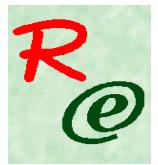


Revista Electrónica "Actualidades
Investigativas en Educación"
E-ISSN: 1409-4703
revista@inie.ucr.ac.cr
Universidad de Costa Rica
Costa Rica

Zapata Rotundo, Gerardo J.; Canet Giner, M. Teresa
PROPUESTA METODOLÓGICA PARA LA CONSTRUCCIÓN DE ESCALAS DE MEDICIÓN A
PARTIR DE UNA APLICACIÓN EMPÍRICA
Revista Electrónica "Actualidades Investigativas en Educación", vol. 8, núm. 2, mayo-agosto, 2008,
pp. 1-26
Universidad de Costa Rica
San Pedro de Montes de Oca, Costa Rica

Disponible en: <http://www.redalyc.org/articulo.oa?id=44713044004>

- ▶ Cómo citar el artículo
- ▶ Número completo
- ▶ Más información del artículo
- ▶ Página de la revista en redalyc.org



Actualidades Investigativas en Educación

Revista Electrónica publicada por el
Instituto de Investigación en Educación
Universidad de Costa Rica
ISSN 1409-4703
<http://revista.inie.ucr.ac.cr>
COSTA RICA

**PROPUESTA METODOLÓGICA PARA LA CONSTRUCCIÓN DE
ESCALAS DE MEDICIÓN A PARTIR DE UNA APLICACIÓN EMPÍRICA**

METHODOLOGICAL PROPOSAL FOR THE CONSTRUCTION OF
MEASURING SCALES FROM AN EMPIRICAL APPLICATION

Volumen 8, Número 2
pp. 1-26

Este número se publicó el 30 de agosto 2008

Gerardo J. Zapata Rotundo.
M. Teresa Canet Giner

La revista está indexada en los directorios:

[LATINDEX](#), [REDALYC](#), [IRESIE](#), [CLASE](#), [DIALNET](#), [DOAJ](#), [E-REVIST@S](#),

La revista está incluida en los sitios:

[REDIE](#), [RINACE](#), [OEI](#), [MAESTROTECA](#), [HUASCARAN](#)

Los contenidos de este artículo están bajo una licencia [Creative Commons](#)



PROPIUESTA METODOLÓGICA PARA LA CONSTRUCCIÓN DE ESCALAS DE MEDICIÓN A PARTIR DE UNA APLICACIÓN EMPÍRICA

METHODOLOGICAL PROPOSAL FOR THE CONSTRUCTION OF
MEASURING SCALES FROM AN EMPIRICAL APPLICATION

*Gerardo J. Zapata Rotundo¹
M. Teresa Canet Giner²*

Resumen: Uno de los principales inconvenientes en cualquier investigación científica es cuantificar adecuadamente los conceptos o fenómenos objetos de estudio; es decir, saber si se está midiendo la variable o constructo latente que realmente se quiere medir y si se efectúa con la rigurosidad metodológica requerida. Esto es aún más relevante cuando se abordan investigaciones con un enfoque cuantitativo. En este sentido, algunos conceptos con cierto grado de complejidad para su estudio y valoración requieren de instrumentos de medición debidamente diseñados que aseguran, además de su eficacia operacional, su fiabilidad y validez. Es así como en el presente trabajo se propone una metodología conformada por once pasos para la construcción de una escala de medición, y cuya aplicación práctica se llevó a cabo a través del diseño de un instrumento para medir el constructo latente formalización del comportamiento en las organizaciones en dos dimensiones que la explican: formalización organizativa y formalización profesional. De esta manera, se plantean distintos aspectos que el investigador debe considerar en la construcción de una escala, incluyendo, entre estos, el análisis de su calidad a través de diferentes métodos tales como: la fiabilidad, la validez de contenido y la validez de constructo.

Palabras Claves: CONSTRUCCIÓN DE ESCALAS DE MEDICIÓN, VALIDEZ Y FIABILIDAD, FORMALIZACIÓN DEL COMPORTAMIENTO.

Abstract: One inconvenient of every scientific research is to quantify properly the concepts or phenomena object of study; that is, to have the certainty that we are measuring the variable or latent construct that we really want to measure, and if that process of measurement has the required methodological accuracy. This is more relevant when we have to develop a quantitative research. Consequently, we can affirm that some complex concepts require a particular study and evaluation through measurement instruments accurately designed in order to assure its operative effectiveness together with its reliability and validity. Hence, this research work proposes an eleven-step methodology for the development of a measurement scale. The practical use of this methodology is showed through the design of an instrument that measures the latent construct "behavior formalization in organizations" by two different dimensions: organizational formalization and professional formalization. That way we make explicit the different aspects that a researcher has to take into account when developing a measurement scale, including an analysis of its quality through different methods: reliability analysis, content validity analysis and construct validity analysis.

Key Words: MEASUREMENT SCALES DEVELOPMENT, VALIDITY AND RELIABILITY, BEHAVIOR FORMALIZATION.

¹ Doctor en Organización y Dirección de Empresa por la Universidad de Valencia, España. Profesor Asociado, Departamento de Administración, Decanato de Administración y Contaduría de la Universidad Centroccidental "Lisandro Alvarado", Barquisimeto-Venezuela. Dirección electrónica: zapager@yahoo.com.

² Doctora en Ciencias Económicas y Empresariales. Profesora Titular, Departamento de Dirección de Empresas de la Facultad de Economía de la Universidad de Valencia-España. Dirección electrónica: teresa.canet@uv.es

Artículo recibido: 30 de octubre, 2007

Aprobado: 6 de agosto, 2008

1. Introducción

Mediante el proceso de medición podemos obtener ciertos conocimientos cuantificables sobre las características de un concepto, un *constructo* o una variable latente; por ejemplo, las tendencias del mercado, la calidad de un producto o el comportamiento de las variables de diseño organizativo, en respuesta a las condiciones impuestas por los factores contextuales. De este modo, medir un fenómeno resulta indispensable para toda actividad científica ya que se asignan números (valores)³ a los objetos observados de manera tal que representen de forma adecuada las característica que tienen (Nunally y Bernstein, 1995).

En este sentido, la idea fundamental cuando se mide un *constructo* latente es que el proceso dé como resultado el valor verdadero de lo que se cuantifica, o que proporcione información lo más cercana posible a la realidad. En este orden de ideas, la presente investigación se centra en el desarrollo y aplicación de una metodología que recoja todos y cada uno de los aspectos necesarios para la construcción de un instrumento de medición eficaz, dentro del marco de una investigación de carácter cuantitativo⁴. Esto se constituye en nuestra propuesta central y, por tanto, en el objetivo principal de este trabajo. A tal efecto, se toma como caso de estudio la construcción de una escala para medir la variable o *constructo* latente *formalización del comportamiento* en grandes empresas españolas, considerando dos dimensiones que la explican: *formalización organizativa* y *formalización profesional*.

Para cumplir con el objetivo de la presente investigación, la hemos dividido en tres partes principales. La *primera* se centra en el desarrollo del marco teórico alrededor de la variable *formalización del comportamiento* y las implicaciones que tiene en el diseño organizativo. Esta parte del trabajo es fundamental, ya que de ella se definen los atributos y dimensiones del concepto objeto de estudio que se requiere conocer con cierto detalle para poder concretar los ítems de la escala que lo cuantifican.

En la *segunda parte*, se abordan las consideraciones teóricas y metodológicas que dan forma a nuestra propuesta central y que consiste, como se ha comentado, en el desarrollo de una metodología para la construcción de instrumentos de medición de variables o *constructos* latentes dentro del marco de una investigación de carácter cuantitativo. Y en la *tercera parte* se recoge el estudio empírico-cuantitativo, donde se cubren todos los aspectos

³ Paréntesis nuestro.

⁴ Con lo cual se enmarca dentro del paradigma epistemológico positivista. La *investigación cuantitativa* está orientada a recoger información objetivamente mesurable (Miquel et al., 1997), y persigue la obtención de información representativa del conjunto de la población objeto de estudio (Bigné, 2000). La investigación cuantitativa tiene una visión de la realidad como algo creado o preexistente a nuestra conciencia, susceptible de ser percibida por los sentidos, se puede ver, oír y palpar (Domínguez, 2004). Es decir, se ubica dentro de la concepción *ontológica realista u objetiva*

de la metodología propuesta. Aquí se incluyen los métodos estadísticos necesarios para evaluar la calidad del instrumento de medida. La investigación culmina con un apartado dedicado a las conclusiones y reflexiones finales.

Finalmente, y desde el punto de vista metodológico, consideramos que la presente investigación tiene un doble carácter: *descriptivo* y *empírico-cuantitativo*. El primero debido a que se basa en una amplia revisión de la literatura, la cual permitió abordar con la profundidad adecuada todos y cada uno de los aspectos y conceptos relacionados con la construcción de un instrumento de medición -escala de medida-, derivándose de ello la metodología propuesta. Y por otra parte, tiene un carácter empírico-cuantitativo ya que se utilizaron una serie de métodos estadísticos que permitieron el análisis y evaluación de la calidad de una escala de medición, basada en un conjunto de datos obtenidos en el trabajo de campo realizado sobre una población de grandes empresas españolas.

2. Aspectos teóricos sobre la formalización en la organización: características y dimensiones

La formalización es una de las variables con mayor influencia en la determinación de las características orgánicas, mecánicas y burocráticas de una organización. Es por ello que ha sido centro de estudio en un amplio número de investigaciones relacionadas con la Teoría de la Administración, y en especial aquellas dedicadas a explicar el comportamiento de las variables de diseño organizativo frente a las condiciones impuestas por los factores de contingencia.

En los estudios clásicos de Hage (1965), Hage y Aiken (1967) y Pierce y Delbecq (1977) se asocia a la formalización con la proporción de trabajos codificados y el uso de reglas que lo regulan. Explican, entonces, que la codificación es el grado mediante el cual un trabajo en una posición determinada está claramente especificado y descrito, representando esto el nivel de *estandarización*. Dentro de esta definición, la *observación de las reglas* se refiere al grado en el que los ocupantes de una posición determinada en la organización son supervisados de conformidad con unos estándares establecidos, pudiéndose señalar un rango de variación tolerado por las reglas definidas para cada trabajo. Es decir, la formalización se constituye en un mecanismo de control burocrático.

Bajo las ideas anteriores, en los trabajos seminales Pugh et al. (1968), Pugh et al. (1969), Pugh (1973) y Pugh y Hickson (1976), se define a la formalización como el grado en el que las reglas, los procedimientos, las instrucciones y las comunicaciones están escritas y

sirven para gobernar la actuación de los miembros de la organización. En este sentido, la formalización incluye: 1) la declaración de los procedimientos de las operaciones diarias, las reglas que los regulan y los roles de cada uno de los cargos; 2) los procedimientos diseñados para dirigir los procesos de toma de decisiones; y 3) los procedimientos para la comunicación de la información a través de la estructura organizativa.

En esa misma línea del grupo de Aston, De la Fuente (1997), Daff (2000), Gibson et al. (2001), Donaldson (2001) y Robbins (2005) reconocen a la formalización como la medida en la que la organización depende de reglas y procedimientos escritos para predeterminar las actividades que se realizan en ella. En este sentido, la organización burocrática está unida por normas, reglamentos, jerarquía de autoridad, rutinas, procedimientos estandarizados y otros mecanismos que tratan de cubrir todas sus áreas con el propósito fundamental de estandarizar, evaluar y controlar el comportamiento y desempeño de sus miembros (Child, 1972; Mansfield, 1973; Grinyer y Ardekani, 1981; Gibson, et al., 2001; Donaldson, 2001).

Muy especialmente, y en relación con las formas burocráticas, la variable formalización está correlacionada de forma positiva con el tamaño de la organización (Pugh et al., 1968; Pugh et al., 1969; Blau, 1970; Pugh, 1973; Mansfield, 1973; Grinyer y Ardekani, 1981; Baker y Cullen, 1993; Hall, 1996; Daff, 2000; Donaldson, 2001). De esta forma, cuando hablamos de empresas de gran tamaño se hace necesario sustituir de manera importante la supervisión y coordinación directa por mecanismos más impersonales basados en la definición de las tareas y responsabilidades para cada puesto de trabajo, y en una buena cantidad de procedimientos correspondientes a las actividades diarias que orientan los procesos de toma de decisiones.

Ahora bien, en una definición más amplia y con una aportación mucho mayor desde el punto de vista práctico, Moreno-Luzón et al. (2001) argumentan que el concepto de formalización debe incluir: a) las reglas no escritas que dirigen y controlan el comportamiento de la organización; b) los procedimientos y tareas o procesos estandarizados, no documentados, que como consecuencia de órdenes explícitas o de la costumbre determinan la manera de llevar a cabo las tareas; y c) las reglas y procedimientos escritos a los cuales han de ajustarse las actividades de la empresa. Esta definición anterior tiene un sentido más amplio en comparación con el concepto tradicionalmente formulado, puesto que abarca *“las reglas de obligado cumplimiento, la estandarización de procesos y los diferentes procedimientos, tanto si los mismos están reflejados en manuales de procedimientos o en otros documentos, como si no lo están”* (Moreno-Luzón et al., 2001, p. 217). En síntesis,

identificamos una primera dimensión de la formalización del comportamiento que es diseñada por la empresa y la denominamos *formalización organizativa*.

Por otra parte, se puede identificar otra dimensión de la formalización que se corresponde a las características y naturaleza de cada profesión (Hall, 1996; Peris et al., 2001; Moreno-Luzon, et al., 2001). Es decir, aquella relacionada con los especialistas formados en instituciones de educación superior que se integran al sistema de dirección y operaciones de la empresa e incorporan estándares, métodos y procedimientos o formas de comportamiento, los cuales sirven de guía en la actuación del personal cualificado o que ejerce un trabajo profesional.

En resumen, estas dos clases de formalización que hemos denominado *formalización organizativa* –en términos de la definición clásica– y *formalización profesional*, pueden estar presentes con mayor o menor incidencia en cualquier tipo de organización, definiéndose como un *constructo multidimensional*.

3. Consideraciones teóricas y metodológicas en la construcción de una escala de medición

Cuando en una investigación de carácter cuantitativo se plantea la posibilidad de comprobar la viabilidad de un modelo teórico a través del contraste de una serie de hipótesis, el investigador se ve en la necesidad de desarrollar una *escala de medición* con el propósito de obtener información acerca de las variables o *constructos latentes* que forman parte del modelo propuesto, y así poder “*estimar la naturaleza del concepto de forma fiable y válida*” (Sarabia, 1999, p. 343)⁵.

Por lo general, en los modelos teóricos se plantean un conjunto de relaciones entre variables no directamente observables y cuantificables de la realidad, pero que sí lo pueden ser a través de un grupo de variables manifiestas, ítems o indicadores. Así, una *escala de medición* “*constituye un conjunto de ítems, frases o preguntas que permiten medir el nivel alcanzado por un atributo o concepto determinado* (ej. la formalización organizativa)⁶ *no directamente observable en un objeto* (ej. una empresa)” (Vila et al., 2000, p. 1)⁷. En este

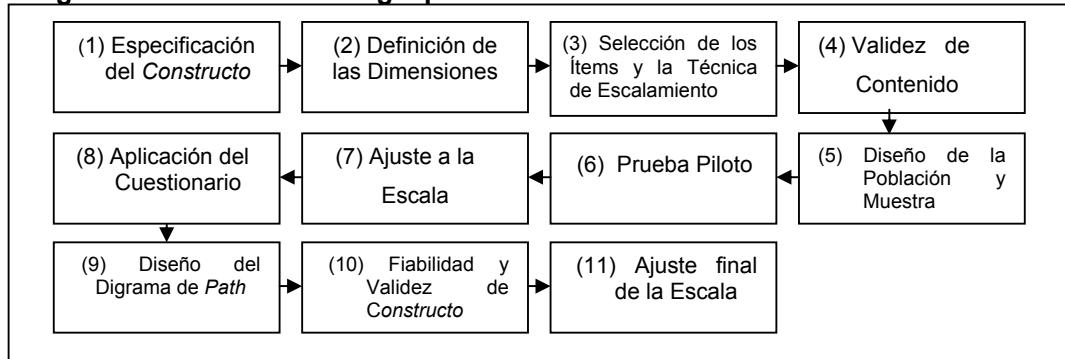
⁵ Una *variable* o *constructo* es *latente* en cuanto a que no es directamente observable en un momento determinado, e implica la existencia de una relación entre ella y los ítems de la escala.

⁶ Paréntesis nuestro.

⁷ La *medida* se puede definir como “un conjunto de reglas que permiten asignar números a los objetos observados de tal forma que representen de manera adecuada la cantidad de un determinado atributo” (Nunnally y Bernstein, 1995). Los indicadores o ítems de la escala se definen también como *variables manifiestas*.

sentido, en la presente investigación se propone una metodología de once pasos para la construcción de escalas de medición. La misma, se representa en la siguiente figura 1.

Figura 1 Metodología para la Construcción de una Escala de Medición



Fuente: Elaboración propia.

Vemos en la figura anterior, que el punto de inicio –pasos 1 y 2– de la metodología, consiste en identificar el dominio de los *constructos* o variables latentes del modelo teórico o de las hipótesis por contrastar⁸. Lo fundamental aquí es realizar una amplia revisión bibliográfica que permita identificar y profundizar en todas aquellas características, elementos y dimensiones del concepto bajo estudio. Así, una vez identificados sus atributos, el investigador, en el tercer paso de la metodología propuesta –ver figura 1–, procede a generar un conjunto de ítems que constituirán la escala de medición o instrumento de levantamiento de información. Para ello se deberá tomar como referencia la revisión de la literatura, los atributos y dimensiones de las variables evaluadas, escalas elaboradas con anterioridad por otros investigadores, y la propia experiencia y conocimientos del investigador. Además, en esta tercera etapa se pasa por seleccionar la técnica de escalamiento, entre las más utilizadas: 1) formato Stapel; 2) diferencial semántico; y 3) técnica aditiva de Likert.

Al culminar las tres etapas anteriores, se obtiene una primera versión de la escala de medición o cuestionario, y seguidamente se le asigna un valor numérico (escalamiento) que represente la medida para cuantificar las características de las variables latentes. Ahora bien, estas variables por lo general presentan características que las hacen más o menos

⁸ Churchill (1999) se refiere a las características del dominio como las propiedades de la variable o fenómeno objeto de medición.

complejas o poco observables, viéndose la conveniencia de considerar en la escala múltiples ítems de manera tal que se pueda conocer con mayor precisión sus atributos⁹. Esta situación puede producir inicialmente un excesivo número de ítems difíciles de ser manejados en forma eficiente durante el proceso de levantamiento de información, por consiguiente, se hace necesario llevar a cabo un primer procedimiento de depuración de la escala a través de la *validez de contenido*.

La validez de contenido busca, de acuerdo con lo señalado por Devlin et al. (1993), que la escala de medición o cuestionario cumpla con las siguientes características: a) induzca a un mínimo de respuestas sesgadas; b) sea fácil de entender e interpretar; c) sea fácil de administrar o aplicar en el trabajo de campo; y d) posea capacidad de discriminar.

En este sentido, y en términos generales, con la estimación de la validez se busca conocer si lo que se está midiendo es realmente la variable o *constructo latente* que se pretende medir (Vila et al., 2000; Martínez y Hernández, 2006). Por tanto, una escala es válida cuando realmente sus ítems cuantifican con exactitud lo que están evaluando. Así, la *validez de contenido* es contrastada por el juicio subjetivo de uno o varios investigadores; no obstante, a pesar de ese carácter subjetivo es una evaluación sistemática, sirve para apreciar la correspondencia entre los ítems individuales y los *constructos* estudiados a través de valoraciones hechas por expertos del área en cuestión. Para llevarla a cabo, se cuenta con diferentes procedimientos tales como la evaluación individual por especialistas, la técnica de Delphi y la aplicación de pre-test. Sin embargo, es oportuno destacar que la decisión final de aceptar las opiniones de los expertos o decidir su uso es facultad del propio investigador.

Identificados los ítems de la escala de medición, a continuación en el paso 5 –ver figura 1– el investigador procede a seleccionar la población y muestra que será utilizada en el estudio empírico, y sobre la cual busca generalizar los resultados. En este sentido existe una amplia literatura especializada acerca de su manejo estadístico, por esta razón no constituye un tema central por desarrollar en la presente investigación.

Posterior al diseño de la población y muestra, se debe realizar una prueba piloto –paso 6, ver figura 1– sobre un grupo reducido de unidades de análisis que forman parte de la población seleccionada. Para que esta sea efectiva, es necesario anexar al cuestionario una

⁹ Miquel et al. (1997) señalan de modo general que las escalas con un sólo ítem son más vulnerable a los errores aleatorios que si se suman las valoraciones de una escala de varios ítems, puesto que las posibles fluctuaciones aleatorias se neutralizan. Es decir, el error aleatorio tendería a compensarse y a corregirse al promediar un gran número de mediciones repetidas (Batista y Coenders, 2000); por tanto, las escalas de múltiples ítems suelen tener más fiabilidad.

guía de observaciones donde los encuestados emiten su opinión en relación con su contenido y estructuración. Con los resultados obtenidos de esta prueba, se procede a realizar los cambios sugeridos –paso 7, ver figura 1–, los cuales pueden ser modificaciones, eliminaciones o incorporaciones de nuevos ítems.

Una vez realizados los ajustes necesarios, la siguiente actividad consiste en aplicar el cuestionario –paso 8, ver figura 1– a la población y/o muestra seleccionada con el propósito de obtener los datos relativos al estudio empírico. Estos datos serán la base para evaluar, a través de métodos estadísticos, la calidad de la escala de medición.

En relación con la calidad de la escala, cabe destacar que esta se constituye en un instrumento útil para la investigación científica si cumple con una serie de propiedades *psicométricas: fiabilidad y validez de constructo*. Esto tiene que ver con el hecho de que cuando se mide un concepto se busca que el proceso dé como resultado un valor verdadero, o lo más cercano a la realidad. Al respecto, la “medición de los diversos conceptos de las ciencias sociales tiene lugar en un complejo contexto donde existen muchos factores que afectan al atributo o fenómeno medido, incluyendo el proceso de medición en sí” (Sánchez y Sarabia, 1999: 364), factores que se convierten en posibles fuentes de errores de medición¹⁰.

Asimismo, una escala de medida tiene una serie de características como el número de preguntas o ítems, su grado de complejidad y tema que aborda, entre otras, que pueden incrementar el riesgo de posibles errores inherentes a cualquier medición¹¹. Por tanto, uno de los objetivos del investigador es tratar de reducir el error de medida, con el propósito de que la escala refleje los valores verdaderos del fenómeno estudiado. En consecuencia se hace necesario abordar adicionalmente a la validez de contenido, el análisis de la *fiabilidad y validez de constructo*¹².

Sin embargo, antes de evaluar la calidad de la escala, es preciso diseñar, tal como se indica en el paso 9 de la metodología propuesta –ver figura 1–, el *diagrama de path* del modelo teórico. Este se convertirá en la base esencial para determinar tanto los valores de los estadísticos correspondientes a la *fiabilidad y validez de constructo*, como de aquellos

¹⁰ El *error de medida* es el grado en el que los valores observados no son representativos de los valores verdaderos (Hair et al., 1999).

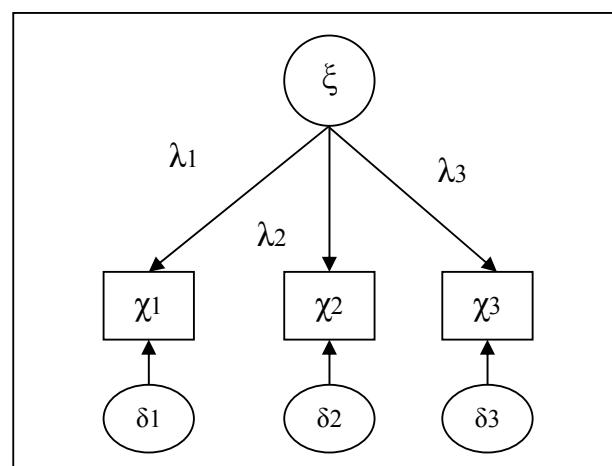
¹¹ Churchill (1999, p. 448-452) ofrece una serie de fuentes que dan origen a diferencias en los valores obtenidos en la medición: 1) los individuos encuestados pueden tener verdaderamente distintas opiniones en relación con las variables estudiadas; 2) la cultura o el estilo de vida de las personas encuestadas; 3) diferencias causadas por factores transitorios del individuo entrevistado –estado psicológico y de salud–; 4) factores situacionales alrededor del individuo entrevistado; y 5) diferencias en la manera de aplicar el cuestionario por parte del entrevistador.

¹² Podemos estar en presencia de dos tipos principales de errores de medición: *sistemático* y *aleatorio*. El *error sistemático* afecta a la medición de manera constante cada vez que esta se realiza. Por otra parte, el *error aleatorio* se debe a problemas transitorios ocurridos durante la medición y por tanto no es constante.

necesarios para estimar la bondad del ajuste del modelo de medida objeto de estudio. Por eso, el *diagrama de path* constituye la “*forma más común de representar las relaciones entre distintos ítems de la escala y el concepto* (variable o *constructo* latente evaluado)¹³ que se busca cuantificar” (Vila et al., 2000, p. 4).

A continuación en la siguiente figura 2, se representa un *diagrama de path* con su correspondiente simbología.

Figura 2 Diagrama de Path de un Modelo de Medida



Un modelo de medida describe la naturaleza de las relaciones entre el *constructo* o variable latente que forma parte del modelo teórico y las variables manifiestas o ítems de la escala que lo miden. Así, en la figura 2, la variable latente se representa con el símbolo ξ en el círculo. El símbolo χ , dentro de los recuadros, indica los ítems¹⁴ que conforman la escala. Por otro lado, la flecha representa la relación causal entre el *constructo* (ξ) y los ítems (χ_i). Y el símbolo δ dentro del óvalo, expresa el error de medida que puede estar presente en cualquier tipo de investigación científica. Finalmente, el símbolo λ en las flechas muestra los valores de las cargas factoriales o coeficientes estandarizados de la relación causal entre el *constructo* y los ítems de la escala¹⁵.

¹³ Paréntesis nuestro.

¹⁴ Denominada también *variables manifiestas*.

¹⁵ Las cargas factoriales son las correlaciones entre cada variable manifiesta o ítems de la escala y el factor o *constructo* latente objeto de estudio, indicando su grado de correspondencia (Hair et al., 1999).

Una vez definido el diagrama de *path* del modelo de medida, el siguiente paso de la metodología consiste en calcular la *fiabilidad* y la *validez de constructo* –paso 10, ver figura 1–. En general, mientras que la *validez* busca asegurar que el *constructo* medido sea realmente el que se pretende evaluar, la *fiabilidad* está dirigida a asegurar que la medición se haga de forma consistente y homogénea (Miquel et al., 1997; Malhotra, 2004).

Uno de los métodos para calcular la *fiabilidad* es a través del coeficiente o *alpha* de Cronbach (1951)¹⁶. Este estadístico asume “que la escala está compuesta por elementos homogéneos los cuales miden las mismas características, y la consistencia interna de la escala puede evaluarse mediante la correlación existente entre todos sus elementos (ítems de la escala)”¹⁷ (Pardos y Ruiz, 2002, p. 596). Es lógico pensar que los ítems o indicadores de una misma escala han de estar positiva y significativamente correlacionados, puesto que evalúan la misma realidad o concepto.

El valor del coeficiente *alpha* de Cronbach varía entre 0 y 1, y en la medida que aumenta o se acerca a 1 indica una mayor fiabilidad, tendiendo a incrementar su valor con el aumento en el número de ítems. En relación con este último punto, algunos autores plantean la existencia de una relación positiva y significativa entre el número de ítems utilizados en la escala y su fiabilidad (Churchill y Peter, 1984; Churchill, 1999; Malhotra, 2004).

Por lo general, y de acuerdo con Malhotra (2004), un valor del coeficiente *alpha* de Cronbach 0,6 o menor, indica una fiabilidad no satisfactoria. Por su parte, Hair et al. (1999) señalan que el acuerdo general sobre el límite inferior para el *alpha* es de 0,70, pudiendo bajar a 0,60 en los casos de investigaciones exploratorias. Cabe notar que en el caso donde hay implicadas en un modelo teórico más de una variable latente que explican un mismo concepto o *constructo*, se hace recomendable calcular el índice de fiabilidad compuesta¹⁸.

Siguiendo las ideas anteriores, la *fiabilidad* solo muestra que sus ítems, al estar muy correlacionados, miden la misma variable latente. Sin embargo, la presencia de una fiabilidad adecuada no asegura que la variable latente cuantificada sea la correcta, es decir, que resulte válida. Por tanto, además de realizar la validez de contenido ya tratada en el paso 4 de la metodología y el cálculo de la fiabilidad, es necesario abordar el análisis de la *validez de constructo*. Esta es más difícil de evaluar y reconocer puesto que no se realiza por procedimientos subjetivos, como es el caso de la validez de contenido, sino cuantificando las

¹⁶ Existen otros métodos para determinar la fiabilidad, por ejemplo: a) la prueba-reprueba; b) el método de fiabilidad absoluta; y c) el índice de fiabilidad compuesta.

¹⁷ Paréntesis nuestro.

¹⁸ La *fiabilidad compuesta* es importante realizarla puesto que el *alpha* de Cronbach para cada factor no tiene en cuenta la influencia sobre la fiabilidad del resto de variables *latentes* o factores (Vila et al., 2000).

correlaciones con el *constructo* latente o concepto bajo estudio a través de métodos estadísticos. Tiene relación principalmente con el problema de saber qué es lo que realmente cuantifica el instrumento de medición (Sánchez y Sarabia, 1999; Hair et al., 1999), constituyéndose en la medida por medio de la cual los ítems de la escala evalúan con precisión el fenómeno estudiado. En relación con lo anterior, existen diferentes métodos para hacer operativa la *validez de constructo*, entre ellos: a) la validez convergente y b) la validez discriminante.

La *validez convergente* está presente cuando existe una alta y significativa correlación entre los ítems de la escala con las variables latentes o conceptos que evalúan (Peter y Churchill, 1986; Hair et al, 1999; Malhotra, 2004); es decir, “*cuando se emplean distintos instrumentos para medir un mismo constructo* (distintos ítems para una misma variable latente)¹⁹, *y estos instrumentos están fuertemente correlacionados*” (Vila et al., 2000, p. 20). Por su parte, la *validez discriminante* viene dada por la baja correlación entre los ítems utilizados para evaluar una variable y otros que se suponen no miden esa misma variable (Spector, 1992; Sánchez y Sarabia, 1999; Malhotra, 2004)²⁰, puesto que estos últimos realmente cuantifican conceptos diferentes. Esto es, la escala no sirve para medir una variable distinta al propósito para la cual fue diseñada²¹. En el epígrafe 4, se hará la aplicación práctica de la *fiabilidad y validez de constructo*.

Finalmente determinadas las propiedades *psicométricas* de la escala de medición –*fiabilidad y validez de constructo*–, se pasa a realizar una última depuración –paso 11 de la metodología, figura 1– con el objetivo de excluir los ítems o indicadores que no satisfacen los requisitos mínimos de calidad exigidos por los métodos estadísticos. Lo anterior se realiza con el propósito de mantener los datos recogidos en el estudio empírico que representen con mayor precisión la realidad del fenómeno estudiado y, por tanto, alcanzar así un adecuado análisis e interpretación de los resultados.

¹⁹ Paréntesis nuestro.

²⁰ En sintonía con lo señalado, Sánchez y Sarabia (1999, p. 381) se refieren a la validez discriminante “como al hecho de que una escala solo cuantifica las características del concepto que trata de medir y no otras características de cualquier otro concepto”.

²¹ Si el *modelo inicial de medida* es inadecuado, entonces se eliminarán o reasignarán variables manifiestas o indicadores a los factores, hasta que se alcance un adecuado ajuste.

4. Estudio empírico y análisis de los resultados: aplicación de la metodología propuesta

Con base en la propuesta realizada en el apartado anterior, el presente epígrafe tiene como objetivo llevar a cabo la aplicación práctica de la metodología para construir una escala de medición en una investigación de carácter cuantitativo. Para tal efecto, hemos considerado como caso de estudio empírico el concepto no directamente observable: *formalización del comportamiento* en las organizaciones.

4.1 Definición de la escala de medición

De acuerdo con la metodología propuesta, las primeras actividades por realizar en el diseño de una escala –pasos 1 y 2, figura 1–, consisten en determinar las características y dimensiones que explican los *constructos* o variables latentes a ser evaluadas, para luego definir los ítems que posibiliten su cuantificación. De este modo, y en el caso que nos ocupa, se ha recurrido a la revisión de un conjunto de investigaciones de carácter empírico y teórico que abordan el tema de la formalización en las organizaciones. En el cuadro 1, se recogen las investigaciones antes citadas.

Cuadro 1 **Trabajos Relacionados con la Formalización en las Organizaciones**

Formalización Organizativa	Hage (1965); Hage y Aiken (1967); Pierce y Delbecq (1977); Pugh et al. (1968), Pugh et al. (1969); Pugh (1973) y Pugh y Hickson (1976); Child (1972); Mansfield, 1973; Grinyer y Ardekani, 1981; Daff (2000); Gibson et al. (2001); Donaldson (2001); De la Fuente et al. (1997); Moreno-Luzón et al. (2001).
Formalización Profesional	Moreno-Luzón et al. (2001); Peris et al. (2001) y Hall (1996).

Tal como se planteó en el marco teórico –epígrafe 2–, la formalización del comportamiento tiene un carácter subjetivo y complejo, y se explica a través de dos dimensiones: *formalización organizativa* y *formalización profesional*. Ambas constituyen en sí mismas dos variables latentes, y requieren para su evaluación de un conjunto de ítems bien diferenciados que muestren su grado o nivel de desarrollo en una organización. En este sentido, se diseñó inicialmente una escala formada por 13 indicadores y se decidió por un escalamiento o formato de medida tipo Likert de cinco puntos, donde el valor 1 significa “muy en desacuerdo”, 2= “en desacuerdo”, 3= “a veces”, 4= “de acuerdo” y 5= “muy de acuerdo”.

Seguidamente, y de acuerdo con el paso 4 de la metodología propuesta –figura 1–, se analizó la *validez de contenido* con el propósito de reducir la escala, si fuera el caso, a los ítems que reflejaran de manera más adecuada los valores de la variable formalización, e hiciera posible su aplicación en el estudio de campo que para tal efecto se llevó a cabo. Por esta razón, la escala se sometió a una revisión por un grupo de expertos en el área organización de empresas pertenecientes al Departamento de Dirección de Empresas de la Universidad de Valencia, España. De su resultado y observaciones, se pasó de una escala inicial de 13 ítems a una de 9. Cabe destacar que para hacer operativo este procedimiento, se entregó al grupo de expertos una hoja de valoración al cuestionario con el propósito de que describieran las observaciones a que dieran lugar.

En el siguiente cuadro 2, se ofrece la escala definitiva una vez sometida a la *validez de contenido*.

Cuadro 2 Escala de Medición de la Formalización Organizativa y Profesional

V1: El perfil de cada puesto de trabajo está claramente especificado.
V2: Las tareas y los procedimientos de cada puesto de trabajo están claramente especificados.
V3: Las reglas y las normas especifican cuándo y quién puede llevar a cabo los procedimientos de trabajo para la producción de los bienes y/o servicios, están claramente establecidas.
V4: Los procedimientos o mecanismos de toma de decisiones, coordinación y control están claramente establecidos (escritos o no). Por ejemplo: planes, programas, presupuestos, órdenes de producción, autorizaciones de venta, entre otros.
V5: El trabajo operativo está sujeto a reglas y procedimientos estándares.
V6: Los objetivos y metas de la organización están claramente especificados, y son transmitidos a toda la organización.
V7: El trabajo de los profesionales cualificados en la empresa (con estudios de diplomado, licenciatura o ingeniería), se caracteriza por el cumplimiento de las normas y protocolos de la profesión.
V8: El nivel de cualificación requerido para incorporarse a mi empresa, en general, implica que se posean estudios Universitarios.
V9: La forma de actuar propia de los profesionales cualificados de la empresa, influye en el conjunto de actividades de diseño y producción de los bienes y/o servicios.

En el cuadro anterior, los ítems de V1 a V6 miden la *formalización organizativa*, y de V7 a V9 la *formalización profesional*.

Una vez definida la escala de medición, y de acuerdo con el paso 5 de la metodología propuesta –figura 1–, se procedió a seleccionar la población sobre la cual se pretende hacer las inferencias. La misma se llevó a cabo a través de la base de datos *Sistema de Análisis de Balances Ibéricos* (SABI)²², considerándose como población las grandes empresas españolas pertenecientes a todos los sectores de su economía –carácter multisectorial²³. Así, y de acuerdo con el criterio elegido, la base de datos suministró un total de 1.709 grandes empresas.

Posteriormente, y siguiendo los pasos 6 y 7 de la metodología, se llevó a cabo un pre-test o prueba piloto, el cual consistió en el envío del cuestionario a un grupo de empresas pertenecientes a la población seleccionada –un total de 15–, surgiendo de sus opiniones cambios pocos significativos y cuya valoración, en general, fue positiva. Así, y una vez realizadas las actividades anteriores, se procedió al envío postal del cuestionario a las empresas de la población objeto de estudio, siendo dirigido a la alta dirección. Se materializa esta parte de la investigación con el trabajo de campo realizado entre los meses de marzo a julio de 2005. Esta última actividad corresponde al paso 8 –aplicación del cuestionario– de la metodología propuesta. El cuadro 4 permite apreciar la ficha técnica de la investigación empírica.

Cuadro 4 Ficha Técnica de la Investigación Empírica

Población	1.709 grandes empresas españolas
Tamaño de la muestra	118
Nivel de confianza	90%
Error muestral	±7,33%
Procedimiento de muestreo	Muestreo no aleatorio – <i>opinático o intencional</i> –
Período de realización	Marzo-julio de 2005
Tipo de instrumento	Cuestionario estructurado, dirigido a la alta dirección.

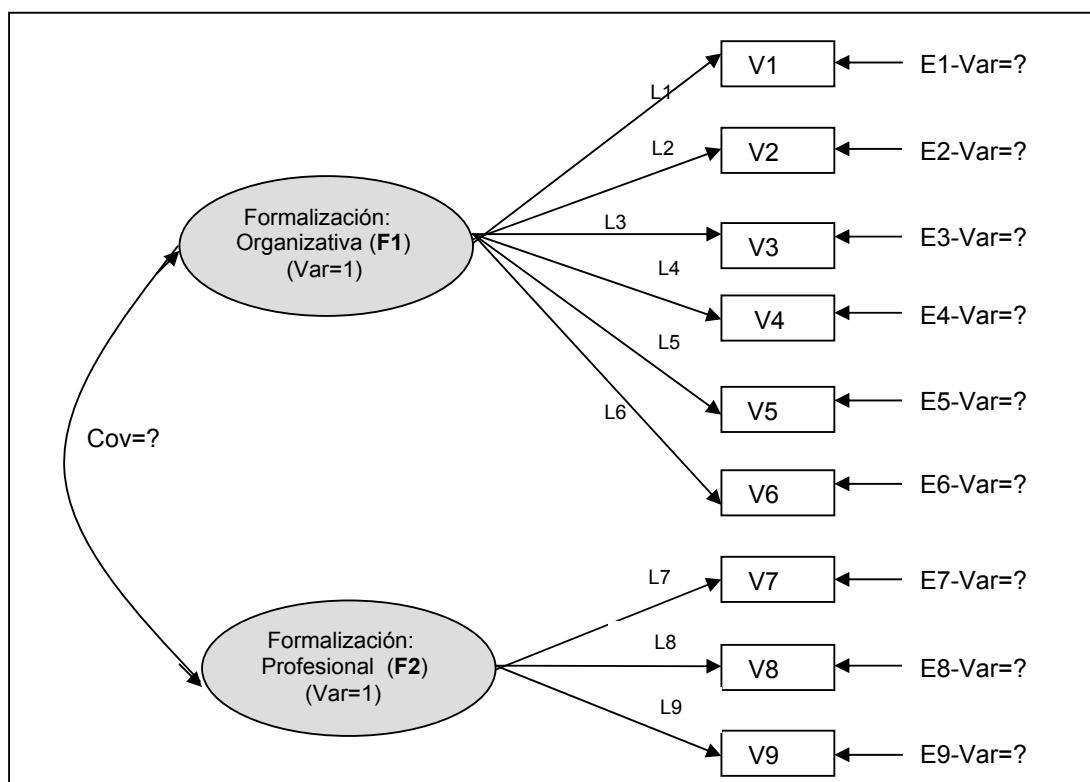
²² Base de datos de empresas españolas.

²³ Para determinar el tamaño de las empresas se siguieron las recomendaciones de la Comisión Europea del 06 de mayo de 2003, publicadas en el Diario Oficial de la Unión Europea No. L-124, estableciéndose como criterio de selección: *las empresas de cualquier sector de la economía española con un volumen de facturación mayor de 50 millones de euros y que ocupen a un número igual o mayor a 250 empleados*.

4.2 Evaluación de la Fiabilidad y Validez de las Escalas de Medición

Siguiendo las recomendaciones de la metodología propuesta en el presente trabajo, seguidamente se procede elaborar el diagrama *path* que representa el modelo de medida de la variable formalización. Este servirá de base para el cálculo, a través del análisis factorial confirmatorio, de los estadísticos que determinarán la bondad del ajuste del modelo, así como de aquellos necesarios para realizar la *validez de constructo*. A continuación, en la figura 3 se representa el modelo teórico de medida del *constructo latente formalización del comportamiento en las organizaciones*.

Figura 3 **Modelo de Medida Inicial para el Análisis Factorial Confirmatorio de la Formalización del Comportamiento: Organizativa (F1) y Profesional (F2)**



Fuente: Elaboración propia

Como hemos utilizado el programa estadístico EQS para estimar el modelo de medida, este posee las siguientes convenciones: a) La letra “F”, identifica las variables latentes o factores (sean independientes o no); b) la letra “V”, las variables manifiestas o ítems que forman parte de la escala de medida; c) la letra “E”, se utiliza para los términos de error de las variables manifiestas o ítems de la escala; d) la letra “L”, representa el valor de las cargas factoriales o coeficientes estandarizados; e) COV, indica el grado de covarianza compartida entre los dos factores (F1 y F2); y f) VAR, significa la varianza de cada factor independiente.

Seguidamente, pasamos a calcular la *fiabilidad* y la *validez del constructo* –paso 10 de la metodología propuesta– de la escala de medición de la formalización. Así, para verificar la *fiabilidad* utilizamos dos métodos: consistencia interna o *alpha* de Cronbach y el índice de *fiabilidad compuesta* de Fornell y Larcker (1981). Sin embargo, antes de calcularlas se realizó un análisis *factorial exploratorio* con el propósito de comprobar nuestros planteamientos teóricos iniciales, en cuanto a que la formalización se explica por dos dimensiones o factores –formalización organizativa y formalización profesional–, los cuales son medidos por ítems claramente diferenciados. En este sentido, los resultados del análisis factorial exploratorio confirmaron nuestros planteamientos; es decir, la formalización está integrada por dos factores que explican el 69,18 % de la varianza total de los datos²⁴. Los resultados se recogen en el cuadro 5, y representan la matriz de componentes principales rotados.

Cuadro 5 Matriz de Componentes Rotados

Ítems de la Escala	Componentes	
	F1	F2
V1	0,785	-0,142
V2	0,808	-0,321
V3	0,811	-0,322
V4	0,737	-0,277
V5	0,709	-0,087
V6	0,615	0,076
V7	0,508	0,512
V8.	0,468	0,697
V9	0,387	0,684

²⁴ La medida de adecuación muestral, test de KMO, toma un valor de 0,805, y la prueba de esfericidad de Bartlett: $\chi^2 (36 gl)= 423,144$, $p=0,000$. Los valores individuales de la diagonal de la matriz de correlación anti-imagen son todos superiores a 0,5. Valores que permiten la aplicación del análisis factorial exploratorio a la muestra. Se utilizó el método de extracción análisis de componentes principales y el método de rotación de normalización Varimax con Kaiser, convergiendo en tres iteraciones.

Vemos en el cuadro 5 anterior que se forman dos grupos de ítems los cuales evalúan dos factores –variables latentes, F1 y F2–. Los ítems de V1 a V6 convergen sus cargas factoriales con mayor peso hacia el factor *formalización organizativa* (F1), y V7 a V9 lo hacen hacia la *formalización profesional* (F2). De este modo, estas agrupaciones responden a una coherencia conceptual sustentada en el marco teórico de la presente investigación.

En cuanto a la fiabilidad de la escala en sí, ha sido calculada por medio del *alpha* Cronbach y haciendo uso del programa estadístico SPSS. Sus resultados se recogen en el cuadro 6.

**Cuadro 6 Fiabilidad de la Escala de Medición de la
Formalización:
Organizativa F1) y Profesional (F2)**

Ítems	Media de la escala si se elimina el ítem	Varianza de la escala si se elimina el ítem	Correlación ítem-total corregida	<i>Alpha</i> si se elimina el ítem
V1	19,98	7,965	0,677	0,829
V2	20,03	7,738	0,755	0,813
V3	20,03	7,912	0,765	0,811
V4	19,81	8,799	0,677	0,831
V5	19,82	9,419	0,591	0,846
V6	19,93	9,004	0,465	0,869
α de Cronbach Formalización Organizativa (F1): 0,858				
V7	7,03	2,095	0,458	0,647
V8	7,44	1,657	0,538	0,551
V9	6,81	2,051	0,520	0,575
α de Cronbach Formalización Profesional (F2): 0,687				

Como puede apreciarse, el coeficiente *alpha* de Cronbach para la escala que mide el factor F1 o variable latente formalización organizativa alcanza un valor de 0,858, cifra adecuada en relación con los parámetros establecidos por la literatura. Se observa también en el cuadro 6 que al excluirse el ítem V6, se incrementa su fiabilidad a 0,869; sin embargo, no consideramos esta opción puesto que la pequeña mejoría observada en la fiabilidad, no justifica la posible pérdida de información debido a la eliminación de un ítem.

Por otra parte en el mismo cuadro 6, la fiabilidad correspondiente a la formalización profesional (F2) es 0,687, es decir, un valor algo menor al mínimo recomendado por la literatura de 0,7 pero muy próximo a él. No obstante, opinamos que con ese nivel de la

fiabilidad se puede mantener la escala debido al carácter especialmente exploratorio que tiene en esta investigación.

Finalmente, y en vista de que la variable *formalización del comportamiento* se explica a través de dos variables latentes o factores cuantificados por escalas diferentes, se procedió a calcular el índice de fiabilidad compuesta –IFC–. En este sentido, la fiabilidad compuesta –IFC1– para la escala de la formalización organizativa es de 0,864, mientras que para la formalización profesional –IFC2– de 0,693.

Vemos, pues, el primer índice de fiabilidad compuesta (IFC1) muy superior al mínimo de 0,7 recomendado por la literatura, reafirmándose la fiabilidad. No obstante, la fiabilidad compuesta para la formalización profesional (IFC2) se mantiene algo baja, tal como fue detectado en el análisis del coeficiente Cronbach, pero superior a este último y más próximo a 0,7. En síntesis, y de acuerdo a los resultados obtenidos, consideramos mantener todos los ítems que miden la formalización profesional.

El cálculo de la fiabilidad debe ser un procedimiento indispensable en cualquier investigación de carácter cuantitativo. Sin embargo, queda una pregunta pendiente: ¿se mide el *constructo* o concepto que realmente se quiere medir? De esta pregunta, entonces, se deriva la necesidad cierta de verificar la *validez* de una escala. Así, la validez se analizó por medio de la *validez de contenido* y la *validez de constructo*. La primera, se llevó a cabo en el paso 4 de la metodología –ver figura 1–, mediante la participación de un grupo de expertos en el área de organización de empresa de la Universidad de Valencia, España. Por otra parte, la validez de constructo se verificó a través del *enfoque convergente*; recordando que la escala de la formalización en sus dos dimensiones es válida solo si evalúa realmente el grado de formalización en la empresa y no otra variable de diseño organizativo para la cual no fue diseñada²⁵.

Un manera de calcular la *validez convergente*, es a través del análisis del test *t* de las cargas factoriales de los ítems que cuantifican a cada una de las variables latentes o *constructos*, de manera tal que si cargan de forma positiva y significativa sobre el *constructo*, entonces se confirma la validez convergente (Anderson y Gerbing, 1988; Bagozzi y Yi, 1988; y Kim et al., 2003). Asimismo, Hair et al. (1999) sugieren un nivel de carga factorial igual o superior a 0,40.

El cálculo de las cargas factoriales se realizó a través del método estadístico análisis factorial confirmatorio del modelo de medida del conjunto de *constructos* objeto de estudio,

apoyándonos para su estimación el programa EQS 6.1. En este sentido, existe validez convergente si todas las cargas factoriales de las variables manifiestas o ítems que miden el mismo *constructo* son estadísticamente significativas (Anderson y Gerbing, 1988).

Dicha estimación se llevó a cabo por el método de máxima verosimilitud –maximun likelihood–, analizándose conjuntamente algunos estadísticos que indican la calidad o la bondad del ajuste del modelo de medida, tales como el χ^2 (*chi cuadrado*) y otros *ad hoc* propuestos por diferentes autores (Marsh et al, 1988; Bentler, 1990; Bentler y Wu, 1993)²⁶. Siendo los estadísticos más utilizados: *Normed Fit Index* –NFI–, *Non-Normed Fit Index* –NNFI–, *Comparative Fit Index* –CFI–, y LISREL: *Goodness of Fit Index* –GFI– y *Adjusted Goodness of Fit Index* –AGFI–. Estos estadísticos deben alcanzar valores superiores o muy cercanos a 0,9 para considerar adecuada la bondad del ajuste del modelo (Bentler, 1990; Schumacker y Lomax, 2004; Uriel y Aldás, 2005).

De esta manera en la primera estimación del modelo de la variable formalización en sus dos dimensiones –formalización organizativa y formalización profesional–, los estadísticos de la bondad del ajuste resultaron muy insatisfactorios ya que NFI, NNFI, LISREL GF y LISREL AGFI alcanzaron valores por debajo 0,9 –ver cuadro 7–. En respuesta a estos resultados, el Test del Multiplicador de Lagrange²⁷ sugirió establecer la relación causal entre el indicador V6 –grado en el que los objetivos y metas de la organización están claramente especificados– con el factor F2 –*formalización profesional* (ver figura 3). Sin embargo, esta sugerencia no la consideramos pertinente puesto que no se encontró ningún tipo de sustento teórico para fundamentarla. En todo caso, se decidió eliminar la relación entre V6 y F1 –*formalización organizativa*– y se procedió a realizar una segunda estimación del modelo de medida.

En esta segunda estimación, los estadísticos de la bondad del ajuste mejoraron de manera significativa. En el cuadro 7, se comparan los estadísticos de la bondad del ajuste correspondiente a cada una de las estimaciones realizadas.

²⁵ Nos referimos, por ejemplo, a las variables centralización y/o descentralización, especialización, formalización, ámbito de control, sistemas de incentivos, entre otras.

²⁶ El estadístico χ^2 (*chi-square*) de bondad del ajuste de un modelo de medida, permite contrastar la hipótesis nula de que el modelo es correcto. Valores altos de χ^2 en comparación con los grados de libertad (*gl*), es un indicador de un mal ajuste. Así, un modelo correcto predice adecuadamente la realidad, es decir, conduce a diferencias reducidas y aleatorias entre las varianzas y covarianzas observadas y las implicadas por el modelo (Batista y Coenders, 2000). Esta hipótesis debería poder ser aceptada, es decir, un χ^2 no significativo ($P>0,05$).

²⁷ El contraste del multiplicador de Lagrange, permite evaluar la mejoría en los estadísticos de la bondad del ajuste del modelo de medida al añadirle una relación causal o una nueva covarianza. (Uriel y Aldás, 2005).

Cuadro 7 Comparación de los Estadísticos de la Bondad del Ajuste del Modelo de Medida Formalización

	χ^2	GL	NFI	NNFI	CFI	LISREL GFI	LISREL AGFI
1ra. estimación	65,25	26	0,851	0,865	0,902	0,890	0,809
2da. estimación	38,25	19	0,901	0,921	0,946	0,924	0,860

Vemos en el cuadro anterior que la relación entre el *chi-square* y los grados de libertad disminuye entre la primera y la segunda estimación. En el primer caso, esa relación es de 2,51 (65,25/26) y en el segundo de 2,01 (38,25/19); es decir, muy próximo al valor máximo de 2 sugerido por la literatura. De igual forma, se observa en la segunda estimación que los estadísticos NFI, NNFI, CFI y LISREL GFI son superiores a 0,9 y el LISREL AGFI cercano a ese valor. Además, hay que indicar que el Test del Multiplicador de Lagrange y el Test Wald²⁸ no sugirieron cambios adicionales en las relaciones establecidas en el modelo de medida. En resumen, eliminando la relación entre el indicador V6 y el factor F1 (ver figura 3) el modelo de medida alcanza una adecuada bondad del ajuste.

Por otra parte, como una manera de verificación, se realizó nuevamente un análisis factorial exploratorio para determinar si hubo algún cambio en la conformación de los grupos de ítems y, en consecuencia, de los factores o dimensiones que explican el *contructo* latente formalización del comportamiento. Ciertamente, los resultados siguen coincidiendo con nuestros planteamientos teóricos iniciales, en cuanto a que los ítems o indicadores de la formalización están integrados en dos factores tal como se recoge en el cuadro 8. De igual manera, para observar el efecto sobre la fiabilidad de la eliminación del indicador V6, se calculó nuevamente el α de Cronbach. Dicho estadístico tuvo un incremento de 0,858 a 0,869; es decir, una leve mejoría.

En vista de los resultados anteriores y que el programa EQS, a través del Test Lagrange y el Test Wald, no sugirió ningún otro cambio en el modelo de medida, se procedió a evaluar la *valididad convergente* por medio del Test *t* de las cargas factoriales basándonos también en el método estadístico análisis factorial confirmatorio. Debemos indicar que para estos cálculos se excluyó el ítem V6 de la escala tal y como lo sugirió el Test de Lagrange en la primera estimación de la bondad del ajuste. En el cuadro 8 se muestran los resultados del análisis factorial confirmatorio.

²⁸ El contraste de Wald se aplica para cuestionarse si deberían suprimirse algunos de los parámetros existentes en el modelo de medida (Uriel y Aldás, 2005).

Cuadro 8 Análisis Factorial Confirmatorio de la Escala de Medición de la Formalización: Organizativa (F1) y Profesional (F2)

Parámetro	Carga Factorial No Estandarizada	Valor de t	Carga Factorial Estandarizada
L1 (V1) → F1	0,653***	9.598	0,780
L2 (V2) → F1	0,717***	11.365	0,874
L3 (V3) → F1	0,655***	10.735	0,842
L4 (V4) → F1	0,441***	7.825	0,672
L5 (V5) → F1	0,356***	6.961	0,613
L7 (V7) → F2	0,478***	5.803	0,606
L8 (V8) → F2	0,641***	6.591	0,700
L9 (V9) →< F2	0,501***	6.253	0,658
Significatividad:		Estadísticos de la bondad del ajuste:	
* =p<0,05 para t>1,960		- NFI= 0,901	- NNFI= 0,921
** =p<0,01 para t>2,576		- CFI= 0,946	
***=p<0,001 para t>3,291		- LISREL GFI= 0,924	- LISREL AGFI= 0,860
		- Standardized RMR= 0,059	- χ^2 (gl=19)=38,25, p=0,005

En el cuadro anterior se observa que el valor del estadístico *t* es significativo para $p>0,001$ en todos los casos, y las cargas factoriales estandarizadas superiores a 0,4 –mínimo sugerido por Hair *et al.* (1999). En síntesis, con estos resultados se demuestra la validez convergente de las escalas que cuantifican la variable formalización del comportamiento en sus dos dimensiones, quedando definitivamente conformada la escala para la formalización organizativa por 5 ítems, y por 3 la correspondiente a la formalización profesional.

En el siguiente cuadro 9, se ofrece un resumen de la depuración de las escalas una vez realizado el cálculo de la fiabilidad y la validez de *constructo*.

Cuadro 9 Resultados de la Depuración de las Escalas de Medición

Variables Latentes (Factor)	Ítems antes de depurar	Depuración de ítems en la fiabilidad	Depuración de ítems en la validez convergente	Total ítems final
Formalización Organizativa (F1)	V1,V2,V3,V4, V5,V6	No se eliminan	Se elimina V6	5
Formalización Profesional (F2)	V7,V8,V9	No se eliminan	No se eliminan	3

5. A MODO DE CONCLUSIONES

La construcción de una escala de medición –instrumento de levantamiento de datos– responde a un conjunto de actividades que requieren de una rigurosidad metodológica con el propósito de que se alcancen adecuados niveles de validez y fiabilidad, y que, por tanto, quede habilitada para la investigación científica. En este sentido, la eficacia de un instrumento de medida está unida al efecto esperado de ella y esto depende esencialmente de dos aspectos fundamentales: la experiencia del investigador para dirigir y conducirse en el trabajo de campo, y la rigurosidad con la que se aborde la metodología para diseñar adecuadamente una escala.

Al respecto, el punto central del presente trabajo ha girado en torno a la propuesta de una metodología compuesta por once pasos, dentro del marco de una investigación de carácter cuantitativo, para la construcción de una escala de medición de variables o *constructos latentes*. Así, la metodología desarrollada se hizo operativa tomando como caso de estudio la escala de medición de la formalización del comportamiento en las organizaciones en sus dos dimensiones: *formalización organizativa* y *formalización profesional*.

Cabe destacar que, aunque dicha escala se sometió a un proceso riguroso respecto a su validez de contenido, al aplicar los métodos estadísticos para verificar su fiabilidad y validez de *constructo*, específicamente la validez convergente, se observó que uno de sus ítems, en concreto el V6 (Los objetivos y metas de la organización están claramente especificados - escritos o no- y son transmitidos a toda la organización), no cumplió con los parámetros estadísticos mínimos de calidad requeridos por la literatura especializada. En consecuencia, se excluyó ese ítem puesto que la información que aportaba podría convertirse en un factor de distorsión de la calidad de la escala y en posible fuente importante de errores de medición que, en consecuencia, podrían afectar negativamente el análisis e interpretación de los resultados.

De lo anterior se desprende la necesidad de que en una investigación de naturaleza cuantitativa se tomen en cuenta un conjunto de consideraciones técnicas y metodológicas para el diseño de una escala de medición, las cuales están destinadas a cuantificar variables complejas y no observables directamente de la realidad. Estas consideraciones no solo responden a las aportaciones conceptuales y teóricas más relevantes sobre el tema en cuestión, y que culmina con la aplicación de la validez de contenido, sino que además deben ir acompañadas de una evaluación posterior al trabajo de campo que permita completar de forma exitosa el ciclo de diseño de la escala de medición o cuestionario.

Lo dicho en el párrafo anterior, pasa por considerar una serie de métodos estadísticos que proporcionen una garantía razonable de que las variables objeto de estudio están cuantificadas por instrumentos suficientemente fiables y válidos; esto es, que el instrumento de levantamiento de información mida el constructo que realmente se espera de él. Siendo el objetivo último lograr un modelo teórico estadísticamente aceptable a los efectos del posterior análisis de los resultados, y que sea, además, una expresión bastante aproximada de la realidad. Por ello, la evaluación de la *fiabilidad* y *validez* de *constructo* deben convertirse en procedimientos y requisitos casi indispensables, y en práctica común, para llevar a cabo una investigación de carácter cuantitativo con adecuados niveles de rigurosidad y calidad científica.

Por otro lado, esta investigación tiene una aportación importante ya que proporciona una escala para evaluar la variable de diseño organizativo *formalización del comportamiento* en sus dos dimensiones: formalización organizativa y formalización profesional. Esta proposición ha sido suficientemente comprobada mediante la aplicación de un análisis factorial exploratorio, y la demostración de sus propiedades *psicométricas* a través del cálculo de los indicadores estadísticos asociados a la fiabilidad y validez.

Es relevante destacar el hecho de que no se encontró en la literatura investigación alguna de carácter cuantitativo dedicada a la evaluación de la formalización profesional en una organización y, por ende, de un cuestionario para tal fin; de ahí, en parte, el carácter y aportación original del presente trabajo²⁹.

Finalmente, se destaca la necesidad de continuar en esta línea de investigación para seguir profundizando en el desarrollo de técnicas y métodos que permitan elaborar instrumentos de medición cada vez más fiables y válidos. El presente estudio va dirigido a encontrar caminos para alcanzar tales propósitos.

REFERENCIAS

- Anderson, James y Gerbing, David. (1988). Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. **Psychological Bulletin**, 103 (3), 411-423.
- Batista, Joan y Coenders, Gallart. (2000). **Modelos de Ecuaciones Estructurales. Cuadernos de Estadística**. Madrid: Editorial La Muralla, S.A.

²⁹ No estamos haciendo referencia a la amplia literatura existente sobre la generación y gestión del conocimiento en las organizaciones.

- Bagozzi, Richard y Yi, Youjae. (1988). On the Evaluation of Structural Equation Models. **Journal of the Academy of Marketing Science**, **16**, 74-94.
- Baker, Douglas y Cullen, John. (1993). Administrative Reorganization and Configurational Context: The Contingency effects of Age, Size, and Change Size. **Academy of Management Journal**, **36** (6), 1251-1277.
- Bentler, Peter. (1990). Comparative Fit Indexes in Structural Models. **Psychological Bulletin**, **107** (2), 238-246.
- Bentler, Peter y Wu, Erick. (1993). **EQS/Windows User's Guide**. Los Angeles: BMDP Statistical Software.
- Bigné, Enrique. (2000). **La Investigación en Marketing: Etapas del Proceso**. Madrid, España.
- Blau, Peter. (1970). A Formal Theory of Differentiation in Organizations. **American Sociological Review**, **35** (2), 201-218.
- Child, John. (1972). "Organization Structure and Strategies of Control: A Replication of Aston Study". **Administrative Science Quarterly**, **18** (2), 163-177.
- Churchill, Gilbert y Peter, Paul. (1984). Research Design Effects on the Reliability of Rating Scales: A Meta-Analysis. **Journal of Marketing Research**, **21** (4), 360-375.
- Churchill, Gilbert. (1999). **Marketing Research: Methodological Foundation**. The Dryden Press, seventh edition.
- Comisión Europea. (2003). **Diario Oficial de la Unión Europea**. No. L-124: 36-41 (20/05/2003).
- Cronbach, Lee J. (1951). Coefficient alpha and the internal structure of test. **Psychometrika**, **16**, 297-334.
- Daft, Richard. (2000). **Teoría y Diseño de la Organización**. México: International Thomson Editores.
- De la Fuente, Juan, García-Tenorio, Jesús, Guerras, Luís y Hernández, Juan. (1997). **Diseño Organizativo de la Empresa**. Madrid: Editorial Civitas, S.A.
- Devlin, Susan; Dong, H. y Brown, Marbue. (1993). Selecting a scale for measuring quality. **Marketing Research**, **5** (3), 12-17.
- Domínguez, D. (2004). **Foro: Investigación Cuantitativa y Cualitativa. ¿Disyuntiva Onto-Epistemológica?**. Centro de Experimentación para el Aprendizaje Permanente (CEPAP).
- Donaldson, Lex. (2001). **The Contingency Theory of Organizations**. Thousand Oaks: Sage Publications.

- Fornell, Claes y Larcker, David. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. **Journal of Marketing Research**, **18** (1), 39-50.
- Gibson, James; Ivancevich, John y Donnelly, James. (2001). **Las Organizaciones. Comportamiento, Estructura y Procesos**. Santiago de Chile: McGraw-Hill.
- Grinyer, Peter y Ardekani-Yasay, Masoud. (1981). Strategy, Structure, Size and Bureaucracy. **Academy of Management Journal**, **24** (3), 471-486.
- Hage, Jerald. (1965). An Axiomatic Theory of Organizations. **Administrative Science Quarterly**, **10** (3), 289-320.
- Hage, Jerald y Aiken Michael. (1967). Relationship of Centralization to Other Structural Properties. **Administrative Science Quarterly**, **12** (1), 72-93.
- Hair, Joseph; Anderson, Ralph; Tatham, Ronald y Black, William. (1999). **Análisis Multivariante**. Madrid: Prentice Hall.
- Hall, Richard. (1996). **Organizations, Structures, Processes and Outcomes**. Nueva York: Prentice Hall International, Inc..
- Kim, Chulmin; Kim, Sounghie, Im, Subin y Chin, Changhoon. (2003). The effect of attitude and Perception on Consumer Complaint Intentions. **The Journal of Consumer Marketing**, **20** (4/4), 352-370.
- Mansfield, Roger. (1973). Bureaucracy and Centralization: An Examination of Organizational Structure. **Administrative Science Quarterly**, **18** (4), 477-488.
- Malhotra, Naresh. (2004). **Marketing Research: An Applied Orientation**. Upper Saddle River (NJ): Pearson Education.
- Marsh, Herbert; Balla, John y McDonald, Roderick. (1988). Goodness-of-Fit Indexes in Confirmatory Factor Analysis: The Effect of Sample Size. **Psychological Bulletin**, **103** (3), 391-410.
- Martínez Lorente, Ángel y Hernández Espallardo Miguel. (2006). La Fiabilidad y la Validez en las Escalas de Medida de Modelos de Dirección de Operaciones. **XVI Congreso Nacional de ACEDE**. Septiembre, 2006. Valencia , España.
- Miquel, Salvador; Bigné, Enrique; Antonio, Cuenca; Miquel, María y Levy, Jean-Perry. (1997). **Investigación de Mercados**. Madrid: McGraw-Hill/Interamericana de España.
- Moreno-Luzón, María; Peris, Fernando y González, Tomás. (2001). **La Gestión de la Calidad y Diseño de Organizaciones: Teoría y Estudio de Casos**. Madrid: Prentice Hall.
- Nunnally, Jum y Bernstein, Ira. (1995). **Teoría Psicométrica**. México: McGraw-Hill/Interamericana de México, S.A.

- Pardo, Antonio y Miguel Ruiz. (2002). **SPSS 11. Guía para el Análisis de Datos.** Madrid: McGraw-Hill Hispanoamericana.
- Peris, Fernando, González, Tomás y Méndez, Manuel. (2001). Organización, Diseño Organizativo y Conocimiento. Modelo Conceptual y modelo Analítico. **Quaderns de Treball, Facultat d' Economia de la Universitat de València, 116.**
- Peter, Paul y Churchill Gilbert. (1986). Relationships Among Research Design Choices and Psychometric Properties of Rating Scales: A Meta-Analysis. **Journal of Marketing Research, XXIII,** 1-10.
- Pierce, Jon y Delbecq, André. (1977). Organization Structure, Individual Attitudes and Innovation. **Academy of Management Review, 2** (1), 27-37.
- Pugh, Derek S; Hickson, D. J.; Hinings, C.R. y Turner, C. (1968). Dimensions of Organizational Structure. **Administrative Science Quarterly, 13** (1), 65-103.
- Pugh, Derek S.; Hickson, D. J. y Hinings, C. R. (1969). An Empirical Taxonomy of Structure of Work Organizations. **Administrative Science Quarterly, 14** (1), 115-126.
- Pugh, Derek S. (1973). Does Context Determine Form?. *Organizational Dynamics. Organization Theory.*
- Pugh, Derek S, y Hickson, D. J. (1976). **Organizational Structure in Its Context: The Aston Programmed I.** Saxon House Studies.
- Robbins, Stephen. (2005). **Administración.** México: Editorial Pearson Educación, México D.F.
- Sánchez, Manuel y Sarabia, Francisco. (1999). **Metodología para la Investigación en Marketing y Dirección de Empresas. Validez y Fiabilidad de Escalas.** Madrid: Ediciones Pirámide, S.A.
- Sarabia, Francisco. (1999). Construcción de Escalas de Medida. **Metodología para la Investigación en Marketing y Dirección de Empresas,** Capítulo 13. Madrid: Ediciones Pirámide, S.A.
- Schumacker, Randall y Lomax, Richard. (2004). **A Beginner's Guide to Structural Equation Modeling.** Nueva Jersey: Lawrence Erlbaum Associates, Inc.
- Spector, Paul E. (1992). **Summated rating scale construction: An introduction.** Sage University Paper Series on Quantitative Applications in the Social Sciences.
- Uriel, Ezequiel y Aldás, Joaquín. (2005). **Análisis Multivariante Aplicado.** Madrid: Thomson Editores, España.
- Vila, Natalia; Kuster, Iñés y Aldás, Joaquín. (2000). Desarrollo y Validación de Escalas de Medidas en Marketing. **Quaderns de Treball, 104.**