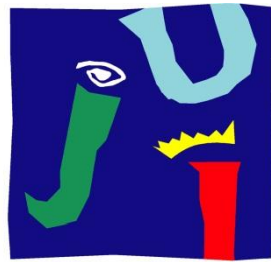


**UNIVERSITAT JAUME I DE CASTELLÓ**

**FACULTAD DE CIENCIAS HUMANAS Y SOCIALES**

**DEPARTAMENTO DE CIENCIAS DE LA COMUNICACIÓN**

**MEMORIA PARA LA OBTENCIÓN DEL GRADO DE DOCTOR**



**UNIVERSITAT  
JAUME·I**

**La reputación global corporativa y la equivalencia métrica de los índices de medición: una propuesta empírica a partir del análisis del índice *Reptrak™ Pulse*.**

**Doctorando: Ángel Alloza Losana  
Directores: Dra. Mayte Benloch Osuna  
Dr. Enrique Carreras Romero**

Castellón, 2015



## RESUMEN/ABSTRACT

Esta investigación pretende poner a prueba un indicador global de reputación corporativa. La tesis trata de comprobar la validez y robustez métricas de un indicador global de reputación concreto, el denominado *Reptrak TM Pulse* con el fin de poder incorporarlo a los criterios de medición y toma de decisiones en el ámbito de la organizaciones empresariales y de las instituciones en general.

En los últimos años, el concepto de la “reputación corporativa” está recabando un interés cada vez mayor tanto en el mundo académico como en el profesional. La reputación podría establecerse como uno de los indicadores no financieros con mayor potencial para convertirse en el más relevante en el mundo de las organizaciones, para influir en la estrategia, demostrando su vinculación directa a la creación de valor, aportando una perspectiva que concilie el corto y largo plazo, favoreciendo una visión *multi-stakeholder* y el alineamiento de todos los *stakeholders* internos y externos de la organización con su estrategia.

La reputación corporativa que está en la base de la creación de valor, desencadena los comportamientos de apoyo o de rechazo de los grupos de interés hacia las organizaciones.

Estamos ante un cambio de paradigma en el modelo económico que ha imperado hasta la actualidad y que se ha sustentado en una estrategia de maximización de beneficios para los accionistas, una estrategia de corto plazo y unos indicadores de éxito apalancados exclusivamente en indicadores financieros, un modelo económico que podría denominarse "Economía financiera". El nuevo paradigma en el que nos encontramos recibe diferentes nombres como "Capitalismo de largo plazo"; "Economía de los Intangibles" o el que hemos adoptado: "Economía de la Reputación".

Las empresas e instituciones están inmersas en la "Economía de la reputación" y necesitan con urgencia un modelo de gobernanza y de planificación estratégica dónde se deben complementar los indicadores financieros tradicionales con indicadores no financieros en especial necesitan disponer de un indicador global de reputación.

Este cambio de paradigma requiere del fortalecimiento conceptual, empírico y métrico de ese indicador global de reputación que debería alcanzar el mismo grado de aceptación, respeto y uso generalizado que tienen hoy los indicadores financieros.

Esta investigación pretende demostrar la validez y robustez de *Pulse* para ayudar a su incorporación a la estrategia de las empresas e instituciones y utilizar este indicador global de reputación para poder navegar con mayor éxito en la nueva "Economía de la reputación".

Palabras clave/key words: *reputación, Pulse, intangibles, métricas, cuadros de mando, stakeholders, modelos de ecuaciones estructurales, análisis de invarianza.*

## AGRADECIMIENTOS

La pasión por aprender, el deseo de investigar para descubrir nuevas realidades y por compartir los conocimientos es el motor que me ha movido a emprender un largo viaje a lo largo de toda la vida, un viaje con esfuerzo y dificultades marcado siempre por una enorme perseverancia.

Este trayecto que ha durado 4 años no lo he hecho en solitario. No podría haberlo iniciado ni continuado sin la inspiración, la motivación, la generosidad y la ayuda imprescindible de algunas personas que quiero mencionar y a las que quiero volver a agradecer el haberme guiado, acompañado y confiado en mi capacidad para emprender este viaje que generó en mi tanta incertidumbre en sus inicios.

En primer lugar quiero recordar a dos personas que ya no están con nosotros: Rafael López Lita que nos dejó el 25 de diciembre de 2011, a quien tuve el honor y el privilegio de tener como amigo y como maestro desde hace más de 20 años, cuándo me invitó a participar como profesor en el Máster de Publicidad de la UCE. Rafael López Lita me persuadió para iniciar este proyecto, empezando por hacer el Máster en Nuevas Tendencias y Procesos de Innovación en Comunicación que él dirigía y de cuyos profesores tanto he aprendido. Quiero mencionar en especial a Magdalena Mut que me acompañó y guió, junto a Mayte Benlloch, en el trabajo final del Máster que fue el paso previo y necesario para este trabajo cuyo tema central fue una inspiración de Rafael.

Mi recuerdo también para Ángel Falquina, otra de las personas clave en mi vida, a quien tanto quise y a quien tanto tengo que agradecer, que falleció el 4 de septiembre de 2003 su creatividad y espíritu explorador me han ayudado a encauzar este proyecto.

En segundo lugar, mi agradecimiento a María Teresa Benlloch y a Enrique Carreras, co-directores de este trabajo, sin cuya guía y ayuda constante y perseverante nunca habría llegado al final del trayecto.

A Mayte Benlloch por la paciencia, perseverancia, revisiones críticas y por todo el trabajo que para ella ha supuesto y que para mí ha tenido un impacto indispensable.

A Enrique con quién tanto he trabajado y compartido en estos años, en la redacción junto a Ana Carreras del libro que publicamos, *Reputación Corporativa*, un proceso de investigación y de redacción que fue determinante para definir con absoluta precisión el foco de esta tesis. Gracias por llevarme de la mano en el proceloso y complejo mundo de la matemática y la estadística más avanzada en la que Enrique es un gran maestro.

A mi equipo de trabajo, investigación y generación de conocimiento compartido, en la fundación Corporate Excellence-Centre for Reputation Leadership, Saida García, Clara Fontán, Augusto Leiva, Beatriz Magro y Anna Ramzina son el motivo por el que cada mañana abordo con ilusión y pasión este magnífico proyecto de cooperación, investigación aplicada, y avances en gestión de intangibles con la colaboración activa de



más de 80 profesores universitarios en todo el mundo. Estos trabajos han alimentado y enriquecido con una perspectiva empírica y rigurosa este proyecto de investigación.

y por último , mi agradecimiento , Leonard Ponzi, director de investigación del Reputation Institute y a Charles Fombrun y sin cuya generosidad y confianza nunca hubiera sido posible abordar este proyecto de investigación con las bases de datos internacionales necesarias a las que he podido tener acceso.



## INDICE TESIS DOCTORAL

### 0 INTRODUCCION. LA REPUTACIÓN EN EL ENTORNO ACADÉMICO E INSTITUCIONAL

INTRODUCCIÓN	11
0-1 LA NECESIDAD DE CONTAR CON INDICADORES Y MÉTRICAS DE REPUTACIÓN	17
0-2 LÍNEAS DE ACTUACIÓN BASADAS EN LAS PRÁCTICAS EMPRESARIALES Y EN LA INVESTIGACIÓN APLICADA	22
0-3 LA NUEVA FUNCIÓN DE GESTIÓN DE LA REPUTACIÓN CORPORATIVA EN LAS ORGANIZACIONES	30
0-4 OBJETIVOS Y FORMULACIÓN DE HIPÓTESIS	32
0-5 HACIA LA HIPÓTESIS DE LA INVESTIGACIÓN	37
0-6 FORMULACIÓN DE LA HIPÓTESIS	39
0-7 ESQUEMA Y DESCRIPCIÓN DE LA INVESTIGACIÓN	49
0-8 FUENTES DE LA INVESTIGACIÓN	52

### CAPÍTULO 1. LÍNEAS DE ACTUACIÓN ACADÉMICAS Y LÍNEAS DE ACTUACIÓN EMPRESARIALES.

APROXIMACIÓN A LAS MÉTRICAS DE REPUTACIÓN DESDE EL ÁMBITO EMPRESARIAL	53
1-1 LAS MÉTRICAS DE LA REPUTACIÓN CORPORATIVA	55
1-2 LA APROXIMACIÓN DIRECTA E INDIRECTA A LAS MÉTRICAS DE REPUTACIÓN	60
1-3 LAS MÉTRICAS DE APROXIMACIÓN INDIRECTA: LA REPUTACIÓN RACIONAL	70
1-4 LOS LÍMITES DE LA APROXIMACIÓN INDIRECTA	81
1-5 LAS MÉTRICAS DE APROXIMACIÓN DIRECTA: LA REPUTACIÓN GLOBAL	83
1-6 OTRAS MÉTRICAS DE LA REPUTACIÓN: AUDITORÍAS Y RANKINGS	91
1-7 LIMITACIONES DE LAS AUDITORÍAS Y LOS RANKINGS DE REPUTACIÓN PARA LOS OBJETIVOS DE NUESTRA INVESTIGACIÓN	95
1-8 IMPORTANCIA ESTRATÉGICA DE MÉTRICAS DIRECTAS SOBRE LAS INDIRECTAS	96
1-8 EL ÍNDICE PULSE COMO MÉTRICA DE APROXIMACIÓN DIRECTA A LA REPUTACIÓN: JUSTIFICACIONES DE ESTA INVESTIGACIÓN	99

### CAPÍTULO 2. FUNDAMENTACIÓN METODOLÓGICA.

2-1 CONSIDERACIONES GENERALES. FUNDAMENTOS METODOLÓGICOS	106
2-2 REVISIÓN DE LA HIPÓTESIS	111
2-3 METODOLOGÍA GENERAL. DISEÑO METODOLÓGICO	114
2-3.1 JUSTIFICACIÓN DE LA MUESTRA ESCOGIDA	115
2-4 EL ANÁLISIS DE LAS FUENTES DE DATOS	119
2-4.1 LAS ENCUESTAS PROCEDENTES DEL GLOBAL REPTRAKTM PULSE	120
2-4.2 MÉTODO DE COLECCIÓN Y SELECCIÓN DEL INFORMANTE	120

2-4.3	TAMAÑOS DE MUESTRA, PONDERACIÓN, NIVEL DE CONFIANZA Y MÁRGENES DE ERROR	122
2-4.4	TAMAÑOS Y MÁRGENES DE ERROR PARA LA BBDD EN ESPAÑA	124
2-4.5	BASE DE DATOS: REPRESENTATIVA DE LA POBLACIÓN DE EEUU	124
2-4.6	BASE DE DATOS REPRESENTATIVA DE LA POBLACIÓN DE INTERNAUTAS DE CHINA	125
2-4.7	DESCRIPCIÓN DE LA MUESTRA UTILIZADA EN ESPAÑA	126
2-4.8	DESCRIPCIÓN DE LA MUESTRA UTILIZADA EN EEUU	127
2-4.9	DESCRIPCIÓN DE LA MUESTRA UTILIZADA EN CHINA	128
2-5	TRATAMIENTO DE LA INFORMACIÓN PERDIDA EN LAS BASES DE DATOS	130
2-6	FORMACIÓN DE GRUPOS DE COMPARACIÓN PARA ANÁLISIS DE INVARIANZA	131
2-7	CUESTIONARIO PARA OBTENER EL ÍNDICE PULSE	133

### **CAPÍTULO 3. TÉCNICAS DE ANÁLISIS UTILIZADAS EN LA INVESTIGACIÓN.**

	SOBRE LAS TÉCNICAS DE ANÁLISIS	135
3-1	ANÁLISIS FACTORIAL EXPLORATORIO (EFA)	139
3-2	LOS MÉTODOS DE ECUACIONES ESTRUCTURALES (SEM)	148
3-3	MÉTRICAS FORMATIVAS VERSUS MÉTRICAS REFLEXIVAS	162
3-4	EL ANÁLISIS FACTORIAL CONFIRMATORIO (CFA)	171
3-5	EL PROBLEMA DE LOS MODELOS SEM: LA EQUIVALENCIA	176
3-6	TÉCNICAS PARA COMPROBAR LA UNI-DIMENSIONALIDAD DEL ÍNDICE PULSE	179
3-7	TÉCNICAS PARA COMPROBAR LA VALIDEZ CONVERGENTE	179
3-8	EL ANÁLISIS DE INVARIANZA POBLACIONAL	182
3-9	LA INVARIANZA PSICOMÉTRICA E INVARIANZA DE LAS MEDIAS	193

### **CAPÍTULO 4. RESULTADOS Y CONTRASTE DE LAS PRUEBAS MATEMÁTICAS DE VERIFICACIÓN DE LAS HIPÓTESIS.**

4-1	LA ESTIMACIÓN DE ADECUACIÓN MÉTRICA DEL ÍNDICE PULSE	196
4-2	VERIFICACIÓN DE LAS PROPIEDADES MÉTRICAS DEL ÍNDICE PULSE	204
4-3	LA VALIDEZ CONVERGENTE DEL ÍNDICE PULSE Y SU FIABILIDAD	206
4-4	EL TIPO DE RELACIÓN ENTRE LOS CUATRO ÍTEMS Y EL PULSE	208
4-5	EL GRADO DE INVARIANZA POBLACIONAL DEL ÍNDICE PULSE	215

### **CAPÍTULO 5. SÍNTESIS DE LAS PRUEBAS MATEMÁTICAS Y ESTADÍSTICAS Y CONTRASTE DE LAS HIPÓTESIS.**

5-1	SÍNTESIS DE LAS PRUEBAS MATEMÁTICAS Y ESTADÍSTICAS	262
5-2	CONTRASTE DE LAS HIPÓTESIS	265
5-3	LIMITACIONES DE LOS RESULTADOS DEL CONTRASTE DE LAS HIPÓTESIS	267

### **CAPÍTULO 6. CONCLUSIONES.**

6-1	CONCLUSIONES PRINCIPALES	268
6-2	CONCLUSIONES MÁS RELEVANTES	282
6-2	SÍNTESIS DE LAS CONCLUSIONES	286

## **CAPÍTULO 7. LÍNEAS FUTURAS DE INVESTIGACIÓN.**

7-1	PRINCIPALES LÍNEAS DE INVESTIGACIÓN	288
-----	-------------------------------------	-----

## **CAPÍTULO 8. GLOSARIO Y GUÍA MATEMÁTICA Y ESTADÍSTICA.**

8-1	PROPIEDADES PSICOMÉTRICAS DE UNA ESCALA MULTIDIMENSIONAL	292
8-2	MODOS DE OPERATIVIZAR LA VARIABLE LATENTE: REFLEXIVO Y FORMATIVO	293
8-3	ANÁLISIS DE INVARIANZA MULTIGRUPO	294
8-4	NOTAS SOBRE LA IMPLANTACIÓN DE UN ANÁLISIS DE INVARIANZA	297
8-5	TÉCNICAS DE MODELADO DE ECUACIONES ESTRUCTURALES (SEM)	304

## **CAPÍTULO 9. BIBLIOGRAFÍA.**

9-1	BIBLIOGRAFÍA CONSULTADA	307
-----	-------------------------	-----

## **ANEXOS.**

1	CÁLCULOS ESTADÍSTICOS REALIZADOS	324
---	----------------------------------	-----



## **0. INTRODUCCIÓN: LA REPUTACIÓN EN EL ENTORNO ACADÉMICO E INSTITUCIONAL.**

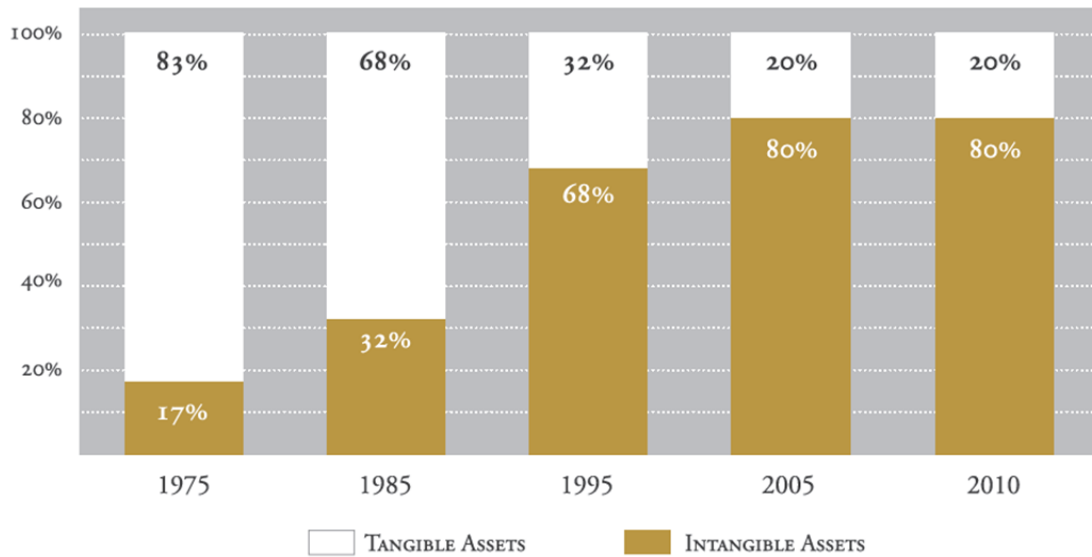
### **INTRODUCCIÓN.**

En los últimos 30 años se ha consolidado la tendencia de incremento constante del valor de los recursos y activos intangibles en las empresas.

Medir y gestionar estos recursos y activos intangibles nunca han sido tan importante como ahora. Representan una proporción cada vez más grande del valor de las empresas, en la medida en la que los resultados financieros y la rentabilidad responden cada vez más a la fortaleza de la relaciones y la confianza que depositan los grupos de interés en una organización, en las percepciones de una diferenciación duradera entre una empresa y otra, en la gestión del talento en las organizaciones, el flujo del conocimiento y de la información. Al mismo tiempo son cada vez menos relevantes para la consecución de los resultados económicos los recursos físicos y tangibles.

Para las empresas esta tendencia se convierte en urgencia cuando analizamos cómo han cambiado las fuentes de creación de valor y la proporción que los intangibles ocupan hoy en el valor presente y sobre todo en la proyección de su valor futuro (Álvarez, 2012b).

La urgencia de las empresas para adaptarse a esta nueva realidad queda patente en el cambio radical que vemos en los datos de los años 1975 hasta finales de la década de los 80, cuando los intangibles representaban únicamente el 20% del valor total de las organizaciones y lo comparamos con la situación actual en la que el 80% del valor agregado reside en los activos y recursos intangibles mientras que sólo el 20% corresponde a las estructuras y bienes tangibles (INTEGRATED REPORTING COUNCIL, 2013).



Source: Ocean Tomo

Gráfico nº 0. Proporción tangible e intangible del valor agregado de las empresas. Fuentes: Consultation draft of the international framework Integrated Reporting Initiative, 2013

Esta realidad contundente contrasta con la forma en la que se siguen gestionando las empresas en el mundo. La importancia y los recursos en los máximos órganos de gestión de las empresas se otorgan a las funciones que en el pasado se ocupaban de la parte de valor tangible de las organizaciones y aseguraban su éxito en el pasado (funciones de producción, logística, operaciones, área financiera, recursos humanos, etc.) siguen siendo las funciones que forman parte de la coalición dominante, las que detienen el poder y definen la estrategia de la empresa mientras que los gestores de los intangibles (teóricamente responsables del 80% del valor) no participan del mismo nivel de relevancia, recursos, herramientas, indicadores, métricas y poder de decisión que se dedican a los recursos tangibles (Van Riel, Alloza, et al. 2013; Mut, 2013).

Dentro de este contexto, la reputación corporativa es uno de los recursos intangibles de mayor relevancia, cuya gestión es de gran importancia y no hay gestión posible sin tener métricas e indicadores sólidos, consensuados como estándares, aceptados por los principales actores del mercado (empresas, inversores y reguladores) e internamente por la alta dirección de la empresa (Alloza, 2012b, 27; Álvarez, 2012a).



Esta es la situación que justifica la realización de esta investigación que pretende ayudar a avanzar en la definición de un indicador de reputación, riguroso desde el punto de vista académico, y al mismo tiempo razonablemente fácil de implantar en la empresa, en el menor tiempo posible. Un indicador que alcance un consenso y una legitimación suficiente como para ocupar una posición relevante en los cuadros de mando, junto con los tradicionales indicadores contables y financieros, que se utilizan para definir la estrategia y evaluar el grado de su cumplimiento en cualquier organización.

El fenómeno de la reputación corporativa requiere de estas métricas e indicadores consolidados puesto que está entrando de forma imparable en la agenda de los asuntos más relevantes para empresas y las instituciones en todo el mundo. Son seis los argumentos que están impulsando esta tendencia:

1. Los riesgos reputacionales
2. Los procesos de internacionalización
3. La necesidad de crear y mantener una diferenciación duradera en el tiempo y difícil de copiar.
4. La vinculación entre la reputación del país/lugar de origen de la empresa y su reputación y la de sus productos y servicios.
5. La necesidad de adoptar una visión de largo plazo y una estrategia que responda a las expectativas de los múltiples *stakeholders* de la empresa.
6. La necesidad consolidar la legitimación de la empresa que le permita conseguir y mantener sus licencias para operar.

Veamos cada uno de estos argumentos con mayor detenimiento y para esto citamos aquí los textos procedentes del trabajo de fin de Máster de Alloza (2012, 20-21) y del libro *Reputación Corporativa* (Carreras, Alloza, Carreras, 2013, 31-32).

*"1. El primer argumento fue la toma de conciencia de la importancia de los riesgos reputacionales para la supervivencia de las empresas (Alloza, 2012). El interés empresarial por la reputación corporativa se inició a comienzos de siglo partiendo de casos de crisis reputacionales como Enron, Tyco, Ahold, Parmalat y Arthur Andersen, que son paradigmas del efecto devastador del*

*llamado riesgo reputacional. Como consecuencia, empresarios de todo el mundo se dieron cuenta de que una reputación mal gestionada puede poner en peligro la continuidad empresarial (Fombrun, Ch., Riel, C.B.M. 2004, 155- 175).*

*2. El segundo argumento que ha puesto a la reputación en la agenda de las empresas es la internacionalización. Las empresas de cualquier país, incluidas las empresas españolas compiten en un mercado que tiende a ser cada vez más global. Incluso las firmas que se resisten a operar en un mercado local, se ven cada vez más obligadas a afrontar los retos de la competencia internacional de empresas y marcas que inundan los mercados más locales. Si sales fuera, no importa, ellos entrarán en tu reino (Alloza, A. 2002; Álvarez, C, Benlloch, M., Olivares, F. 2012).*

*3. El tercer argumento es la búsqueda de una diferenciación sostenible. En un contexto de progresiva globalización, las empresas españolas ya no pueden competir de manera sostenible únicamente por precio y, a largo plazo, ni tan siquiera por la calidad de la oferta (Anhold, S. 2002, 230). Los productos y servicios tienden a ser cada vez más parecidos y su calidad es una condición necesaria pero no suficiente. La verdadera diferenciación para conseguir una ventaja competitiva sostenible está en la marca (Alloza, A. 2010), en las experiencias que la marca entrega (Alloza et al. 2004) y en el estilo de relación, compromiso y alineamiento (Alloza, 2006; Oliver y Alloza, 2009) que permite establecer con todos sus stakeholders (Oroval, 2012 y Alloza, A. 2001,207-264). El profesor Cees Van Riel (2012) entiende que el éxito de la gestión empresarial descansa en su capacidad de diferenciarse a partir del fortalecimiento de las relaciones con los principales stakeholders, clientes, empleados e inversores y la consiguiente generación de valor compartido para la organización y para los propios stakeholders. En palabras llanas el autor defiende que la empresa necesita “construir esa diferenciación no copiable y ganarse la confianza y legitimidad de los grupos de interés para atraer clientes, talento, capital y licencias para operar”. Este punto de vista otorga un lugar central a la métrica de la reputación*

*corporativa que es el elemento básico sobre el que descansa la monitorización de la confianza.*

*4. El cuarto argumento radica en la vinculación de las empresas y las marcas a sus países de origen. El país de origen de un producto o servicio representa una parte significativa de su fortaleza o debilidad. La marca país actúa como respaldo de confianza para las marcas de producto o servicio provenientes del mismo (Alloza, A, 2009). Se produce un trasvase de atributos entre los que se asigna al país y los que se vinculan a las marcas de producto, a las empresas (Anholt, S., 2002,59-60). Este trasvase puede funcionar tanto en sentido positivo como negativo. Una marca o empresa procedente de un país percibido como sofisticado encontrará pocos problemas en ser percibida como sofisticada, pero por el mismo motivo una marca procedente de un país percibido como informal tendrá muchas dificultades para ser percibida como seria (Anholt, S. 2002,59-60). Las marcas y empresas exitosas y con buena reputación contribuyen a mejorar la percepción y reputación de la marca país. Esto funciona en especial con las grandes corporaciones que tienen gran visibilidad. Este efecto de las grandes empresas, a su vez, ayuda a mejorar la reputación de otras marcas y empresas más pequeñas con el mismo origen (Alloza, A. 2009).*

*5. El quinto argumento se basa en la idea de ayudar a las empresas y las organizaciones a adoptar un marco de actuación de largo plazo, con una visión multi-stakeholder. Gestionar la marca país y la reputación significa hablar de creación de valor a largo plazo, implica asumir un nuevo rol de la empresa al servicio de los grupos de interés, donde sólo la creación de valor compartido y equilibrado permitirá un crecimiento sostenible y capaz de generar una mayor cohesión social (Alloza, A., 2011,44-53).*

*6. El sexto argumento defiende la necesidad de toda organización de tener legitimidad suficiente para operar. La legitimidad de una empresa en el pasado era una cuestión puramente jurídica, hoy se pone el foco en la legitimación social que conceden los grupos de interés a las organizaciones que evalúan y como consecuencia de esa evaluación positiva o negativa,*

*adoptan conductas de apoyo u oposición (Carreras et al., 2013). La legitimidad social abarca a todos los grupos de interés capaces de aportar valor o de destruir valor. A la hora de explicar la legitimidad de una corporación, todos los grupos de interés entran en juego (Walker 2010). En consecuencia va más allá del cliente como principal fuente de valor para la empresa, reconociendo la contribución de otros colectivos como los empleados, los accionistas y la comunidad (Balmer y Greysser 2006; Bromley 2002; Vidaver-Cohen 2007; McMillan et al. 2004). El nuevo paradigma de la reputación amplía el viejo modelo de lealtad del cliente hacia un modelo más general basado en la lealtad del “stakeholder” (Carreras et al. 2013,320). La legitimación social de las organizaciones se entiende en términos de los argumentos que llevan a los stakeholders a apoyar o rechazar a un proveedor de productos y servicios. Precisamente, una de las aportaciones más recientes en el campo de la reputación ha consistido en formalizar teóricamente los principales factores que fomentan la decisión de apoyar o rechazar a una institución<sup>11</sup>.*

A partir de la teoría de la legitimación evaluativa (Suchman, 1995, 571-610), Carreras, Alloza y Carreras han formulado las cinco áreas principales que pueden llegar a despertar actitudes de apoyo en las audiencias y que constituyen el contenido empírico del constructo: área pragmática (evaluación del beneficio racional esperado); área de relación (evaluación de los vínculos personales de confianza con la empresa); área de capacidad y competencia (evaluación de la capacitación de la empresa); área de liderazgo (evaluación de la calidad de los directivos y su liderazgo); y finalmente área moral y de responsabilidad que evalúa el comportamiento moral y responsable de la organización (Carreras et al, 2013).

Esta clasificación es muy similar a la que se emplea en el modelo Reprtrak™ Pulse (Ponzi et al. 2011, 18), en el que el componente racional del modelo corresponde a la evaluación de siete grandes áreas de desempeño de la empresa que incluyen la

---

<sup>11</sup> Se trata de una aportación sustancial porque hasta la fecha no se sabía muy bien si las diferentes métricas captaban todos los motivos por los que las personas llegamos a desarrollar una disposición favorable o desfavorable hacia una institución.

calidad de la oferta comercial, la innovación, el entorno de trabajo/empleados, los resultados financieros así como el área de liderazgo (calidad de la gestión directiva), y el área moral y responsable (ética, buen gobierno y responsabilidad social). La clasificación se aproxima también de forma muy clara a la que emplea Justo Villafañe en su forma de entender los componentes de la reputación (Villafañe, 2004: pp.45).

### **0.1. LA NECESIDAD DE CONTAR CON INDICADORES Y MÉTRICAS CONSOLIDADAS DE REPUTACIÓN.**

A pesar del peso que tienen los seis argumentos que hemos señalado como pilares que demuestran la importancia de la reputación para la sostenibilidad y el éxito de una organización en el contexto de la economía actual, la difusión y disponibilidad de herramientas para medir y gestionar la reputación, en todos sus grupos de interés clave es más bien pobre.

Este desconocimiento por parte de las empresas y la carencia de una metodología consolidada para medir la reputación ha favorecido una gran proliferación de métricas y propuestas disponibles hoy para los gestores de las empresas. Sin embargo, la proliferación de métodos de medición de reputación deteriora su credibilidad, dificulta su comprensión y limita la capacidad del gestor para establecer un diálogo y una posición de poder e influencia equiparable al resto de funciones directivas que gobiernan la estrategia y la toma de las decisiones relevantes en la empresa.

Esta situación produce inevitablemente una gestión deficiente y limita tanto la correcta evaluación de los riesgos que ponen en peligro a la organización, como la identificación de las oportunidades de crecimiento y éxito que podría generar potencialmente una gestión excelente de la reputación.

En este contexto se sigue otorgando una importancia excesiva a la visión contable tradicional que se focaliza casi exclusivamente en el control de los activos tangibles puesto que la prioridad, los recursos dedicados y la relevancia para la

gestión se restringe muchas veces únicamente a aquello que se mide y solo mide aquello que comprende bien (Álvarez, 2012a) .

Por tanto los indicadores que dominan y configuran la estrategia empresarial quedan limitados a indicadores financieros y contables tradicionales que se aceptan como estándares que nadie cuestiona.

No obstante, la reputación sigue siendo el mejor exponente del recurso intangible de mayor valor estratégico con claros déficits de comprensión, medición y gestión que podría desarrollar todo su potencial si conseguimos avanzar en investigaciones como la que presentamos, que buscan consolidar una métrica de reputación que refleje con exactitud los sentimientos de admiración, respeto y confianza de los grupos de interés concernidos (empresas, inversores, reguladores)

Como después veremos en el marco teórico, estos sentimientos son responsables de los comportamientos de apoyo o de rechazo de los grupos de interés hacia la empresa y es la base sobre la que se podría construir una nueva forma de entender el rol de la empresa y transformar el modelo económico con elevado sesgo financiero que rige en la actualidad.

Algunos profesionales, entre los que se encuentra Alloza, y algunas empresas, agrupadas en asociaciones como el Foro de Reputación Corporativa, el Instituto de Análisis de los Intangibles o el más reciente laboratorio de ideas Corporate Excellence Centre for Reputation Leadership, defienden la idea de la necesidad de un cambio de paradigma en el modelo económico que ha imperado hasta la actualidad y que se sustenta en una estrategia de maximización de beneficios para los accionistas, una visión de corto plazo y unos indicadores de éxito apalancados exclusivamente en ítems financieros y contables. Esa visión limitada al corto plazo sólo explica cómo se ha ganado dinero en el pasado pero que no ayuda a la empresa a identificar claramente cómo se va a generar riqueza en el futuro.

El nacimiento del término específico “economía de la reputación” fue acuñado por primera vez en la decimoquinta edición de la conferencia internacional: "International Conference on Corporate Reputation, Brand, Identity and Competitiveness" en 2011 en Nueva Orleans (Estados Unidos). Se trata de un

término que ilustra el nuevo contexto económico y de competitividad en el que cambia el rol de la empresa y la relación de poderes tradicionales; implica entender que el poder hoy está en manos de todos los grupos de interés (Montañés et al., 2011, 43).

El poder efectivo se inclina a favor o en contra de las empresas en función de la confianza y de la buena reputación que otorgan estos grupos de interés a cada empresa e institución.

Este cambio de paradigma requiere, una vez más, del fortalecimiento conceptual, científico y métrico de los indicadores no financieros como el indicador de reputación corporativa que debería alcanzar el mismo grado de aceptación, respeto y uso generalizado que tienen hoy los indicadores financieros (Barton, D. 2011).

La justificación principal de esta investigación radica precisamente en la necesidad de contar, para la reputación corporativa, con un indicador general de reputación que pueda tener suficiente rigor y validez y que al mismo tiempo consiga ser ampliamente aceptado y utilizado por la comunidad empresarial, por los analistas e inversores así como por el regulador.

Si conseguimos avanzar en esta dirección, la reputación podría implantarse como uno de los indicadores no financieros con mayor potencial para convertirse en el más relevante en el mundo de las organizaciones, tanto privadas como públicas para influir en la estrategia, aportar una perspectiva de largo plazo y una visión *multi-stakeholder*.

Las métricas de reputación en este contexto se configuran como un requisito indispensable para asegurar el retorno económico y financiero de la buena gestión empresarial.

Siguiendo este argumento, las métricas de reputación deberían formar parte del cuadro de mando al más alto nivel de las organizaciones. Un cuadro de mando en el que se complementen los indicadores financieros convencionales con indicadores no financieros, como la reputación, la marca, la satisfacción de los clientes, el compromiso de los empleados y el grado en que empleados y clientes

están dispuestos a prescribir/recomendar una empresa y su oferta comercial (Alloza, 2012d).

A nuestro juicio, la incorporación de una métrica de reputación permitirá a las empresas y a las instituciones superar la visión de corto plazo que ofrecen los indicadores financieros potenciando una visión de largo plazo y un enfoque *multi-stakeholder* que posibiliten cambios profundos en la estrategia empresarial. Cambios orientados a consolidar un nuevo modelo de empresa sostenible, generador de diferenciación duradera, capaz de alinear a sus *stakeholders* internos y externos en torno a su estrategia de negocio (Alloza, 2013).

Esta sensibilidad ya está entrando en el mundo de la empresa. Según los resultados de una investigación global realizada en 2012 sobre la importancia relativa de los indicadores (KPI) en la gestión empresarial (tabla nº 0), para los directores de comunicación corporativa contar con un indicador de reputación parece ser una de las prioridades más importantes junto con el indicador de satisfacción de clientes.

Table 7: Most Important Key Performance Indicators (KPIs) in Organizations

KPIs	Early Phase	Advanced Phase	Significant Differences
Brand Awareness	19.0%	11.1%	NS
Brand Consideration	10.5%	11.1%	NS
Brand Preference	19.6%	26.7%	NS
Corporate Reputation	47.1%	57.8%	*
Customer Retention (or Loyalty)	30.4%	37.8%	NS
Customer Satisfaction	55.6%	57.8%	NS
Employee Alignment	4.2%	17.8%	*
Employee Engagement	48.4%	40.0%	NS
Media Share of Voice and/or Favorability	22.5%	11.1%	*
Net Promoter Score	12.4%	8.9%	NS
Word of Mouth (or Buzz)	7.2%	2.2%	NS

\* Differences are significant at  $p < .1$ ; NS=not significant

Tabla nº 0. Indicadores más importantes en las organizaciones. Source: Reputation Institute Global CRO Study, June 2012 (polling 351 senior reputation leaders at 318 major companies around the world about how their company perceives and manages reputation)



Estas razones vuelven a justificar la necesidad de esta investigación que ambiciona profundizar una de las métricas de reputación corporativa disponibles en la actualidad. En concreto trataremos de evaluar y analizar en qué medida uno de los índices presentes hoy en la literatura académica y en el mercado de los proveedores de métricas de reputación, se trata del "RepTrak <sup>TM</sup> Pulse", en adelante "Pulse" que podría tener, si se validan las hipótesis de esta investigación, la capacidad de dotar a las empresas y las instituciones de una tecnología válida y accesible para gestionar su reputación global.

## **0.2. LÍNEAS DE ACTUACIÓN EMPÍRICAS Y BASADAS EN LAS PRÁCTICAS EMPRESARIALES Y EN LA INVESTIGACIÓN APLICADA.**

Como señalamos al inicio de este capítulo, la segunda línea de trabajo en métricas de reputación que se ha puesto en marcha en los últimos años para tratar de obtener todo el potencial que encierra la reputación, ha correspondido a las prácticas empresariales y a la investigación aplicada.

Esta línea de actuación ha complementado y se ha nutrido de la investigación académica que hemos revisado en el epígrafe anterior, y en los casos que vamos a citar, impulsa y aporta recursos económicos e intelectuales para el desarrollo de la investigación académica.

Como han señalado los profesores Carreras y Alloza (Carreras, et al. 2013, 34-35), algunas de las experiencias más innovadoras han nacido en España hace diez años, con la creación de grupos empresariales que trabajan bajo el marco de modelos de “co-opetencia”, colaboración público-privada y mediante alianzas con algunos de los referentes académicos y científicos (Alloza, 2011, 44-53).

*"Los casos más relevantes en este campo han sido protagonizados por algunas de las mayores empresas que cotizan en el IBEX 35 y por algunas grandes empresas públicas y privadas en España.*

*Han sido tres las iniciativas empresariales que han tenido mayor impacto en este campo en España:*

- *El Foro de Reputación Corporativa (2002 - 2011).*
- *El Instituto de Análisis de los Intangibles (2004 - 2011).*
- *Corporate Excellence – Centre for Reputation Leadership (2011)."*

## **EL FORO DE REPUTACIÓN CORPORATIVA (2002-2011).**

El proceso de internacionalización de las grandes empresas españolas, a comienzos de la pasada década, ayudó a la toma de conciencia por parte de sus gestores del peso y relevancia de la reputación corporativa y de los intangibles. La experiencia durante los años 90 en América Latina, y a comienzos del 2000 en nuevos mercados con una fuerte competitividad basada en grandes marcas, como Estados Unidos y Europa, demostraban la validez de la tesis que defiende que construir una buena reputación es fundamental para competir con éxito en los mercados globales (Alloza y Noya, 2004, 11.).

El Foro de Reputación Corporativa (fRC) nace en 2002 de la mano de cuatro grandes empresas españolas (Agbar, BBVA, Telefónica y Repsol), Alloza es uno de los profesionales fundadores de esta asociación que se organiza para trabajar juntos constituyendo el primer grupo empresarial de esta naturaleza en el mundo. Su objetivo consistía en entender, medir y gestionar la reputación corporativa, y demostrar empíricamente su contribución al negocio y a la creación de valor. El proyecto se desarrolló rápidamente y llegó a agrupar a empresas muy relevantes en la economía española como Abertis, Adif, Agbar, BBVA, Critería, Danone, Ferrovial, Unión Fenosa, Iberdrola, Iberia, Renfe, Repsol, Sol Meliá, Telefónica, entre otras.

## **EL INSTITUTO DE ANÁLISIS DE INTANGIBLES (2004-2011).**

Mientras las empresas avanzaban en los modelos y herramientas de gestión de la reputación corporativa, los estudios académicos sobre otros intangibles se multiplicaban manifestando la ausencia de un consenso generalizado.

Ante la necesidad de crear una doctrina sólida sobre el valor de los intangibles y su gestión, desde España y junto a las empresas del fRC y otras organizaciones relevantes (entre las que había consultoras, escuelas de negocio, administración pública...) se creó en 2004 el Instituto de Análisis de los Intangibles (IAI), Alloza fue fundador y director general del IAI. El objetivo del IAI era dotar a las áreas de gestión de intangibles de instrumentos y parámetros de medición y monitorización de los activos intangibles, e integrarlos en los procesos de valoración y gestión que necesitan implantar las organizaciones (IAI publicaciones, 2004-2010).

Otra de las iniciativas promovidas desde el IAI fueron los premios de investigación sobre intangibles para impulsar la investigación en el ámbito académico; y el lanzamiento de sus actividades de formación en intangibles en 2007 del Programa Ejecutivo en Gestión y Análisis de Intangibles en colaboración con la Escuela de Organización Industrial (EOI), que desde su primera edición se convirtió en el referente de la formación de más de cien profesionales dedicados a su gestión.

El último hito del IAI fue el lanzamiento y publicación en 2008 del estudio sobre la reputación de España en el mundo a partir de la metodología *CountryRep* desarrollada por el Reputation Institute y en colaboración con el investigador principal en esta materia del Real Instituto Elcano, Javier Noya. Se publicó anualmente esta investigación del 2008 al 2011. Actualmente, este modelo de medición ha sido adoptado por el Alto Comisionado de la Marca España como uno de los métodos de medición que nutre al Observatorio de la Marca España que dirige el Real Instituto Elcano que cuenta con la colaboración técnica de Corporate Excellence (el autor de esta tesis es el investigador designado por el Real Instituto Elcano, Alloza, 2009).

La iniciativa empresarial más reciente, la creación de Corporate Excellence – Centre for Reputation Leadership (CE) ha integrado en su creación en 2011, a

todas las actividades y a la gran mayoría de las empresas que fundaron el Foro de Reputación Corporativa (fRC) y el Instituto de Análisis de los Intangibles (IAI), junto con sus redes y alianzas con expertos e instituciones académicas, el fRC y el IAI se han disuelto formalmente y han pasado a formar parte de CE.

Las empresas pronto se dieron cuenta de que la fórmula más efectiva para avanzar en la gestión de la reputación, la marca, la comunicación, los asuntos públicos y las métricas, suponía aunar esfuerzos en una nueva y única iniciativa empresarial en la que se integrasen las actividades y las empresas del Foro de Reputación Corporativa y del Instituto de Análisis de los Intangibles, junto con sus redes de expertos, consultoras y de instituciones académicas. De esta forma, en 2011, el Foro de Reputación Corporativa, con una perspectiva más empresarial, y el Instituto de Análisis de Intangibles, con una perspectiva más académica, se unen formando Corporate Excellence – Centre for Reputation Leadership.

**CORPORATE EXCELLENCE CENTRE FOR REPUTATION  
LEADERSHIP (2011).**

Se constituye en 2011 como un laboratorio de ideas sin ánimo de lucro en torno a la gestión de la reputación, la marca, la comunicación, los asuntos públicos, las métricas y la formación más avanzada en estas materias. Se trata de una fundación empresarial, declarada de interés público y adscrita al Protectorado del Ministerio de Cultura y Educación del Gobierno de España con el nombre "Fundación del Foro para la Reputación, Marca y Comunicación Corporativas y para los Intangibles" .

La fundación fue creada por BBVA, CaixaBank, Iberdrola, Repsol, Santander y Telefónica (como patronos natos) y a la que se suman relevantes empresas públicas y privadas como Adif, Agbar, Bankinter, Correos, Danone, El Corte Inglés, Gas Natural Fenosa, Mapfre, Renfe. En 2013, se suman a esta iniciativa Banco Popular y DKV Seguros Médicos, en 2014 el Grupo Volkswagen e Ibercaja en 2015. En conjunto, este grupo de empresas emplean a más de 1 millón de personas, están presentes en 100 países y suman una capitalización bursátil que representa más del 70% del IBEX 35. Alloza es director y secretario general de esta fundación.

Como se afirma en los Estatutos fundacionales (2011), el nombre de este laboratorio de ideas expresa la idea de innovación que encierra la gestión de la reputación. Es un centro de liderazgo reputacional (“Centre for Reputation Leadership”), que, mediante la gestión de la reputación, consigue transformar a las empresas en organizaciones excelentes (“Corporate Excellence”).

La reputación, entendida como el "conjunto de evaluaciones colectivas –suscitadas por la organización en sus grupos de interés– que motivan conductas capaces de generar valor" (Carreras, Alloza, (2013), se ha convertido en el camino que lleva a la excelencia. El nombre de este laboratorio de ideas expresa la idea de innovación que encierra la gestión de la reputación. Es un centro que promueve el liderazgo reputacional (Centre for Reputation Leadership) que, mediante la gestión de la

reputación, consigue transformar a las empresas en organizaciones excelentes (Corporate Excellence).

Las empresas que han decidido liderar por su buena reputación son aquellas que han introducido en su gestión una dinámica de escucha activa y de incorporación de las expectativas de sus grupos de interés. Someterse al escrutinio de los, a lo largo del tiempo y considerando de forma transversal a todas las áreas de gestión empresarial que fortalecen la reputación, es un proceso que inevitablemente conduce a la puesta en marcha de planes de mejora, de transformación interna que conducen la excelencia empresarial. Fieles a esta visión, las empresas que conforman esta organización afirman su posicionamiento público al emplear el lema corporativo *Leading by Reputation*, que responde a los retos de competencia del nuevo ciclo económico de la economía de los intangibles y de la reputación, en el que las empresas y las organizaciones competirán por ganarse la confianza y el favor de sus grupos de interés.

La visión de CE es: " Que todas las organizaciones a través de la gestión excelente de los intangibles contribuyan a mejorar la sociedad".

Corporate Excellence continúa, tras tres años de actividad, esforzándose en conseguir ser una referencia profesional, académica y ética en el ámbito internacional para ayudar a todas las empresas e instituciones a fortalecer su reputación mediante la gestión integrada de los intangibles y tiene los siguientes objetivos:

- Fortalecer la competitividad de las empresas a través de la gestión excelente e integrada de los intangibles.
- Demostrar la rentabilidad económica y financiera de los activos intangibles desarrollando métricas y modelos de gestión.
- Promover la cooperación público-privada para contribuir a la recuperación de la confianza.
- Fortalecer la cooperación y la buena sintonía entre el mundo académico y el mundo empresarial

El método de trabajo elegido e implantado por Corporate Excellence, se basa en la colaboración público-privada, en el establecimiento de alianzas con el mundo académico, con los mejores profesionales y consultores. Este modelo puede considerarse como una experiencia única que aporta valor para la generación de conocimiento y de innovación aplicados. Esta iniciativa empresarial está dando lugar a la creación de grupos similares en otros países del mundo que se inspiran en este modelo colaborativo y de co-opetencia creado en España.

Corporate Excellence genera conocimiento aplicado introduciendo nuevos modelos y herramientas de gestión de los intangibles en las organizaciones y compartiendo conocimiento en torno a la gestión de intangibles con el resto del tejido empresarial y el mundo académico en seis áreas de actividad: Reputación, Marca, Comunicación, Asuntos Públicos, Métricas y Formación. Es por tanto un modelo empresarial avanzado de "*Business intelligence*" pero conservando su identidad como organización sin ánimo de lucro (Alloza, 2012).

Podemos visualizar la historia cronológica de la evolución del Foro de Reputación Corporativa, el Instituto de Análisis de Intangibles y su integración final en Corporate Excellence Centre for Reputation Leadership, las instituciones en las que Alloza ha desarrollado y sigue llevando a cabo su labor investigadora.

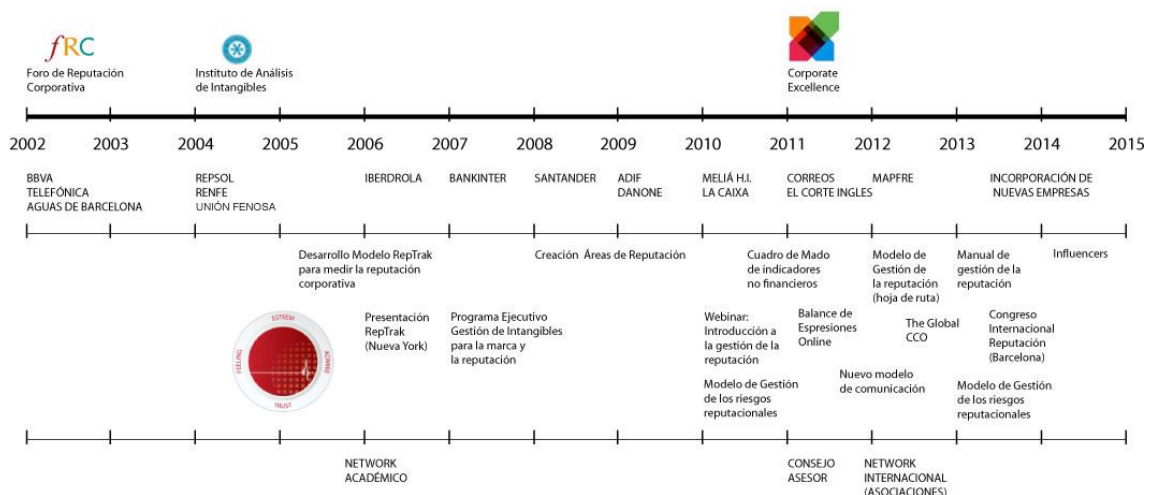


Figura 0.6. Fuente: Corporate Excellence. Historia de la organización.



En esta labor de investigación destaca la permanente colaboración con el ámbito académico desde el inicio de esta trayectoria en 2002. El modelo de investigación adoptado desde el principio ha sido el de transferencia de conocimiento entre el mundo académico y el mundo empresarial.

En 2014 esta labor conjunta se ha traducido en una colaboración activa de transferencia de conocimiento y de proyectos de investigación en reputación y otros intangibles de Alloza con las siguientes universidades (facultades de Comunicación y afines) y escuelas de negocio en España y otros países:

Universidades en España: Universidad de Alicante, Universidad Autónoma de Barcelona, Universidad de Cádiz, Universidad Cardenal Herrera Oria Ceu de Valencia, Universidad Complutense de Madrid, Universidad de Deusto, Universidad Internacional de Cataluña, Universidad de Extremadura, Universidad Francisco de Vitoria, Universitat Jaume I, Universidad de Málaga, Universidad de Murcia, Universidad de Navarra, Universidad de Nebrija, Universidad Pontificia de Salamanca Universidad San Pablo CEU, Universidad de Sevilla, Universidad de Valladolid, Universidad de Vigo.

Universidades en otros países: Pontificia Universidad Católica de Chile, Universidad de Plymouth, Universidad de São Paulo, Universidad Sergio Arboleda, Instituto Tecnológico de Monterrey (TEC).

Escuelas de Negocios en España: Esade, IESE, Instituto de Empresa, Esic, Deusto Business School.

El caso de Corporate Excellence ha sido recogido por el profesor Paul Argenti como un modelo de éxito en la buena gestión de los intangibles y de la reputación corporativa (Argenti, 2014, 59-72).

Este esfuerzo se ha complementado con el que realiza la asociación de profesionales más activa en este campo en España, DIRCOM (asociación de directivos de la comunicación), en Estados Unidos, la Arthur Page Society y en el ámbito internacional, la Global Alliance for public relations and communication management que agrupa a más de 160.000 profesionales.

### **0.3. LA NUEVA FUNCIÓN DE GESTIÓN DE LA REPUTACIÓN CORPORATIVA EN LAS ORGANIZACIONES.**

Estas iniciativas empresariales nacidas y sostenidas a lo largo del tiempo por la convicción de las empresas sobre la importancia estratégica de la reputación están generando en las organizaciones un marco organizativo y funcional desde el que poder gestionar la reputación, dentro del organigrama de las organizaciones, en el lugar que debe corresponder a su importancia.

*"Desde hace poco más de dos décadas, las Direcciones de Comunicación han sido los órganos encargados de asumir la tarea de medir y gestionar la reputación corporativa en un número cada vez mayor de grandes organizaciones e instituciones de todo tipo (RI, Reputation Leaders Study, 2014). Se trata de áreas de gestión, herederas de los departamentos de prensa que han evolucionado y redefinido el rol de la comunicación y del resto de intangibles estratégicos (Alloza et al. 2007). En la actualidad esta función se está consolidando en el mundo de la empresa a la vez que adquiere una importancia creciente (Arthur Page Society, 2012)". (Carreras et al, 2013, 32).*

Así la Dirección de Comunicación asume progresivamente nuevas funciones que se corresponden con cuatro elementos clave para determinar el éxito futuro de las empresas: la reputación, la marca corporativa, la comunicación, los asuntos públicos, la responsabilidad social (Benavides, 1997) y las métricas (Zerfass et al. 2014; Argenti et al. 2005; Argenti, 2014; Van Riel, Alloza et al. 2013).

La gestión excelente de estos intangibles está llamada a tener mayor relevancia en el proceso de transformación actual de las organizaciones que deseen afrontar con éxito el nuevo entorno competitivo de la "economía de los intangibles y de la reputación" (Alloza et al. 2007, 113-118).

Sin embargo, la posición todavía dominante hoy, en la mayoría de las empresas, establece que el máximo órgano directivo de una organización decide la estrategia empresarial y que ésta, a su vez y como en cascada, inspira la estrategia de las áreas subalternas entre las que figura la función de comunicación. Hasta la fecha, la práctica empresarial normalizada ratifica esta posición. La existencia la

dirección de comunicación y su presencia en los comités de dirección se da solamente en algunas grandes corporaciones, más sensibles a los nuevos retos (Mut, 2011, Mut et al. 2013).

En este sentido, la necesidad de gestión del intangible reputación que va a resultar clave en el éxito corporativo, exige que se desarrollen instrumentos y escalas métricas destinadas a medir de forma fiable, válida y sensible la trayectoria de este recurso intangibles tan importante. Como han apuntado algunos autores, la reputación como instrumento de gestión va a resultar “clave tanto para la formación como para la ejecución de la estrategia corporativa” (Fombrun y Van Riel, 1997, 5-13).

El director de comunicación, como responsable de la comunicación con todos los *stakeholders* (Cornelissen, 2009) debe conocer sus aspiraciones, demandas y expectativas y así contribuir eficazmente tanto a la formulación como a la implantación y la comunicación de la estrategia corporativa.

Este enfoque implicaría la necesidad de fortalecer la figura del director de comunicación o del directivo de primer nivel que tenga la responsabilidad de gestionar los intangibles de una organización, la tendencia en el contexto de las grandes corporaciones globales es que esta tarea recae cada vez más en las direcciones de comunicación (Zerfass, 2014; SpencerStuart y Weber Shandwichk 2014), pero sobre todo, dotarle de las herramientas de medida y evaluación de la reputación para legitimar su participación activa en la toma de decisiones corporativas de mayor relevancia.

Esta limitación también la encontramos en el mundo académico. En muchas universidades y escuelas de negocios, las materias sobre estrategia de comunicación o sobre reputación corporativa se imparten en departamentos de marketing y, en algunos casos, en departamentos de la especialidad, pero raramente en el ámbito de la política de empresa (Alameda, 2006). De la misma forma en las escuelas de negocios, la gestión de intangibles está prácticamente ausente de la formación ejecutiva, la destinada a los futuros directivos de primer y segundo nivel de las organizaciones e instituciones.

#### **0.4. OBJETIVOS Y FORMULACIÓN DE HIPÓTESIS.**

##### **OBJETIVOS.**

El indicador "RepTrak™ Pulse", en adelante Pulse es el objeto de estudio de esta investigación. Dado que el Pulse es una métrica latente, se precisa demostrar que es plenamente interpretable, es decir que las variaciones registradas reflejan cambios reales en los sentimientos globales de reputación y no cambios en la estructura de los colectivos implicados en la medida.

Vamos a poner a prueba el grado de **invarianza poblacional** por **sexo, edad y nivel de estudios** del índice de reputación global RepTrak™ Pulse para la población general, en tres contextos culturales distintos, España, EEUU y China.

Se trata esencialmente de poner a prueba la invarianza poblacional de Pulse y es por este motivo por el que utilizaremos bases de datos de Pulse pertenecientes a distintos países (España, EEUU y China), a diferentes segmentos de población (ambos sexos, diferentes edades y distintos niveles de estudios/clases sociales), así como referenciados a diferentes empresas de los sectores de banca y telecomunicaciones.

La selección de los tres países, los diferentes segmentos de población y los dos sectores de actividad obedecen a la intención de poner a prueba las hipótesis de la investigación en contextos socio-culturales y de madurez empresarial en la gestión de la reputación muy distintos (por este motivo la elección de estos tres países en concreto), de la misma forma la selección de banca y telecomunicaciones responde al objetivo de analizar dos sectores de actividad muy diferentes. Sin embargo, en un diseño ideal hubiéramos debido incluir un mayor número de países y de continentes para tener la máxima diversidad cultural posible en las muestras analizadas. Se trata de una limitación de la investigación debida exclusivamente a la disponibilidad de datos.

De la misma forma, la elección de dos sectores de actividad (banca y telecomunicaciones) obedece a los mismos motivos. En un diseño ideal hubiéramos debido incluir un mayor número de sectores de actividad para tener la

máxima varianza posible en las muestras analizadas. Se trata de una limitación de la investigación debida exclusivamente a la disponibilidad de los datos.

Como mencionamos en el apartado anterior, uno de los problemas del estado de desarrollo actual de las métricas de reputación es la carencia de un índice global que represente la reputación general de la compañía, institución, ciudad, país, etc.

Necesitamos una métrica o constructo único que sintetice la actitud global que mantienen todos los segmentos de población con independencia del sexo, la edad o la clase social, en lugar de contar, como hasta ahora, con las escalas racionales, con una métrica de reputación para cada segmento, de tal forma que la organización debería utilizar tantos indicadores de reputación como segmentos de población con los que se relaciona.

Aunque algunas de las escalas racionales de reputación han verificado su equivalencia semántica como es el caso del “Reputation Quotient” (Groeland, 2002) o para la escala de carácter corporativo (Davies et al. 2004), estos índices de reputación racional no pueden ser integrados en un constructo único que exprese la reputación global porque no cumplen con la equivalencia métrica (Chun y Davies 2006; Walsh et al. 2009), de esta forma las mediciones de reputación que se obtienen con cada segmento de población resultan ambiguas porque su comparabilidad no está garantizada.

En el caso de Pulse, la situación es diferente. Investigaciones recientes han verificado que algunos índices globales como el Pulse o el de Highhouse, han demostrado tener equivalencia semántica, pero sólo se han demostrado evidencias parciales de su potencial equivalencia métrica (Carreras et al. 2013). A este respecto, la investigación sobre Pulse de Leonard Ponzi, Charles Fombrun y Naomi Gardberg resulta prometedora, al haber demostrado la validez del índice Pulse en diecisiete países. Sin embargo, todavía no se ha demostrado la equivalencia métrica de Pulse, con lo que su integración en una medida única en los cuadros de mando de las empresas todavía está por verificar empíricamente.

Nuestro estudio busca precisamente poner a prueba la hipótesis de la equivalencia métrica de Pulse. De esta forma, si se corroboraran las hipótesis e investigación,

las organizaciones podrían contar con un único indicador global de reputación con equivalencia métrica para la población general, (1) con independencia del sexo, edad o nivel educativo de las personas que evalúan a una organización; (2) con independencia de las diferencias culturales (población general de distintos países) y (3) con independencia del sector de actividad de la organización a la que juzgan esta personas en términos reputacionales.

El objetivo de esta tesis es profundizar en esta línea para ofrecer una métrica global de reputación. Este es un requisito indispensable para poder integrar en el cuadro de mando del alto nivel de las organizaciones (el que establece los indicadores clave para diseñar y evaluar los resultados y recompensar a los directivos y empleados por el grado de cumplimiento de la estrategia general de la organización), un único indicador de la reputación de la organización.

Hemos seleccionado Pulse que ha validado en países y regiones culturales muy diversas (Ponzi et al. 2011), pero que hasta la fecha no ha sido sometido a la prueba de equivalencia métrica de una forma rigurosa que permita defender la solidez y validez de este indicador ante los diversos estamentos académicos y empresariales con el fin de que pueda ganarse la aceptación generalizada y un uso extenso por parte de la organizaciones (con el mismo nivel de aceptación que tienen algunos de los indicadores financieros considerados como estándares "de facto" por la comunidad empresarial, por los inversores y por una parte significativa de la academia).

La situación actual del índice Pulse es que ha sido validado parcialmente, pero no se ha demostrado su capacidad para reflejar de forma no ambigua las variaciones reales en las actitudes del *stakeholder*. A este propósito se dedica la presente investigación, se trata de analizar la "interpretabilidad" de este índice clave para la gestión actual de la reputación corporativa.

Los motivos por los que hemos seleccionado Pulse para esta investigación son por tanto múltiples:

- Se trata del indicador global de reputación que se ha implantado en más países (35 países), sectores de actividad y número de empresas en el mundo.
- Es un indicador accesible, en términos económicos y logísticos para cualquier empresa u organización puesto que se obtiene mediante la simple aplicación de un cuestionario cerrado muy breve que contiene sólo cuatro preguntas a las que se responde en una escala de siete puntos y el resultado se plasma en una escala de 1 a 100 puntos. La puntuación obtenida por una empresa expresa la fortaleza de su reputación en esta escala de 100 puntos.
- Es objeto de múltiples publicaciones en medios económicos y en publicaciones generalistas y académicas en forma de ranking de reputación en 15 países, midiendo 18 sectores de actividad y centenares de empresas. Las bases de datos disponibles son superiores a las de cualquier otro indicador global de reputación.
- Pulse ha sido objeto de investigaciones aplicadas en la gestión empresarial y en los programas de investigación y desarrollo realizados por Corporate Excellence Centre for Reputation Leadership, por parte de Alloza y de su equipo de colaboradores en los últimos ocho años que han conseguido demostrar la relación causal entre las mejores en la puntuación global de este indicador y los incrementos en el negocio en empresas de diferentes sectores de actividad (Estévez, 2011, 22-23).
- La actividad investigadora del Alloza y sus acuerdos de colaboración con otros investigadores, en particular con el equipo del profesor Ponzi del Reputation Institute, le ha permitido tener acceso a las bases de datos de Pulse de distintos países, sectores de actividad y segmentos de población para poder realizar esta investigación.
- Pulse ha sido utilizado también en los últimos cinco años para la medición de la reputación de los países por parte del equipo de investigación de Corporate Excellence. Alloza ha colaborado también con el equipo del Alto Comisionado de la Marca España y con el equipo investigador del Real

Instituto Elcano, a partir del año 2013, con el fin de incorporar a las métricas de seguimiento de la marca y de la imagen de España que utiliza el Gobierno y la Oficina del Alto Comisionado la metodología Reptar y el indicador Pulse aplicado a la medición de la reputación de España. Finalmente la decisión ha sido favorable y el Real Instituto Elcano ha adoptado desde el año 2014, entre las distintas métricas de seguimiento de la imagen y la marca España, la metodología Reprakt<sup>TM</sup> y el indicador agregado Pulse, el primer informe ha sido publicado el 3 de octubre de 2014<sup>2</sup>.

Estos motivos determinan que Pulse tiene el mayor potencial, frente a otros indicadores de reputación para convertirse en un referente que genere la necesaria amplia aceptación en el mundo empresarial y en el de los analistas e inversores para ser considerado como un estándar de facto para medir y ayudar a gestionar la reputación corporativa.

---

<sup>2</sup> [http://www.realinstitutoelcano.org/wps/portal/web/rielcano\\_es/contenido?WCM\\_GLOBAL\\_CONTEXT=/elcano/elcano\\_es/zonas\\_es/reputationinstitute-elcano-reputacion-de-espana-en-el-mundo-countryretrack-2014#.VE73YfmG-So](http://www.realinstitutoelcano.org/wps/portal/web/rielcano_es/contenido?WCM_GLOBAL_CONTEXT=/elcano/elcano_es/zonas_es/reputationinstitute-elcano-reputacion-de-espana-en-el-mundo-countryretrack-2014#.VE73YfmG-So)



## **0.5. HACIA LA HIPÓTESIS DE LA INVESTIGACIÓN.**

A nuestro juicio se trata de una investigación esencial para determinar el alcance y el potencial real del indicador Pulse que mide la reputación global para poder utilizarlo en la gestión, protección y fortalecimiento de la reputación corporativa.

La posibilidad de evaluar la invarianza demográfica del Pulse ayudará a los gestores a interpretar mejor la evolución real del índice, su comparación entre países y grupos socio-demográficos y medir con mayor precisión los efectos causales de las intervenciones y acciones implementadas para proteger o mejorar la reputación de las organizaciones sin las limitaciones que supondría la necesidad de disponer de sub-indicadores de reputación para cada uno de los segmentos de población con los que interactúa la empresa (segmentos de clientes, no clientes, etc.).

En efecto, el Pulse ha sido diseñado como una variable latente, se trata de un indicador que se construye de forma reflexiva. Este tipo de medidas conllevan un riesgo que conviene evaluar: se trata de analizar en qué medida los diferentes colectivos socio-demográficos entienden la reputación de forma diferente, las variaciones en dicho índice podrían estar reflejando tanto las pretendidas variaciones reales de la actitud hacia la corporación o la institución, como variaciones en la composición poblacional de los evaluadores. De esta forma, un ascenso en la reputación podría no estar reflejando una mejoría en la actitud favorable hacia la empresa sino un cambio poblacional, con colectivos que entienden la reputación de forma distinta.

Numerosos autores proponen que las escalas o índices latentes deberían demostrar el grado en que son insensibles a la composición de población, al menos por sexo y edad, de forma que las medidas obtenidas reflejen lo más fielmente posible las variaciones en la actitud real y no los cambios demográficos. Según Roger Millsap (2011), el análisis de "invarianza poblacional" es una prueba fuerte que debería acompañar cualquier medida psicométrica que representa un rasgo latente, como en el caso del Pulse.

Aunque los estudios de Leonard Ponzi y colegas han verificado en el 2011, la validez de la utilización del índice PULSE en diferentes contextos culturales, aun no ha quedado demostrada su “invarianza poblacional”. Se trata de una cuestión esencial para la interpretación adecuada de los resultados de la aplicación del índice y para la extensión generalizada del uso de este indicador sintético de reputación por parte de las empresas y las instituciones.

Esto es lo que queremos investigar en esta tesis como veremos en el epígrafe siguiente.

## 0.6. FORMULACIÓN DE LA HIPÓTESIS.

Precisamente nuestra investigación trata de dilucidar este problema verificando el grado en que el índice Pulse no varía en función del sexo, edad y nivel de estudios en diferentes países, y en diferentes sectores de actividad, a la hora de evaluar la reputación de las empresas (en términos más técnicos se trata de verificar la invarianza poblacional del índice Pulse).

Se trata de poner a prueba la hipótesis que nos permita afirmar o desmentir que Pulse alcanza un grado suficiente de invarianza poblacional por sexo y edad, para la población general, de modo que las puntuaciones obtenidas con Pulse para cualquier empresa reflejan lo más fielmente posible las actitudes subyacentes y no se deben a que pueda influir en el resultado la composición de la población entrevistada.

La hipótesis se puede formular de una forma sencilla y fácil de comprender:

*" Pulse mide la reputación global de una empresa u organización con independencia del sexo, edad o clase social de la población que sea entrevistada sobre esta cuestión" (invarianza poblacional).*

Sub-hipótesis:

- 1. "La puntuación resultante de Pulse refleja fielmente la fortaleza de esa reputación" (equivalencia y fortaleza métricas).*
- 2. "Los individuos de una población entienden lo mismo por reputación" (equivalencia semántica y contenido emocional).*

Desde un punto de vista matemático, Pulse es un constructo numérico expresado en una escala de 1 a 100 puntos, que se construye con las respuestas de la población entrevistada mediante la aplicación de un cuestionario estructurado con contiene las 4 preguntas con las que se calcula el indicador Pulse:

1. "Es una empresa en la que confío"
2. "Es una empresa que admiro"
3. "Es una empresa sobre la que tengo buena impresión"
4. "Tiene buena reputación"

En terminología más técnica y matemática la hipótesis principal y las sub-hipótesis de esta investigación se pueden formular de la siguiente forma:

Hipótesis principal:

- Vamos a intentar demostrar o refutar la "invarianza poblacional de Pulse".

Debido a que el índice Pulse adopta la forma de una variable latente evaluada por la combinación lineal de esos cuatro atributos, dispuestos en forma reflexiva <sup>3</sup>, antes de proceder a verificar el grado de invarianza poblacional, conviene comprobar previamente si cumple o no con dos requisitos matemáticos que harían de Pulse un indicador con equivalencia y fortaleza métricas. Estas pruebas de comprobación de las sub-hipótesis se realizan antes de abordar las correspondientes a la hipótesis principal, si no se cumplieran tendríamos que descartar la hipótesis principal.

Sub-Hipótesis:

1. La escala de Pulse cumple con la condición de uni-dimensionalidad interna.
2. La escala de Pulse tiene validez convergente.

---

<sup>3</sup> En los apartados "las propiedades psicométricas del índice reflexivo Pulse" y "El sistema completo: Los sub-sistemas de medida y estructural", desarrollamos este concepto.

Por tanto las primeras comprobaciones matemáticas y estadísticas van a poner a prueba las **propiedades métricas** de Pulse para demostrar o a refutar hasta qué grado Pulse supera las siguientes pruebas de verificación:

1. **Uni-dimensionalidad.** Que en el índice Puse subyace una sola dimensión que mide la reputación con independencia de que se realicen 4 preguntas distintas, todas miden el mismo concepto. La técnica más adecuada para identificar la cantidad de dimensiones subyacente a un conjunto de ítems/preguntas es el método de extracción por componente principales.
2. **Validez convergente.** La segunda cuestión es determinar el grado de convergencia de los cuatro ítems que demostraría que se está midiendo lo mismo en contextos culturales distintos (en los tres países testados). Esta parte requiere de la utilización de una tecnología diferente, el análisis factorial confirmatorio.

Si se verifican estas dos propiedades métricas entonces intentaremos comprobar la hipótesis principal:

3. **Invarianza poblacional.** La fundamentación de la invarianza poblacional (hipótesis principal de la investigación) se hará comprobando si Pulse cumple con los requisitos de equivalencia métrica de la escala. Este análisis consiste en la formulación y aplicación secuenciada de una serie de 4 modelos estadísticos “anidados” cuya contrastación va corroborando el nivel de invarianza alcanzado por el índice. Normalmente la secuencia de pruebas de verificación o modelos anidados es la siguiente:

3.1. La equivalencia configural. Los grupos de sexo, edad y nivel estudios utilizan los mismos ítems para construir el índice Pulse. Si se demuestra la equivalencia configural, todos los individuos con independencia del sexo, edad o nivel de estudios / clase social estarían utilizando los mismos términos para crear su puntuación global latente.

3.2. La equivalencia métrica de los ítems. Las cargas de cada ítem son las mismas en cada uno de los *stakeholders*. En este sentido, el peso de cada uno los cuatro términos sería el mismo para los todos los grupos de interés.

3.3. La equivalencia de la varianza del Pulse, que indicaría el grado en que el Pulse provoca el mismo rango de dispersión entre los diferentes grupos poblacionales.

3.4. La invarianza residual o estricta. Los residuos de los ítems son los mismos en los tres segmentos de población, indicando que la proporción explicada por el índice global en cada ítem es la misma en los tres segmentos de población.

Tipo de relaciones que se establecen entre los cuatro ítems del Pulse. Además de las propiedades métricas del Pulse conviene analizar el tipo de relaciones que se establecen entre los cuatro ítems que lo forman y con el índice reflexivo resultante (Pulse).

En principio cabría diferenciar entre modelos co-genéricos, en los que las cargas factoriales de cada ítem difieren entre colectivos, los modelos tau-equivalentes, en los que la contribución de cada ítem es la misma por población y por último, la relación paralela en la que las unicidades de los ítems son las mismas.

La puesta a prueba de las hipótesis de esta tesis requiere de la aplicación de diversas técnicas estadísticas y matemáticas entre las cuáles vamos a utilizar las técnicas de modelado de ecuaciones estructurales (SEM) que nos permitan aceptar o refutar la hipótesis.

El número y rigor de estas pruebas matemáticas es el necesario para demostrar con solidez y sin cuestionamiento la hipótesis principal de la investigación.

En el gráfico siguiente expresamos visualmente lo que hemos enunciado en este apartado.

### Objetivo de la investigación:

Objetivo de la investigación: Dado que el Pulse es una métrica latente, se precisa poner a prueba hasta qué grado se puede afirmar o refutar que es un indicador plenamente interpretable, es decir que las **variaciones registradas en este indicador reflejan cambios reales en la reputación global de una empresa o institución** y no cambios en la estructura de los segmentos de población que componen las muestras seleccionadas a las que se aplica mediante encuesta un cuestionario estructurado con cuatro ítems para calcular el Pulse, la métrica latente de reputación global.

Determinar el grado de **invarianza poblacional** por **sexo, edad y nivel de estudios** del índice de reputación global RepTrak™ Pulse para la población general, en tres contextos culturales distintos, España, EEUU y China

La investigación del grado de invarianza poblacional se desarrolla en tres fases en cada país:

#### 1. Medición de la adecuación del índice Pulse en tres países /contextos culturales distintos:

Comprobación de la adecuación de la métrica en España, EEUU y China

#### 2. Verificación de propiedades métricas de Pulse:

- Comprobación de la Unidimensionalidad
- Comprobación de la Validez convergente
- Comprobación de la Equivalencia semántica y prevalencia emocional

#### 3. Verificación de la invarianza poblacional del Pulse:

- Comprobación del grado de "invarianza psicométrica"
- Comprobación de la invarianza de los valores medios del Pulse

Gráfico nº 0.1. Objetivos y Pruebas de verificación de las hipótesis y su-hipótesis.

Las hipótesis de investigación se desarrollan en pruebas de verificación estadística a contrastar, con sus técnicas y umbrales de aceptación.

La primera condición que debería cumplir Pulse es la de adecuación del indicador para estimar la reputación global en tres países/contextos culturales diferentes. Las pruebas matemáticas y estadísticas a aplicar son las siguientes:

## HIPÓTESIS MATEMÁTICA H0. ADECUACIÓN MÉTRICA DE PULSE EN DISTINTOS CONTEXTOS GEOGRÁFICOS/CULTURALES (TRES PAÍSES).

- Análisis Factorial Confirmatorio (ACF).
- Estimación por máxima verosimilitud.
- Tratamiento de la no normalidad multivariada con las técnicas de Bootstrap y cálculo del p-valor por Bollen-Stein.

Para verificar la hipótesis de adecuación métrica se deben superar estas tres pruebas estadísticas.

## HIPÓTESIS MATEMÁTICA H1: UNI-DIMENSIONALIDAD DEL PULSE

- *H1.1. Uni-dimensionalidad de los 4 indicadores Pulse. Queremos comprobar si en el análisis por componentes principales, la varianza extraída por el primer factor es predominante, por encima del 80% en los tres países.*
- *H1.2. Validez **convergente** de los atributos. Formada a su vez por cuatro sub-hipótesis:*
  - *H1.2.1. Comprobación de la **fiabilidad** de cada uno de los cuatro atributos es al menos del 0,5 (Bagozzi y Yi 2012) en los tres países*
  - *H1.2.2. Cálculo del **índice de consistencia interna**, alfa de Cronbach, que debería ser superior al 0,6 (De Vellis 2012), en los tres países.*
  - *H1.2.3. Cálculo del **índice del promedio de la varianza extractada** (AVE) que debería ser superior al 0,5 (Fornell y Larker 1981) en los tres países.*
  - *H1.2.4. Cálculo del **índice de Dillon-Goldstein o fiabilidad del componente**, que debería ser superior al 0,6 (Bagozzi y Yi, 1988), en los tres países.*
- *H1.3. **Relación co-genérica de los indicadores con el constructo**. En los tres países se tendría que verificar que los coeficientes que vinculan los cuatro atributos con el Pulse son significativamente diferentes.*



Si conseguimos demostrar estas cualidades métricas del índice Pulse en las dos condiciones que enunciamos al principio de este epígrafe abordaríamos la verificación o refutación de la hipótesis principal.

Si se confirman estos requisitos previos, entonces las pruebas matemáticas y estadísticas que vamos a enunciar a continuación para comprobar la invarianza poblacional del Pulse por sexo, edad y nivel de educación. Estas pruebas deberán ser corroboradas o rechazadas en los tres países para los tres segmentos de población contemplados para poder aceptar como válida la hipótesis principal de la investigación.

## **HIPÓTESIS MATEMÁTICA H2: INVARIANZA POBLACIONAL DEL PULSE POR SEXO**

- *H2.1. Invarianza **configural por sexo**. El modelo Pulse estimado simultáneamente en varones y mujeres debe ser conforme a datos. El test sobre la función de ajuste global debe ser no significativo (el p-valor del  $\chi^2 > 0,05$ ).*
- *H2.2. Invarianza de **cargas por sexo**. En cada uno de los tres países se debe cumplir al menos una de las siguientes condiciones: (1) el modelo Pulse estimado simultáneamente en varones y mujeres bajo la igualdad de cargas respectivas debe ser conforme a datos (el p-valor del  $\chi^2 > 0,05$ ) o bien (2) la pérdida de ajuste respecto del modelo configural deber ser no significativa (  $\chi^2_{dif}$  con p-valor  $> 0,05$ ).*
- *H2.3. Invarianza de **la varianza extractada por el factor Pulse**. En cada uno de los tres países se debe cumplir al menos una de las siguientes condiciones: (1) el modelo Pulse estimado simultáneamente en varones y mujeres bajo la igualdad de cargas y varianza del Pulse, respectivas, debe ser conforme a datos (el p-valor del  $\chi^2 > 0,05$ ) o bien (2) la pérdida de ajuste respecto del modelo verdadero debe ser no significativa (  $\chi^2_{dif}$  con p-valor  $> 0,05$ ).*
- *H2.4 Invarianza **de los residuos de los ítems por sexo**. En cada uno de los tres países se debe cumplir al menos una de las siguientes condiciones:*

(1) el modelo Pulse estimado simultáneamente en varones y mujeres bajo la igualdad de cargas, varianza del Pulse e igualdad de residuos, respectivos, debe ser conforme a datos (el p-valor del  $\chi^2 > 0,05$ ) o bien (2) la pérdida de ajuste respecto del modelo verdadero debe ser no significativa ( $\chi^2_{dif}$  con p-valor  $> 0,05$ ).

### **HIPÓTESIS MATEMÁTICA H3: INVARIANZA POBLACIONAL DEL PULSE POR EDAD**

*H3.1 Invarianza configural por edad, El modelo Pulse estimado simultáneamente en jóvenes y adultos debe ser conforme a datos. El test sobre la función de ajuste global debe ser no significativo (el p-valor del  $\chi^2 > 0,05$ ).*

- *H3.2 Invarianza de cargas por edad. En cada uno de los tres países debe cumplir al menos una de las siguientes condiciones: (1) el modelo Pulse estimado simultáneamente en jóvenes y adultos, bajo la igualdad de cargas respectivas debe ser conforme a datos (el p-valor del  $\chi^2 > 0,05$ ) o bien (2) la pérdida de ajuste respecto del modelo configural debe ser no significativa ( $\chi^2_{dif}$  con p-valor  $> 0,05$ ).*

- *H3.3. Invarianza de la varianza extractada por el factor Pulse. En cada uno de los tres países debe cumplirse al menos una de las siguientes condiciones: (1) el modelo Pulse estimado simultáneamente en jóvenes y adultos, bajo la igualdad de cargas y varianza del Pulse, respectivas, debe ser conforme a datos (el p-valor del  $\chi^2 > 0,05$ ) o bien (2) la pérdida de ajuste respecto del modelo verdadero debe ser no significativa ( $\chi^2_{dif}$  con p-valor  $> 0,05$ ).*

- *H3.4 Invarianza de los residuos de los ítems por edad. En cada uno de los tres países se debe cumplir al menos una de las siguientes condiciones: (1) el modelo Pulse estimado simultáneamente en jóvenes y adultos bajo la igualdad de cargas, varianza del Pulse e igualdad de residuos, respectivos, debe ser conforme a datos (el p-valor del  $\chi^2 > 0,05$ ) o bien (2) la pérdida de ajuste respecto del modelo verdadero debe ser no significativa ( $\chi^2_{dif}$  con p-valor  $> 0,05$ ).*

**HIPÓTESIS MATEMÁTICA H4: INVARIANZA POBLACIONAL DEL PULSE POR NIVEL DE ESTUDIOS.**

- *H4.1 Invarianza configural por nivel de estudios. El modelo Pulse estimado simultáneamente en los grupos con estudios altos y bajos debe ser conforme a datos. El test sobre la función de ajuste global debe ser no significativo (el p-valor del  $\chi^2 > 0,05$ ).*
- *H4.2 Invarianza de cargas por edad. En cada uno de los tres países debe cumplirse al menos una de las siguientes condiciones: (1) el modelo Pulse estimado simultáneamente en estudios altos y bajos, bajo la igualdad de cargas respectivas debe ser conforme a datos (el p-valor del  $\chi^2 > 0,05$ ) o bien (2) la pérdida de ajuste respecto del modelo configural debe ser no significativa (  $\chi^2_{dif}$  con p-valor  $> 0,05$ ).*
- *H4.3. Invarianza de la varianza extractada por el factor Pulse. En cada uno de los tres países se debe cumplir al menos una de las siguientes condiciones: (1) el modelo Pulse estimado simultáneamente en estudios bajos y altos, bajo la igualdad de cargas y varianza del Pulse, respectivas, debe ser conforme a datos (el p-valor del  $\chi^2 > 0,05$ ) o bien (2) la pérdida de ajuste respecto del modelo verdadero debe ser no significativa (  $\chi^2_{dif}$  con p-valor  $> 0,05$ ).*
- *H4.4 Invarianza de los residuos de los ítems por nivel de estudios. En cada uno de los tres países se debe cumplir al menos una de las siguientes condiciones: (1) el modelo Pulse estimado simultáneamente en estudios bajos y altos, bajo la igualdad de cargas, varianza del Pulse e igualdad de residuos, respectivos, debe ser conforme a datos (el p-valor del  $\chi^2 > 0,05$ ) o bien (2) la pérdida de ajuste respecto del modelo verdadero debe ser no significativa (  $\chi^2_{dif}$  con p-valor  $> 0,05$ ).*

**HIPÓTESIS MATEMÁTICA H5: INVARIANZA POBLACIONAL DEL PULSE POR GRUPOS HOMOGÉNEOS DE SEXO, EDAD Y NIVEL DE ESTUDIOS.**

Hipótesis de invarianza por grupos homogéneos de sexo, edad y nivel de estudios. Esta serie de hipótesis se probará en España y EEUU. En China, debido a la

distribución concentrada del nivel de estudios en la muestra utilizada en el estudio, sólo se comprobará la combinación conjunta de sexo y edad.

- *H5.1. Invarianza configural por grupos homogéneos de sexo, edad y nivel de estudios. En los tres países se debe verificar que el test sobre la función de ajuste del modelo de grupos homogéneos es no significativa (el p-valor del  $\chi^2 > 0,05$ ).*
- *H5.2. Invarianza de cargas por grupos homogéneos. En cada uno de los tres países se debe cumplir al menos una de las siguientes condiciones: (1) el modelo Pulse de grupos homogéneos estimado simultáneamente, bajo la igualdad de cargas respectivas debe ser conforme a datos (el p-valor del  $\chi^2 > 0,05$ ) o bien (2) la pérdida de ajuste respecto del modelo configural debe ser no significativa (  $\chi^2_{dif}$  con p-valor  $> 0,05$ ).*
- *H5.3 Invarianza de la varianza del factor Pulse por grupos homogéneos. En cada uno de los tres países se debe cumplir al menos una de las siguientes condiciones: (1) el modelo Pulse de grupos homogéneos estimado simultáneamente, bajo la igualdad de cargas y varianza Pulse, respectivas debe ser conforme a datos (el p-valor del  $\chi^2 > 0,05$ ) o bien (2) la pérdida de ajuste respecto del modelo configural debe ser no significativa (  $\chi^2_{dif}$  con p-valor  $> 0,05$ ).*
- *H5.4 Invarianza de los residuos de los ítems del pulse. La igualdad en los residuales en cada indicador para cada grupo homogéneo no debe producir una pérdida significativa de ajuste. El  $\chi^2_{dif}$  respecto del modelo configural debe presentar un p-valor  $> 0,05$ .*
- *H5.4 Invarianza de los residuos de los indicadores Pulse por grupos homogéneos. En cada uno de los tres países se debe cumplir al menos una de las siguientes condiciones: (1) el modelo Pulse de grupos homogéneos estimado simultáneamente, bajo la igualdad de cargas, varianza Pulse y residuos, respectivos debe ser conforme a datos (el p-valor del  $\chi^2 > 0,05$ ) o bien (2) la pérdida de ajuste respecto del modelo configural debe ser no significativa (  $\chi^2_{dif}$  con p-valor  $> 0,05$ ).*

## 0.7. ESQUEMA Y DESCRIPCIÓN DE LA INVESTIGACIÓN.

La investigación que aquí se presenta se ha desarrollado desde finales del año 2010 hasta el momento actual, septiembre de 2014.

Las fases principales se recogen en el cronograma en la tabla siguiente.

PERIODO	OBJETIVO	PROCESO
2010 - 2012	Estado de la cuestión  Diseño de la investigación	<ul style="list-style-type: none"><li>- Realización del Máster en Nuevas Tendencias y Procesos de Innovación en Comunicación</li><li>- Revisión bibliográfica</li><li>- Realización y presentación del trabajo de investigación : "Métricas de Reputación Corporativa. Fundamentos conceptuales, metodologías y aplicaciones para la gestión en el marco de la función de Dirección de Comunicación Corporativa de las organizaciones".</li></ul>
2012 - 2013	Diseño del marco teórico y del encuadre estadístico y matemático	<ul style="list-style-type: none"><li>- Revisión bibliográfica específica</li><li>- Trabajo de fin de Máster</li><li>- Publicación del libro "Reputación Corporativa" (Carreras, Alloza, Carreras, 2013).</li><li>- Preparación del soporte informático y analítico</li></ul>
2014	Resultados y conclusiones	<ul style="list-style-type: none"><li>- Recepción y tratamiento estadístico de las bases de datos</li><li>- Aplicación de las pruebas estadísticas y matemáticas</li><li>- Contraste de hipótesis</li><li>- Nuevas líneas de investigación</li></ul>

Tabla 01. Cronograma de la investigación.

Siguiendo esta secuencia se ha organizado la investigación en 9 capítulos - además de esta Introducción y de un Anexo final.

La introducción contiene la justificación de la investigación, el interés del tema elegido y la motivación personal, académica y de gestión para llevar a cabo la investigación, propuesto los objetivos y enunciadas las hipótesis que guiarán la investigación.

El capítulo uno describe, el marco teórico, la fase de estudio documental de esta investigación. Una vez que en La primera parte de la investigación corresponde al marco teórico de la investigación en el que se ha realizado la revisión científica del concepto y de las métricas de reputación, desde el punto de vista académico y desde el punto de vista de la praxis empresarial. Se exponen las razones que llevaron a seleccionar el índice global de reputación RepTrak TM PULSE como variable focal de esta investigación.

La segunda parte de este documento, en el capítulo dos se recoge la fundamentación metodológica y el método seleccionado de análisis secundario de las bases de datos utilizadas, así como una descripción detallada de las características y ficha técnica completa de las muestras que contienen las bases de datos de los tres países seleccionados para la investigación y que serán sometidas a análisis.

El tercer apartado, capítulo tres, está dedicado a presentar las principales herramientas metodológicas y técnicas matemáticas y estadísticas utilizadas en la investigación. El análisis de componentes principales, ha servido para identificar las dimensiones subyacentes a los cuatro ítems del Pulse, el análisis factorial confirmatorio ha permitido implantar el índice en los seis contextos analizados correspondientes a los tres países y dos sectores de actividad empresarial. Después se presenta una introducción a los modelos de ecuaciones estructurales, en el que el análisis multi-grupo es una técnica especializada para probar las hipótesis centrales de la investigación como es la invarianza que denominaremos psicométrica de la escala RepTrak TM PULSE.

.En el capítulo cuatro se presentan los resultados y contraste de las pruebas matemáticas de verificación de las hipótesis.

En el capítulo cinco, se procede a contrastar cada una de las hipótesis y sub-hipótesis de la investigación y se sintetizan de las pruebas matemáticas y estadísticas.

El capítulo seis de conclusiones se extraen las conclusiones de la investigación y sus consecuencias en la aplicación empírica del índice RepTrak TM PULSE a la realidad de la gestión de la reputación en las organizaciones.

El capítulo siete propone líneas de investigación futuras.

El capítulo ocho recoge un glosario y guía matemática y estadística.

El capítulo nueve incluye la bibliografía utilizada y, el Anexo cierra el texto de la investigación con todos los cálculos matemáticos y estadísticos realizados durante la investigación y sus resultados correspondientes.

## **0.8. FUENTES DE LA INVESTIGACIÓN.**

Las fuentes específicas que fueron consultadas para la realización de esta investigación, principalmente en la fase documental y en la del diseño de la investigación proceden de cuatro fuentes:

1º. Fuentes personales: investigaciones del trabajo final del máster en Nuevas Tendencias y Procesos de Innovación en Comunicación e investigaciones personales realizadas en los últimos seis años desde la responsabilidad de dirección de estrategia de comunicación, marca, reputación corporativa y métricas para el Grupo BBVA; así como las investigaciones realizadas en el Foro de Reputación, el Instituto de Análisis de los Intangibles y la fundación Corporate Excellence - Centre for reputation Leadership. Las investigaciones y publicaciones realizadas por Alloza durante estos años han sido el principal valor aportado para el marco empírico y el marco analítico de esta investigación.

2º. Fuentes bibliográficas, procedentes de investigaciones, libros, informes, manuales, guías, cursos realizados por docentes, investigadores y profesionales del mundo empresarial e institucional, asociaciones, fundaciones y centros de investigación y desarrollo que han servido para acotar el objeto de estudio a la reputación y los principales activos y recursos intangibles de las organizaciones.

3ª. Artículos de revistas científicas y profesionales relacionadas con el ámbito de la reputación corporativa, las funciones de gestión de la comunicación corporativa y de los intangibles.

4º. Recursos bibliográficos en la red: revistas electrónicas, artículos de opinión, blogs, actas de jornadas, seminarios y congresos, materiales didácticos audiovisuales, y tesis doctorales alineados con el objeto de estudio de esta tesis.



## CAPITULO 1. MARCO TEÓRICO. LÍNEAS DE ACTUACIÓN ACADÉMICAS Y LÍNEAS DE ACTUACIÓN EMPRESARIALES.

### APROXIMACIÓN A LAS MÉTRICAS DE REPUTACIÓN DESDE EL ÁMBITO EMPRESARIAL.

Las líneas de trabajo en métricas de reputación que se han puesto en marcha en los últimos años para cumplir con estos retos y tratar de obtener todo el potencial que encierra la reputación y la gestión de la confianza han consistido esencialmente en dos tipos de actuaciones que fueron recogidas en la obra reciente escrita por Alloza y los profesores Carreras (Carreras, Alloza, y Carreras, 2013) :

*"1.1. Líneas de actuación basadas en la investigación y los trabajos académicos.*

*1.2. Líneas de actuación empíricas y basadas en las prácticas empresariales y en la investigación aplicada.*

*La primera línea ha generado investigaciones y avances en el concepto de comunicación, de reputación corporativa, sus métricas correspondientes y las implicaciones organizativas que conlleva. Esta tesis se incardina dentro de esta línea de investigación.*

*La segunda línea ha originado la creación de asociaciones, fundaciones, laboratorios de ideas o "think tanks" y otros órganos de la sociedad civil que han congregado a distintas empresas para unir fuerzas.*

*La fuerte demanda institucional y el interés de las organizaciones y las empresas por la reputación ha llevado al mundo académico a desarrollar una amplia oferta de herramientas, métricas y escalas, destinadas a obtener medidas de este recurso intangible de valor estratégico". En investigaciones previas del doctorando que han servido como punto de partida de esta investigación, hemos revisado más de 20 métricas diferentes sólo dedicadas a medir la reputación corporativa (Carreras et al. 2013) y otras tantas destinadas a constructos afines como la confianza*

organizacional, el prestigio corporativo, la identidad corporativa e incluso una escala destinada a evaluar la “desconfianza institucional”.

Con el objeto de clarificar la panoplia de métricas e índices disponibles para un lector seguiremos el planteamiento propuesto en la obra antes citada, la doble aproximación métrica.

Excluimos en esta investigación los rankings y auditorías de reputación porque nos hemos centrado en métricas destinadas a obtener las actitudes del propio *stakeholder*. En este sentido los rankings y auditorías de reputación, si bien han tenido un papel relevante en el desarrollo de las métricas y teoría de la reputación, no son relevantes para los objetivos de esta investigación. Para tener una información detallada de las métricas basadas en los rankings y auditorías de reputación puede consultarse el texto de Carreras, Alloza y Carreras 2013, 351-366).

La literatura científica reconoce dos modos básicos<sup>4</sup> para aproximar la medida de la reputación corporativa, la directa o global y la indirecta o analítica (Fombrun et al. 2000; Goldberg y Hartwick 1990; Davies et al. 2004). Vamos a sintetizar cada uno de estos dos enfoques y en los epígrafes siguientes desarrollaremos cada uno de ellos con mayor detalle.

---

<sup>4</sup> Cuando se pretende medir fenómenos psicológicos complejos como la inteligencia, la motivación, pero también la satisfacción o la reputación, utilizamos la teoría del constructo psicológico. Un constructo es un concepto complejo, con varias dimensiones, que explica la emergencia de una serie de conductas. Por ejemplo, el constructo “lealtad psicológica” se conceptualiza como el compromiso que mantiene una persona con un proveedor que hace que prefiera los productos de esa empresa, la recomiende a amigos y/o tolere precios ligeramente superiores a los de sus competidores.

## 1.1. LAS MÉTRICAS DE LA REPUTACIÓN CORPORATIVA.

En los últimos años, el concepto de la “reputación corporativa” está recabando un interés cada vez mayor tanto en el mundo académico como en el profesional. En una revisión reciente, Kent Walker ha llegado a identificar hasta 1559 artículos, la mayoría con una antigüedad menor de 15 años (Walker 2010, 358-359). El éxito del fenómeno se debe a la abundante evidencia empírica de relación entre reputación y viabilidad de la organización. Una empresa con buena reputación tenderá a atraer inversores, talento, clientes y conservará o ampliará su licencia para operar (Alloza, 2012, 33-35).

La reputación corporativa parece estar convirtiéndose en uno de los recursos más prometedores para la gestión empresarial. El reconocimiento de su importancia estratégica y su contribución a la generación de valor se ha debido a la convergencia de una serie de factores entre lo cuáles podríamos destacar los siguientes (Alloza, 2012) :

1. Las crisis reputacionales que hacen desaparecer las grandes empresas.
2. La internacionalización de las empresas.
3. La búsqueda de una diferenciación sostenible.
4. La vinculación de las empresas y las marcas a la reputación de sus países de origen.
5. La divulgación de los rankings de reputación.
6. La explosión de los medios sociales en Internet que hace más vulnerable y volátil el valor de las compañías en el mercado.
7. La demanda urgente por parte de las empresas de indicadores no financieros para mejorar la estrategia y el buen gobierno de las mismas.
8. La necesidad para las empresas y las organizaciones de adoptar un marco de actuación de largo plazo, una visión multi-*stakeholder* y un mayor compromiso de responsabilidad corporativa.

En este contexto, la gestión de los activos y recursos intangibles como son la marca, la comunicación, la reputación y los asuntos públicos se ha convertido en una de las mayores oportunidades de creación y protección de valor para cualquier

institución o empresa. El Integrated Reporting International Initiative Council destaca que hoy el 80% del valor total de una organización reside en sus recursos y activos intangibles frente a lo que ocurría hace unas décadas<sup>5</sup>. La gestión excelente de los intangibles configura una gran oportunidad para aquellas organizaciones y líderes que sean capaces de adaptarse con éxito y rapidez a este nuevo contexto<sup>6</sup>.

La gestión de la reputación parece permitir la transformación y mejora permanente e introduce un modelo de innovación que afecta tanto en la forma de gestionar los asuntos corporativos, como en la manera de actuar y de comunicar para poder presentarse ante los grupos de interés con discursos coherentes y consistentes, que generen la credibilidad y confianza necesarias para cumplir con los objetivos de estrategia y de negocio (Alloza, 2012, 35 - 38).

La reputación aparece en este nuevo entorno como un concepto que es doblemente innovador.

En primer lugar es una innovación en la historia de la gestión empresarial que es incorporada a la agenda de la empresa hace tan solo diez años (Alloza, 2012, 40).

En segundo lugar, la incorporación de la reputación como concepto psicosocial a la práctica empresarial y a la institucional produce también innovación en la medida en que genera una transformación permanente en las organizaciones que se comprometen con su gestión.

En este sentido la gestión de la reputación establece una dinámica mediante la cual la organización se dota de herramientas de investigación y de escucha para conocer las percepciones, opiniones y las evaluaciones que los *stakeholders* hacen de ella.

---

<sup>5</sup> Integrated Reporting International Council [http://www.theiirc.org/wp-content/uploads/2013/12/IIRC-PP-Yearbook-2013\\_PDF4\\_PAGES.pdf](http://www.theiirc.org/wp-content/uploads/2013/12/IIRC-PP-Yearbook-2013_PDF4_PAGES.pdf)

<sup>6</sup> Corporate Excellence (2013), "What makes a Chief Communications Officer Excellent?" Ediciones CE .

Este mecanismo le permite identificar los “gaps” entre la realidad de la organización, la realidad comunicada y la realidad finalmente percibida y valorada por estos *stakeholders*.

La gestión permanente de la reputación incorpora al hacer y al comunicar de la empresa este diagnóstico que es en esencia un mecanismo generador de innovación que suscita una transformación permanente de la empresa como respuesta a las expectativas y demandas, cada vez más exigentes, de sus grupos de interés.

La reputación sería una fuente de innovación creadora de valor en la medida en - que como afirma Joan Costa (2009, 17-19) sólo crea valor la innovación social, la que es capaz de hacer que se fortalezcan las relaciones de confianza y afinidad de los grupos de interés con las organizaciones.

La fortaleza de estas relaciones con los *stakeholders* responde al reconocimiento que estos le otorgan. El reconocimiento es conocido por la organización gracias a los sistemas de escucha activa y de mejora continua que algunas organizaciones ponen en marcha para someterse reiteradamente al escrutinio de los grupos de interés. Estos métodos de escucha, basados en encuestas a los grupos de interés permiten conocer las áreas específicas donde poner en marcha planes de acción y de mejora que afectan a los procesos que subyacen a las experiencias que la empresa entrega a sus *stakeholders* en todos sus puntos de contacto, y en consecuencia, permite establecer un camino de transformación y de mejora continua que conducen a la excelencia.

Esta nueva perspectiva de gestión parece estar configurando una nueva relación empresa-sociedad al tejer puentes entre las distintas expectativas sociales y la toma de decisiones dentro de la empresa. Es una forma de incorporar herramientas de inteligencia social como la gestión de la reputación con el fin de ayudar a las organizaciones a tomar decisiones alineadas con las expectativas y exigencias de sus grupos de interés. Esta forma de entender la gestión de la empresa, conectada con las expectativas de los ciudadanos (Alloza, 2014, 50-55) es congruente con la

mirada del "capitalismo de largo plazo" al que se refiere Dominic Barton<sup>7</sup>, o Michael Porter<sup>8</sup> con el concepto de "valor compartido".

La reputación, así entendida, puede considerarse como una moneda con dos caras: puede ser una fuente de creación de valor pero también existen riesgos reputacionales que deterioran de forma acelerada la propia existencia de la organización actuando así como una clave poderosa de destrucción de valor.

Esto sucede cuando se rompe el equilibrio adecuado de respuesta satisfactoria a las demandas de los *stakeholders* que son quienes hoy detienen realmente el poder (Montañés, 2011, 40-44).

Una organización que gestiona su reputación puede reaccionar de forma rápida y efectiva tanto para proteger el valor, identificando y mitigando el riesgo reputacional como para incrementar el valor, gracias a las oportunidades de fortalecimiento de la reputación.

A medida que la reputación se fortalece genera para la organización una diferenciación sostenible en el tiempo y como consecuencia desencadena los comportamientos favorables por parte de sus grupos de interés que fortalecen de esta forma la creación de riqueza.

Las empresas que más rápido se adaptan a los cambios, es decir, las que más capacidad de escucha y reacción han creado, son las que obtendrán mayores flujos de caja futuros (The Copenhagen Charter, 1999).

La reputación permitiría fortalecer las relaciones de confianza y afinidad con los grupos de interés. Pero no debemos olvidar que también existen riesgos de erosión de la reputación que son fuente de destrucción de valor.

La creación y protección sostenible del valor de las empresas y organizaciones en el largo plazo, diferenciado y generador de confianza requiere velar por la buena

---

<sup>7</sup> BARTON, Dominic (2011): «Capitalism for the long term», en *Harvard Business Review*.

<sup>8</sup> PORTER, M. y KRAMER, M. (2011): «Creating shared value», en *Harvard Business Review*, n.4, pp. 1-13.

gestión de estos activos y recursos intangibles con el mismo nivel de importancia, excelencia y rigor que las empresas y organizaciones vienen otorgando a sus activos y recursos económicos. Esto es especialmente importante cuándo llega el momento de diseñar e implantar la estrategia empresarial, momento en el que los activos y recursos intangibles muchas veces no son incorporados con el peso adecuado (Alloza, 2012, 36-38).

El ex presidente de Corporate Excellence - Centre for Reputation Leadership, Luis Abril, prologó el primer libro de la Biblioteca de esta Fundación escrito por el profesor Cees Van Riel (2012) con una idea inquietante, a la vez que retadora, el nuevo paradigma empresarial entiende el éxito de la gestión por su capacidad de diferenciarse a partir del fortalecimiento de las relaciones con los principales *stakeholders*, clientes, empleados e inversores y la consiguiente generación de valor compartido para la organización y para los propios *stakeholders*. En palabras llanas del presidente, la empresa necesita “ganar... confianza para atraer clientes, talento y capital”.

Este nuevo paradigma otorga un lugar central a la métrica de la reputación corporativa que es el elemento básico sobre el que descansa la monitorización de la confianza.

En la última década se han sentado las bases y los fundamentos de las métricas de reputación.

## **1.2. LA APROXIMACIÓN DIRECTA E INDIRECTA A LAS MÉTRICAS DIRECTAS DE REPUTACIÓN.**

La aproximación directa consiste en solicitar a una persona que valore una compañía de forma global (Davies et al. 2004, 126). Por el contrario, la aproximación indirecta consistiría en preguntar por aquellos aspectos que llevan a las personas a establecer ese juicio global de reputación (Fombrun et al. 2000, 242). Desde esta perspectiva, se preguntaría a la persona por los determinantes básicos del apoyo social, que motivan a las personas a comprometerse con conductas de apoyo hacia la organización. La agregación de dichas respuestas daría lugar a una medida indirecta de reputación corporativa. Cuanto mejor puntúen en las diferentes dimensiones de apoyo, mayor coherencia en las razones del prestigio y más alta la reputación, cuanto peor, baja reputación; la ambivalencia mostraría a una persona indecisa, sin expresar un compromiso claro. Según el “principio de agregación” (Ajzen y Fishbein 2005, 181-182) la suma de respuestas a los diferentes motivos reflejan razonablemente bien la disposición general porque tienden a cancelar la unicidad de cada uno de los aspectos por separado. Devorah Vidaver-Cohen encontró una correlación de 0,88 entre la medida agregada de la escala RepTrack™, medida indirecta, y la del índice Pulse, medida global, obteniendo una prueba empírica de que ambas métricas estaban midiendo el mismo constructo (Vidaver-Cohen 2007, 281).

La reputación indirecta, en cuanto conjunto de evaluaciones sobre el comportamiento de la organización debe ocupar la posición de antecedente del modelo, con lo que facilitaría los puntos fuertes y débiles desde la perspectiva del *stakeholder*, mientras que la aproximación global debería ocupar la posición de salida o focal porque representa los sentimientos globales que activan la conducta.

Encontramos tres argumentos:

- a. Un indicador global integra también otros indicadores relevantes.
- b. Un indicador global dispara conductas de apoyo o de rechazo.
- c. Un indicador global tiene capacidad predictiva.



Vamos a ver con mayor detalle cada uno de estos tres argumentos.

a. Las medidas globales de un constructo canalizan los efectos de las medidas analíticas. Los autores que hemos citado como defensores de la aproximación indirecta (Fombrun et al. 2000, 242) suelen coincidir en que las variables sumarias de satisfacción, confianza o reputación suelen canalizar los efectos de las escalas indirectas que reúnen los motivadores que las generan. Davies y colegas utilizaron la satisfacción global para demostrar que abarcaba la mayor parte de la variabilidad de la escala de carácter corporativo (Davies et al. 2004, 137). Walsh y Beatty comprobaron la validez de su escala reputación “CBR” asociándola a cuatro constructos globales, satisfacción, lealtad, confianza y recomendación (Walsh y Beatty, 2007, 138) y, como ya hemos comentado, Vidaver-Cohen demostró una correlación muy alta entre el Reprakt™, medida indirecta y el Pulse, medida global (Vidaver-Cohen, 2007).

b. Las medidas globales son los disparadores de la conducta. En situaciones cotidianas, las personas activamos antes los juicios globales, más intuitivos, que los juicios deliberativos, llenos de detalles y matices. Según Jonathan Haidt, “las evaluaciones afectivas ocurren de manera tan generalizada, rápida y automática que se siente como una parte integral de la percepción (Haidt, 2001, 819). Para Russell Fazio el juicio global es un mecanismo mental eficiente para reducir el conflicto previo a la toma de una decisión, así una persona puede tomar decisiones rápidas activando automáticamente los juicios globales, sin tener que pasar por reflexiones pormenorizadas. De hecho una persona poco informada podría emitir un juicio global sobre una empresa, con la misma convicción que un respondiente altamente informado, aunque en realidad no lo esté (Carreras et al. 2013, 338). Este fenómeno también es visible desde las encuestas. En ocasiones, las personas no son capaces de aportar información de detalle, aunque responden con pocas dudas sobre las cuestiones generales. Estos juicios globales expresan la presencia de actitudes con fuerte carga emocional y esa carga emocional actuaría como disparador de la conducta (Bagozzi, 1992, 178-204; Lazarus, 2006, 9-46).

c. Las medidas globales son sumarias y muestran mayor capacidad predictiva sobre la conducta que las medidas indirectas. Gustafsson y colegas afirman que las

evaluaciones globales no sólo captan las razones sino que predicen mucho mejor los efectos en las conductas que las escalas indirectas del performance (Gustafsson et al. 2005, 216). El estudio de Dabhokar, Shepherd y Thorpe aportan una evidencia adicional sobre la ventaja de utilizar medidas globales para mediar entre las medidas indirectas y la intención de conducta. Los consumidores evalúan diversos aspectos relacionados con la calidad del servicio, pero también han generalizado una idea global de calidad que no es exactamente la suma de sus componentes y que tiene mayor capacidad predictiva sobre las intenciones de compra (Dabhokar et al. 2000, 166). También Richard Bagozzi apoya esta idea cuando plantea la necesidad de mediar los efectos de las evaluaciones del performance sobre la intención a través de mediadores globales emocionales como la satisfacción (Bagozzi 1992, 188). Por último, Leonard Ponzi y colegas han resaltado el mismo papel mediador de la versión en global de reputación o como dicen “en formato reducido” (Ponzi et al. 2011, 16).

Estos tres argumentos refuerzan la idea de que el juicio global de reputación debe ocupar una posición intermedia entre las evaluaciones de atributos propias de las escalas indirectas y las conductas que activarían dichas evaluaciones.

La escala de reputación indirecta, aclararía cuáles son las razones que llevar a un sujeto a expresar una actitud positiva o negativa. Estas evaluaciones son altamente informativas, facilitan la diagnosis y son las responsables de la aparición del juicio global, pero es ese juicio global el que, una vez formado, activa la conducta. En múltiples ocasiones de la vida cotidiana, la simple percepción de una marca corporativa “activa automáticamente la evaluación (global) asociada a este nombre de empresa... y tal activación... puede ejercer una influencia sobre... la conducta” (Fazio, 2007, 610-611). En este sentido decimos que la reputación global se constituye en el *KPI*<sup>9</sup> (*Key Performance Indicator*) o "indicador clave" que aglutina y contiene en su expresión matemática, las razones que subyacen en una actitud y que predice mejor la conducta favorable, o desfavorable de los individuos hacia una empresa.

---

<sup>9</sup> KPI ó “indicador del performance, clave”, que debe ser monitorizado en los cuadros de mando.

De lo anterior se puede concluir que las métricas indirectas poseen capacidad diagnóstica porque proporcionan las razones que llevan a una persona a considerar una buena o mala reputación. A su vez, las métricas globales careciendo de dicha capacidad, resultan excelentes para sintetizar el nivel de reputación que siente el sujeto.

Por tanto, ambas aproximaciones presentan un perfil complementario, diagnóstico versus síntesis y deben formar parte del modelo de lealtad.

El modelo de lealtad del *stakeholder* desvela cuáles son los mecanismos que transforman las evaluaciones realizadas por un sujeto en conductas favorables, creadoras de valor, hacia una empresa u organización.

Veamos este modelo de lealtad con mayor detenimiento. En los años sesenta y contra todo pronóstico, el mundo académico sólo había encontrado una débil relación entre actitud y conducta (Wicker, 1969, 50-52). Parecía como si los individuos acababan haciendo lo que no pensaban o sentían. Años después se resolvió la paradoja incorporando dos elementos nuevos a la explicación del comportamiento:

1. Kothandapani descubrió que actitud y conducta siempre media la intención. No basta con que los sujetos evalúen de forma positiva o negativa un objeto, la actitud actúa generando una intención de conducta, en la que el sujeto planifica y considera si está a su alcance desplegar una acción y sólo cuando ha formado esa intención es cuando se traduce en conducta real (Kothandapani, 1971, 321-333). Por esa razón los modelos que explican la decisión de compra, incluyen todos como variable clave la medición de la “intención de conducta” en sus diversas expresiones: intención de recomendar, de volver a comprar, de invertir, de postularse como candidato para trabajar en la empresa, etc.
2. Izeck Ajzen y Martin Fishbein aportaron otro argumento a la hora de explicar el fracaso de las investigaciones previas para demostrar la relación entre actitud y conducta. Estos autores se percataron de que esas

investigaciones anteriores cometían un error teórico al tratar de asociar una actitud general con una conducta concreta. Por ejemplo, la calidad general de los servicios<sup>10</sup> y la intención de comprar un servicio específico porque en dicha relación interfieren contingencias y situaciones particulares que no explican la influencia de la actitud general.

*"Ambos conceptos que se evalúan deben establecerse en un mismo nivel de generalidad. Según el "principio de compatibilidad" (Ajzen y Fishbein 2005, 51-56), la percepción de calidad en una empresa ejerce su influencia un rango de conductas suficientemente variado como para evaluar sus efectos, como son la recompra, la recomendación, la sensibilidad a precios altos, la credibilidad publicitaria, el tiempo de permanencia como cliente..." (Carreras et al. 2013, 218-219).*

La teoría general que ha dado lugar al desarrollo del modelo explicativo de lealtad del cliente ha sido la teoría de la acción razonada<sup>11</sup> (TAR) de los autores antes citados Ajzen y Fishbein. Se trata de una teoría general que se ha verificado en diversos campos de la conducta humana, explicando con éxito desde las decisiones de practicar deporte o dejar de fumar hasta las decisiones de compra en el ámbito de consumo.

En síntesis, la teoría viene que decir que una persona se decide a realizar una conducta porque la considera atractiva o beneficiosa (componente de la utilidad), siente que debe hacerlo (componente normativo o moral) y se ha formado la intención de llevarla a cabo. Posteriormente Richard Bagozzi (1992) incorporaría el elemento emocional como mediador necesario entre la actitud y la intención. Todos los constructos se presentan en el mismo plano de generalidad.

---

<sup>10</sup> Sobre la conceptualización de la calidad percibida de los servicios como una actitud, puede consultarse Carreras y González 2012, pp. 351-352.

<sup>11</sup> El texto de Carreras, Alloza y Carreras (2013) cita la teoría de la acción planeada, que es un desarrollo posterior de Izeck Ajzen en 1991. En aras de la simplificación, aquí sólo mencionamos la teoría de la acción razonada.

## MODELO GENERAL DE EXPLICACIÓN DE UNA DECISIÓN

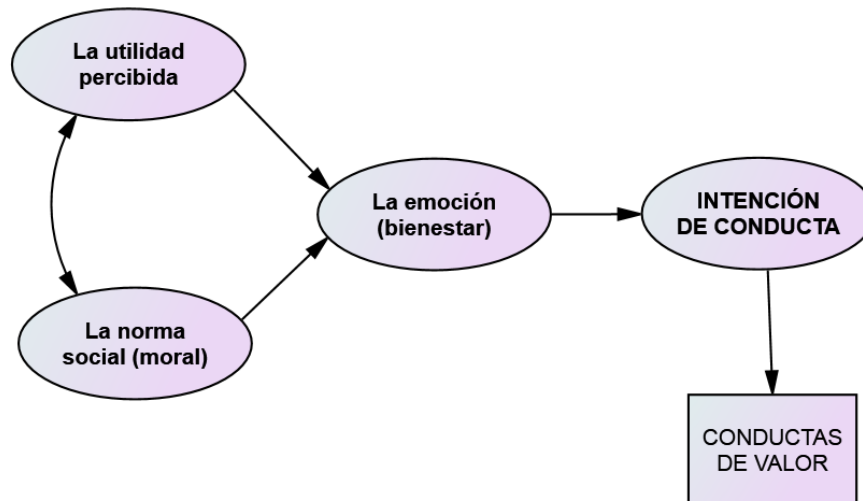


Figura 0.1. Modelo de explicación de la decisión basado en las teorías de la acción razonada (Ajzen y Fishbein, 2005) y la de la emoción (Lazarus, 2006).

El modelo aplicado a los comportamientos de consumo plantea que la decisión de recompra o lealtad de un cliente se explica a través de un proceso deliberativo en el que se identifican tres planos, los determinantes o evaluaciones de la experiencia de consumo, la sensación de bienestar resultante y posteriormente la variable resultante, que es la intención de volver a hacer negocios con la empresa.

Con este bagaje teórico, Alloza y los profesores Carreras formularon un nuevo modelo: "el modelo de lealtad del *stakeholder*"<sup>12</sup>. Este modelo adquiere un carácter más general que los modelos de lealtad del cliente porque contiene la diversidad de fuentes del apoyo a una organización que pueden conseguir, no sólo por parte de los clientes sino por cualquier otro colectivo y a la vez, integra los diferentes mecanismos de mediación que se han aportado, como la satisfacción, la confianza o la vinculación afectiva. Este nuevo modelo podría facilitar el

---

<sup>12</sup> En el capítulo 5 del libro se relacionan la serie de investigaciones que han validado parcialmente el modelo.

diagnóstico, la planificación y coordinación de diversas estrategias de gestión de los activos y recursos intangibles de las empresas y en particular de la reputación corporativa.

La lealtad psicológica del cliente se concibe como un compromiso general con una empresa. Richard Oliver la ha definido como “un compromiso sostenido, para comprar o apoyar consistentemente en el futuro, un servicio o un producto, provocando una compra repetitiva de la misma marca o conjunto de marcas, a pesar de las influencias situacionales y esfuerzos de marketing que tienen el potencial de causar el cambio o la baja” (Oliver, 1997,392). El compromiso de un *stakeholder* hacia una empresa (Freeman, 2010, 63-65) reflejaría la presencia de una actitud compleja basada en la información favorable que posee sobre dicha empresa en las siete dimensiones que son más relevantes en base al modelo RepTrak™ (oferta comercial, trata bien a sus empleados, es ética y está bien gobernada, es responsable socialmente, es innovadora, tiene buenos gestores y buenos resultados económicos). El buen desempeño percibido en estas siete dimensiones despertaría satisfacción y/o confianza y que se traducen en la consolidación de la decisión de adoptar comportamientos favorables: comprar acciones, productos, solicitar empleo o recomendar a dicha empresa.

Siguiendo este razonamiento, el compromiso de un *stakeholder* puede ser medido y gestionado a partir de un modelo que contenga los tres planos explicativos: las evaluaciones que actúan como antecedentes, los procesos de mediación y la formación de la intención de la conducta de valor (figura nº 0.2).

## EL MODELO DE LEALTAD DEL STAKEHOLDER

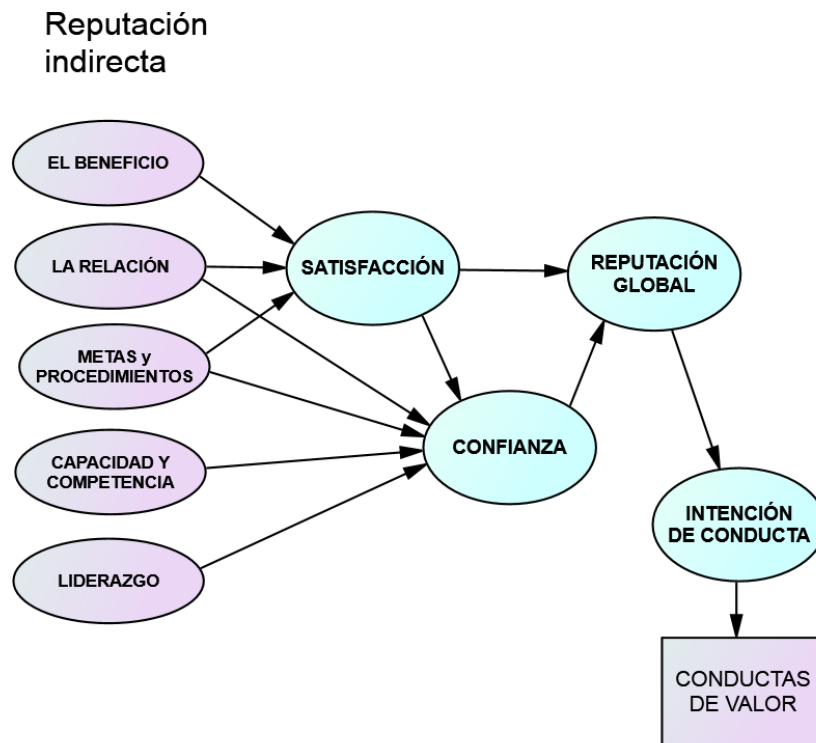


Figura nº 0.2. Modelo de lealtad del *Stakeholder*. Esquema simplificado, basado en el texto de Carreras et al. 2013, 320.

El plano de los antecedentes estaría representado por las evaluaciones de la reputación indirecta, que contienen todas las dimensiones que representan el valor de la empresa para cada *stakeholder*.

El modelo de lealtad está formulado como un modelo causal de impactos funcionales<sup>13</sup> que permite identificar el perfil de factores que maximizan la generación de valor para cada *stakeholder* y para cada empresa.

Los determinantes van activando el segundo plano, los mecanismos de mediación. Aunque en la obra citada (Carreras et al., 2013) se mencionan cinco mediadores,

---

<sup>13</sup> La metodología para estimar estos modelos son las ecuaciones estructurales o modelos SEM (Structural Equation Modelling)

en aras de la concreción sólo comentaremos los tres más relevantes, la satisfacción global, la confianza y la reputación global<sup>14</sup>.

La satisfacción<sup>15</sup> se ha definido como la evaluación retrospectiva de la experiencia de consumo que genera bienestar o malestar en el cliente (Oliver, 1999, 34). Otro tipo de sentimiento general desencadenado por las evaluaciones del comportamiento de la empresa es la confianza. La confianza se entiende como una “expectativa optimista” que es necesaria para establecer una transacción con cualquier organización (Hosmer, 1995, 379-403. Ambos conceptos son extrapolables a los colectivos de empleados, clientes, accionistas, etc.<sup>16</sup>.

Conviene destacar que los antecedentes de ambos sentimientos no son necesariamente los mismos, ni tienen el mismo impacto para los distintos grupos de interés. Esto explica que durante la última década del siglo, las investigaciones sobre satisfacción y confianza hayan avanzado en paralelo, como líneas rivales en la investigación de la lealtad (Carreras et al. 2013, 293-295).

La “escuela de la satisfacción” ha estado auspiciada por Richard Oliver (1997), Claes Fornell, y colegas (1996) o el equipo de Michael Johnson (2001), mientras que la de la confianza ha sido abordada por autores como Morgan y Hunt (1994, 20- 38), Garbarino y Johnson (1999, 70-87) y Jones (1996, 4-25). Lo cierto que es la escuela de la satisfacción ha enfatizado el componente del beneficio (calidad de servicios y valor), mientras que la confianza se ha centrado más en los componentes de las relaciones humanas, la competencia y el comportamiento ético.

Más recientemente, una serie de autores (Harris y Goode 2004, 139-158; Quero-Guervilla, 15-36 y Cater y Zabkar 2008, 1-13) han intentado integrar ambos

---

<sup>14</sup> Los otros dos mecanismos son la “vinculación afectiva” con la empresa y el “compromiso calculativo”, más racional.

<sup>15</sup> Se trata de la “satisfacción generalizada con la empresa”, recordemos que según el principio de compatibilidad, los constructos deben moverse en el mismo plano de generalidad.

<sup>16</sup> Los clientes, empleados o inversores potenciales también activan sentimientos de satisfacción y confianza, la diferencia se encuentra en que el nivel de información que manejan es mucho menor y menos fiable porque procede de la experiencia comunicada.



sentimientos, pero sin proponer una estructura común antecedente, como es el caso de la propuesta de Alloza y de los profesores Carreras.

Los efectos de ambos sentimientos en la actitud intencional de los *stakeholders* han quedado demostrados en numerosas investigaciones (Carreras et al. 2013). Sin embargo parece dibujarse una jerarquía entre ellos. La satisfacción sería el antecedente de la confianza porque la experiencia estaría actuando como refuerzo del optimismo que implica la confianza (Bowden 2009, 63–74; Quero-Gervilla y Ventura-Fernández 2009, 15-36; Sirdeshmuk et al. 2002, 15-37; Suki 2011, 1207-1210.). Estos resultados son congruentes con la teoría de la lealtad del *stakeholder* porque la confianza se considera como una expectativa optimista y por tanto estaría más próxima a la conducta que la satisfacción que sólo es balance afectivo retrospectivo.

Los índices de reputación indirecta y de reputación global abren y cierran los procesos de mediación. La reputación indirecta contiene todos los factores que motivan la satisfacción, la confianza o ambos a la vez y por tanto se ubica como el antecedente de los procesos de mediación, mientras que el índice global de reputación reúne toda la variación de la satisfacción y la confianza situándose como la resultante final del proceso (v. figura nº 0.4). Recordemos que la reputación global incorpora a todos los componentes evaluativos, que son los antecedentes comprobados de la satisfacción y de la confianza.

Aunque las evidencias empíricas a favor de este modelo de lealtad del *stakeholder* son aún parciales (Carreras et al. 2013, 300-302; Money et al. 2010, 197-211; Nettet y Helgesen 2009, 327-345; Sung y Yang 2009, 787–811.; Walsh et al. 2009, 187-203), parece razonable admitir que el indicador global de reputación podría incorporar los efectos de estos mecanismos de mediación. Leonard Ponzi ha señalado que Pulse incorpora los sentimientos generales asociados a la reputación global de las empresas como son la confianza, la buena impresión (cercano a la satisfacción), la admiración y el respeto (Ponzi et al. 2011, 15-17). El índice global de reputación Pulse incluiría los sentimientos y evaluaciones de los distintos *stakeholders* para predecir la actitud intencional, que es la responsable final de la conducta efectiva.

### **1.3. LAS MÉTRICAS DE APROXIMACIÓN INDIRECTA: LA REPUTACIÓN RACIONAL.**

Las escalas racionales de reputación son aproximaciones indirectas que identifican e integran los elementos que generan y sobre los que descansa la reputación de una organización. Son propias de los procesos deliberativos que participan en la formación de la actitud y su principal virtud, como ya hemos comentado, es la capacidad de diagnosticar los puntos fuertes y débiles de la reputación. Contar con estas dimensiones de la reputación permite a las organizaciones implementar modelos de gestión multifuncionales avanzados (Alloza, 2008, 371-378.).

Dentro de este apartado veremos el modelo “Reputation Quotient”, (Fombrun et Al, 1996; Fombrun et Al. 2000, 241 -255) y su reformulación más reciente y avanzada en el Reptrak™ (Alloza y Martínez, 2007, 113-128; Fombrun, 2006, 165-170), la “escala del carácter corporativo” de Davies y colegas (2004, 125-146), la “reputación basada en el cliente” (CBR) y la escala formativa de Sabrina Helm (2005, 95-109).

Las escalas racionales de reputación miden el componente cognitivo de la actitud de las audiencias hacia las empresas y en este sentido contienen información suficiente para ofrecer las claves que aumenten el valor de la compañía en éstas audiencias.

El problema radica en que diagnosis y comparabilidad son dos propiedades que se encuentran en tensión, dentro de cualquier escala. De hecho se produce un efecto generalizado que determina que cuanto mayor sea la capacidad diagnóstica de una escala, menor será la posibilidad de que sea comparativa entre distintas poblaciones y ámbitos geográficos.

La comparación precisa moverse en un plano demasiado general para localizar dimensiones válidas en todos los colectivos, pero cuando descendemos a cuestiones más específicas que proporcionen mayor detalle a la diagnosis, las escalas van perdiendo las dimensiones generales y la comparación se vuelve más difícil. De hecho los autores que han apostado por escalas más concretas, han comprobado que incumplen el principio de la equivalencia semántica y/o métrica

entre las distintas audiencias (Chun y Davies, 2006, 138-146; Gardberg, 2006, 39-61; Helm, 2005, 95-109, ; Lloyd 2011, 221-233; Schwaiger y Wiedmann, 2004, 46-71; Walsh y Beatty, 2007, 127-143).

Conseguir un buen nivel diagnóstico y que cuente con propiedades comparativas parece una cuestión de equilibrio óptimo, habría que identificar las dimensiones que siendo de aplicación general a los distintos grupos de interés tengan suficiente concreción como para identificar problemas específicos y por tanto poder ser utilizadas para la gestión práctica de los procesos, actuaciones y formas de comunicación que se implantan en la organización para proteger o mejorar su reputación (Alloza, 2005, 6-8.).

## LA ESCALA “RQ” DE FOMBRUN (2000).

A principios de siglo, Fombrun, Gardberg y Server desarrollaron el “Reputation Quotient (RQ)” (2000), una escala indirecta destinada a medir la reputación diseñada para que fuera lo suficientemente general como para medir la reputación en distintos grupos de interés, sectores de actividad y países.

El método de investigación empleado para el desarrollo de esta escala se basó en la realización de una fase de investigación cualitativa inicial. Los autores realizaron un estudio cualitativo que se basó en la realización de cinco grupos de discusión con directivos, viajeros de líneas aéreas, consumidores de informática y alumnos de escuelas de negocio. Esta fase permitió identificar un listado de 32 ítems asociados a la buena o mala reputación de una empresa, que después agruparon en ocho categorías o dimensiones teóricas. En la segunda fase de cuantificación de los resultados, procedieron a validar la escala mediante un análisis factorial exploratorio aplicado a los datos obtenidos a partir de una encuesta realizada con clientes de 20 empresas distintas.

La escala final de reputación quedó estructurada en seis dimensiones son sus correspondientes atributos:

- Atractivo emocional: Tengo un buen "*feeling*"/sentimiento de la compañía, admiro y respeto esa compañía, confío en esa compañía.
- Productos y servicios: La compañía está detrás de los productos y servicios, desarrolla productos y servicios innovadores, ofrece productos y servicios de alta calidad, tiene una buena relación calidad precio ("*value for money*").
- Visión y liderazgo: Excelente liderazgo, clara visión de su futuro, reconoce y aprovecha las oportunidades del mercado.
- Cultura y entorno de trabajo: Bien gestionada, parece una buena compañía para trabajar, parece que tiene buenos empleados.
- Responsabilidad social: Apoya buenas causas, responsable con su entorno, trata bien a las personas de su comunidad.

- Resultados financieros: Sólido historial de rentabilidad, bajo riesgo de inversión, supera a sus competidores, buenas perspectivas de crecimiento futuro.

Posteriormente la escala se validó mediante una encuesta en el sector de la informática y se comprobó que la escala “RQ” se articulaba en dos grandes dimensiones, ambas con alta fiabilidad: El “atractivo emocional” y los cinco componentes restantes (Fombrun et al. 2000, 254).

La escala demostró ser aplicable más allá del contexto norteamericano. E. Groeland probó el RQ en Holanda (Groeland, 2002, 305-315). Todas las dimensiones del RQ fueron reconocidas por los entrevistados como palancas de reputación, pero aparecieron dos dimensiones adicionales la “representatividad carismática” del líder y el “origen holandés” de la empresa. En el caso holandés, la validez de contenido no se verificó plenamente porque requiere la presencia de esas dos dimensiones adicionales.

La escala ha sido corroborada en Holanda (Groeland, 2002, 305-315), Alemania (Walsh y Wiedmann, 2004, 412-420) y Reino Unido (Gardberg, 2006, 39-61).

El RQ no ha estado exento de críticas. Manfred Schwaiger plantea que la escala confunde en un mismo plano las dimensiones indirectas con el indicador global (Schwaiger et al. 2009, 46-71). En otra investigación en la que participó Alloza junto a Luis Carlos Martínez se pone de manifiesto la debilidad que representa asumir la equi-ponderación a la hora de realizar el cálculo de la medida agregada. Este procedimiento oculta el hecho de que algunos aspectos pueden tener mayor incidencia en la formación de la reputación, conocimiento que resulta estratégico para su gestión (Alloza y Martínez, 2007, 4).

Las carencias detectadas en esta escala llevaron a su reformulación y perfeccionamiento en una nueva métrica, el RepTrak™ .

## LA ESCALA REPTRAK™.

La escala de reputación “RepTrak™” es el resultado de una iniciativa española para mejorar la escala RQ (Alloza y Martínez 2007:4: 1; Fombrun, 2006). Nace de una colaboración entre una asociación de grandes empresas españolas, el “Foro de Reputación Corporativa” (hoy Corporate Excellence) y una firma de consultoría internacional especializada en la medición y gestión de la reputación, el “Reputation Institute”.

De esta forma, en el año 2004 se inicia un amplio programa de investigación cualitativa en España y 8 países de América Latina, para conseguir crear una métrica cuyo objetivo central es la implantación de un modelo de gestión de la reputación (Alloza y Martínez, 2007, 2), con un metodología cuantitativa, basada en recogida continua de la información o tracking continuo, un método de recogida de información continua que fue creado por Gordon Brown, fundador de la consultora de investigación de mercados, Millward Brown Int. a principios de los años 80 (Alloza et al. 2000, 819-857).

A partir de entrevistas en profundidad a directivos y de grupos de discusión con consumidores de diferentes sectores en España y Latinoamérica (8 países), se obtuvo una lista de ítems relacionados con la reputación y mediante una encuesta online en España a la que se aplicó un análisis factorial exploratorio, se obtuvo la escala “RepTrak™” con 21 ítems agrupados en siete dimensiones. A diferencia del “RQ”, la investigación del RepTrak™ se centró en determinar cuál era la estructura de pesos de cada uno de los atributos y de cada una de las dimensiones en la reputación global, es decir en el índice Pulse (Alloza y Martínez 2007, 10).

La escala RepTrak™ se estructura en siete dimensiones racionales:

- La oferta de productos y servicios. Representado con cuatro indicadores (calidad de productos y servicios, garantiza sus productos y servicios, relación calidad-precio y orientación al cliente)

- Liderazgo. Cuatro ítems (buena gestión, líderes fuertes y respetados, visión clara de futuro y potencial de crecimiento futuro)
- Finanzas. Dos ítems (genera beneficios para sus propietarios y buenos resultados)
- Trabajo. Tres indicadores (paga a empleados de forma justa, seguridad de sus empleados y buen lugar para trabajar)
- Ciudadanía. Tres ítems (apoya causas sociales, protege el medioambiente y contribuye al desarrollo del país)
- Innovación. Tres indicadores (adaptación fácil al cambio, lanzamiento regular de nuevos productos, impulsa a empleados a generar ideas)
- Gobierno Corporativo. Dos indicadores (no realiza un uso indebido de su poder, no se involucra en negocios no éticos)

La escala RepTrak™ contiene algunas dimensiones que construyen reputación que no estaban presentes en la antigua escala “RQ”, como son la innovación, la ciudadanía o el gobierno corporativo.

El “RepTrak™” es una escala cognitiva multi-*stakeholder*, que se ha sido implantada para diferentes colectivos poblacionales en numerosos países por el Reputation Institute, una firma de consultoría especializada en reputación, propietaria intelectual del “RepTrak™”, fundada por los profesores Charles Fombrun y por Cees Van Riel (Gardberg 2006, 39-61; Groeland, 2002, 305-315; Fombrun y Pan 2006, 165-170).

Las siete dimensiones de la escala de RepTrak™, constituyen lo que podríamos denominar el marco de medición racional de la reputación, para nuestra investigación es de especial relevancia mencionar los estudios empíricos y modelos matemáticos que demuestran la relación entre esta escala racional y el indicador global y emocional de reputación Pulse.

Se ha calculado la contribución de cada una de las 7 dimensiones de RepTrak™ y sus 26 atributos subyacentes en el fortalecimiento o debilitamiento de la reputación global de una empresa medida con el indicador Pulse.

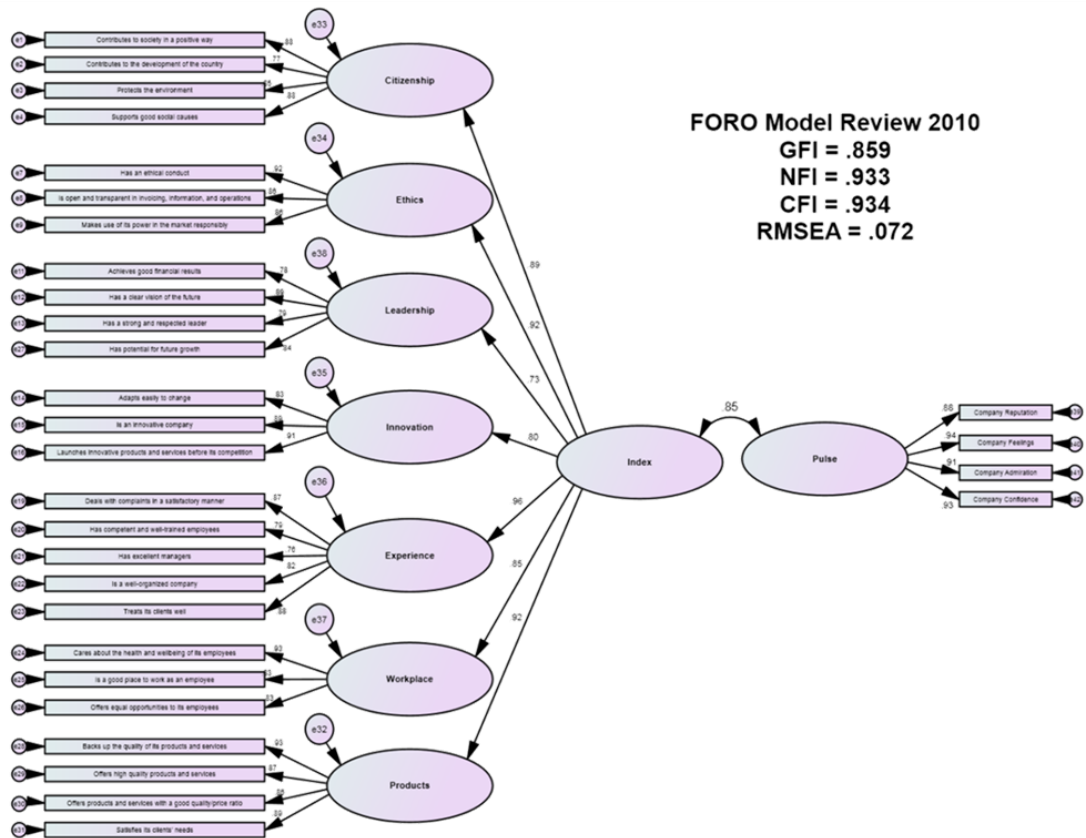


Figura 0.3. Pesos de los atributos y las dimensiones en el reputación Index (componente racional y Pulse). Fuente: Foro de Reputación Corporativa, hoy Corporate Excellence, 2006.



## LA ESCALA DEL CARÁCTER CORPORATIVO (DAVIES ET AL. 2004).

En el 2004, Gary Davies, Rosa Chun, Rui Vinhas da Silva y Stuart Roper desarrollaron una escala de reputación racional que se basa en la “metáfora de la personificación”. Este enfoque considera que la organización tiene rasgos humanos, como intenciones, deseos y estilos de comportamiento específicos. De este modo un observador podría llegar a percibir “el carácter corporativo” de una empresa. Según Davis y colegas, la tendencia a percibir a las empresas como si tuvieran una personalidad definida, ofrece la ventaja de estimular el vínculo afectivo del *stakeholder* con la corporación (Davies et al. 2004,127).

En la revisión literaria, los autores encontraron cuatro rasgos de personalidad corporativos que han sido asociados a una organización: Su carácter agradable o desagradable, el grado de eficiencia o ineficiencia, la orientación a la innovación versus la tradición y el grado de neocriticismo o inestabilidad emocional en las relaciones con sus *partners* internos y externos. Con este bagaje teórico, realizaron entrevistas en profundidad a directivos y grupos de discusión con consumidores de diferentes sectores productivos y países. Mediante análisis factorial exploratorio sobre los datos de una encuesta a varios *stakeholders*, llegaron a estructurar la “escala del carácter corporativo” que contiene 49 ítems, agrupados en 16 facetas integradas en 7 dimensiones. En una tercera fase, se procedió a la validación de la escala propuesta, mediante análisis factorial confirmatorio.

Dos años más tarde, en una investigación independiente se demostró validez nomológica de la escala al comprobar el impacto de las dimensiones de la escala en la satisfacción del cliente y del empleado (Chun y Davies, 2006, 138-146).

La escala se despliega en tres niveles: Dimensiones, facetas y atributos. Las dimensiones “amabilidad”, “trabajo”, “chic” y “crueldad” se desarrollan en facetas más específicas, pero las dimensiones “Informalidad” y “machismo” son unidimensionales y no se despliegan en facetas.

El objetivo de la escala del carácter corporativo es llegar a identificar el estilo de conducta que perciben las audiencias.

La escala ha sido implantada en clientes y empleados y ha permitido numerosas investigaciones comparando las diferencias de percepción entre los empleados y los clientes de una misma organización, la reducción de esta distancia incrementaría la buena reputación de la organización (Chun y Davies, 2006, 138-146) y ha sido probada en otros contextos culturales como el asiático (Chetthamrongchai, en 2010, 209–219).

## LA ESCALA “CBR” DE WALSH Y BEATTY (2007).

Gianfranco Walsh y Sharon Beatty presentaron en el 2007 una escala para evaluar la reputación específica en el cliente. Los autores consideran que las percepciones, intereses y valores de las audiencias y *stakeholders* son diferentes y en ocasiones enfrentados (Helm, 2011, 95-109; Lloyd, 2011, 221-233; Gardberg, 2006, 39-61) y por tanto sería más lógico construir métricas racionales específicas para cada *stakeholder*, como es el caso de la escala de Reputación Basada en el cliente “CBR” (Walsh y Beatty, 2007, 129).

La investigación cualitativa preliminar llegó a identificar 60 ítems. La lista fue filtrada por un grupo de jueces expertos quienes eliminaron los ítems ambiguos o poco claros y por último un grupo de alumnos de doctorado, evaluó la aplicabilidad de cada atributo al concepto reputación. Al final la relación quedó reducida a un total de 39 ítems.

Como en los casos anteriores y a partir de los datos de una encuesta online a tres sectores, banca, pequeños comercios y restauración de comida rápida, la estructura de la escala fue identificada mediante análisis factorial exploratorio y después validada mediante técnicas confirmatorias. La escala definitiva quedó organizada en cinco dimensiones que agruparon a 31 ítems:

- Dimensión 1: Orientación al cliente
  - Tiene empleados que se preocupan por las necesidades del cliente
  - Tiene empleados que atienden a clientes con cortesía
  - Manifiesta su preocupación por sus clientes
  - Se toma en serio los derechos de los clientes
  - Se esfuerza por tratar a sus clientes de manera justa
  - Toma en serio los derechos del cliente
  
- Dimensión 2: Buen empleador
  - Parece una buena compañía para trabajar
  - Parece que trata bien a su gente
  - Parece tener un excelente liderazgo

- Tiene un equipo directivo que parece prestar atención a las necesidades de sus empleados
- Parece tener buenos empleados
- Parece mantener altos estándares en la forma en que trata a las personas
- Parece estar bien gestionada
- Dimensión 3: Empresa confiable y sólida financieramente
  - Tiende a superar a los competidores
  - Parece reconocer y aprovechar las oportunidades de mercado
  - Parece que tiene sólidas perspectivas de crecimiento futuro
  - Parece que es una buena inversión
  - Parece tomar decisiones económicamente racionales
  - Mantiene un control exacto del beneficio
  - Lo está haciendo bien desde el punto de vista financiero
  - Parece tener una visión clara de su futuro
  - Parece ser consciente de su responsabilidad en la sociedad
- Dimensión 4: La calidad del producto y servicio
  - Ofrece productos y servicios de calidad
  - Es una empresa sólida y confiable
  - Respalda/responde los servicios que ofrece
  - Desarrolla servicios innovadores
  - Ofrece productos y servicios con buena relación calidad-precio
- Dimensión 5: Responsabilidad social y ambiental
  - Parece hacer un esfuerzo para crear nuevos puestos de trabajo
  - Reduciría sus ganancias para asegurar un medio ambiente limpio
  - Parece ser ambientalmente responsable
  - Parece apoyar buenas causas sociales.

La escala “CBR” fue probada mediante un análisis de invarianza en tres países, EEUU, Alemania y Reino Unido, para verificar si era equivalente en diferentes culturas. Los resultados fueron pobres porque sólo quedó corroborada la invarianza configural (Walsh et al, 2009, 928), concepto matemático que más adelante definiremos con exactitud. En consecuencia, la comparación internacional de la escala es relativamente débil.

#### 1.4. LOS LÍMITES DE LA APROXIMACIÓN INDIRECTA.

De las diferentes escalas indirectas examinadas, en el trabajo publicado por los profesores Carreras y Alloza se afirma que *"la escala RepTrak™ sería una de las más completas ya que contiene las razones básicas que llevan a un sujeto a apoyar a una organización (Carreras et al. 2013, 409)"*.

La escala del carácter corporativo es útil cuando el objetivo no es medir la reputación sino diagnosticar el estilo de relación de la empresa con sus grupos de interés. La escala CBR está limitada a un único *stakeholder*, el cliente.

La virtud principal de las escalas indirectas o cognitivas es la capacidad diagnóstica. Sin embargo, cuando son utilizadas para medir el nivel total de reputación en un colectivo resultan poco apropiadas por tres razones:

1. La medida indirecta no mide la reputación de todo el colectivo observado. Algunos *stakeholders* pueden tener una idea clara de reputación y sin embargo no disponer de información suficiente sobre la Compañía, con lo que sus respuestas serían muy escasas y en todo caso, incapaces para producir a partir de esa información una medida agregada válida de la reputación global de la organización.
2. Estas escalas conciben la reputación como una actitud eminentemente cognitiva, basada en procesos deliberativos. Se trata de evaluaciones de atributos y dimensiones que pueden ser contra-argumentadas, tienen un carácter racional, pero son procesos mentales en los que predomina la ambivalencia cognitiva (Ajzen, 2001, 27-58)
3. Enrique Carreras, Ángel Alloza y Ana Carreras han planteado la posibilidad de que la reputación sea una actitud activada automáticamente ante la percepción de una compañía, que actúa como *"priming"* o desencadenante de la evaluación global, sin procesamiento cognitivo (Ajzen, 2001, 27-58; Fazio, 2001, 115-141). En este caso obviamente, las escalas más cognitivas no recogerían este fenómeno reputacional global.

Para concluir, podemos afirmar que las diferencias entre ambas aproximaciones, la indirecta y la directa parece apoyarse en varias evidencias empíricas:

- Deborah Vidaver-Cohen (2007, 115-141), en su estudio sobre reputación de las escuelas de negocio demostró que ambas aproximaciones no miden lo mismo. El Reprak™, prototipo de escala racional sólo llegaba a explicar el 77,4% de la varianza total de la reputación (Vidaver-Cohen 2007, 281).
- La mayoría de investigaciones otorgan mayor poder predictivo a los índices globales de reputación que a las escalas racionales (Carreras et al. 2013, 204).
- Como ya hemos comentado, los índices globales están captando las opiniones de personas que mantienen una idea general de reputación pero que son incapaces de verbalizar esta evaluación con los detalles y pormenorización que requieren las escalas racionales. En consecuencia se puede afirmar que las evaluaciones globales directas captan mejor la actitud general de los diferentes *stakeholders*.

## **1.5. LAS MÉTRICAS DE APROXIMACIÓN DIRECTA: LA REPUTACIÓN GLOBAL.**

Ya hemos avanzado que las métricas globales cumplen una función sumaria importante para medir el grado de apoyo que tiene una organización entre sus grupos de interés. Leonard Ponzi y colegas (2011, 18-20) han profundizado en las ventajas que tienen los índices globales de reputación, con un formato reducido, de bajo coste, porque requieren cuestionarios mínimos, y son simples para el entrevistado porque exigen un bajo nivel de procesamiento cognitivo. A estas ventajas se podría añadir su capacidad para captar todas las fuentes de variación, tanto los procesos deliberativos como los que proceden de actitudes que se activan automáticamente.

En este sentido la aproximación global sería una fórmula más apropiada para resumir el grado de compromiso y apoyo por parte de los diferentes colectivos hacia la organización.

Este constructo evalúa directamente los sentimientos generales asociados a la reputación tiene por tanto una carga emocional muy significativa. Como señalan Leonard Ponzi y colegas, las interpretaciones cognitivas de la reputación han contribuido a oscurecer su significado directo (Ponzi et al. 2011, 16). El significado del término reputación en castellano e inglés, es muy similar y se refiere a la opinión que se tiene de algo o alguien en términos de respeto, prestigio o aprecio, basado en su calidad o conducta anterior. Por tanto, la reputación de una empresa combina dos ideas globales, el prestigio o aprecio adquirido por la confianza y credibilidad que da su forma de ser y de proceder. Charles Fombrun ya había anticipado que estos sentimientos suelen ser similares entre los distintos grupos de interés (Fombrun y Van Riel, 2004, 175-176).

El consumidor puede estar más interesado en la fiabilidad de los productos, el inversor podría mirar más por la ética y buen gobierno o por el rendimiento financiero y el empleado por la confianza, pero al final, todos los elementos o aspectos considerados por las escalas racionales desembocan en el fortalecimiento o debilitamiento de los mismos sentimientos que se traducen en tener una “buena

impresión”, un sentimiento de “confianza” o de "admiración y respeto" hacia una organización.

Se describen a continuación algunos de los índices de reputación que se han desarrollado, entre los que se encuentra el Pulse, objeto de estudio de esta investigación.

### **EL ÍNDICE DE BROWN (1995).**

Una de las primeras formulaciones globales de reputación contemplada en la literatura científica corresponde al índice de Steven Brown (1995, 170-181). En este caso, el propósito de la investigación no fue construir un índice global de reputación sino evaluar los efectos de la actitud del vendedor sobre la actitud del cliente, para ello Steven Brown introdujo en el modelo una escala global de reputación con la esperanza de que controlara posibles efectos espurios en la relación, como la imagen de la compañía. Su interés en la reputación fue más bien circunstancial y por esa razón no se detuvo demasiado en su construcción y validación.

La “reputación percibida de una compañía” fue definida como "la percepción de los compradores de una empresa en términos del grado de conocimiento, fiabilidad, confianza, reputación, credibilidad y mejor posición" (Brown, 1994, 172). La escala adoptó la forma de 6 ítems con modalidades de respuesta adjetivales bipolares de 7 posiciones. Resultó una escala altamente fiable, con un "alpha de Cronbach" superior al 0,90, se trata de una prueba estadística que mide la fiabilidad de una escala de evaluación semántica.

Desde un punto de vista métrico, el índice de Brown se compone de 6 ítems generales, organizados en forma reflexiva:

- El nivel de conocimientos y competencia del empleado
- La fiabilidad de las respuestas del empleado
- La confianza



- La credibilidad
- El prestigio
- La mejor posición

Como ya se señaló en la obra publicada por los profesores Carreras, Alloza y Carreras (Carreras et al. 2013, 414-434), se realizó una encuesta por correo con un 44% de retorno a 1.000 compradores de varios sectores (manufactura, servicios, instituciones gubernamentales y organizaciones sin ánimo de lucro...) en diversas localizaciones geográficas.

Mediante análisis factorial confirmatorio se verificaron las propiedades de convergencia y discriminación entre los constructos, (Farrell y Rudd, 2009) demostrando que el factor general de reputación alcanzaba suficiente validez de constructo (convergencia y discriminación).

La reputación global de Brown aparece de forma equilibrada entre los elementos que generan confianza, como la credibilidad, la propia confianza y la predictibilidad y los elementos propios de la admiración como la notoriedad, la reputación y la preeminencia. Esta combinación entre valor y confianza va a ser una constante en las sucesivas definiciones globales de la reputación.

## LA REPUTACIÓN GLOBAL DE HIGHHOUSE (2009).

Scott Highhouse, Alison Broadfoot y Jennifer E. Yugo proponen una versión más teórica de la métrica global. Conciben la reputación como un recurso intangible que “aumenta la satisfacción y lealtad del cliente, la atracción y retención del empleado, el valor de la empresa y el atractivo para el inversor ” (Highhouse et al. 2009,782) y ese recurso intangible adopta la forma de un juicio evaluativo global sobre una empresa, estable en el tiempo y que es compartido por diferentes *stakeholders* (Highhouse et al. 2009, 783).

En la fase exploratoria un grupo de expertos evaluaron un conjunto de ítems genéricos y los sometieron a un análisis de varianza para retener los ítems que fueran entendidos de forma muy similar entre distintos jueces expertos y que fueran estables a lo largo del tiempo. Los tres ítems retenidos fueron los siguientes (Highhouse et al, 2009,785):

- Esta compañía tiene una reputación excelente.
- Esta compañía es admirada y respetada.
- Esta compañía está entre las mejores.

Se ha señalado que el problema con el índice de reputación de Highhouse y colegas es que el efecto de la minimización en la varianza inter-ítems ha podido producir unos ítems demasiado similares, que giran en torno a la idea de prestigio (Carreras et al. 2013). Al final, los tres ítems seleccionados apuntan a una única dimensión, el prestigio. El problema con este índice es que el campo semántico es muy limitado o ambiguo y que no reúne a los diferentes sentimientos que se asocian a la reputación. Por otro lado, es admisible la crítica de Lars Bergkvist y John Rossiter cuando afirman que la aproximación multi-indicador es superflua en incluso contraproducente cuando se trata de un solo indicador (Bergkvist y Rossiter 2007, 176).

## EL ÍNDICE DE REPUTACIÓN DE CASALÓ, FLAVIÁ Y GUINALÍ (2008).

El índice global propuesto por Luis Casaló, Carlos Flavia, Miguel Guinalí en el contexto del comercio electrónico nos ofrece un ejemplo relativamente reciente de medida multi-indicador basada exclusivamente en la idea directa de reputación.

Aunque los autores conciben la reputación en términos de “credibilidad”, cuando proponen los ítems que compondrán el constructo global, utilizan el mismo término de reputación en diferentes situaciones. En efecto, al comienzo de su artículo, afirman que “la reputación puede ser considerada desde una perspectiva más general asociándola a la credibilidad de la organización. En este caso la reputación sería la consecuencia de la comparación entre lo que la compañía promete y lo que efectivamente cumple”<sup>17</sup> (Casaló et al. 2008, 327).

Desde un punto de vista métrico, parece que los autores están optando por los estudios de credibilidad corporativa que miden el grado en que los individuos “sienten que la compañía tiene conocimiento o competencia para cumplir con sus declaraciones y si pueden confiar en que dicen o no la verdad”<sup>18</sup> (Newell y Goldsmith 2001, 235). Sin embargo, a la hora de seleccionar los ítems, se apartan del dominio teóricamente establecido que basaría la reputación global en el desarrollo de tres cuestiones como las de ser percibidos como experto, la integro y honesto (Newell y Goldsmith, 2001, 238). En su lugar, proponen cuatro aproximaciones a la buena reputación:

- Esta *website* tiene buena reputación.
- Esta *website* tiene buena reputación comparada con otras *websites* rivales.
- Esta *website* tiene la reputación de ofrecer buenos servicios.

---

<sup>17</sup> Traducción directa: “Nevertheless, reputation may also be considered from a more global perspective, associating it with the credibility of the organization. In this case, reputation would be the consequence of the comparison between what the company promises and what it eventually fulfils”

<sup>18</sup> Traducción directa: “(Corporate Credibility is the extent to which consumers...) feel that the firm has the knowledge or ability to fulfill its claims and whether the firm can be trusted to tell the truth or not”

- Esta *website* tiene la reputación de ser justa (*fair*) en las relaciones con sus usuarios.

Desde un punto de vista estrictamente métrico, el índice de Casaló y colegas presenta buenas propiedades psicométricas, con una alta uni-dimensionalidad, el 86,3% de la varianza total común, una buena validez convergente demostrada por índices de fiabilidad del componente superiores al 0,6 y AVE por encima del 0,5 y una buena validez discriminante, estas puntuaciones superan los estándares fijados para estas pruebas estadísticas permitiendo así afirmar que el resultado de las pruebas es sólido y riguroso (Casaló et al. 2008, 335). Sin embargo, en la fundamentación de la validez de contenido, el índice global es más bien pobre y no desarrolla el dominio teórico propuesto de la credibilidad y su relación con la reputación. Además, el hecho de basarse en la idea ingenua de buena reputación equivale a dejar indeterminados los sentimientos globales que subyacen a esta noción ingenua de la buena reputación.

## EL ÍNDICE PULSE DE PONZI, FOMBRUN Y GARDBERG (2011).

El índice Pulse fue diseñado para captar el contenido de los sentimientos generales que supuestamente configuran el concepto global de reputación y por tanto, la métrica resultante debería tener un campo semántico más acotado y un componente emocional significativo.

Este índice podría tener ventajas considerables frente a los descritos en las líneas anteriores.

La idea primigenia del índice surgió en los primeros estudios de la escala “*Reputation Quotient*” de Charles Fombrun. Dentro de la escala, Alloza se percató de que aparecía un factor distinto a los componentes más racionales que denominó “atracción emocional<sup>19</sup>” (Fombrun et al 2000, 254).

En el 2011, el equipo de Fombrun concibe este factor emocional como un índice global de reputación que es la síntesis de los componentes de la escala racional. En el texto de Leonard Ponzi se señala la conveniencia de “deslindar las palancas de la reputación (escala racional) del propio constructo (índice global)” (Ponzi et al. 2011, 16). En este punto coinciden con Scott Highhouse y colegas que consideran la reputación como una impresión global duradera.

Para descubrir cuáles son los sentimientos globales asociados a la reputación, los autores emplearon técnicas cualitativas de investigación basadas en grupos de discusión. Cuando los entrevistados fueron interrogados por lo que les sugería la idea de una reputación general (*overall reputation*), las repuestas fueron nucleándose en torno a los cuatro temas que forman los indicadores de la reputación global (Ponzi et al 2011, 23):

- Es una compañía que me transmite buena impresión.
- Es una compañía en la que confío.
- Es una compañía que admiro y respeto.
- Es una compañía que tiene buena reputación.

---

<sup>19</sup> El término en inglés “emotional appeal”

El índice Pulse ha sido la métrica más utilizada hasta el momento de la redacción de esta investigación por empresas e instituciones en el mundo, en diferentes *stakeholders* de los cinco continentes (Fombrun y Pan, 2006, 166; Ponzi et al, 2011, 20-21).

El índice parece haber demostrado disponer de buenas propiedades psicométricas de convergencia, uni-dimensionalidad y discriminación respecto del constructo de generación de intención de conducta favorable por parte de los *stakeholders*. En este sentido el resultado más prometedor se encuentra en su capacidad para explicar las conductas de apoyo en diferentes países y grupos de interés (Ponzi et al. 2011, 28).

Según los autores antes citados, el Pulse parece contener los ingredientes necesarios para constituirse en una buena referencia directa de la reputación corporativa, superando las limitaciones del resto de índices que hemos descrito.

A diferencia del índice de Brown, los ítems contemplados son sentimientos asociados a la idea global de reputación. Respecto del índice de Highhouse, capta mejor la heterogeneidad de los sentimientos asociados y con respecto al de Casaló y colegas, representa mejor el contenido de esa idea global. Por otro lado, el Pulse presenta la ventaja de la disponibilidad de información y bases de datos para diferentes sectores de actividad y países.

## **1.6. OTRAS MÉTRICAS DE LA REPUTACIÓN: AUDITORÍAS Y RANKINGS.**

Es relevante mencionar una serie de métricas de reputación corporativa que han tenido eco en la literatura y en la praxis empresarial, pero en nuestra investigación específica no resultan relevantes, no se trata de una observación crítica sobre estas métricas, que cuentan con una amplia literatura científica y un uso extendido en el ámbito empresarial, sino simplemente una constatación de su menor interés para los objetivos concretos de esta tesis.

Se trata de las llamadas métricas “objetivas” de reputación, los rankings de reputación y las auditorías (Carreras et al. 2013, 345).

Este tipo de métricas "objetivas" de la reputación coinciden en objeto y método. Parten del supuesto de objeto: la reputación corporativa es un fenómeno real que debe ser captado de la forma más objetiva posible mediante algún sistema de auditorías llevadas a cabo por expertos o por profesionales especializados con amplia información sobre la organización a evaluar. Las referencias que se consideran más relevantes en este ámbito son las dos siguientes.

### **MERCO -MONITOR EMPRESARIAL DE REPUTACIÓN CORPORATIVA (VILLAFAÑE, (2001).**

La definición del profesor Justo Villafañe defiende este enfoque metodológico cuando afirma que “la finalidad del proyecto **MERCO** o **Monitor Empresarial De Reputación Corporativa** fue crear una medida lo más fiable y objetiva posible, sobre el reconocimiento del comportamiento corporativo excelente” (Villafañe, 2004, 122). En consecuencia, dos rasgos son esenciales:

- La reputación implica un reconocimiento social y ese mismo reconocimiento es un fenómeno que puede ser captado de la forma más objetiva y fiable posible.
- Los sistemas de medida deben evitar el sesgo subjetivo mediante el uso de agencias/auditores independientes o mediante la consulta limitada a los *stakeholders* altamente especializados e informados (Cravens et al. 2003, 210).

## EL ÍNDICE DE REPUTACIÓN DE CRAVENS Y COLEGAS (2003)

Es una medida agregada obtenida por agencias y auditoras especializadas, basada en los juicios de valor de profesionales expertos sobre los componentes esenciales de la reputación corporativa. En consecuencia intenta captar de forma fiable el grado en que una empresa contiene objetivamente los ingredientes necesarios para generar una buena reputación en sus grupos de interés, lo que según los autores le da al índice mayor “credibilidad e importancia” (Cravens et al. 2003, 210). Las dimensiones que deben evaluar los auditores son:

- Calidad y valor de los productos/servicios (calidad y valor de los productos para el cliente, reconocimiento de la marca).
- Empleados (formación, imagen que presentan, tiempos medios de permanencia, número de solicitantes de empleo...)
- Relaciones externas (relaciones con proveedores, competidores, inversores, competidores, cuidado del medio ambiente, acción social y de filantropía).
- Innovación y creación de valor (desarrollo de nuevos productos, adecuación a las necesidades del cliente, retención de cliente).
- Viabilidad y fortaleza financiera (auditoría sobre informes financieros).
- Estrategia (riesgo de las elecciones estratégicas, sistemas de control e incentivos para mantener la empresa en una buena posición estratégica).
- El clima ético (la cultura, posibles violaciones éticas con repercusiones en grupos de *stakeholders*, existencia y aplicación de códigos éticos de comportamiento y buen gobierno en la organización).
- Obligaciones y responsabilidades (los compromisos adquiridos con los empleados, los procesos implementados, los sistemas de comunicación e información y la configuración de la organización, que podría afectar a la viabilidad de la organización).



**MERCO** ha propuesto un sistema de auditorías y evaluaciones en los que intervienen colectivos altamente informados como directivos, analistas financieros, asociaciones de consumidores, sindicatos y periodistas especializados.

Las dimensiones que generan reconocimiento serían las siguientes <sup>20</sup>:

- Los resultados financieros (beneficio, solvencia y calidad de la información económica).
- La calidad de la oferta comercial (valor del producto, de la marca, servicio al cliente).
- Reputación interna (calidad laboral, ética y profesionalidad, identificación con el proyecto de empresa).
- Ética y Responsabilidad Corporativas (compromiso empresarial, con la comunidad y responsabilidad social y medioambiental).
- Dimensión internacional de la empresa (países en los que opera, volumen de negocio en el extranjero, alianzas internacionales).
- Innovación (inversiones en I+D, nuevos productos/servicios e innovación y cambio).

La metodología **MERCO** sigue el principio de que la evaluación debe proceder de la “población mejor informada”. Los evaluadores sólo consideran el área que es de su especialidad, así los resultados financieros son considerados por los analistas financieros, la ética y responsabilidad corporativas, por las ONG’s y los periodistas, la calidad de la oferta comercial por las asociaciones de consumidores y la reputación interna por los sindicatos. Los directivos de empresas valoran todas las dimensiones.

El índice **MERCO** representa una métrica para obtener una representación lo más completa posible del reconocimiento social de la empresa, que intenta superar las carencias de los rankings de reputación que nacieron en el ámbito de las revistas del ámbito económico/financiero.

---

<sup>20</sup> “<http://www.merco.info/es/countries/4-es/rankings/2/pages/5-metodologia>”

## OTROS RANKINGS

### **-El Ranking AMAC (American Most Admired Companies).**

Algunas prestigiosas revistas del mundo de los negocios se han dedicado a elaborar “rankings de reputación” de las empresas de mayor tamaño. Uno de los primeros “Rankings” fue diseñado e implementado en los años ochenta por la revista *Fortune*: El Ranking AMAC<sup>21</sup>. El procedimiento es relativamente sencillo: se solicita a los participantes, CEO’s, directivos y analistas financieros que valoren las 10 empresas que ellos consideran más admiradas en su sector respecto de ocho cuestiones enlazadas con el prestigio de una empresa (Schultz et al. 2001, 27):

- Calidad del Management.
- Calidad de los productos.
- Innovación.
- Inversiones a largo plazo.
- Solidez financiera.
- Capacidad para atraer, desarrollar y mantener personal.
- Responsabilidad hacia la comunidad y el medio ambiente.
- Inteligencia en la utilización de los recursos corporativos.

Estos enfoques sobre la reputación, como las auditorías o los Rankings no van a ser objeto de nuestra investigación. Los hemos citado porque forman parte del mundo de las métricas de la reputación, pero no son objeto de esta investigación que se centra en otro tipo de métricas que miden la actitud colectiva de TODOS los *stakeholders* (incluyendo a la opinión pública, los clientes y clientes potenciales, los empleados y los accionistas).

---

<sup>21</sup>Fortune: “empresas más admiradas de América” (AMAC)

### **1.7. LIMITACIONES DE LAS AUDITORÍAS Y LOS RANKINGS DE REPUTACIÓN PARA LOS OBJETIVOS DE NUESTRA INVESTIGACIÓN.**

Las métricas de reputación basadas en auditorías y rankings tienen un valor claro para los gestores de la reputación, sin embargo para los objetivos concretos que perseguimos en nuestra tesis (evaluar para contribuir a la consolidación de la métrica de la reputación y su integración en la gestión), el enfoque de auditorías y rankings presenta algunas limitaciones que nos llevaron a descartar este enfoque.

1. El reconocimiento de una empresa establecido por observadores externos, no explica la aparición de las conductas de apoyo o de rechazo si no es de forma indirecta. Los resultados de estas auditorías y rankings activan las conductas de los sujetos, únicamente cuando “son asumidos” por estos. Sólo cuando los individuos prestan suficiente atención y son persuadidos por estos mensajes, es decir, sólo cuando estos mensajes generan o modifican sus actitudes es cuando el valor generado por la buena reputación puede emerger. En consecuencia estamos interesados en las medidas directas de reputación por parte de los *stakeholders* y no en sus fuentes de información como las auditorías o los rankings.
2. El reconocimiento establecido objetivamente tiene efecto en actitudes y comportamientos solo cuando es subjetivamente internalizado por los actores sociales que a la organización le interesa movilizar favorablemente (Benavides, 1993, 82-83). La subjetividad se vuelve un elemento esencial. La conducta de valor de los *stakeholders*, la compra del cliente, la cartera de acciones o el compromiso del empleado, no dependen de la fiabilidad del dato de reconocimiento, ni de su publicitación, sino de la forma subjetiva en que la información y el juicio sobre la empresa sea o no sea considerada, interpretada y asumida de forma subjetiva y personal .

## **1.8. IMPORTANCIA ESTRATÉGICA DE LAS MÉTRICAS DIRECTAS SOBRE LAS INDIRECTAS.**

Una de las ideas que se desprenden de los capítulos anteriores tiene que ver con la preeminencia de las medidas directas sobre las indirectas como variable focal sumaria para la gestión de la reputación corporativa. En efecto, la revisión científica parece preferir las métricas directas de reputación sobre las indirectas o racionales. Elaboramos aquí los tres argumentos apuntados en el texto de Carreras, Alloza y Carreras (2013):

1. Las medidas sumarias de reputación canalizan los efectos de las medidas más analíticas.

Los autores suelen coincidir en que las variables sumarias como se ha demostrado en los estudios de satisfacción, confianza o reputación, suelen canalizar los efectos de las escalas indirectas, de carácter más cognitivo que contienen los argumentos que generan estos sentimientos globales (Ryan et al. 1995, 12). Davies y colegas utilizaron un proxy de reputación global, la satisfacción global para demostrar la validez predictiva de la escala de carácter corporativo (Davies et al. 2004, 137). Walsh y Beatty comprobaron la validez de su escala reputación basada en el cliente (CBR) contra cuatro constructos globales, satisfacción, lealtad, confianza y recomendación (Walsh y Beatty, 2007, 138) y Vidaver-Cohen demostró una correlación muy alta entre las dimensiones y atributos del Reprak™ y la reputación global Pulse (Vidaver-Cohen, 2007, 281).

2. Las medidas sumarias son más exhaustivas y reúnen información de un mayor número de sujetos, tanto los que poseen una información detallada como los que sólo disponen de ideas o impresiones genéricas (Carreras et al., 2013).

En el texto del libro "Reputación Corporativa" tuvimos ocasión de comprobar que la actitud de reputación tenía un doble origen, como una actitud de tipo deliberativo donde la persona maneja información detallada o como una actitud con ideas y sentimientos globales, pero con escaso nivel de información. En la

revisión identificamos dos fenómenos que apuntan la presencia de este segundo tipo de actitud global de reputación:

- La activación automática de las actitudes (Fazio 2001).
- La reputación como “estatus eventual”, idea introducida por Wesley Shrum (1988) y rescatada recientemente por Donald Bergh y colegas (2010).

Ambos mecanismos son de orden actitudinal y por tanto desencadenan las conductas generadoras de valor para la empresa, pero no precisan de elaboración cognitiva y en este sentido son captadas de forma deficiente por la reputación racional. Aquí es donde la reputación global obtiene su ventaja como constructo sobre la versión cognitiva. Una persona poco informada puede emitir un juicio global sobre una empresa, con la misma convicción que un respondiente altamente informado. En el primer caso la actitud podría tener una carga afectiva mayor activada por los estereotipos sobre dicha empresa, mientras que en el segundo, la convicción tendría una base más racional, pero en ambos casos, la actitud desata una conducta de valor que favorece a la empresa como se afirma en Carreras, Alloza y Carreras (2013).

Obviamente los constructos globales están captando variación de elementos no contemplados en las escalas cognitivas y que también pueden incidir en la formación de la conducta de valor del *stakeholder*.

3. Las medidas sumarias muestran mayor capacidad predictiva sobre la conducta que las medidas indirectas.

Anders Gustafsson, Michael Johnson e Inger Roos afirman que las evaluaciones globales no sólo captan las razones sino que también predicen mucho mejor los efectos sobre las conductas que las escalas indirectas del performance (Gustafsson et al. 2005, 216). En este sentido, la reputación global captaría mejor los efectos de las escalas racionales (Vidaver-Cohen, 2007, 290) y a la vez predice mejor los efectos de la conducta actitudinal.

El estudio de Pratibha Dabhokar, David Shepherd y Dayle Thorpe aporta una evidencia adicional sobre la ventaja de utilizar medidas globales para mediar entre las medidas indirectas y la conducta. Los consumidores evalúan diversos aspectos relacionados con la calidad del servicio, pero también han generalizado una idea global de calidad que no es exactamente la suma de sus componentes y que tiene mayor capacidad predictiva sobre las intenciones de compra (Dabhokar et al. 2000, 166).

También Richard Bagozzi apoya esta idea cuando plantea la necesidad de mediar los efectos de las evaluaciones del performance sobre la intención a través de mediadores globales emocionales como la satisfacción (Bagozzi, 1992, 188). Leonard Ponzi y colegas también han resaltado el papel mediador de la versión global de reputación o como dicen la reputación “en formato reducido” del Pulse (Ponzi et al. 2011, 16).

En consecuencia concluimos que la reputación global cumple una función sumatoria de las razones deliberativas, capta otras activaciones automáticas provocadas por la percepción y evaluación de la compañía y muestra mejores propiedades predictivas sobre la conducta que desencadena en el *stakeholder*.

### **1.9. EL ÍNDICE PULSE COMO MÉTRICA DE APROXIMACIÓN DIRECTA A LA REPUTACIÓN : JUSTIFICACIONES DE LA INVESTIGACIÓN.**

En la revisión de la literatura académica, hemos examinado la conveniencia de utilizar las medidas globales de reputación como constructos o variables sumarias que expresan los sentimientos de confianza, admiración, afinidad emocional y respeto que despiertan las compañías en las audiencias y que miden la actitud general que mejor predice la conducta de valor por parte de dichos grupos de interés.

También hemos considerado la conveniencia de seleccionar el índice Pulse como medida de referencia entre otras diferentes alternativas existentes como las de Scott Highhouse y colegas (2009) o la de Steven Brown (1995) por razones tanto teóricas, porque capta una mayor variabilidad de sentimientos globales asociados a la evaluación general, como prácticas, debido a la alta disponibilidad de datos de Pulse en diferentes países.

En el caso del Pulse, la identificación de las emociones y sentimientos asociados al constructo fue explorado mediante investigación cualitativa con grupos de discusión con diferentes *stakeholders* (Ponzi et al. 2011).

Una vez identificadas las cuatro variables que representan el constructo de reputación global Pulse, se procede a construir una escala que represente la medida del constructo, ahora denominada variable latente. Normalmente las respuestas ofrecidas por cada persona en cada uno de los cuatro ítems, son agregadas<sup>22</sup> en una puntuación única que representa el nivel de reputación global para cada persona.

El segundo y tercer requisitos tienen que ver con las propiedades métricas que deben cumplir el conjunto de variables que forman la escala que mide el constructo. A este tipo de validez métrica, se le reconoce como “validez de constructo”. Una escala tiene validez de constructo si contiene los fenómenos observables que se suponen relacionados con el constructo, se relaciona con otros

---

<sup>22</sup> La agregación suele ser ponderada, según requisitos empíricos o teóricos. Para un examen detallado de los diferentes métodos de agregación, véase DiStefano et al. 2009

fenómenos observables en la forma en que teóricamente se espera que lo haga y no se relaciona con otros fenómenos observables con los que no debe hacerlo. La validez de constructo se demuestra mediante los cálculos de tres Sub-tipos:

(1) Validez convergente, grado en que los indicadores de un mismo constructo convergen o están altamente correlacionados entre sí.

(2) Validez discriminante o grado en que los indicadores de distintos constructos están menos relacionados, que los indicadores del mismo constructo.

(3) validez nomológica o grado en que un constructo se comporta con otros constructos en la forma en que se supone teóricamente que debe hacerlo (Shiu et al. 2001).

El objeto de nuestra tesis consiste en determinar el grado de invarianza del índice Pulse y en este sentido, de los tres requisitos asociados a un constructo nos centraremos en demostrar el primer tipo de validez, el denominado “validez convergente” (grado en que los indicadores de un mismo constructo convergen o están altamente correlacionados entre sí). Antes de establecer la invarianza de la combinación lineal de los cuatro atributos que componen el índice, necesitamos demostrar que convergen en un mismo concepto latente y que éste tiene un carácter unidimensional.

Las organizaciones necesitan poder contar con un indicador global de reputación que complemente a los indicadores financieros que están presentes desde hace décadas en estos cuadros de mando de alto nivel.

Sería útil poder emplear en la gestión de la reputación de las organizaciones un indicador global sintético, aceptado por un amplio consenso y fácil de comprender, que pudiera incorporarse a los cuadros de mando que utiliza la alta dirección para fijar la estrategia y evaluar su grado de cumplimiento (Kaplan y Norton, 1996, 76).

Este indicador global de reputación debería ser simple, aplicable en cualquier país, a cualquier sector de actividad, válido para los diferentes segmentos socio-económicos. Debe poder ser utilizado con independencia del sexo, la edad o la clase social de los individuos que se relacionan con la empresa, accesible y



utilizable por cualquier empresa con independencia de su tamaño, sector de actividad o lugar geográfico/contexto cultural en el que opere.

Las numerosas investigaciones destinadas a desarrollar una escala de reputación corporativa se han basado en la aproximación racional que consiste en medir los argumentos o razones que despiertan reputación, más que obtener medidas globales del sentimiento de reputación. Dentro de estas coordenadas tenemos la escala del carácter corporativo (Davies et al., 2004), la escala formativa de reputación (Helm, 2005), la reputación basada en el cliente (CBR de Walsh y Beatty, 2007) y otras numerosas propuestas de indicadores y rankings que han defendido el carácter exclusivamente racional de la reputación.

Desde el punto de vista métrico, el problema de estas escalas racionales es que no parecen cumplir con los requisitos de la invarianza métrica, necesarios para interpretar de forma no ambigua las medidas obtenidas (Carreras et al. 2013; Chun y Davies, 2006; Gardberg, 2006; Groeland, 2002; Lloyd, 2011; Walsh y Wiedmann, 2004). La invarianza métrica, también denominada psicométrica o invarianza de la medida, pretende establecer el grado en que los diferentes colectivos entienden y usan los ítems el constructo del mismo modo.

Sin equivalencia métrica no podemos estar seguros de que el indicador de reputación seleccionado es entendido y usado de la misma forma por los diferentes grupos de sexo, edad y sustrato social de una misma población.

En este sentido, si no se cumple la equivalencia tendríamos tantas medidas de reputación como segmentos de población o *stakeholders* tenga la organización o al menos éstas no serían comparables entre sí. Estos diferentes segmentos socio-económicos establecerían sistemas de valores e intereses que darían lugar a esquemas cognitivos diferentes y por tanto a reputaciones diferentes en cada uno de los segmentos para una misma organización.

La falta de equivalencia métrica que han demostrado las diferentes escalas e racionales de reputación supone una clara limitación para la gestión operativa de la reputación en las empresas.

Alloza se ha enfrentado a las limitaciones de las escalas racionales de reputación en sus tareas, durante más de diez años, como responsable de la gestión de la reputación de una de las principales entidades bancarias españolas en todos los países dónde está presente este banco (Dirección global de reputación corporativa del Grupo BBVA, 2002-2011)

Estas investigaciones realizadas en estos años, junto con esta investigación académica y los trabajos de otros investigadores están avanzado con el objetivo de intentar verificar la posibilidad de utilizar un indicador integral/global de reputación, dotado de un claro componente emocional, como podría ser el indicador Pulse que vamos a someter a prueba en este trabajo. Este indicador será objeto de análisis y su validez y rigor, serán puestos a prueba empírica en esta tesis para comprobar si cumple con las hipótesis que defendemos como requisitos necesarios de la equivalencia semántica y la equivalencia métrica que podrían darse en el Pulse.

En este sentido, el Ángel Alloza ha desarrollado en los últimos 8 años, en su rol de gestor de la reputación, en colaboración con amplio equipo de matemáticos, un amplio programa de investigación aplicada que ha permitido la creación de un cuadro de mando para la marca y la reputación corporativas que ha conseguido demostrar empíricamente algunas relaciones causales entre las mejoras en los indicadores de marca y los de reputación y algunas variables directas e indirectas de generación de negocio (Estévez, 2009, 18-19). El indicador de reputación que hemos empleado en estas investigaciones empíricas ha sido precisamente el Pulse. En la figura siguiente representamos conceptualmente el cuadro de mando con el que se ha podido demostrar empíricamente un efecto positivo entre los incrementos de Pulse y el impacto positivo en los indicadores del negocio, en sectores como banca, seguros, energía, grandes almacenes, gran consumo y telecomunicaciones (los datos reales son confidenciales, propiedad de las empresas pertenecientes a Corporate Excellence).

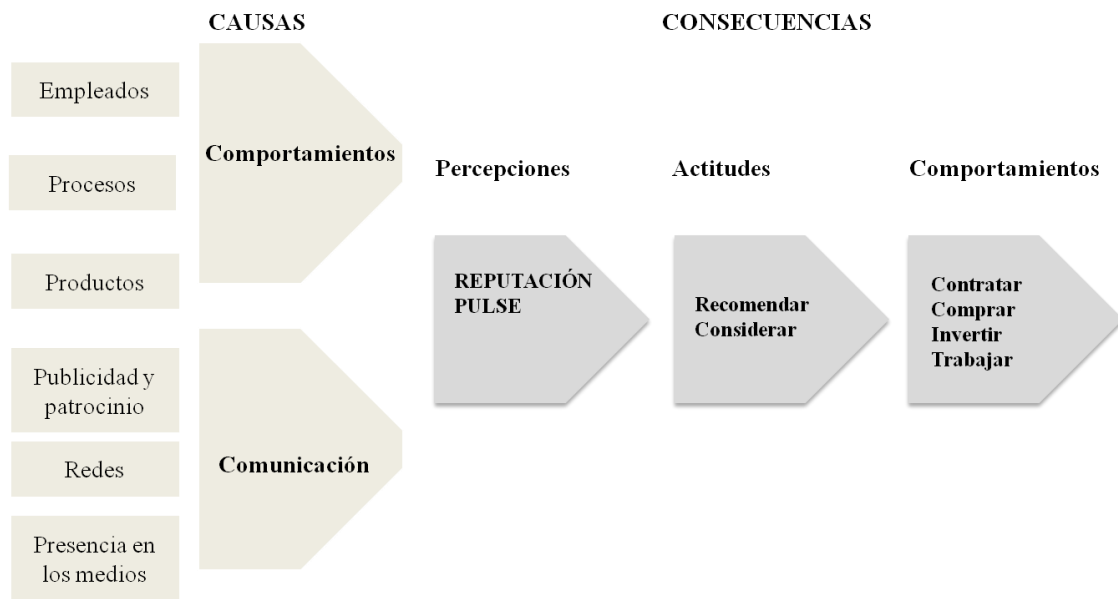


Figura nº 0.4. Adaptación del cuadro de mando de reputación de Corporate Excellence 2014.

Estos mismos efectos entre Pulse y otras variables de generación directa e indirecta de negocio o de capitalización bursátil también ha sido investigado y demostrado empíricamente por Fombrun y su equipo (Ponzi et al. 2011, 28).

En el gráfico siguiente se representa la relación positiva entre reputación y propensión a recomendar una empresa y sus productos/ servicios para una muestra de 120 empresas de distintos sectores de actividad.

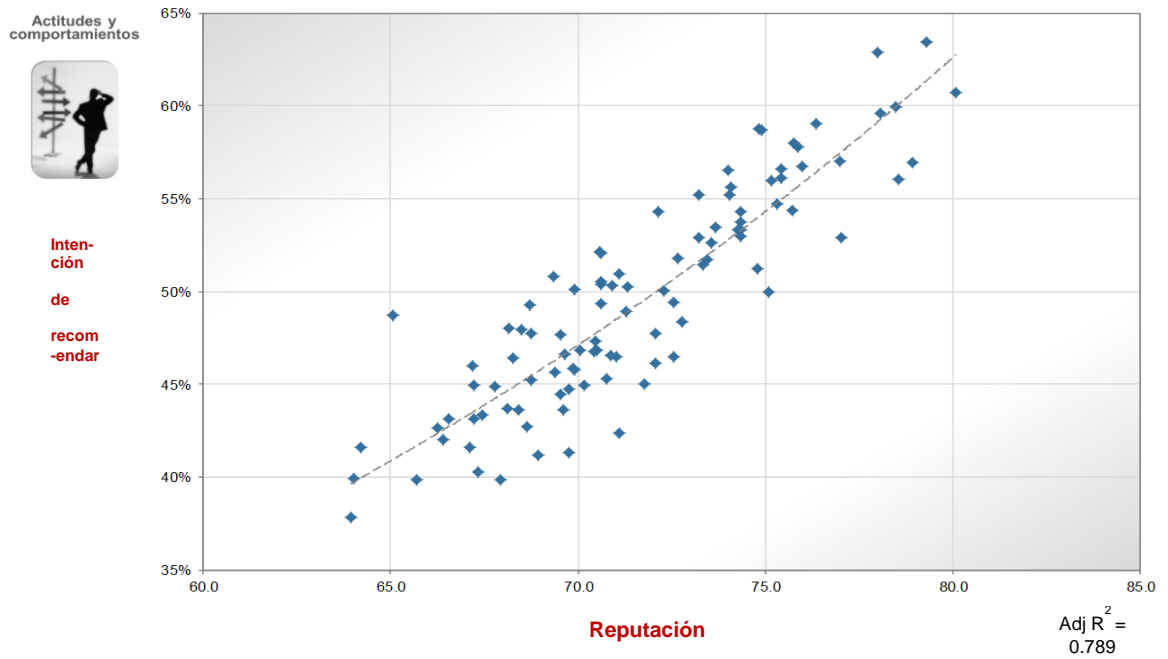


Figura nº 0.5. Correlación entre reputación y disposición a recomendar la empresa y sus productos/servicios. Fuente: Reputation Institute, 2012.

Podemos ver también la comparación de las empresas con mejor reputación del Standard & Poor Index frente a la cotización media de empresas incluidas en este índice bursátil.



Figura nº 0.6. Evolución valor en bolsa S&P 500 de las empresas con mejor reputación, Fuente Reputation Institute, 2013.

El análisis del comportamiento favorable en cotización bursátil de las empresas con mejor reputación, evaluadas por tener un elevado índice Pulse es una línea de investigación recurrente del equipo de Fombrun en el Reputation Institute.

Sabemos por tanto por estas investigaciones que existe una relación positiva entre el Pulse y los comportamientos favorables de los individuos (clientes y no clientes de una empresa) pero nos queda por demostrar que Pulse es válido tanto para hombres, mujeres, distintos grupos de edad y de clase social. Hoy sólo hemos demostrado el efecto positivo de Pulse sobre los comportamientos de la población pero este dato empírico no es suficiente para consolidar a Pulse como un indicador global de reputación que no pueda ser cuestionado por los académicos o los gestores de las organizaciones (trabajos internos, confidenciales, no publicados de Alloza, el equipo de investigación de Corporate Excellence y el equipo de matemáticos de Conento, 2012-2014). Necesitamos demostrar con suficiente rigor académico, la equivalencia semántica y métrica de Pulse.

Este paso es necesario puesto que buscamos evaluar a Pulse como posible indicador global de reputación, someter a contraste las distintas hipótesis como su invarianza poblacional con el fin de fortalecer su aceptación y aumentar el grado de consenso de Pulse, frente a otros múltiples indicadores de reputación existentes en el mercado. Este objetivo requiere el máximo rigor en la puesta a prueba de nuestras hipótesis para la defensa del indicador ante cualquier público: científico, académico, empresarial. Este es el motivo por el que hemos aplicado un número exhaustivo de pruebas matemáticas y estadísticas que consideramos necesario para conseguir este objetivo de fuerte validación.

Esta es esencialmente la justificación por la que Ángel Alloza aborda este trabajo de tesis.

### 2.1. CONSIDERACIONES GENERALES. FUNDAMENTOS METODOLÓGICOS.

El capítulo dos presenta todas las características técnicas de la metodología utilizada en la etapa empírica, tanto en la obtención de la información como en su análisis y tratamiento matemático y estadístico, para la realización de esta tesis doctoral.

La revisión teórica presentada en el capítulo uno es necesaria para concretar y delimitar el campo de estudio empírico concreto que hemos definido para esta investigación. Como señalan Busquet, Medina y Sort el encuadre teórico es esencial para la investigación científica así como un conocimiento experto del tema a tratar estando esta segunda condición cubierta por la experiencia y práctica profesional y académica de Alloza en el campo específico de la reputación corporativa (Busquet et al., 2006, 125-126).

La revisión teórica realizada aporta datos, información, resultados de investigaciones previas sobre el objeto de estudio que permiten generalizar más allá de los datos específicos analizados en un proceso de inferencia lógica. Sin embargo es necesario ir más allá de las conclusiones teóricas para contrastar de manera empírica el campo de investigación que hemos delimitado. Esta necesidad de introducir la verificación empírica ha determinado la selección de técnicas cualitativas y cuantitativas que permitan la obtención de los datos relevantes para el contraste, verificación o rechazo de las hipótesis de esta investigación.

El concepto de ciencia, como ha señalado Fernández Uría, consiste en " un conjunto estructurado y jerarquizado de leyes y teorías, que formulan los científicos, siguiendo un conjunto de pasos que hay que dar para la obtención de un fin, es decir, mediante la aplicación de un método científico.

Para Ferrater Mora, el método científico se determina por dos rasgos fundamentales, la racionalidad y la objetividad:

1. La racionalidad se fundamenta por su carácter crítico, por la búsqueda de criterios de verificación de sus enunciados e hipótesis y también por su aspiración a la coherencia, al buscar la conexión de las ideas dentro de un sistema lógico en el que no se generen contradicciones.

2. La objetividad es el criterio que cualifica un conocimiento como "conocimiento científico", en este sentido la objetividad está en la esencia del método científico.

No existe consenso sobre si existe un único método científico o se trata de una multiplicidad de métodos, Sierra Bravo (1984, 45) afirma que "se pueden distinguir tantas clases de métodos como tipos de actividades a que se refieren", citado por Balcells i Junyent (1999, 33).

En este mismo sentido, Bunge (2000, 38-42) nos dice que todas las ciencias comparten, de forma genérica y amplia un *método científico general*, si bien cada disciplina empleará técnicas diferentes en función de los problemas y objetos de estudio a los que se enfrente.

A pesar de este aparente falta de consenso, sí se puede afirmar que se distinguen algunos enfoques metodológicos específicos que se adecuan a los objetivos e hipótesis de cada investigación como han clarificado Busquet, Medina y Sort (2006, 26), concretando más, Hernández, Fernández y Baptista (2006) mencionan dos grandes clasificaciones: los enfoques cuantitativo (predominio de la lógica deductiva) y cualitativo (predominio de la lógica inductiva) como los dos paradigmas dominantes en la investigación científica en el campo de las ciencias sociales. Nos hablan de las similitudes entre ambos enfoques que supone que ambas técnicas se apalancan en cinco fases similares y relacionadas entre sí:

1. Realizan observación y evaluación de fenómenos
2. Establecen hipótesis o ideas a partir de la observación y evaluación realizadas
3. Demuestran el grado en que las hipótesis o ideas tienen fundamento
4. Revisan tales hipótesis o ideas sobre la base de las pruebas empíricas o del análisis realizado

5. Proponen nuevas observaciones y evaluaciones para esclarecer, modificar y fundamentar las hipótesis o ideas o incluso para generar otras

En nuestro caso para la elección de los métodos empleados, se han tenido en cuenta tres puntos de vista: a) el problema a investigar, b) las experiencias previas personales y profesionales del investigador, c) a quién se dirige este trabajo y el objetivo operativo se persigue, que se explica a continuación.

a) Desde la perspectiva del problema a investigar:

La investigación ha requerido una extensa revisión teórica y documental que incorpora tanto el marco teórico de la investigación más general sobre la reputación y su medición y revisión documental más concreta de los resultados de las investigaciones sobre los indicadores globales de reputación y en particular de los estudios previos realizados hasta la fecha con Pulse con sus avances y las limitaciones. Esta fase teórica y documental ha sido necesaria para fijar con precisión el foco concreto de la investigación empírica realizada con Pulse en este trabajo.

b) Desde la perspectiva de las experiencias previas personales y profesionales del investigador.

Durante el periodo de 2000 a 2014, desde el punto de vista de la generación de conocimiento y de la investigación empírica, Alloza ha participado en la fundación y en la gestión de las tres organizaciones, como miembro co-fundador del Foro de Reputación Corporativa (fRC), como responsable del área de métricas del fRC y como director general del Instituto de Análisis de los Intangibles (IAI) y como director y secretario general de la fundación, Corporate Excellence Centre for Reputation Leadership. Desde el punto de vista de la gestión empresarial de la reputación, durante el periodo de 2000 a 2011 Ángel Alloza ha sido máximo responsable global de la gestión, las métricas y sistemas de medición de la reputación, la marca corporativa, de la estrategia de comunicación del grupo bancario BBVA, desde la central en Madrid y con responsabilidad para todos los países y unidades de negocio.



En este sentido la metodología de investigación empleada en el capítulo 2 del marco teórico en la parte de revisión documental de la líneas de actuación empresariales, se ha basado en técnicas cualitativas, en concreto en el estudio de casos y en la narrativa como forma de investigar en este caso.

La metodología basada en el caso es pertinente y especialmente adecuada para abordar esta investigación, dado que es Alloza quién aporta su experiencia como “observador participante”, una técnica de investigación cualitativa aplicada con frecuencia en el estudio de casos por científicos sociales y por profesionales de las ciencias sociales y de la comunicación como es nuestro caso (Balcells i Junyent, 1994, 376 - 378).

Este elemento de la investigación se enmarca en el *action research* (Reason y Bradbury, 2001) y el *clinical research* (Schein, 2001, Miller y Crabtree, 2000). Se trata de métodos basados en casos prácticos que dotan a este documento de algunos apartados considerados como los adecuados por algunos autores y que fueron incorporados en el capítulo dos, en la parte correspondiente a las líneas de actuación empresariales (Young, 1960, 294): a) una introducción preliminar para identificar los datos; b) exposición de los problemas que se formulan; c) exposición de los métodos y técnicas empleados y d) registro de los datos más importantes que describen el caso relatado.

Finalmente la técnica de la narrativa ha sido el método de investigación empleado en ese capítulo dos. Esta técnica tiene su origen en los estudios antropológicos realizados en los Estados Unidos en la década de 1920 y han vuelto a utilizarse en la actualidad por la importancia que parece haber adquirido la investigación sociológica centrada en las “historias de vida” o narrativa de la vida cotidiana (Balcells i Junyent, 1994, 379).

De esta manera uno de los enfoques metodológicos utilizados en para la investigación se ha apalancado en el estudio de casos en base a la narrativa organizativa utilizada tanto en el proceso de investigación como en el de la presentación de los resultados en el capítulo dos (Czarniawska, 1998 y 2004).

c) Desde la perspectiva de a quién se dirige este trabajo y el objetivo operativo se persigue.

Se trata de difundir los resultados de esta investigación prioritariamente a la comunidad profesional de gestores de la reputación dentro de las organizaciones, empresas e instituciones, así como al resto de directivos gestores de las funciones más importantes en las empresas y al conjunto de inversores y reguladores.

Se persigue fortalecer el consenso sobre la importancia de medir la reputación como paso necesario para su gestión. En concreto se pretende poner el énfasis en un aspecto concreto de esta métrica, el indicador global de reputación Pulse. La posibilidad de contrastar o refutar las hipótesis de validación de la invarianza poblacional de Pulse tiene especial relevancia puesto que el objetivo principal de los gestores de la reputación es utilizar un indicador global de reputación con el suficiente rigor científico para que no sea cuestionado internamente en la organización o externamente por parte de terceros (académicos, expertos, reguladores, etc.).

El fin último de facilitar la aceptación por parte de la comunidad empresarial de su inclusión en los cuadros de mando de alto nivel de las organizaciones, complementando al resto de indicadores más relevantes (financieros y no financieros <sup>23</sup>) para influir en la estrategia, aportar una perspectiva de largo plazo y una visión multi-*stakeholder*.

El enfoque metodológico más adecuado para conseguir este objetivo es el cuantitativo. Como afirman Marzal y López Lita (apuntes asignatura metodología de investigación, 2011), la adopción de este enfoque cuantitativo exige una revisión de la literatura previa más profunda que la que demandan los enfoques cualitativos. En nuestro caso esta revisión documental ha jugado un papel crucial como guía de la investigación, algo que ha sido fundamental para acotar la hipótesis principal y las hipótesis secundarias, así como para el diseño de toda la investigación.

---

<sup>23</sup> El otro indicador es la “satisfacción del cliente”, pieza angular del desarrollo de los sistemas de gestión de la calidad.

## 2.2. REVISIÓN DE LA HIPÓTESIS.

Una vez analizado en la revisión documental del capítulo uno en el que se revisó el marco teórico, incluyendo tanto las líneas de investigación académicas como la empresariales se considera necesario volver a revisar la validez y pertinencia de las hipótesis, enunciada en el capítulo introductorio

Esta hipótesis nos sirvió como guía para orientar el trabajo de documentación del marco teórico y de esta misma forma va a determinar cuál es el diseño metodológico más adecuado del trabajo de obtención de la información y de la selección de las técnicas de análisis estadístico y matemático más pertinentes para demostrar o refutar la hipótesis de esta investigación. Como señala Ruiz,

*"en cualquier área del conocimiento científico el interés radica en poder plantear hipótesis,(...) para alcanzar una comprensión más amplia y profunda del origen, desarrollo y transformación de aquello investigado (...), a la ciencia le interesa confrontar sus verdades con la realidad concreta ya que el conocimiento, como se ha dicho, no puede considerarse acabado, definitivo, tiene que ajustarse continuamente, en menor o mayor grado según el área de que se trate, a la realidad concreta la cual se encuentra en permanente cambio en permanente cambio" (Ruiz L. Ramón, 2006, 131).*

Una vez que hemos analizado la hipótesis, desde esta perspectiva y a la luz de la revisión del marco teórico, se ha considerado que sigue siendo válida dadas las evidencias y argumentos aportados por la revisión documental recogida en el capítulo uno:

1. Las conclusiones de la etapa documental ponen en evidencia las posibilidades que tienen las escalas de reputación indirecta, que descubren cuáles son las razones que llevan a un sujeto a expresar una actitud global positiva o negativa hacia una organización y que se complementan con la evaluación global de la reputación.

2. Estas evaluaciones son altamente informativas, facilitan la diagnosis y son las responsables de la aparición del juicio global, y es ese juicio global el que, una vez formado, activa la conducta, los comportamientos favorables o de destrucción de valor.

3. Las métricas globales careciendo de dicha capacidad de diagnóstico, resultan sin embargo, excelentes para sintetizar el nivel de reputación que siente el sujeto. El índice global de reputación reúne toda la variación de la satisfacción y la confianza situándose como la resultante final del proceso. La reputación global incorpora de esta forma todos los componentes evaluativos, que son los antecedentes comprobados de la satisfacción y de la confianza.

4. De los índices globales de reputación revisados en la etapa documental el RepTrak™ Pulse parece incorporar los sentimientos generales asociados a la reputación global de las empresas como son la confianza, la buena impresión (cercano a la satisfacción), la admiración y el respeto. Este índice global de reputación incluiría los sentimientos y evaluaciones de los distintos *stakeholders* para predecir la actitud intencional, que es la responsable final de la conducta efectiva.

5. El índice RepTrak™ Pulse parece haber demostrado disponer de buenas propiedades psicométricas de convergencia, uni-dimensionalidad y discriminación respecto del constructo de generación de intención de conducta favorable por parte de los *stakeholders*. En este sentido el resultado más prometedor de la revisión documental realizada se encuentra en sus posibilidades para explicar las conductas de apoyo en diferentes países y grupos de interés

Se pone ahora de manifiesto la necesidad de avanzar en la investigación para comprobar empíricamente o para refutar las cualidades psicométricas que harían de RepTrak™ Pulse un indicador global de reputación con suficiente validez y rigor métrico.

La potencialidad de este indicador global de reputación requieren someterlo a una investigación empírica y a una serie de pruebas estadísticas y matemáticas que no permitan verificar o refutar la hipótesis planteada y las sub-hipótesis que son necesarias para su contraste.

Recordamos la hipótesis planteada,

***" Pulse mide la reputación global de una empresa u organización con independencia del sexo, edad o clase social de la población que sea entrevistada sobre esta cuestión" (invarianza poblacional).***

Sub-hipótesis:

***1. "La puntuación resultante de Pulse refleja fielmente la fortaleza de esa reputación" (equivalencia y fortaleza métricas).***

***2. "Los individuos de una población entienden lo mismo por reputación" (equivalencia semántica y contenido emocional).***

Como señala Visauta, *"la hipótesis es una anticipación, en el sentido de que propone ciertos hechos o relaciones que pueden existir, pero que todavía no conocemos o no hemos comprobado que existen"* (Visauta, 1989, 110).

Una vez que hemos validado la continuidad en el proceso de investigación y la validez de la hipótesis gracias al contraste de la etapa documental se ha considerado necesario abordar la utilización de técnicas de investigación cuantitativas y herramientas de análisis estadístico y matemático con el fin de verificar o refutar la hipótesis formulada en esta tesis doctoral.

### **2.3. METODOLOGÍA GENERAL. DISEÑO METODOLÓGICO.**

Metodológicamente, el proyecto adopta la forma de un análisis secundario, sobre datos ya obtenidos mediante encuesta a la población general en países distintos y para sectores de actividad empresarial diferentes, por el equipo de investigación y métricas del Reputation Institute, dirigido por su director de investigación y métricas, Profesor Leonard Ponzi que ha tenido la gentileza de facilitarnos las bases de datos que han sido sometidas un proceso que garantiza el anonimato por sujeto entrevistado y por empresa para la realización de esta investigación, asegurando así la confidencialidad de los datos de origen.

Estas bases de datos han sido procesadas por el programa AMOS 20, que dispone de algoritmos para estimar un sistema de ecuaciones estructurales basadas en estructuras de covarianzas. Una de las herramientas analíticas del programa es el análisis multi-grupo que analiza la invarianza de una escala entre los diferentes colectivos socio-demográficos, homogeneizados por sexo, edad y nivel de estudios. Este tipo de análisis es capaz de identificar el grado en que los 4 atributos o ítems que componen el Pulse resultan invariantes ante poblaciones con características socio-demográficas diferentes.

Para poner a prueba los objetivos de la investigación, la equivalencia semántica y métrica de Pulse necesitamos ponerlo a prueba en contextos culturales diferentes, por ello, se han seleccionado tres países con contexto socio-culturales diferentes, como son España, EEU y China.

Como ya hemos comentado, la equivalencia métrica exige demostrar que Pulse es válido con independencia del sector de actividad de las empresas, por este motivo se han contrastado los datos con dos sectores de actividad también diferentes, el bancario y las telecomunicaciones y se han estudiado dentro de cada sector a las empresas líderes en cada país y sector, un total de 69 empresas.

Estas 69 empresas que han sido evaluadas por la población general/opinión pública utilizando un índice idéntico en todos los casos: PULSE con sus 4 ítems/atributos.

Para demostrar la equivalencia semántica y métrica necesitamos demostrar que Pulse es válido en para todos los segmentos de edad, clase social y sexo. Por este motivo la investigación utiliza una muestra representativa de la población general de cada uno de los tres países compuesta por hombres y mujeres con edades comprendidas entre los 18 y los 64 años de todos los estratos socioeconómicos. Las explicaciones y justificaciones técnicas de las técnicas de muestreo se han incluido en el capítulo 2 de fundamentación metodológica del trabajo empírico realizado.

### **2.3.1. JUSTIFICACIÓN DE LA MUESTRA ESCOGIDA.**

Con el fin de poder generalizar los datos obtenidos a toda la población se han seleccionado muestras probabilísticas de la población general que son representativas de las poblaciones de los tres países en los que se ha realizado la investigación empírica ( España, EEUU y China). La composición de las muestras en los tres países representa al universo de referencia: la población de individuos entre 18 y 64 años, ambos sexos y todas las clases sociales quedan representadas en las muestras.

La selección de los tres países, los diferentes segmentos de población y los dos sectores de actividad obedecen a la intención de poner a prueba las hipótesis de la investigación en contextos socio-culturales y de madurez empresarial en la gestión de la reputación muy distintos (por este motivo la elección de estos tres países en concreto), de la misma forma la selección de banca y telecomunicaciones responde al objetivo de analizar dos sectores de actividad muy diferentes. En los tres países fue posible tener acceso a bases de datos robustas, representativas de las poblaciones a analizar.

La naturaleza de los datos es por tanto cuantitativa, se trata de datos numéricos que se han incorporado a las tres bases de datos, eliminando las referencias personales de los sujetos entrevistados para garantizar el anonimato necesario pero conservando sus registros de sexo, edad, y nivel de estudios a nivel estadístico agregado y anónimo.

Dado que el objetivo principal de esta investigación es poner a prueba el índice global de reputación Pulse para verificar o refutar *su invarianza poblacional o grado de equivalencia métrica del índice global Pulse* entre los grupos de edad, sexo y nivel de estudios.

Este es el motivo por el que hemos seleccionado diferentes países, sectores de actividad e individuos de diferente sexo, edad y estatus social:

- 3 países con contextos socio-culturales y con niveles de desarrollo en la gestión de la reputación empresarial diferentes (España, EEUU y China). La elección de estos tres países concretos obedece por tanto a incorporar tres contextos socio-culturales muy diferentes en tres continentes distintos y con grados de maduración en la gestión de la reputación también diversos (China dónde la gestión de la reputación empresarial es prácticamente inexistente, EEUU como el país del continente americano más avanzado en esta materia y España que tiene un nivel medio en Europa en cuanto a la gestión de esta materia, valoraciones que proceden de la investigación realizada en 2014 por el REPUTATION INSTITUTE: "Reputation Leaders Study").
- 2 sectores de actividad/industrias diferentes en cada país: banca y telecomunicaciones. Excepto en China dónde solo hemos utilizado el sector banca por falta de disponibilidad de bases de datos de otros sectores de actividad. El motivo de selección de banca y telecomunicaciones ha obedecido a las grandes diferencias en los modelos de negocio de estas industrias/sectores de actividad.
- 3 variables socio-demográficas: sexo, edad y nivel de estudios. Siendo el nivel de estudios una variable estrechamente vinculada a la clase social/estatus social de los individuos para Hoffmeyer-Zlotnik, Jürgen (2008).

Desde el punto de vista metodológico, el índice Pulse, se presenta como un constructo psicológico que no sólo puede ser observado indirectamente a través de un conjunto de cuatro atributos asociados a él. Técnicamente adopta la forma de



una “variable latente” que subyace al comportamiento observado de los cuatro atributos que se integran dentro de su dominio de contenido semántico.

Según Nunnally y Berstein, cuando postulamos la reputación global como un constructo, debemos diagnosticar su validez resolviendo tres cuestiones (Nunnally y Berstein, 1997, 112):

1. Especificar el dominio de contenido de variables relevantes.
2. Determinar el grado en que las variables seleccionadas miden el constructo propuesto y no otro constructo
3. Determinar si las relaciones entre las métricas de diferentes constructos son consistentes con la teoría sustantiva.

El primer requisito alude al criterio de la validez de contenido. Según Seth Noar, una escala tiene validez de contenido cuando todas las dimensiones relevantes para el constructo forman parte de la escala (Noar, 2003, 627). Al menos hay dos formas básicas para garantizar que el dominio empírico del constructo está bien representado por la escala métrica, la “validez representacional” que garantiza que las dimensiones teóricas relevantes están presentes en la escala y la “validez aparente<sup>24</sup>” un grupo de expertos juzga si las dimensiones son esenciales en dicho concepto (Rossiter, 2002).

En este sentido necesitamos probar o refutar para el índice Pulse compuesto por 4 atributos:

- Que los distintos grupos de sexo, edad y nivel estudios de una población utilizan los mismos indicadores del índice Pulse. Si se demuestra la equivalencia configural, los grupos sociales estarían utilizando los mismos términos a la hora de crear su puntuación global de reputación. Con el fin de añadir rigor estadístico a esta prueba no nos hemos limitado a ponerlo a prueba en un único país,

---

<sup>24</sup> “Face validity”

sino que hemos seleccionado 3 países diferentes para comprobar o refutar esta hipótesis.

- Que el significado de los 4 atributos que componen el Pulse es el mismo para los todos los segmentos poblacionales. Esta prueba de la supuesta equivalencia métrica de los ítems se ha sometido al rigor de testarlo en 3 países diferentes con idiomas muy diferentes y con 2 sectores de actividad diferentes (para comprobar si los significados de los 4 atributos que componen la reputación global Pulse son los mismos en tres idiomas diferentes y a la hora de evaluar a empresas/sectores de actividad dispares).

Estos son los motivos que justifican haber seleccionado para nuestra investigación: 3 muestras representativas de la población de 3 países distintos, 3 segmentos de población (sexo, edad y estatus social) y 2 sectores de actividad diferentes en los que se ha medido la reputación global utilizando el mismo índice con idénticas preguntas (Pulse).

El segundo motivo de justificación de las muestras escogidas es puramente operativo y no obedece a criterios científicos o técnicos, se trata de las limitaciones de disponibilidad de bases de datos con un número suficiente de casos (en el apartado de limitaciones de los resultados recogemos este punto como una de las limitaciones de la investigación).

En el capítulo siguiente detallamos las características específicas de las muestras poblacionales seleccionadas en cada uno de los 3 países que componen las bases de datos que hemos utilizado para realizar las pruebas estadísticas y matemáticas que nos permitirán aceptar o rechazar las hipótesis de la investigación.

## **2.4. EL ANÁLISIS DE LAS FUENTES DE DATOS.**

El capítulo dos describe las características de la metodología de colección de datos empleada en la etapa empírica y detalla las especificaciones técnicas de los diseños de muestra, de cuestionario y organización de los trabajos de campo que han dado lugar a las bases de datos para la investigación de esta tesis doctoral.

Dado que el objetivo principal de esta investigación se centra en verificar el potencial del índice de reputación Pulse como constructo único que sintetice la actitud global que mantienen todos los segmentos de población con independencia del sexo, la edad o la clase social, en lugar de tener que contar, como hasta ahora, con las escalas racionales, con una métrica de reputación para cada segmento.

Recordemos que detrás del objetivo operativo de esta investigación se encuentra la necesidad de defender la adopción por parte de las empresas, organizaciones e instituciones de un único indicador global de reputación con equivalencia métrica y fortalecer el argumento de la necesidad de incorporar este indicador global de reputación (Pulse) en los cuadros de mando que utilizan las empresas para fijar y evaluar el cumplimiento de su estrategia.

La investigación de esta tesis busca precisamente esta verificación empírica del Pulse. Para cumplir con este objetivo se han seleccionado las metodologías de investigación y las técnicas estadísticas y matemáticas que se han considerado necesarias y suficientes con un elevado nivel de exigencia para verificar o refutar las hipótesis planteadas.

El diseño de la investigación adopta, como ya hemos señalado la forma de un análisis secundario a partir de las bases de datos individuales procedentes de encuestas destinadas a obtener medidas del índice Pulse en los siguientes países y sectores para la población general con sujetos comprendidos entre los 18 y 64 años de edad.

#### **2.4.1. LAS ENCUESTAS PROCEDENTES DEL GLOBAL REPTRAK™ PULSE.**

En consecuencia, el método de recogida de la información ha sido la encuesta probabilística a la población general, con la aplicación de un cuestionario estructurado y cerrado que incluye las cuatro preguntas del Pulse, las preguntas sobre intención de conducta y las correspondientes a las variables socio-demográficas, sector de actividad de la empresa, sexo, edad y nivel de educación alcanzado por el entrevistado. Este cuestionario forma parte de una serie de investigaciones internacionales realizadas anualmente por el Reputation Institute, dirigidas por el Profesor Leonard Ponzi, esta serie recibe el nombre de "GLOBAL REPTRAK PULSE".

La fecha de realización de las entrevistas en los tres países ha sido la misma: los trabajos de campo se ha llevado a cabo durante los meses de enero y febrero del año 2012.

#### **2.4.2. MÉTODO DE COLECCIÓN Y SELECCIÓN DEL INFORMANTE.**

Las tres BBDD proceden de encuestas online realizadas sobre paneles de población con muestras probabilísticas representativas de la población general. Para asegurar la representatividad de la muestra se aplicaron técnicas de ponderación por sexo y edad cuyos pesos corresponden a los datos estadísticos oficiales de cada país. La técnica de recogida de la información se basa en la aplicación de un cuestionario estructurado y cerrado cumplimentado online por los entrevistados. Con el fin de asegurar mejor el control de la consistencia de los datos todas las entrevistas han sido realizadas por el mismo proveedor de campo. Las muestras proceden de los paneles de población general disponibles en cada país.

La técnica para localizar al respondiente ha sido el "*screening*", o técnica de selección que permite seleccionar un subconjunto de la población general que cumple con los requisitos y objetivos de la investigación. El "*screening*" o selección se lleva a cabo aplicando preguntas filtro para averiguar si se trata de un

entrevistado cualificado para responder. Cada entrevistado debe cumplir con los siguientes requisitos para ser elegible para la entrevista:

- Filtro de participación: el sujeto entrevistado no debe haber participado en ninguna investigación realizada por el Reputation Institute en los últimos tres meses previos a la entrevista.

- Filtro de Familiaridad: a los potenciales entrevistados se les presenta una lista aleatoria con un máximo de 20 empresas a evaluar. Para cada empresa, se les aplica la siguiente pregunta: ¿En qué medida conoce usted a las siguientes empresas? ¿Diría usted sobre que la conoce bastante, la conoce algo, sólo la conoce de oídas o que no la conoce?. Las respuestas se clasifican en 4 posibles alternativas:

- No la conoce.
- La conoce sólo de oídas.
- La conoce algo.
- La conoce bastante.

Únicamente las personas que contestan "la conoce algo o la conoce bastante" son elegibles para continuar con el proceso de la entrevista.

- Filtro de los atributos de Pulse: el índice Pulse se compone de 4 atributos evaluados en una escala numérica de 1 a 7 puntos, para que un sujeto sea elegible para la entrevista debe contestar al menos a 3 de los 4 atributos, es decir que sólo se admite que pueda contestar "no sabe/ no contesta" a 1 de los 4 atributos.

### **2.4.3. TAMAÑOS DE MUESTRA, PONDERACIÓN, NIVEL DE CONFIANZA Y MÁRGENES DE ERROR.**

Con el fin de asegurar la representatividad estadística de las muestras sobre el universo de referencia (población general de cada país), las muestras de sujetos que superan los filtros de selección anteriores se someten a un proceso de ponderación por criterios de sexo y edad en cada país. Se utilizan datos oficiales de los censos de cada país para efectuar estas ponderaciones. No se pondera ninguna otra variable socio-demográfica.

Los intervalos de edades son los siguientes: 18-24, 25-34, 35-44, 45-64. Se emplea un proceso de ponderación denominado "*rim weighting*" (o más formalmente ajuste proporcional iterativo).

Se trata de un algoritmo para ponderar las muestras en sus datos marginales con los correspondientes a los censos oficiales de población para asegurar la buena representatividad de la muestra que se empleará en los análisis estadísticos y matemáticos posteriores.

Se utiliza esta técnica de ponderación porque reduce los posibles sesgos de las muestras, reduce al mínimo el factor de corrección empleado. al tiempo que asegura la extrapolación de los resultados a la población en general de cada país.

Los tamaños de la muestra para cada país, aseguran la precisión de los resultados en cada país y han sido las siguientes:

- España: 4905 sujetos entrevistados.
- EEUU: 5218 sujetos entrevistados.
- China: 939 sujetos entrevistados.

Los resultados de medición del Pulse a lo largo de los años y con grandes bases de datos procedentes de diferentes países y sectores de actividad han permitido establecer algunos patrones generales sobre la distribución de las puntuaciones globales del índice Pulse que expresamos en los dos siguientes gráficos facilitados por Leonard Ponzi para esta tesis:

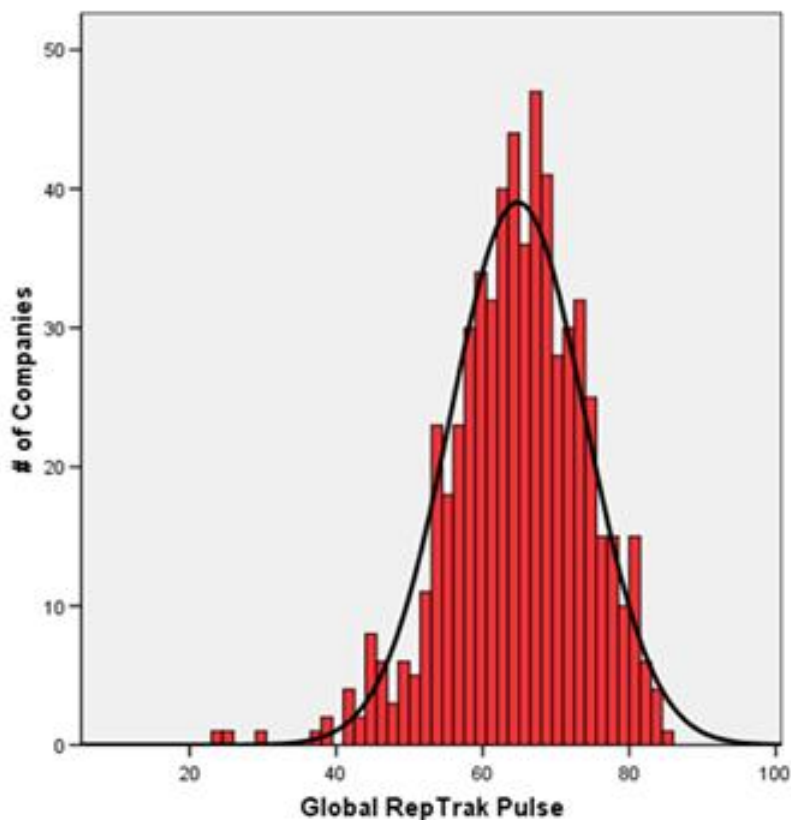


Figura nº 0.7. Distribución del Índice Pulse en el mundo (todos los países y todos los sectores de actividad). Gráfico proporcionado por Leonard Ponzi para esta investigación (2014).

Este análisis agrupa las bases de datos de todos los países. En el gráfico se percibe que las muestras proceden de una distribución teórica normal con esperanza y varianzas determinadas que han servido de base para establecer una escala normativa, de 1 a 100 puntos, que sirve para clasificar las puntuaciones globales de reputación en 5 niveles que van desde: una reputación global excelente (Índices Pulse que superan los 80 puntos), hasta una reputación global muy baja (por debajo de 40 puntos).

Excellent/Top Tier	above 80
Strong/Robust	70 – 79
Average/Moderate	60 – 69
Weak/Vulnerable	40 – 59
Poor/Lowest Tier	below 40

Tabla nº 0.2. Clasificación en categorías de los diferentes intervalos de reputación.

#### 2.4.4. TAMAÑOS Y MÁRGENES DE ERROR PARA LA BBDD EN ESPAÑA.

España	
n	4905
Varianza	0,25
ET <sup>2</sup>	3,8416
margen de error	
0,014	

El tamaño de muestra para España ha sido de 4905 entrevistas válidas para el análisis. Con un nivel de confianza del 95% (1,96 errores típicos de la distribución muestral del estadístico) y para la estimación de proporciones con máxima variabilidad ( $p=q=0,5$ ), el margen de error bajo el supuesto de muestreo aleatorio simple y población infinita<sup>25</sup> es de  $\pm 1,4$  %.

Aunque el panel son internautas, la representatividad queda garantizada por el alto índice de penetración del internet en los hogares españoles que, según datos del INE para mayo del 2012 era prácticamente del 70%<sup>26</sup> y el esquema de ponderación administrado.

Tabla nº 1. Cálculo del margen de error global para la BBDD de España

#### 2.4.5. BASE DE DATOS: REPRESENTATIVA DE LA POBLACIÓN DE EEUU.

EEUU	
n	5218
Varianza	0,25
ET <sup>2</sup>	3,8416
margen de error	
0,012	

En la BBDD de Estados Unidos se han utilizado 5218 entrevista. Con un nivel de confianza del 95% (1,96 errores típicos) y para la estimación de proporciones con máxima variabilidad ( $p=q=0,5$ ), el margen de error bajo el supuesto de muestreo aleatorio simple y población infinita es de  $\pm 1,4$  %. En EEUU la representatividad queda garantizada por la alta penetración del internet en el hogar, 74,8% y el

esquema de ponderación<sup>27</sup>.

Tabla nº 2. Cálculo del margen de error global para la BBDD de EEUU.

---

<sup>25</sup>  $e = \sqrt{\frac{ET^2 * S_p^2}{n}}$  Siendo: “e” el margen de error, “ET” los errores típicos, “ $S_p^2$ ”, la varianza del estimador y “n” el tamaño de muestra.

<sup>26</sup> Datos INE Instituto Nacional de Estadística. Encuesta sobre Equipamiento y Uso de Tecnologías de Información y Comunicación en los Hogares. Año 2013.

<sup>27</sup> Source: United State Census Bureau, Current Population Survey, 2012.



#### 2.4.6. BASE DE DATOS REPRESENTATIVA DE LA POBLACIÓN DE INTERNAUTAS DE CHINA.

	<b>China</b>	La BBDD de China dispone de 939 casos válidos. Con un nivel de confianza del 95% (1,96 errores típicos) y para la estimación de proporciones con máxima variabilidad ( $p=q=0,5$ ), el margen de error bajo el supuesto de muestreo aleatorio simple y población infinita es de $\pm 3,2\%$ . La base de datos de China es poco representativa de la población general. En este caso, dada la baja proporción de internautas respecto a la población general del país que en 2012 era inferior claramente al 47,4% <sup>28</sup> , las características socio demográficas de la muestra son específicas de la subpoblación de referencia, los internautas. Veremos estas características especiales en los epígrafes siguiente.
	n	
	939	
<b>Varianza</b>	0,25	
<b>ET<sup>2</sup></b>	3,8416	
	margen de error	
	0,032	

respecto a la población general del país que en 2012 era inferior claramente al 47,4%<sup>28</sup>, las características socio demográficas de la muestra son específicas de la subpoblación de referencia, los internautas. Veremos estas características especiales en los epígrafes siguiente.

Tabla nº 3. Cálculo del margen de error global para la BBDD de China.

<sup>28</sup> El dato 47,4% se corresponde al año 2014, por tanto, en el 2012 debía de ser inferior. Source: Internet Coaching Library. Internet World Stats: usage and Population Statistics 2014.

## 2.4.7. DESCRIPCIÓN DE LA MUESTRA UTILIZADA EN ESPAÑA.

La base de datos para España contiene las repuestas de ciudadanos que han evaluado la reputación global de veinte empresas en el sector bancario, con un promedio de 185 valoraciones por empresa y siete empresas en el sector de las telecomunicaciones, con 171 casos por empresa. Las distribuciones reflejan una variación suficiente por sexo edad y nivel de estudios como para abordar las hipótesis de investigación.

### 2.4.7.1. LAS DISTRIBUCIONES POR SEXO Y EDAD EN ESPAÑA SON LAS SIGUIENTES:

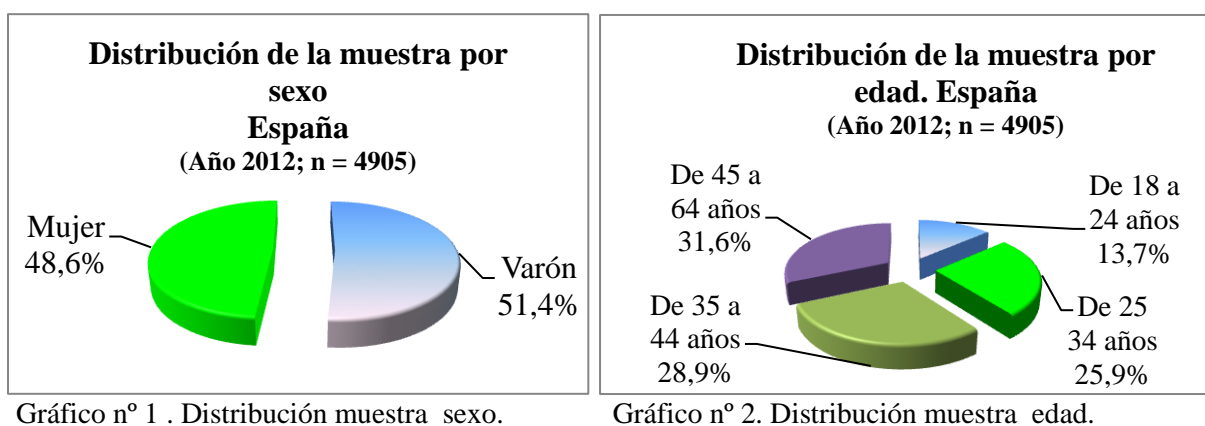


Gráfico nº 1 . Distribución muestra sexo.

Gráfico nº 2. Distribución muestra edad.

### 2.4.7.2. LA DISTRIBUCIÓN POR NIVEL DE ESTUDIOS EN ESPAÑA:

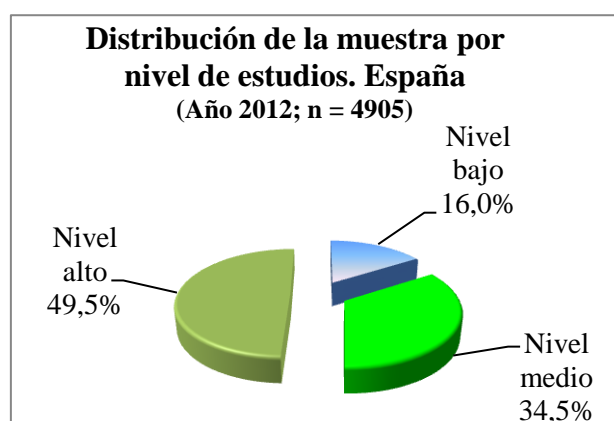


Gráfico nº 3. Distribución población por nivel de estudios.

### 2.4.8. Descripción de la muestra utilizada en EEUU.

La base de datos de EEUU contiene la evaluación de treinta empresas en el sector bancario, con un promedio de 130 valoraciones por empresa y seis empresas en el sector de las telecomunicaciones, con una media de 218 evaluaciones por empresa. También en este caso las distribuciones son variadas por sexo y edad, pero en el caso del nivel de estudios presentan una homogeneidad mayor, predominando los respondientes con estudios altos.

#### 2.4.8.1. DISTRIBUCIONES POR SEXO Y EDAD SON LAS SIGUIENTES EN EEUU.

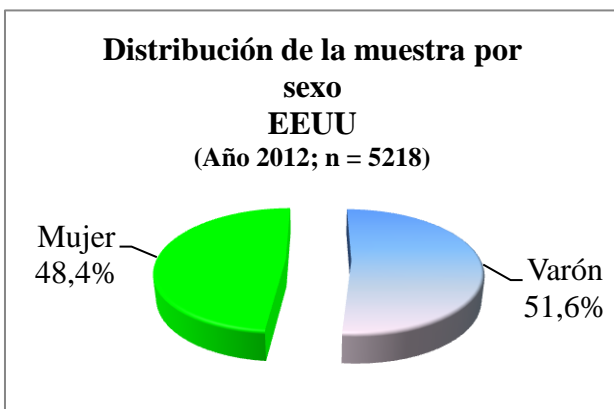


Gráfico nº 4. Distribución sex0. edad.

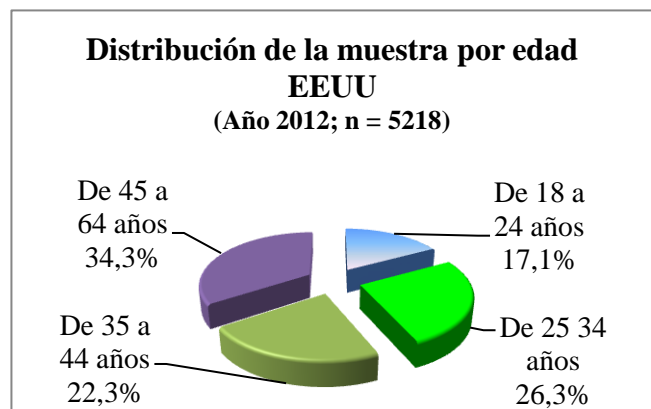


Gráfico nº 5. Distribución

#### LA DISTRIBUCIÓN POR NIVEL DE ESTUDIOS: EEUU.

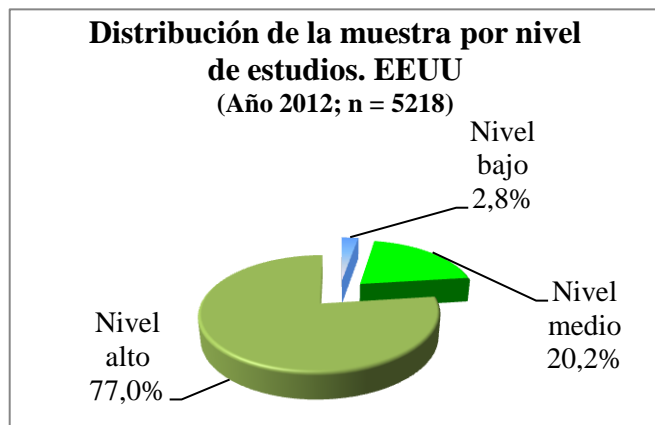


Gráfico nº 6. Distribución de la muestra por nivel de estudios.

#### 2.4.9. DESCRIPCIÓN DE LA MUESTRA UTILIZADA EN CHINA.

En China se han evaluado de seis empresas en el sector bancario, con un promedio de 157 evaluaciones por empresa. No contiene ninguna empresa del sector de las telecomunicaciones por no ser posible disponer de esta información para ese país.

**DISTRIBUCIONES POR SEXO Y EDAD SON LAS SIGUIENTES**

**CHINA:**

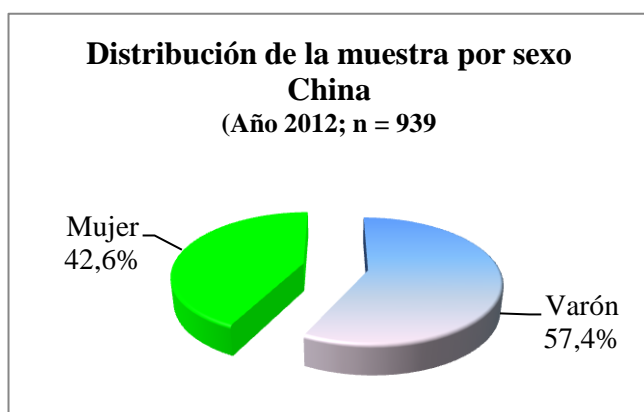


Gráfico nº 7. Distribución sexo.

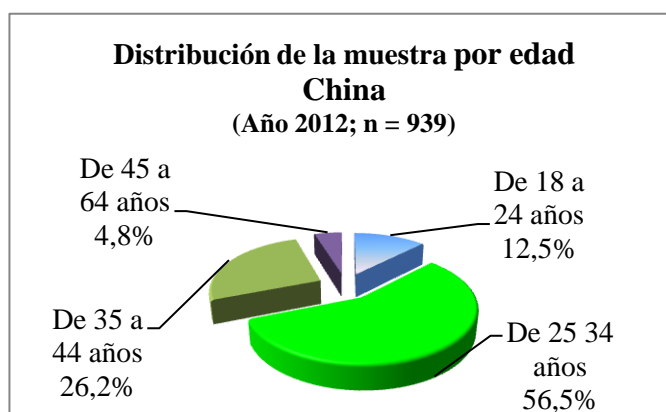


Gráfico nº 8. Distribución edad.

**LA DISTRIBUCIÓN POR NIVEL DE ESTUDIOS EN CHINA:**

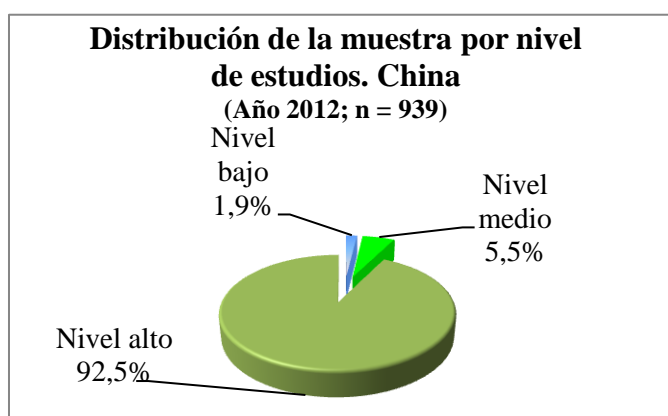


Gráfico nº 9. Distribución de la muestra por nivel de estudios.

Por el contrario en el caso Chino, la muestra se muestra excesivamente homogénea en la variable nivel de estudios, el 92,5% de los respondientes han declarado tener un nivel alto.

Las estructuras de la población que ha evaluado la reputación global de las empresas de los sectores de telecomunicación y banca de población de España y EEUU por sexo y edad son prácticamente idénticas. Aproximadamente el 51% de la población es masculina, más de la mitad de las muestras tienen entre 35 y 65 años (España, el 60,5% y EEUU el 56,6%). El nivel de estudios de las personas que han evaluado ambos sectores es superior en EEUU, mientras España el 49,5% dice tener estudios superiores al bachiller superior, en EEUU ese porcentaje asciende al 77%.

Sin embargo, el perfil demográfico del respondiente en China es completamente distinto. Predominan los varones con un 57,4%, el grupo de personas entre 35 y 65 años es menor del 31,0% y el porcentaje de personas con estudios altos alcanza al 92,5% de los entrevistados. Estas características de la muestra responden a la población a la que representa, no se trata de un error. Tal y como hemos apuntado antes el perfil de internautas se aleja del perfil de la población general. Recordemos que el método de entrevistas es un panel de sujetos a los que se aplican entrevistas con cuestionarios cumplimentados online.

La muestra en China nos permitirá verificar o refutar las hipótesis de invarianza poblacional de Pulse por sexo y edad pero no por nivel de estudios/clase social puesto que la muestra en este segmento de población es homogénea y solo representa a la sub-población con altos niveles de formación/clase social.

En este sentido y dada la estructura de población disponible, la muestra de China no resultará útil para comprobar la invarianza por nivel de estudio y por este motivo no la utilizaremos en ese análisis concreto.

## 2.5. TRATAMIENTO DE LA INFORMACIÓN PERDIDA EN LAS BASES DE DATOS.

En cualquier investigación que utiliza técnicas cuantitativas y un método de encuesta basado en la aplicación de cuestionarios, se produce siempre un porcentaje de sujetos que no contestan a determinadas preguntas. Cuando este porcentaje de respuestas "no sabe/no contesta" (casos con información perdida) supera el 20% en investigaciones para medir la imagen/reputación se considera que se puede producir un sesgo que afectaría a la robustez de los resultados según Schafer, J.L. y Graham, J.W. (2002, 148).

En nuestro caso, la cantidad de casos con información perdida son mínimos, en ningún caso supera el 4%, con lo que el sesgo esperado por la pérdida de estos casos es prácticamente despreciable. La tabla siguiente presenta la distribución de los casos con información perdida en alguna de las variables utilizadas para el análisis, los cuatro ítems del Pulse y la tres variables socio-demográficas.

BBDD	Total casos	Casos completos	Casos perdidos	% perdidos
España	4905	4766	139	0,0283
Estados Unidos	5218	5017	201	0,0385
China	939	932	7	0,0075
<b>Total</b>	<b>11062</b>	<b>10715</b>	<b>347</b>	<b>0,0314</b>

Tabla nº 4 . Casos con información completa en las BBDD de España, EEUU y China

El tipo de análisis matemático/estadístico que vamos a emplear para testar nuestras hipótesis en el tratamiento estadístico de las bases de datos requieren disponer de toda la información completa en los casos analizados, es importante resolver el problema de la falta de información en algunas variables.

Procedimos a eliminar los casos con información incompleta y nos quedamos sólo con los casos con información completa siguiendo el procedimiento "listwise" recomendado por Little, R. J. A. y Rubin, D. B. (1989, 295-296).

## 2.6. FORMACIÓN DE GRUPOS DE COMPARACIÓN PARA ANÁLISIS DE INVARIANZA.

Otros de los tratamientos previos a las tres bases de datos y en este caso con el objeto de disponer de suficiente tamaño muestral en los grupos de comparación, se han creado los siguientes grupos de comparación por países:

ESPAÑA: Grupos básicos sexo, edad, estudios		
Sexo	Varón	2456
	Mujer	2310
	Total	4766
Sector	Bancario	3590
	Telecomunicación	1176
	Total	4766
Edad	34 años o menos	1886
	35 años o más	2880
	Total	4766
Educación	Alto	2399
	Bajo	2367
	Total	4766

ESPAÑA: Grupos desagregados por sector, sexo, edad y estudios						
Sector	Edad	Sexo	Nivel de educación			
			Bajo	Alto		
Bancario	34 años o menos	Varón	380	402		
		Mujer	277	387		
	35 años o más	Varón	613	479		
		Mujer	525	527	Bancario	
	Total bancario			1795	1795	3590
	Telecomunicación	34 años o menos	Varón	117	109	
Mujer			87	127		
35 años o más		Varón	196	160		
		Mujer	204	176	Teleco	
Total telecomunicación			604	572	1176	
<b>Total España</b>					<b>4766</b>	

Tabla nº 5. Datos de España. Tamaños de los grupos de comparación por variables, tabla de la izquierda y por subgrupos homogéneos, tabla de la derecha.

La base de datos de España contiene suficientes efectivos para formar los subgrupos de comparación. Únicamente el colectivo de mujeres jóvenes, con nivel de educación bajo y que han evaluado el sector de las telecomunicaciones contiene 87 casos, el resto supera la centena de casos.

En el caso de la matriz de datos de EEUU encontramos una situación semejante con cuatro colectivos con menos del centenar de casos, los varones y mujeres que han evaluado la reputación de empresas de telecomunicación. El grupo mínimo contiene 69 casos.

EEUU: Grupos básicos sexo, edad, estudios		
<b>Sexo</b>	Varón	2612
	Mujer	2405
	Total	5017
<b>Sector</b>	Bancario	3754
	Telecomunicación	1263
	Total	5017
<b>Edad</b>	34 años o menos	2187
	35 años o más	2830
	Total	5017
<b>Educación</b>	Alto	1136
	Bajo	3881
	Total	5017

EEUU: Grupos desagregados por sector, sexo, edad y estudios				
Sector	Edad	Nivel de educación		
		Bajo	Alto	
<b>Bancario</b>	Varón	34 años o menos	189	678
		35 años o más	200	921
	Mujer	34 años o menos	197	581
		35 años o más	223	765
	Total bancario	809	2945	3754
<b>Telecomunicación</b>	Varón	34 años o menos	84	174
		35 años o más	97	269
	Mujer	34 años o menos	69	215
		35 años o más	77	278
	Total telecomunicación	327	936	1263
<b>Total EEUU</b>				5017

Tabla nº 6. Datos de EEUU. Tamaños de los grupos de comparación por variables, tabla de la izquierda y por subgrupos homogéneos, tabla de la derecha

La BBDD de China, contiene menos efectivos, aunque todos los subgrupos de comparación superan la centena.

CHINA: Grupos básicos sexo, edad, estudios		
<b>Sexo</b>	Varón	535
	Mujer	397
	Total	932
<b>Edad</b>	34 años o menos	643
	35 años o más	289
	Total	932

CHINA: Grupos desagregados por sector, sexo, edad y estudios			
Sector	Edad		
<b>Bancario</b>	Varón	34 años o menos	Bajo
		35 años o más	354
	Mujer	34 años o menos	181
		35 años o más	289
	Total bancario	932	108

Tabla nº 7. Datos de China. Tamaños de los grupos de comparación por variables, tabla de la izquierda y por subgrupos homogéneos, tabla de la derecha.

Según Dennis Jackson, el tamaño de muestra y la fiabilidad de las medidas son los factores que más influyen en detectar modelos mal especificados (Jackson 2007: 71). En concreto, cuando las cargas estandarizadas que enlazan los ítems con su latente son inferiores a 0,60, el tamaño de los grupos mínimos debería aproximarse a los 400 casos, sin embargo, cuando la fiabilidad es alta, por encima de 0,8, los tamaños de



muestra superiores a 200 casos podrían ser apropiados cuando la relación entre el número de casos y los parámetros a estimar es alta (Jackson 2007, 68).

En nuestro caso, el tamaño mínimo de muestra para la combinación de sexo, edad y nivel de estudios para el grupo de mujeres jóvenes con estudios bajos es de 266 efectivos. Este dato unido a la alta fiabilidad del índice Pulse<sup>29</sup> y a la alta relación entre casos y número de parámetros a estimar, dos grados de libertad, hace que los resultados sean concluyentes a la hora detectar modelos mal especificados.

## **2.7. CUESTIONARIO PARA OBTENER EL ÍNDICE PULSE.**

Las técnicas cuantitativas empleadas en nuestro caso se concretan en un diseño técnico de encuestas basado en una recolección de datos mediante la aplicación de un cuestionario estructurado con preguntas cerradas y escalas de respuesta numéricas de 7 posiciones y anclas semánticas extremas a una muestra de sujetos que es representativa de la población de referencia en cada uno de los tres países en los que se llevó a cabo la encuesta.

El cuestionario que se aplicó en todos los casos tiene un mismo origen en inglés. Se tradujo a los dos idiomas. Este cuestionario en su versión en español es el que figura en la página siguiente.

---

<sup>29</sup> Véase el apartado de la validez convergente. En todos los grupos analizados, el índice Pulse presenta un índice de fiabilidad superior al 0,80 propuesto por Dennis Jackson (2007).

## CUESTIONARIO CON LAS 4 PREGUNTAS PARA CALCULAR EL ÍNDICE PULSE:

**Pregunta Constructo PULSE** Ahora le voy a leer unas preguntas que describen la opinión o impresión que se puede tener hacia \_\_\_\_\_ [**INSERTAR EMPRESA 1, EMPRESA 2 Ó EMPRESA 3**].

Por favor, utilice una escala del "1" al "7" donde "1" significa "Completamente en desacuerdo" y "7" significa "Completamente de acuerdo". Puede utilizar cualquier valor intermedio entre 1 y 7 para matizar su respuesta. **ENTREVISTADOR: LEER.**

**ENTREVISTADOR: EVITAR MARCAR "NO SABE" Y "NO CONTESTA". SI LO MENCIONA, SUGERIR: "¿QUIERE DECIR QUE NI ESTÁ DE ACUERDO NI EN DESACUERDO?". SI ES ASÍ, MARCAR CODIGO 4].**

**PROGRAMADOR.- SI DE LAS 4 FRASES DICE NO SABE / NO CONTESTA EN 2 O MAS LA EMPRESA SERA NO VALIDA PARA ESTA ENTREVISTA. Y PODEMOS SELECCIONAR OTRA EMPRESA.**

[ROTAR FRASES:] ↓	Completamente en desacuerdo							Completamente de acuerdo	No sabe (NO LEER)	No contesta (NO LEER)
a. Tiene buena reputación.....	1	2	3	4	5	6	7	8	9	
b. Es una empresa sobre la que tengo buena impresión .....	1	2	3	4	5	6	7	8	9	
c. Es una empresa en la que confío .....	1	2	3	4	5	6	7	8	9	
d. Es una empresa que admiro y respeto .....	1	2	3	4	5	6	7	8	9	

### **SOBRE LAS TÉCNICAS DE ANÁLISIS**

Para alcanzar los objetivos y contrastar las hipótesis de esta investigación se han practicado sobre las 3 bases de datos un elevado número de pruebas matemáticas y estadísticas de considerable especialización y complejidad técnica con el fin de someter nuestra hipótesis a análisis exigentes y rigurosos procedentes de la metodología de ecuaciones estructurales<sup>30</sup> que las demuestren o las refuten.

En los epígrafes siguientes vamos a describir cada uno de los distintos tipos de análisis estadísticos que hemos empleado para demostrar o refutar las hipótesis de esta investigación, cuyos resultados se ofrecen en los capítulos siguientes.

El índice Pulse es una escala psicométrica multidimensional compuesta por cuatro ítems que pretende medir el constructo de reputación corporativa o concepto complejo transformado en medida que debería la presencia o ausencia de un conjunto de comportamientos cuando dicha medida sube o baja. En este sentido y debemos comprobar si la escala de Pulse cumple o no con dos tipos de cualidades o propiedades esenciales en toda escala psicométrica (De Vellis 1991):

(1) la “validez de contenido” .

(2) la “fiabilidad”.

La validez de contenido, de orden teórico, garantiza que el índice Pulse contiene todos los rasgos o dimensiones relevantes del fenómeno (reputación) que pretende medir o “dominio de contenido”.

- La fiabilidad mide el grado en que una medida es consistente en la persona, es decir, mientras tenga un mismo nivel de reputación, las puntuaciones tenderán a ser las mismas. La fiabilidad requiere verificar si se cumplen o no tres dos tipos de validación más concretas, a saber:

---

<sup>30</sup> El análisis factorial exploratorio se puede considerar un análisis tipo SEM en el que todos los factores comparten todos los indicadores.

- La “validez convergente”. Las conductas observables a través de los cuatro ítems de la escala convergen de forma fiable en el constructo que representan (la reputación).
- La “uni-dimensionalidad”. El grado en que los cuatro indicadores reflejan una sola dimensión, la reputación

Si consiguiéramos demostrar que Pulse responde a estos requisitos de validación y cumple con las propiedades psicométricas que se exigen a un índice psicológico multidimensional, entonces podríamos afirmar que las puntuaciones del índice Pulse son teóricamente adecuadas y producen valores fiables y válidos del fenómeno considerado que es una variable latente no observable directamente: la reputación.

Cuando investigamos con variables latentes como la reputación tenemos que seleccionar las técnicas matemáticas más adecuadas para validar o refutar de qué forma los cuatro ítems que componen el índice Pulse construyen esta variable latente. Para llevar a cabo esta comprobación debemos elegir entre dos tipos de enfoques teóricos que exponemos a continuación (enfoque reflexivo o enfoque formativo), justificando cuál de los dos hemos seleccionado para nuestra investigación.

Los ítems pueden relacionarse con la variable latente (factor no observado) de dos modos:

- El **modo reflexivo**. Los ítems son un reflejo de la variable latente no observada. En este caso, el rasgo latente sería el factor común que explica el comportamiento de los ítems que lo representan. Por tanto, la variación de cada ítem se descompone a la variación debida al rasgo latente y una variación residual o única que representa el “error” debido a que cada ítem en cuestión supone una aproximación imperfecta del rasgo latente.
- El **modo formativo**. Los ítems causan o “forman” la variable latente. El contenido o extensión empírica del rasgo latente es reflejado exactamente por la combinación de los ítems.

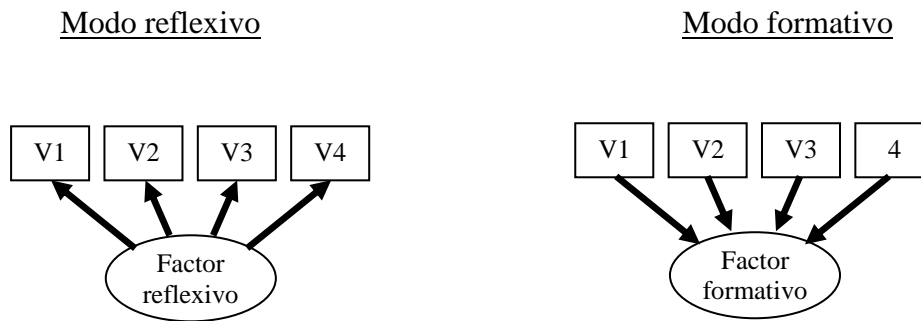


Figura nº 1. Modos reflexivo y formativo de relación con la variable latente.

En nuestro caso, hemos optado por el modo reflexivo. Este modo resulta más adecuado para aproximar constructos psicológicos, como en el caso de la reputación global como afirman varios autores (Carreras et al. 2013; Highhouse et al. 2009; Ponzi et al. 2011; Nessel y Helgesen 2009; Sung y Yang 2009; Walsh et al. 2009).

La técnica estadística adecuada para evaluar variables latentes (reputación) de modo reflexivo es el "análisis de invarianza multigrupo". Se trata de una técnica concreta procedente de las ecuaciones estructurales, destinada a comparar variables latentes medidas en modo reflexivo en colectivos distintos (tres países, y diferentes segmentos de sexo, edad y nivel educativo). Debido a que las variables latentes son constructos no observables si no es a través de los indicadores o ítems que los representan, antes de comparar las evaluaciones alcanzadas en dichos colectivos, se requiere comprobar la invarianza o grado en que estos grupos usan y entienden la métrica del mismo modo.

Neal Schmitt y Goran Kuljanin han expresado esta idea de forma sencilla: una escala viola la invarianza cuando dos individuos de poblaciones diferentes con idéntica puntuación en el constructo (reputación), puntúan de forma distinta los ítems (los 4 ítems del cuestionario Pulse) que lo aproximan (Schmitt y Kuljanin 2008, 211). Dicho de otro modo sólo se verifica esta hipótesis si ambos colectivos "usan e interpretan" de la misma forma los cuatro ítems que expresarían teóricamente la reputación.

La serie de análisis matemáticos y estadísticos que vamos a emplear para demostrar o refutar la hipótesis y sub-hipótesis de esta investigación serán los tres siguientes:

- Un **análisis factorial exploratorio** (EFA) para comprobar la unidimensionalidad de los índices Pulse estimados por sector y país (H1.1).
- Un **análisis factorial confirmatorio** (CFA) para testar la validez y naturaleza del Pulse (de la H1.2.1 a la H1.2.4).
- Un **análisis de varianza multi-grupo** para comprobar las hipótesis sobre el grado de invarianza poblacional por sexo y edad (de la H2.1 a la H4.4).

En los apartados siguientes exponemos en detalle cada una de las técnicas matemáticas y estadísticas y la forma en la que hay que interpretar los resultados de la aplicación de cada una de ellas.

### **3.1. ANÁLISIS FACTORIAL EXPLORATORIO (EFA).**

Esta técnica la hemos seleccionado para comprobar la uni-dimensionalidad del índice Pulse medido por sector y país (hipótesis: H1.1).

En la metodología del desarrollo de escalas, una de las técnicas más utilizadas para comprobar la cantidad de dimensiones que subyacen a los indicadores de una escala es el análisis factorial exploratorio (De Vellis, 1991; Noar, 2003; Sprott et al. 2009). En el caso concreto de las escalas de reputación esta técnica ha sido ampliamente utilizada para identificar las dimensiones subyacentes y sus ítems más relevantes. (Fombrun, 2000; Davies et al. 2004, Walsh y Beatty, 2007).

En realidad las técnicas factoriales constituyen una familia de herramientas estadísticas con propiedades y utilidades diferentes, que conviene diferenciar. En primer lugar distinguiremos entre las técnicas factoriales exploratorias y las confirmatorias (Carreras, 2009).

#### **3.1.2. DIFERENCIAS ENTRE FACTORIALES EXPLORATORIOS Y CONFIRMATORIOS.**

El análisis factorial fue desarrollado para organizar y comprender un grupo de variables relacionadas entre sí. Cuando se dispone de un número de variables observadas covariando entre sí, el análisis factorial permite derivar los factores causales subyacentes que explican la estructura de inter-correlaciones que se puede dar entre esas variables. Jae-On Kim y Charles Müeller proponen la siguiente definición: “El análisis factorial se refiere a una variedad de técnicas estadísticas cuyo objetivo común es representar un conjunto de variables en términos de un número menor de variables hipotéticas”<sup>31</sup> (Kim y Müeller, 1978, 9). Este método de análisis estadístico resulta adecuado para comprobar nuestra hipótesis de uni-dimensionalidad porque permite averiguar si los cuatro indicadores del Pulse están

---

<sup>31</sup> Traducción directa: “Factor analysis refers to a variety of statistical techniques whose common objective is to represent a set of variables in terms of a smaller number of hypothetical variables”.

covariando a partir de un solo factor, es decir que los cuatro ítems representan todos ellos la variable latente de reputación global.

Dentro de la familia factorial, conviene distinguir entre los análisis factoriales exploratorios (AFE) y los análisis factoriales confirmatorios (AFC). El análisis exploratorio sirve para identificar la cantidad de factores que subyacen a un conjunto de variables inter-correlacionadas, permitiendo que todos los factores expliquen todas las variables. Por el contrario, el análisis confirmatorio se utiliza para validar una estructura establecida, porque permite averiguar qué variables se relacionan con determinados factores y con cuales no se relacionan (Kline 1998, 88).

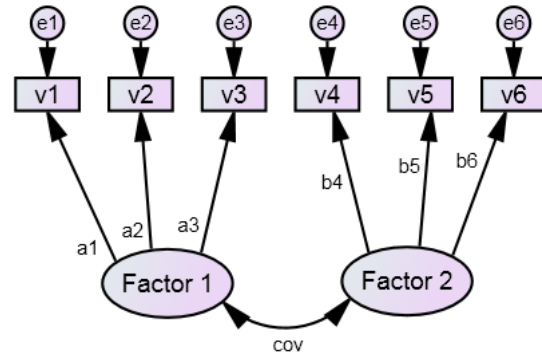
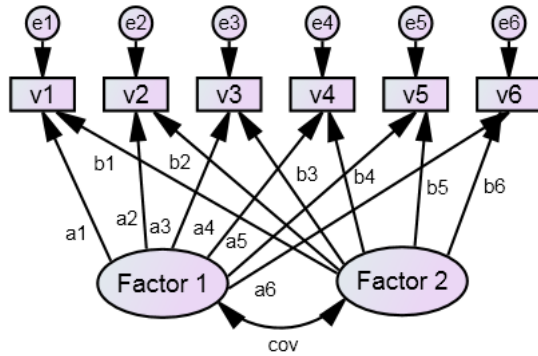
El diagrama siguiente representa las diferencias entre los AFE y los AFC. Supongamos que disponemos de 6 variables, de la “v1” a la “v6” covariando entre sí y que esa covariación puede ser reproducida con dos factores “F1” y “F2”.

El análisis exploratorio estimaría los factores F1 y F2 mediante combinaciones lineales con cargas en todas las variables, “ai” sobre el factor 1 y “bi” sobre el factor 2, de modo que se visualizaría la contribución de cada carga en cada factor, determinando el grado de información que aporta cada variable a los factores. En este sentido decimos que el análisis explora la estructura interna.



Representación diagramática  
diagramática

Representación



Representación matricial

Representación matricial

$$\begin{bmatrix} v1 \\ v2 \\ v3 \\ v4 \\ v5 \\ v6 \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \\ a_4 & b_4 \\ a_5 & b_5 \\ a_6 & b_6 \end{bmatrix} * \begin{bmatrix} F1 \\ F2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{bmatrix}$$

$$\begin{bmatrix} v1 \\ v2 \\ v3 \\ v4 \\ v5 \\ v6 \end{bmatrix} = \begin{bmatrix} a_1 & 0 \\ a_2 & 0 \\ a_3 & 0 \\ 0 & b_4 \\ 0 & b_5 \\ 0 & b_6 \end{bmatrix} * \begin{bmatrix} F1 \\ F2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{bmatrix}$$

Figura nº 2. Diagrama de relaciones entre los AFE y los AFC.

Por el contrario, en el análisis factorial confirmatorio, el analista impone restricciones en la configuración de los factores. En nuestro caso decimos que las contribuciones de las tres primeras variables no contribuyen en la formación del factor 1, este factor sólo es explicado por la covariación de las variables v4, v5 y v6. En forma similar, el factor 2 sería explicado por las tres últimas, sin la intervención del resto.

Fácilmente se comprende que el análisis factorial exploratorio se utiliza para averiguar las dimensiones que subyacentes en un conjunto de variables, mientras que el confirmatorio se utiliza para validar empíricamente una escala ya estructurada (Noar, 2003). El análisis factorial confirmatorio lo examinamos más adelante, ahora continuamos describiendo los rasgos de las técnicas factoriales exploratorias.

El método factorial exploratorio tiene por objeto reproducir la variación total de un conjunto de variables observables en otro conjunto de factores que son combinaciones lineales de las variables anteriores, pero estos factores han sido estimados extrayendo secuencialmente el máximo de varianza posible, de forma que un subconjunto más reducido en número de dichos factores sería capaz de reproducir el máximo de la variabilidad inicial.

El análisis trabaja en dos momentos: El método de extracción de los factores, que estima las combinaciones lineales y el método de rotación, que las transforma en forma interpretable.

### **3.1.3. COMPONENTES PRINCIPALES PARA TESTAR LA UNIDIMENSIONALIDAD.**

Presentamos la formación de los factores según el método por componentes principales (Dunteman, 1989). El análisis obtiene tantos factores como variables en la combinación, en el caso del índice Pulse, con 4 ítems obtendríamos 4 factores que reproducirían la variación total de los 4 ítems.

Los componentes son una combinación lineal ponderada por los pesos de las variables observadas:

$$\alpha_1' = (a_1, a_2, a_3, a_4) \quad y \quad V = \begin{bmatrix} V1 \\ V2 \\ V3 \\ V4 \end{bmatrix}$$

El producto es el componente:  $F1 = \alpha_1'V = a_1V1 + a_2V2 + a_3V3 + a_4V4$

Como tenemos cuatro variables observadas, el análisis por componentes principales extraerá cuatro componentes, al pre-multiplicar la matriz traspuesta de cargas  $A'$  por el vector de variables  $V$ :

Siendo  $\alpha'$  el vector traspuesto de cargas y  $V$  el vector de variables, su producto origina el factor.

$$A' = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}; V = \begin{bmatrix} V1 \\ V2 \\ V3 \\ V4 \end{bmatrix} \text{ producto es el vector "F" de componentes}$$

$$A'V = \begin{bmatrix} a_{11}V1 + a_{12}V2 + a_{13}V3 + a_{14}V4 \rightarrow \text{combinación del F1} \\ a_{21}V1 + a_{22}V2 + a_{23}V3 + a_{24}V4 \rightarrow \text{combinación del F2} \\ a_{31}V1 + a_{32}V2 + a_{33}V3 + a_{34}V4 \rightarrow \text{combinación del F3} \\ a_{41}V1 + a_{42}V2 + a_{43}V3 + a_{44}V4 \rightarrow \text{combinación del F4} \end{bmatrix}$$

El análisis de componentes trata de estimar los valores de la matriz de pesos o auto-vectores A con tres condiciones:

1. Reproduce la variación total original de las variables observables<sup>32</sup>.

La ecuación fundamental del análisis factorial refleja la reproducción total de la varianza de la combinación lineal de las variables observadas “Σ” en función de la matriz de auto-vectores “A” y una matriz auxiliar de auto-valores “Λ”.

$$\Sigma = A\Lambda A'$$

Siendo:

- “Σ” la variación total de las variables observadas o matriz de correlaciones,
- “A”, la matriz pesos, también llamada matriz de auto-vectores”
- “Λ” la matriz de auto-valores.

Mediante el método de Joseph Louis Lagrange que es de obligada y rutinaria utilización cuando se realizan análisis factoriales (Benjumea Acevedo, et al.,2006, 237-240), se crea la matriz adicional de auto-valores que pre y post-multiplicada por la matriz de auto-vectores, reproduce la covariación total “Σ”.

---

<sup>32</sup> La diferencia entre el método de componentes principales y el resto de técnicas factoriales como la de ejes principales, factorización alfa, etc., radica en que el primero tiende a reproducir la varianza total, mientras que en los otros métodos sólo la varianza compartida. En el primer caso, los valores de la diagonal son la unidad y en el resto, estimaciones de la covarianza)

Las matrices de auto-vectores y auto-valores se estiman, en opinión consensuada de los expertos en análisis factorial a partir de la matriz “ $\Sigma$ ”, mediante el algoritmo de Jacobi (Marradi, A. (1981, 110-112).

La estimación de componentes se realiza extrayendo secuencialmente el máximo de varianza hasta alcanzar la varianza total, de forma que los primeros lleguen a captar prácticamente la varianza total, consiguiendo así la reducción del número de factores.

Aplicando el método de Lagrange se estima el vector de cargas del primer componente<sup>33</sup> “F1”, reteniendo el máximo de varianza total<sup>34</sup>:

$$\Sigma\alpha_1 = \lambda_1\alpha_1$$

El valor de “ $\lambda_1$ ” es el que maximiza la varianza de la combinación lineal del 1º componente “ $\alpha_1'X$ ”.

Como el segundo componente debe estar incorrelacionado con el primero, la covarianza entre ambos debe ser cero “ $\alpha_1'\Sigma\alpha_2 = 0$ ”. Se vuelve a aplicar el método de Lagrange introduciendo esta restricción, obteniéndose el segundo auto-valor “ $\lambda_2$ ”, con su respectivo auto-vector de cargas “ $\alpha_2$ ” que vuelve a captar el máximo de la varianza remanente. Así sucesivamente hasta estimar todos los componentes.

2. Los componentes extraídos están incorrelacionados, es decir, estima componentes independientes entre sí.

El ejemplo ficticio siguiente expresa el procedimiento de componentes principales. Supongamos que disponemos de la siguiente matriz de correlaciones observadas “ $\Sigma$ ”, de los 4 indicadores del Pulse obtenidos en una encuesta (Carreras 2009):

---

<sup>33</sup> Con la restricción de que la suma del cuadrado de las cargas ( $\alpha_1'\alpha_1$ ) sea igual a la unidad:  $\Sigma\alpha_1^2=1$

<sup>34</sup> El multiplicador de Lagrange es un cálculo que se utiliza para maximizar o minimizar funciones diferenciables. Se utiliza para encontrar los puntos extremos de la función donde la pendientes es cero (la segunda derivada indicará que es un máximo).

**Matriz observada "S"**

	V1	V2	V3	V4
V1	1	0,779	0,94	0,752
V2	0,779	1	0,805	0,741
V3	0,94	0,805	1	0,787
V4	0,752	0,741	0,787	1

**Matriz de auto-valores "Λ"**

λ1	λ2	λ3	λ4
3,405	0	0	0
0	0,287	0	0
0	0	0,25	0
0	0	0	0,058

**Matriz de auto-vectores "A"**

	F1	F2	F3	F4
V1	0,511	-0,483	-0,268	0,658
V2	0,487	0,097	0,866	0,046
V3	0,52	-0,351	-0,214	-0,749
V4	0,48	0,796	-0,363	0,063

Tabla nº 8. Componentes extraídos.

Demostración:  $S = AA\Lambda'$ . Reproduce la matriz de correlaciones original "S"<sup>35</sup>.

A partir de la matriz de auto-vectores, se obtendría la matriz de cargas factoriales mediante la transformación: Cargas factoriales =  $\alpha_{ij} * \lambda_j^{1/2}$ .

Sabemos que la varianza total de los componentes extraídos es igual a la suma de sus varianzas. Recordemos que los componentes están incorrelacionados y que por tanto, sus covarianzas son cero (Dunteman 1989).

---

<sup>35</sup> Operaciones con matrices:

**Premultiplicamos A por Λ → AΛ**

1,741	-0,14	-0,07	0,038
1,66	0,028	0,217	0,003
1,771	-0,1	-0,053	-0,04
1,635	0,229	-0,09	0,004

**Posmultiplicamos AΛ por A' → AΛA' = Σ**

	V1	V2	V3	V4
V1	1	0,779	0,94	0,752
V2	0,779	1	0,805	0,741
V3	0,94	0,805	1,000	0,787
V4	0,752	0,741	0,787	1,000

Además los componentes son variables estandarizadas, con lo que sus varianzas igualan la unidad. Tenemos en nuestro ejemplo hipotético 4 componentes, por tanto la varianza total será igual a 4.

Los auto-valores detectan la varianza extractada por componente y por tanto su suma iguala la varianza total. Se verifica que la suma de auto-valores iguala la variación total.

Componente	Auto valores	% varianza total
1	3,405	0,851
2	0,287	0,072
3	0,25	0,063
4	0,058	0,014
	Suma	Suma
	4	1

Tabla nº 9. Cantidad de varianza de cada componente.

En nuestro ejemplo, la primera dimensión extraída ha conseguido reproducir 85,1% total de la varianza total, prácticamente reproduce la matriz de las correlaciones originales. En consecuencia, podemos concluir que la variación de las variables es prácticamente reproducida por una sola dimensión.

Aplicando las propiedades anteriores es factible determinar el grado en que un solo componente llega a reproducir las correlaciones originales.

Volviendo a nuestro ejemplo, las matrices de auto-valores y auto-vectores generadas por un solo componente son las que enunciamos en la tabla nº 8 de la página siguiente.

La suma de los residuales cuadráticos arroja una diferencia mínima<sup>36</sup> de 0,148, frente a un error cuadrático total bajo la hipótesis de incorrelación<sup>37</sup> del 11,742, significa que

---

<sup>36</sup> Suma de residuales cuadráticos

<sup>37</sup> Bajo hipótesis de incorrelación tendríamos un error total igual a la varianza encontrada en la muestra, la suma de los coeficientes de determinación.

con la extracción de un factor, el error cuadrático total sería del 1,3%. En este caso podríamos hablar con propiedad de que existe una sola dimensión subyacente.

$$ECT = \frac{0,148}{11,742} = 0,013$$

**Suma de residuales cuadráticos**

	V1	V2	V3	V4	Suma
V1	0,012	0,005	0,001	0,007	0,025
V2	0,005	0,036	0,003	0,003	0,048
V3	0,001	0,003	0,006	0,004	0,015
V4	0,007	0,003	0,004	0,046	0,060
					0,148

**Matriz de auto-valores "Λ" para 1 componente**

	λ1	λ2	λ3	λ4
λ1	3,405	0	0	0
λ2	0	0	0	0
λ3	0	0	0	0
λ4	0	0	0	0

**Matriz de auto-vectores "A" para un componente**

	F1	F2	F3	F4
V1	0,511	0	0	0
V2	0,487	0	0	0
V3	0,52	0	0	0
V4	0,48	0	0	0

**Matriz de correlaciones implicada por un componente**

	V1	V2	V3	V4
V1	0,89	0,849	0,905	0,836
V2	0,849	0,809	0,863	0,797
V3	0,905	0,863	0,921	0,85
V4	0,836	0,797	0,85	0,785

**Matriz de correlación residual: S - S<sup>1º</sup> Dim.**

	V1	V2	V3	V4
V1	0,11	-0,07	0,035	-0,084
V2	-0,07	0,191	-0,058	-0,056
V3	0,035	-0,058	0,079	-0,064
V4	-0,084	-0,056	-0,064	0,215

Tabla nº 10. Matriz de correlaciones.

### 3.2. LOS MÉTODOS DE ECUACIONES ESTRUCTURALES (SEM).

Los métodos de ecuaciones estructurales<sup>38</sup> representan un conjunto de técnicas diseñadas específicamente para trabajar con constructos psicosociales posicionados en modelos causales, también denominados “redes nomológicas” que permiten explicar fenómenos en los que dichos constructos están implicados (Bagozzi y Yi, 2012).

Esta tecnología ha sido empleada en numerosas investigaciones de marketing destinadas a explicar fenómenos como la satisfacción del cliente (Carreras y González, 2013; Fornell et al., 1996; Johnson et al., 2001; Vogel et al., 2008), la lealtad del cliente (Cronin et al., 2000; Harris y Goode, 2004) o la reputación corporativa (Bartikowski y Walsh, 2011; Carreras et al., 2012; Money et al., 2010; Nettet y Helgesen, 2009).

En nuestro caso dos hipótesis centrales de la investigación, la comprobación o refutación de la cualidad psicométrica del índice Pulse y la invarianza poblacional de Pulse, no podrían ser debidamente abordadas si no aplicásemos esta metodología debido a la naturaleza latente de nuestro índice.

De la familia de técnicas que componen los modelos SEM, sólo dos son pertinentes para resolver las hipótesis propuestas, el análisis factorial confirmatorio y el análisis de invarianza multi-grupo.

El análisis factorial confirmatorio (CFA) se emplea para testar las propiedades psicométricas de una escala como nuestro índice Pulse (Bagozzi y Yi 2012, 14). Mediante esta técnica se puede volver a comprobar la uni-dimensionalidad<sup>39</sup> del índice Pulse, además de la validez de constructo. Una vez comprobada la validez métrica del índice Pulse, se puede pasar a demostrar su invarianza poblacional mediante la aplicación de otra técnica SEM, el análisis multi-grupo de invarianza para testar las diferencias en las pautas factoriales (Cheung y Rensvold, 2002; Millsap, 2011).

---

<sup>38</sup>Se les conoce con el acrónimo inglés SEM, “Structural Equation Models”.

<sup>39</sup> La uni-dimensionalidad ha sido establecida mediante el análisis factorial exploratorio, el confirmatorio se emplea para revalidarla (Noar 2003).



En las páginas siguientes presentamos una breve introducción a los modelos SEM para profundizar en las dos técnicas que emplearemos en nuestra investigación: el análisis factorial confirmatorio y el análisis multi-grupo.

Los modelos SEM son sistemas de ecuaciones simultáneas que permiten verificar las relaciones funcionales complejas entre variables latentes formuladas en un modelo causal. Se trata de una tecnología de investigación confirmatoria especialmente diseñada para desarrollar teorías con cierta complejidad en las que los constructos psicológicos implicados pueden estar presentes en varios planos de relaciones causales, como antecedentes, procesos de mediación y variables latentes finales.

Además representa una tecnología versátil porque permite someter a prueba todos los tipos posibles de relaciones funcionales entre variables latentes (Carreras, 2006), a saber relaciones directas (influencia inmediata de una variable sobre otra), indirectas (influencia a través de mediadoras), espurias (detecta relaciones aparentes desvanecidas ante variables anteriores), de moderación (cuando la influencia entre dos variables es distinta ante los niveles de un inter-actor), recíprocas (mutuamente impactadas) y relaciones no analizadas (covarianzas).

Básicamente existen dos grandes familias metodológicas para estimar los llamados modelos causales, las técnicas de reducción de varianza y las técnicas de covarianzas (Bollen, 1989; Carreras, 2006; Anderson y Gerbin, 1988). Las primeras se basan en procedimientos funcionales que minimizan los errores de estimación de los valores individuales, mientras que las segundas no se basan en los casos sino en las covarianzas. Los autores parecen coincidir en que las primeras resultan más apropiadas para resolver problemas de pronóstico, mientras que las segundas parecen más apropiadas para el desarrollo de investigación teórica (O'loughlin y Coenders, 2002). Dado que nuestra investigación es más bien de naturaleza teórica, parece más apropiado utilizar los métodos basados en las estructuras de covarianzas.

Los métodos de ecuaciones estructurales basados en covarianzas descansan en una lógica relativamente sencilla. Dado que cualquier modelo de impactos puede ser algebraicamente transformado en la matriz de varianzas y covarianzas observadas, se calcula la matriz de covarianzas implicada " $\Sigma(\theta)$ " por el modelo y si coincide con la

matriz de covarianzas observada “ $\Sigma$ ”, se concluye que dicho modelo de impactos es conforme a los datos, en caso contrario, la realidad ha rechazado el modelo que se ha planteado como hipótesis a demostrar (Hooper, et al. 2008).

Ahora bien, los datos proceden de una muestra probabilística y los desajustes entre las matrices observada y la implicada podrían proceder de errores de muestreo. Es necesario despejar esta incertidumbre mediante el contraste de hipótesis fundamental de estos métodos:

$$H_0: \Sigma = \Sigma(\theta)$$

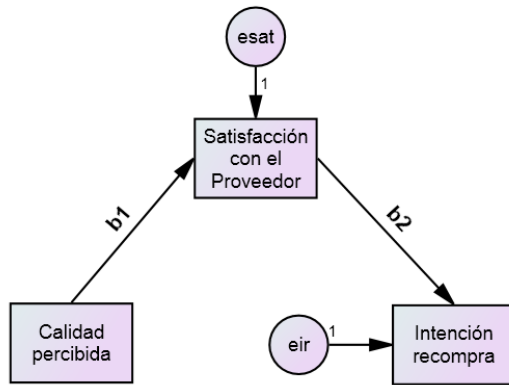
$$H_1: \Sigma \neq \Sigma(\theta)$$

En realidad, estos métodos trabajan en cuatro fases (Bollen, 1989):

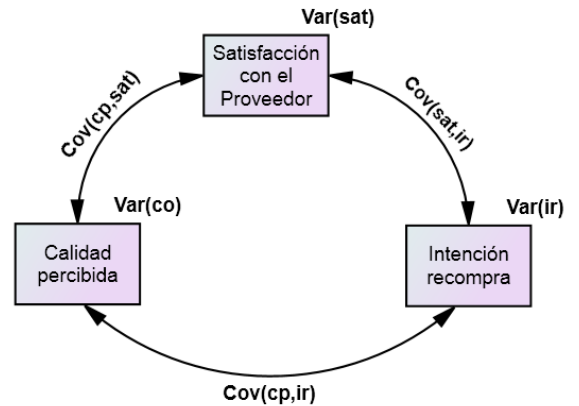
- Propuesta del modelo de impactos. Se dibuja un modelo explicativo de relación entre los constructos.
- Derivación de la matriz implicada por el modelo. Mediante un algoritmo de transformación, se determinan las ecuaciones que darán lugar a la matriz de varianzas y covarianzas implicada por el modelo propuesto. Verificación del problema de la identificación de los parámetros
- Estimación de los parámetros el modelo.
- Validación empírica del modelo propuesto en función de la valoración del grado de ajuste entre ambas matrices, observada e implicada, para tomar la decisión de rechazar o aceptar la  $H_0$ .

Presentamos el esquema con un sencillo ejemplo. Sean tres variables observadas “calidad global” (CG), “satisfacción con el proveedor” (SAT) e “intención de recompra” (IR). Proponemos el siguiente modelo de impactos, la calidad percibida promueve la recompra a través de la satisfacción, pero si esta no se activa, la intención de recompra no se llega a formar, es decir la influencia de la calidad sobre la intención se ejerce, en su totalidad, a través de la satisfacción, por tanto el *path* (el camino) que vincula calidad con intención es cero y no se representa en el modelo. En este caso decimos que el proceso está mediado completamente por la satisfacción.

Modelo de impactos propuesto  
(diagrama)



Modelo de relaciones observadas  
(diagrama)



Ecuaciones:

$$Sat = b1 * CP + e_{sat}$$

$$IR = b2 * Sat + e_{ir}$$

Figura nº 3. Modelo de impactos.

**CÁLCULO DE LA MATRIZ DE COVARIANZAS IMPLICADA (IDENTIFICACIÓN).**

Aplicando las propiedades de los momentos distribucionales, se determina la matriz de covarianzas implicadas.

La matriz de covarianzas observadas:

$$\Sigma = \begin{vmatrix} S_{CP}^2 & S_{CP,SAT} & S_{CP,IR} \\ S_{SAT,CP} & S_{SAT}^2 & S_{SAT,IR} \\ S_{IR,CP} & S_{IR,SAT} & S_{IR}^2 \end{vmatrix}$$

Son las observaciones del modelo, 3 varianzas y 6 covarianzas observables

La matriz de covarianzas implicadas por el modelo de impactos propuesto que reproduce la matriz original.

$$\Sigma(\theta) = \begin{vmatrix} S_{cp}^2 & b_1 Var(cp) & b_1 b_2 Var(cp) \\ b_1 Var(cp) & b_1^2 Var(cp) + Var(e_2) & b_1^2 b_2 Var(cp) + b_2 Var(e_2) \\ b_1 b_2 Var(cp) & b_1^2 b_2 Var(cp) + b_2 Var(e_2) & b_1^2 b_2^2 Var(cp) + b_2^2 Var(e_2) + Var(e_3) \end{vmatrix}$$

El vector de parámetros a estimar “ $\theta$ ” = {b1, b2, e2, e3} sería:

- “b<sub>1</sub>” La pendiente que relaciona la calidad percibida con la satisfacción
- “b<sub>2</sub>” La pendiente que relaciona la satisfacción con la intención de recompra
- “e<sub>2</sub>” El error de satisfacción o variabilidad de satisfacción no explicada por “cp”
- “e<sub>3</sub>” El error de intención o varianza de intención no explicada por satisfacción

Por tanto, tenemos 6 ecuaciones en función de los parámetros del modelo que llegan a reproducir la matriz de covarianzas observadas:

- $S_{cp}^2 = S_{cp}^2$
- $S_{SAT}^2 = b_1^2 Var(cp) + Var(e_2)$
- $S_{IR}^2 = b_1^2 b_2^2 Var(cp) + b_2^2 Var(e_2) + Var(e_3)$
- $S_{CP,SAT} = b_1 Var(cp)$
- $S_{CP,IR} = b_1 b_2 Var(cp)$
- $S_{SAT,IR} = b_1^2 b_2 Var(cp) + b_2 Var(e_2)$

Antes de proceder con la estimación, debemos comprobar que el modelo está debidamente identificado. La identificación tiene que ver con la cantidad de parámetros conocidos y desconocidos en los sistemas de ecuación. Los parámetros conocidos son características de las distribuciones de población, normalmente varianzas y covarianzas de las variables observadas. Mientras que los parámetros desconocidos son los valores que se desean estimar.

Como disponemos de 6 puntos momentos muestrales observados (3 varianzas y 3 covarianzas) y sólo 4 parámetros “ $\theta$ ” a estimar, el modelo está sobre-identificado, lo

que indica un exceso de información para identificación de los parámetros, con lo que no tendríamos una única solución, sino una serie de soluciones posibles que satisfarían las ecuaciones.

En los modelos SEM, la sobre-identificación es una propiedad deseable (Bollen 1989; Kline 1998). De todas las soluciones posibles seleccionamos la que minimiza la matriz residual: “ $\Sigma - \Sigma(\theta)$ ”.

### 3.2.1. ESTIMACIÓN DE LOS VALORES DE LOS PARÁMETROS.

Hay varios métodos de estimación de los parámetros para minimizar la matriz residual. Por tanto, todos los métodos se basan proponen una función con la misma lógica, minimización de la matriz residual:

$$F(S, \hat{\Sigma}(\theta))$$

Para explicitar las fórmulas y cálculos de las funciones de ajuste, remitimos al lector al texto de Kenneth Bollen ampliamente citado en esta investigación. Tan sólo recordaremos que los valores de los parámetros se estiman mediante un conjunto de derivadas parciales de la función de ajuste, respecto de cada valor a estimar, igualado a cero y cuya segunda derivada es positiva, indicando un valor mínimo para la función de ajuste, que es lo que estábamos buscando:

$$\frac{\partial F_{ML}(\theta)}{\partial \theta_i} = 0$$

Como ya hemos comentado, normalmente se producirán varias soluciones de vectores “ $\theta$ ” y mediante procedimientos iterativos<sup>40</sup> se construye la secuencia de vectores posibles  $\theta_1, \theta_2, \dots, \theta_n$  que van minimizando sucesivamente la función de ajuste hasta alcanzar un mínimo preestablecido que representa la mejor solución posible.

---

<sup>40</sup> Kenneth Bollen cita el algoritmo de Newton-Raphson

Los diferentes métodos de estimación difieren según la forma en que consideran la discrepancia. Las funciones de minimización o ajuste están basadas en la matrices de covariación observadas “S”, que son una realización muestral de su correspondiente poblacional “ $\Sigma$ ”. Por el contrario, “ $\hat{\Sigma}(\theta)$ ” es la matriz de covariación estimada mediante el algebra de transformación y cuyos valores han sido estimados conjuntamente por la función que minimiza la matriz residual.

Las funciones de minimización más conocidas son la “ML” (Máxima Verosimilitud <sup>41</sup>), la GLS (mínimos cuadrados generalizados) y la ADF (Distribución libre asintótica) (Carreras, 2010, 138 y ss.).

El método de la máxima verosimilitud estima los parámetros verdaderos, que ha inducido, con mayor probabilidad, las covarianzas que hemos observado en la muestra. Las propiedades de los estimadores obtenidos por máxima verosimilitud son asintóticos, es decir, con muestras superiores a 200 casos son insesgados, el error mínimo cuadrático coincide con la varianza del estimador, consistentes, aumentando la muestra las estimaciones convergen en el valor del parámetro, eficientes, con menor varianza de error y normalidad multivariante.

Los mínimos cuadrados generalizados (GLS) basan sus estimaciones en la minimización de los residuos en las observaciones. Las propiedades asintóticas de los estimadores son, consistencia y normalidad multivariada, pero no suelen ser más eficientes que los estimados por ML.

Ambos métodos de estimación GLS y ML suponen normalidad multivariante y variables continuas.

Michael Browne diseñó un método de estimación de distribución libre asintótica, “ADF”<sup>42</sup> que tiene en cuenta los momentos de cuarto orden (kurtosis) y no sólo los de segundo orden (varianzas y covarianzas) como sus predecesores. El problema es que el algoritmo de estimación supone una gran limitación práctica porque el cálculo de la

---

<sup>41</sup> Procede del acrónimo en inglés “Maximum Likelihood” ó ML

<sup>42</sup> Acrónimo del ingles “Asymptotically Distribution Free” Estimator

matriz de momentos de cuarto orden requiere grandes tamaños de muestra para que las estimaciones sean estables, en ocasiones más de 5000 casos (Browne, 1984).

En nuestra investigación utilizaremos la función por máxima verosimilitud como método de estimación porque es “la función de ajuste más ampliamente utilizada en los modelos de ecuaciones estructurales”<sup>43</sup> (Bollen, 1989,107) pero sobre todo porque ha demostrado ser robusta ante alteraciones no severas de normalidad (West et al., 1995; Curran et al., 1996).

### 3.2.2. CORROBORACIÓN DEL MODELO: LOS ÍNDICES DE AJUSTE.

Una vez se ha alcanzada la mejor solución en los valores de los parámetros, el investigador debe decidir si el modelo se corresponde con la realidad o no. Para ello tienen que resolver la hipótesis fundamental de los modelos causales, si la matriz de covarianzas implicada por el modelo ha conseguido reproducir o no la matriz observada poblacional,<sup>44</sup> dicho en otros términos, el modelo propuesto ha conseguido explicar toda la varianza observada, es decir no es rechazado por los datos. Recordemos que el contraste ha quedado formulado en los siguientes términos:

$$H_0: \Sigma - \Sigma(\theta) = 0$$

$$H_1: \Sigma - \Sigma(\theta) \neq 0$$

El test requiere de un estadístico de contraste que decida si ambas matrices son iguales, con un nivel de significación pre-determinado.

Sabemos que la función del ajuste por máxima verosimilitud “F<sub>ML</sub>” y mínimos cuadrados generalizados “F<sub>GLS</sub>” siguen una distribución de muestreo chi-cuadrado

---

<sup>43</sup> Traducción directa: “... the most widely used fitting function for general structural equation models is the maximum Likelihood function”.

<sup>44</sup> Las diferencias observadas son debidas a un mero error de muestreo. Lo que resuelve la hipótesis de contraste es que la diferencia entre las matrices poblacionales, implicada y observada, es cero.

(Hancock, 2006), con lo que es posible tomar las decisiones en función del valor del estadístico:

$$(n - 1)F_{ML} \approx \chi^2$$

El estadístico “ $(n - 1)F_{ML}$ ” representa el valor empírico en las abscisas de la distribución chi-cuadrado, según los grados de libertad,<sup>45</sup> y el área comprendida entre el dicho valor y más infinito, representa la probabilidad de error al rechazar la  $H_0$  de igualdad de matrices, siendo verdadera. Cuando la probabilidad de cometer error al rechazar la  $H_0$  es grande, superior a 0,05, se toma la decisión de aceptar la hipótesis nula y concluimos que el modelo sí ha conseguido reproducir la variación original y por tanto es conforme a datos, se verificaría por tanto la hipótesis que se ha puesto a prueba.

Los programas que estiman estos modelos (Lisrel, Amos, EQS) suelen aportar el “p-valor” o valor exacto de la probabilidad de cometer error tipo I<sup>46</sup>. Cuando el “p-valor < de 0,05” se concluye que es significativo y la hipótesis nula quedaría rechazada y aceptada la alternativa:  $H_1: \Sigma - \Sigma(\theta) \neq 0.$ , lo que invalidaría el modelo.

Conviene resaltar que en los análisis de ecuaciones estructurales, la interpretación del estadístico de contraste va en sentido contrario al de la investigación tradicional. Cuando es significativo, debemos rechazar el modelo, porque las matrices no han convergido. Por eso el investigador siempre busca aceptar la hipótesis nula.

A pesar de la elegancia metodológica, el test de contraste de la  $\chi^2$  ha suscitado bastante polémica, hasta el punto de resultar poco relevante a la hora de tomar decisiones en la buena parte de las investigaciones (Bollen, 1989; Kline, 1998; Hu y Bentler, 1999).

---

<sup>45</sup> Los grados de libertad equivalen a la diferencia entre el número de estadísticos observados, menos el número de parámetros a estimar.

<sup>46</sup> Probabilidad de equivocarse al rechazar la  $H_0$  y concluir que el modelo no ha conseguido reproducir la matriz observada, cuando en realidad si lo ha hecho.



El problema radica en que el estadístico de contraste “ $F^*(n-1) \approx \chi^2$ ” es altamente sensible al tamaño de muestra y cuando los casos superan los 200, con mucha probabilidad tenderá a mostrarse significativo. De esta forma, discrepancias mínimas o triviales entre las matrices, llevarán al investigador a concluir que el modelo no es conforme a la realidad. De forma inversa, ante tamaños pequeños de muestra, se tiende a concluir que las matrices convergen, cuando en realidad presentan discrepancias poco aceptables. Por esa razón recientemente Bagozzi y Yi han recordado que “debido a que el  $\chi^2$  es sensible al tamaño de la muestra se vuelve difícil conseguir ajustes de modelos satisfactorios, conforme el tamaño de muestra aumenta”<sup>47</sup> (Bagozzi y Yi, 2012, 28).

Para superar este obstáculo, se han diseñado una serie de índices de ajuste global que ayudan en la decisión de aceptar o rechazar la hipótesis nula. Obviamente, estos índices de ajuste están basados en el valor del “ $\chi^2$ ”, pero lo corrigen en función de la parsimonia del modelo <sup>48</sup>, los grados de libertad o la cantidad de discrepancia explicada respecto de un modelo nulo, bajo la hipótesis de variables latentes incorrelacionadas (modelo sin explicación posible).

Siguiendo las recomendaciones de Richard Bagozzi y Yujae Yi, a la hora de aceptar o rechazar un modelo, el investigador debería reparar en cuatro índices: “RMSEA”, “CFI”, “NNFI” y “SRMR” (Bagozzi y Yi, 2012, 28):

- El “RMSEA” o raíz del error promedio cuadrático de aproximación <sup>49</sup>.

$$RMSEA = \sqrt{\frac{(\chi^2 - gl)}{gl(n - 1)}}$$

---

<sup>47</sup> Traducción directa: “Because the  $\chi^2$  is sensitive to sample size, it becomes difficult to achieve satisfactory models fits as the sample size increases”.

<sup>48</sup> La parsimonia alude a modelos los más estrictos posible, muy pocos parámetros a estimar, respecto de los puntos de observación, cuantos más grados de libertad, más parsimonia y más sencillo el modelo.

<sup>49</sup> Acrónimo de “Root Mean Square Error of Approximation” de James Steiger (2007)

El término “ $\chi^2 - gl$ ” indica la cantidad de discrepancia encontrada entre ambas matrices (parámetro de no centralidad), respecto de la esperada<sup>50</sup>, en términos de la complejidad y la cantidad de casos. Cuando RMSEA es cero, el ajuste es completo y se indica con el valor cero.

El índice RMSEA mide el grado desajuste entre las matrices implicada y observada. Cuanto menor su valor, mejor es el ajuste.

- El “CFI” es el índice de ajuste comparado<sup>51</sup>.

$$CFI = 1 - \frac{(\chi^2 - gl)_{propuesto}}{(\chi^2 - gl)_{nulo}}$$

Compara la no centralidad del modelo encontrado respecto de la no centralidad del modelo nulo, cuando todas las latentes están incorrelacionadas (Hooper et al., 2008, 55).

- El índice CFI indica la bondad del ajuste, variando entre cero y 1 (se trunca para que adopte el valor 1), cuanto más cercano a uno, mejor ajuste entre las matrices.
- El “NNFI” es el índice de ajuste no normado <sup>52</sup>.

$$NNFI = \frac{\frac{\chi_{nulo}^2}{gl_{nulo}} - \frac{\chi_{propuesto}^2}{gl_{propuesto}}}{\frac{\chi_{nulo}^2}{gl_{nulo}} - 1}$$

El índice compara la discrepancia del modelo actual respecto de la que podría alcanzar en el modelo nulo, respecto de la cantidad de discrepancia posible del modelo<sup>53</sup>. Ambos modelos están relativizados respecto de la parsimonia o grados de libertad.

---

<sup>50</sup> Recordar que la esperanza de la distribución chi-cuadrado son los grados de libertad:  $E(\chi^2) = gl$ . Cuando  $\chi^2 < gl$ , indicaría que la discrepancia es menor de la esperada y RMSEA se establece a cero.

<sup>51</sup> Acrónimo de “Comparative Fit Index” índice construido por Peter Bentler en 1990.

<sup>52</sup> Acrónimo de “Non-Normed Fit Index”, índice propuesto por Tucker y Lewis en 1973.

Se trata de un test de bondad de ajuste, cuanto mayor sea el valor, mejor es el ajuste. Truncado a la unidad.

- El “SRMR”, es el residual raíz medio-cuadrático estandarizado<sup>54</sup>

$$\text{SRMR} = \sqrt{\frac{2 \sum_{i=1}^p \sum_{j=1}^i \left[ \frac{s_{ij} - \hat{\sigma}_{ij}}{s_{ii}s_{jj}} \right]^2}{p(p+1)}}$$

El índice evalúa la suma de las covarianzas residuales, implicadas menos observadas, estandarizadas respecto del total de las “p” variables observadas del modelo.

Es un índice de desajuste, con el valor cero indicando la ausencia de residuales entre las matrices implicada y observada.

Richard Bagozzi y Yujae Yi (2012, 28) proponen los siguientes umbrales para dar por aceptable un modelo, para aceptar la hipótesis nula fundamental ( $H_0: \Sigma - \Sigma(\theta) = 0$ ):

- RMSEA  $\leq 0.06$
- NNFI  $\geq 0.95$
- CFI  $\geq 0.95$
- SRMR  $\leq 0.08$

Aunque en modelos excesivamente complejos, los umbrales podrían relajarse en los siguientes valores:

- RMSEA  $\leq 0.07$
- NNFI  $\geq 0.92$
- CFI  $\geq 0.93$
- SRMR  $\leq 0.10$

---

<sup>53</sup> La unidad en el denominador significa que es la discrepancia bajo ajuste perfecto, cuando se cumple:  $\chi^2 = E(\chi^2) = gl$

<sup>54</sup> Acrónimo de “standardised root mean square residual” (Kline 1998)

Según Bagozzi y Yi, los cuatro índices podrían minimizar ambos tipos de error tipo I y tipo II. “*Aunque ningún índice en solitario cumple con ambos criterios, el conjunto de los cuatro presentados más arriba, colectivamente proporciona criterios satisfactorios para la evaluación de un modelo global*”<sup>55</sup>(Bagozzi y Yi 2012, 29).

La evaluación de los índices globales debe completarse el análisis de la significación individual de los parámetros estimados. Los autores han propuesto dos procedimientos básicos, los test de contraste individual, los llamados “CR” o “ratios cocientes”<sup>56</sup> y las pruebas más rigurosas o test de anidamiento (Bollen, 1989; Kline, 1998). Como después veremos, las pruebas del anidamiento resultan esenciales para el análisis multi-grupo.

- Los estadísticos de contraste individual.

Según Kenneth Bollen, la estimación por Máxima verosimilitud, permite estimar los errores estándar asintóticos de los parámetros estimados a partir de la inversa de la matriz de información, de forma que el valor estimado de un parámetro respecto de su error típico, representa un estadístico de contraste que se distribuye como una variable aleatoria que sigue una distribución normal. Por ejemplo, en el caso de los *paths* (los caminos) de regresión, esta condición, permite establecer un contraste de hipótesis nula, individual:

$$CR_{\hat{\theta}} = \frac{\hat{\theta} - 0}{S_{\hat{\theta}}}$$

Siendo “ $\hat{\theta}$ ” el parámetro estimado, “0” el verdadero valor poblacional de dicho parámetro y “ $S_{\hat{\theta}}$ ” el error típico obtenido por máxima verosimilitud.

---

<sup>55</sup> Traducción directa: “Although no single index meets all the above criteria, the set of four presented above collectively provides satisfactory criteria for overall model evaluation”

<sup>56</sup> La expresión en inglés es “coefficient ratio”.

El programa AMOS, software que hemos utilizado en esta investigación, asocia a cada parámetro estimado su test de contraste y su “p-valor” que nos permite decidir si un *path* (camino de análisis) resulta o no significativo.

Para estimar y valorar los ajustes diseñados en esta investigación utilizaremos el programa AMOS SPSS versión 20.

Ahora bien, las pruebas anteriores no resultan muy convincentes porque se basan en aproximaciones a los errores estándar de las supuestas distribuciones de probabilidad. Para superar esa debilidad los autores han propuesto una prueba más fuerte para determinar la significación de un parámetro individual. Se trata del test de modelos anidados que veremos en el apartado “La validación del análisis de invarianza: los test de anidamiento”.

### **3.2.3. EL SISTEMA COMPLETO: LOS SUB-SISTEMAS DE MEDIDA Y ESTRUCTURAL.**

En aras de buscar la sencillez expositiva, hemos presentado un ejemplo de ecuaciones simultáneas con variables observables. En realidad el potencial de los modelos estructurales radica en su capacidad para trabajar con variables latentes.

Los modelos de ecuaciones estructurales nacen de la integración de dos tipos de subsistemas de ecuaciones, el de medida y el estructural.

El subsistema de medida configura la forma en que los indicadores se relacionan con sus respectivas variables latentes, mientras que el sistema estructural, describe las relaciones que mantienen entre sí dichas variables latentes. Pero antes de presentar el subsistema de métricas debemos abordar la cuestión previa de la formación de la medida.

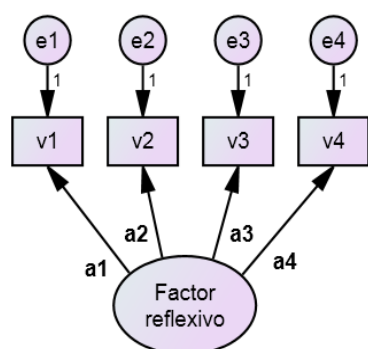
### 3.3. MÉTRICAS FORMATIVAS VERSUS MÉTRICAS REFLEXIVAS.

Hay dos formas básicas en las que pueden relacionarse los indicadores con sus variables latentes, en nuestro caso los cuatro ítems que construyen el índice Pulse con la variable latente de reputación corporativa global. El modo formativo y el reflexivo (Bagozzi y Yi, 2012).

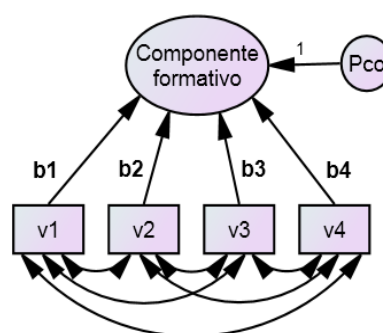
En el modo formativo asumimos que los indicadores “causan la variable latente... y que el constructo es la suma de las variables observadas... (Mientras que en el modo reflexivo) ... *los indicadores son representaciones del constructo*” (Helm, 2005, 97).

En el modo formativo, la variable latente es una combinación lineal perfecta de los indicadores que la forman. En el diagrama y la ecuación de la derecha de la figura siguiente que representan el modo de relación, vemos matemáticamente que el componente “F” es una combinación lineal perfecta de los indicadores que se originan al pre-multiplicar el vector traspuesto de las variables “ $v' = \{v_1, v_2, v_3, v_4\}$ ” por el vector de las pendientes “b” que representa la contribución de cada variable a la formación del constructo.

Modo reflexivo de medida



Modo formativo medida



Por ejemplo, el estatus socioeconómico es concebido como la combinación exacta del nivel de ingresos, educación y ocupación de la persona. La medida está fundamentada sobre bases teóricas y por tanto los indicadores deben ser sustantivamente definidos y no depender de sus propiedades métricas. Los constructos son aceptables si tienen validez de contenido, si sus indicadores han sido seleccionados teóricamente. *“Las medidas con pobre validez de contenido producen resultados empíricos*

*equivocados... y pueden llevar a la aceptación o rechazo erróneos de hipótesis e incluso teorías enteras*”<sup>57</sup> (Rossiter, 2008, 381).

Por el contrario, cuando usamos una métrica reflexiva *“estamos asumiendo que los indicadores son reflejos o representaciones del constructo”* (Helm, 2005, 97).

Matemáticamente aquí no disponemos de una ecuación sino de un vector de ecuaciones “*v*”, tantas como variables observables que representan el constructo. Cada una de ellas es función del mismo constructo que representan y de un error que les es específico, representado por el vector “*ei*”.

Los ítems comparten la variación de la latente, cuanto más inter-correlacionados, mayor varianza pueden llegar a compartir. Ninguna medida por separado es capaz de reflejar completamente el contenido del constructo que sólo se refleja por la combinación común de ellas. Cada ítem observado tiene asociado su error “*ei*” que refleja la variabilidad del ítem ajena al constructo.

Dado su carácter latente, las métricas deben demostrar que poseen la validez de constructo, es decir, las dimensiones que subyacen, el grado de convergencia de los indicadores en el constructo considerado, su discriminación con los indicadores de otros constructos y su capacidad para relacionarse teóricamente con otros constructos<sup>58</sup>.

La mayoría de autores coinciden en aceptar que el modo reflexivo es el más indicado para medir el constructo de reputación global (Carreras et al., 2012; Highhouse et al., 2009; Ponzi et al., 2011; Nettet y Helgesen, 2009; Sung y Yang, 2009; Walsh et al., 2009). Los partidarios de la aproximación formativa a la reputación cognitiva afirman que la métrica global debe ser aproximada reflexivamente (Helm, 2005, 99).

Por tanto adoptaremos el enfoque reflexivo para nuestra investigación.

---

<sup>57</sup> Traducción directa: “Measures with poor content validity produce misleading empirical results ... and can lead to erroneous acceptance or rejection of hypotheses and even entire theories”.

<sup>58</sup> Este problema fue planteado desde un punto de vista teórico en el apartado “El índice Pulse como variable latente”.

### 3.3.1. LAS ECUACIONES DE LOS SUBSISTEMAS DE MEDIDA Y ESTRUCTURAL.

Los modelos de ecuaciones estructurales funcionan con medidas formativas, reflexivas o mixtas, como los modelos MIMIC que combinan ambas (Bagozzi y Yi, 2012; Bollen, 1989; Kline, 1998). Dado que nuestra métrica principal, la reputación global, se aproxima reflexivamente, proponemos el subsistema de medida reflexivo.

Este subsistema reúne las ecuaciones de las variables observadas del modelo. Bollen, siguiendo la nomenclatura de Jöreskog y Söbom, distingue entre variables observadas exógenas “x” y variables observadas endógenas “y”. Proponemos el modelo de la teoría de la acción razonada, presentado por Fishbein y Ajzen en 1975 (Bagozzi, 1992):

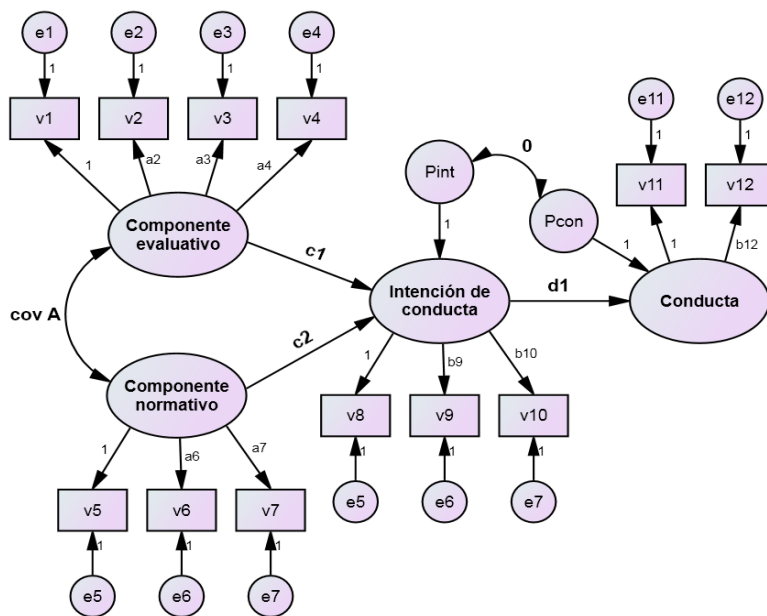


Figura nº 5. Esquema de un modelo de ecuaciones estructurales (modelo basado en la teoría de la acción razonada, Bagozzi 1992, 179).

El modelo contempla las principales tesis de la teoría de la acción razonada. Los determinantes de la conducta son de dos tipos, el componente actitudinal o evaluativo



y el componente normativo. Ambos pueden estar relacionados, como se indica por la covariante que los une. Sin embargo, ambos constructos no son suficientes para desencadenar la conducta, precisan de un elemento disparador que es la intención de conducta. Esto queda reflejado por el proceso de mediación, sin enlaces directos que unan los determinantes con la conducta final. En términos más simples diríamos que una persona tiende a desarrollar una conducta bien porque la valora positivamente y/o bien porque siente que debe hacerlo, pero ambos elementos no se materializarán en una conducta real si en la persona no se forma la intención de actuar, que vuelve reales y materiales las motivaciones teóricas anteriores.

La base empírica u observable del modelo la componen en este caso 12 indicadores. Los siete primeros son exógenos, cuatro reflejan el componente evaluativo (de la v1 a la v4) y los otros tres el normativo (de la v5 a la v7). Los cinco restantes son observables endógenos, agrupados de la “v8” a la “v10” en el componente intención, mientras que los dos últimos, “v11” y “v12” se reservan para aproximar la conducta.

La configuración de cómo se relacionan los indicadores con sus respectivas latentes formando los 12 sistemas de ecuaciones queda explicitada por el subsistema de medida.

Según Bollen “un modelo de medida especifica un modelo estructural que conecta las variables latentes a una o más medidas o variables observables”<sup>59</sup> (Bollen,1989, 182).

El subsistema de medida se compone de dos sistemas de ecuaciones, el de las variables exógenas observables y el de las endógenas.

Subsistema de media exógeno

$$x = \Lambda_x \xi + \delta$$

Subsistema de medida endógeno

$$y = \Lambda_y \eta + \varepsilon$$

Siendo:

“x” Vector de variables observables exógenas

“y” Vector de variables observables endógenas

“ $\Lambda_x$ ” Matriz de cargas de latentes exógenas “ $\xi$ ”

“ $\Lambda_y$ ” Matriz de cargas de latentes exógenas “ $\eta$ ”

“ $\xi$ ” Vector de variables latentes exógenas

“ $\eta$ ” Vector de variables latentes endógenas

---

<sup>59</sup> Traducción directa: “A measurement model specifies a structural model connecting latent variables to one or more measures or observed variables”.

“ $\delta$ ” Vector de residuales de observables exógenas    “ $\varepsilon$ ” Vector de residuales de observables endógenas

En nuestro ejemplo:

$$\begin{bmatrix} v1 \\ v2 \\ v3 \\ v4 \\ v5 \\ v6 \\ v7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ a_2 & 0 \\ a_3 & 0 \\ a_4 & 0 \\ 0 & 1 \\ 0 & a_6 \\ 0 & a_7 \end{bmatrix} * \begin{bmatrix} Cev \\ Cnor \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \\ \delta_7 \end{bmatrix} \quad \begin{bmatrix} v8 \\ v9 \\ v10 \\ v11 \\ v12 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ b_9 & 0 \\ b_{10} & 0 \\ 0 & 1 \\ 0 & b_{12} \end{bmatrix} * \begin{bmatrix} Ic \\ C \end{bmatrix} + \begin{bmatrix} \varepsilon_8 \\ \varepsilon_9 \\ \varepsilon_{10} \\ \varepsilon_{11} \\ \varepsilon_{12} \end{bmatrix}$$

Figura nº 6. sistemas de ecuaciones de las variables exógenas observables y de las endógenas.

En este caso disponemos de un sistema de siete ecuaciones para las variables observables exógenas que representan los dos constructos latentes determinantes del modelo “Cev”, componente evaluativo, “Cnor” componente normativo, y cinco observables endógenas que forman las latentes explicadas, “Ic” intención de conducta y “C” conducta.

Las cargas “ai” y “bi” indican cuanto cambia la variable por unidad de cambio en su correspondiente latente y las matrices de cargas, señalan como se relacionan los indicadores con sus latentes. Para establecer la escala de la variable latente, se fija una carga a un valor, normalmente 1. El componente evaluativo ha quedado fijado a la observada “v1”, el normativo a la “v5”, la intención a la “v8” y la conducta a la “v11”. Obviamente estas cargas están fijas y no serán estimadas.

Los errores de medida “ $\delta_i$ ” y “ $\varepsilon_i$ ”, como ya hemos comentado, representan fuentes de variación ajenas al constructo.

En el ejemplo, los parámetros a estimar serían, las cinco cargas “ai”, las tres “bi”, las siete varianzas de los residuales delta, Var ( $\delta_i$ ) y las cinco de los épsilon, Var ( $\varepsilon_i$ ) (las

covarianzas entre los residuales exógenos y endógenos están restringidas a cero<sup>60</sup>). En total, veintitres parámetros a estimar en los modelos de medida.

Una vez descontaminadas las medidas latentes de las diferentes fuentes de error, se procede a formular el modelo de impactos estructural que determina cómo se relacionan entre sí.

El subsistema estructural de impactos representa “el modelo de variable latente que muestra la influencia de una variable latente con otra”<sup>61</sup> (Bollen, 1989, 319).

En nuestro ejemplo, el modelo estructural está formado por dos ecuaciones que se corresponden con las variables latentes explicadas. La primera “intención de conducta” representa un proceso de mediación que enlaza la influencia de los determinantes evaluación y norma, sobre la variable final, “conducta”. El proceso de mediación se especifica en la matriz de de coeficientes gamma “ $\Gamma$ ”, que deja la última fila restringida a ceros, indicando que las latentes exógenas o determinantes no tienen impacto directo sobre la conducta, dejando que toda la influencia de los determinantes sobre la final se procese a través de la variable intermedia, expresada en la matriz de

---

<sup>60</sup> En realidad son las matrices de varianzas y covarianzas de los residuales, pero las covarianzas han sido restringidas a cero, indicando que los residuales son independientes entre sí.

Matriz de residuales exógenos

$$\begin{bmatrix} V\delta_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & V\delta_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & V\delta_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & V\delta_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & V\delta_5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & V\delta_6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & V\delta_7 \end{bmatrix}$$

Matriz de residuales endógenos

$$\begin{bmatrix} V\varepsilon_8 & 0 & 0 & 0 & 0 \\ 0 & V\varepsilon_9 & 0 & 0 & 0 \\ 0 & 0 & V\varepsilon_{10} & 0 & 0 \\ 0 & 0 & 0 & V\varepsilon_{11} & 0 \\ 0 & 0 & 0 & 0 & V\varepsilon_{12} \end{bmatrix}$$

<sup>61</sup> Traducción directa: “... latent variable model that shows the influence of one latent variable on another”.

coeficientes beta “B”, con el parámetro “d1” que indica el impacto de la intención sobre la conducta<sup>62</sup>.

### El subsistema estructural

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

Siendo:

- “η” El vector de variables latentes endógenas
- “B” La matriz de coeficientes beta que relacionan las latentes endógenas entre sí
- “Γ” La matriz de coeficientes gamma que relacionan las latentes exógenas con las endógenas
- “ξ” El vector de variables latentes exógenas
- “ζ” El vector de perturbadores de las variables latentes endógenas

En nuestro ejemplo, los impactos entre las variables latentes se despliega en las siguientes matrices:

$$\begin{bmatrix} Ic \\ C \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ d1 & 0 \end{bmatrix} * \begin{bmatrix} Ic \\ C \end{bmatrix} + \begin{bmatrix} c1 & c2 \\ 0 & 0 \end{bmatrix} * \begin{bmatrix} Cev \\ Cnor \end{bmatrix} + \begin{bmatrix} Pint \\ Pcon \end{bmatrix}$$

Los parámetros del modelo interno a estimar son dos coeficientes “ci” que enlazan los determinantes con la intención de conducta, un *path* “d1” que vincula las endógenas entre sí, las dos varianzas de los perturbadores endógenas porque su covarianza se ha fijado en cero <sup>63</sup>, que representan la variabilidad de las latentes que no ha sido explicada por el modelo. La presencia de una covarianza entre estos perturbadores, indicaría que hay parte de la varianza no explicada por el modelo que está explicada por otro factor. En total, en el modelo interno, se deberían estimar cinco parámetros.

En resumen, los modelos estructurales completos, con los subsistemas de medida y estructural, estarían formados por ocho matrices de parámetros a estimar agrupados en el vector de parámetros “θ”.

<sup>62</sup> Los ceros en la diagonal de la matriz “B” indica que las latentes no se explican entre sí. Además, el cero del coeficiente de la fila 1, columna 2, indica que no hay una relación recíproca, lo que devolvería el efecto de la conducta sobre la intención.

<sup>63</sup> La matriz  $\begin{bmatrix} V(Pint) & 0 \\ 0 & V(Pcon) \end{bmatrix}$

- Las dos matrices de cargas de las latentes exógenas y endógenas: “ $\Lambda_x$ ” y “ $\Lambda_y$ ”.
- Las dos matrices de varianzas y covarianzas de los residuales: “ $\delta$ ” y “ $\epsilon$ ”.
- La matriz de varianzas y covarianzas de los perturbadores: “ $\zeta$ ”.
- La matriz de coeficientes que enlazan las endógenas entre sí “ $B$ ”.
- La matriz de coeficientes que enlazan las exógenas con las endógenas “ $\Gamma$ ”.
- El vector de varianzas de las variables latentes endógenas,  $\text{Var}(\eta)$ .

Nuestro ejemplo propuesto podría ser estimado porque está sobre-identificado. Por un lado contiene tendría 78 puntos de observación, correspondiendo a las 12 varianzas y 66 covarianzas<sup>64</sup>. Como disponemos de 28 parámetros a estimar, 23 del subsistema externo y 5 del interno, el modelo tendría 28 parámetros a estimar, con 50 grados de libertad.

En síntesis, como ya hemos dicho, estos métodos descansan en una sencilla idea de encaje. La matriz de covarianzas observadas representa la variación total de todos los indicadores entre sí, todos con todos, sin restricciones. Ahora proponemos dos modelos, uno de medida que determina cómo se forman las variables latentes (subsistema de medida o externo) y el otro para determinar el modelo de impactos entre dichas latentes (subsistema estructural o interno). Nuestra hipótesis se vería verificada o refutada si la modelización conjunta consigue reproducir completamente la variación poblacional observada<sup>65</sup> porque esto significaría que el modelo se ajusta a la realidad.

Aplicando los principios del álgebra de covarianzas, transformamos los parámetros del modelo en una matriz de covarianzas implicada por el modelo, y estimamos sus valores mediante un algoritmo que vuelve mínima la matriz residual (diferencia entre

---

<sup>64</sup> La regla para el cálculo de elementos muestrales:  $\frac{\text{variables}(\text{variables}+1)}{2}$ . En nuestro caso:  $\frac{12*13}{2} = 78$

<sup>65</sup> En realidad, que las diferencias encontradas se deban a un error de azar, pero que en la población ha reproducción ha sido completa.

la observada y la implicada: “ $S - \hat{\Sigma}(\theta)$ ”). Por último, comprobamos si hemos sido capaces de reproducir toda la variación original mediante un contraste de hipótesis:  $H_0: \Sigma = \Sigma(\theta)$ . El contraste de hipótesis y los índices de ajuste ayudan a concluir si los datos empíricos, corroboran<sup>66</sup> o rechazan el modelo.

---

<sup>66</sup> En rigor, mejor que “corroboran”, deberíamos decir “no rechazan”.

### 3.4. EL ANÁLISIS FACTORIAL CONFIRMATORIO (CFA).

El análisis factorial confirmatorio es una técnica de los métodos de ecuaciones estructurales que se utiliza para validar la calidad de las escalas de medida. Según Richard Bagozzi y Youjae Yi esta técnica sirve para verificar las dimensiones que subyacen a un conjunto de ítems, para determinar la fiabilidad y la validez de constructo de esas dimensiones (Bagozzi y Yi 2012, 14 y ss). Estas pruebas se corresponden con las de nuestro primer conjunto de hipótesis de investigación.

El modelo factorial confirmatorio (ACF) se origina al dejar covariar todas las latentes entre sí. En efecto, cuando saturamos las relaciones entre las variables latentes enlazándolas con covarianzas, estamos reproduciendo la variabilidad completa del modelo estructural y en este sentido, la posible falta de ajuste entre las matrices observadas e implicada sería exclusivamente imputable al modelo de medida (Anderson y Gerbing, 1988). Esta condición permite evaluar todas las propiedades psicométricas antedichas de las variables latentes propuestas mediante el AFC.

Para verificar las propiedades psicométricas del Pulse proponemos el siguiente modelo factorial confirmatorio:

#### Modelo reflexivo del índice Pulse

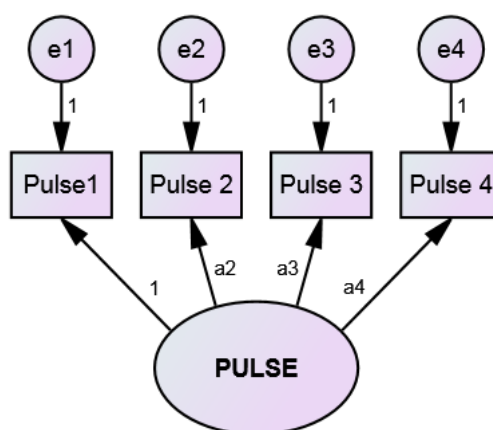


Figura nº 7. Modelo factorial confirmatorio de los cuatro ítems del Pulse.

Las ecuaciones del sistema de medida:

$$v = \Lambda_v \xi + \delta$$

Siendo:

- “v” Vector de variables observables
- “a<sub>x</sub>” Vector de cargas
- “ξ” Factor del Pulse
- “e<sub>i</sub>” Vector de residuales de observables exógenas

Da lugar al siguiente conjunto de ecuaciones:

$$\begin{bmatrix} v1 \\ v2 \\ v3 \\ v4 \end{bmatrix} = \begin{bmatrix} 1 \\ a2 \\ a3 \\ a4 \end{bmatrix} * [Pulse] + \begin{bmatrix} e1 \\ e2 \\ e3 \\ e4 \end{bmatrix}$$

El modelo está sobre identificado y podrá ser estimado. Tiene ocho parámetros a estimar, tres cargas, la varianza de la latente y cuatro varianzas de residuales. El modelo contiene 10 datos observacionales ( $\frac{4*5}{2}$ ), cuatro varianzas, seis covarianzas, con lo que el modelo presenta 2 grados de libertad ( $10 - 8 = 2$ ).

El algebra que transforma el modelo de impactos en la matriz de covarianzas implicadas<sup>67</sup> es el siguiente (Bollen, 1989: 236):

---

<sup>67</sup> La derivación es sencilla. Dado que se trata de variables observadas estandarizadas, la matriz de covarianzas implicada queda simplificada:

$$\begin{aligned} \Sigma(\theta) &= E(xx') = E[(\Lambda_x \xi + e) * (\xi' \Lambda'_x + e')] = \\ &= E(\Lambda_x \xi \xi' \Lambda'_x + \Lambda_x \xi e' + e \xi' \Lambda'_x + ee') = \\ &= \Lambda_x E(\xi \xi') \Lambda'_x + 0 + 0 + E(ee') \end{aligned}$$

Los ceros son debidos a la incorrelación de los residuales con las latentes,  $E(\xi \xi')$  es la matriz de varianza y covarianzas de la endógena y el término  $E(ee')$  es la matriz de varianzas y covarianzas de los residuales,  $\theta e$



$$\Sigma(\theta) = \Lambda_x \phi \Lambda'_x + \Theta_e$$

Siendo:

- “ $\Lambda_x$ ” La matriz de cargas de las observadas
- “ $\phi$ ” La matriz de varianzas y covarianzas de las latentes
- “ $\Theta_e$ ” La matriz de varianzas y covarianzas residuales

### 3.4.1. REPRESENTATIVIDAD Y RELACIÓN DE LOS INDICADORES CON SU VARIABLE LATENTE.

Antes de verificar las propiedades psicométricas del índice Pulse, conviene despejar dos cuestiones previas. La primera tiene que ver con el grado en que los cuatro ítems representan el índice Pulse y la segunda analiza el tipo de relaciones que mantienen dichos cuatro indicadores con el constructo que representan (la reputación global).

- Comprobación de la representatividad de los indicadores del Pulse.

Para comprobar el grado en que los indicadores reflejan el constructo, se realizarán dos pruebas:

- (1) El cálculo de la fiabilidad de cada indicador .
- (2) La in-correlación entre los residuos de sus indicadores.

La fiabilidad de cada indicador señala la cantidad de varianza del valor verdadero que contiene el indicador (Bagozzi y Yi, 2012, 17).

$$\rho_{indicador} = \frac{\lambda_{ind}^2 Var(factor)}{\lambda_{ind}^2 Var(factor) + Var(e_i)}$$

Siendo:

- “ $\lambda_{ind}^2$ ” La carga de la latente sobre el indicador
- “ $e_i$ ” El residual del indicador

Los autores proponen que fiabilidades superiores al 0,7 serían las adecuadas (Bagozzi y Yi, 2012, 17).

### **3.4.2. Ausencia de correlación entre los residuales de los cuatro indicadores del Pulse.**

El análisis no estaría completo si no verificáramos la incorrelación de los residuos de los indicadores. Recordemos que los residuos contienen la varianza de cada uno de los cuatro indicadores ajena al constructo y que ese tipo de varianza puede ser aleatoria o sistemática, proveniente de otro factor común. Como veremos en el capítulo de resultados, este supuesto no se ha cumplido para el índice Pulse en el caso de España.

- El tipo de relación entre los indicadores y el Pulse

Cuando el conjunto de indicadores representa una única dimensión, conviene determinar el tipo de relación que mantienen con el factor común. El estudio de las relaciones entre el conjunto de indicadores y su constructo ayuda a comprender mejor el significado empírico del mismo.

Se distinguen tres tipos de relación (Nunnally y Berstein, 1994; Kline, 1998):

- A. Relación Co-genérica. Es la relación más laxa que mantienen los indicadores con su variable latente (la reputación en nuestro caso). Significa que cada uno contiene la varianza común del factor, pero en grado diferente; de los cuatro indicadores, algunos indicadores lo representarían mejor y contendrían más variabilidad del valor verdadero que otros.
- B. Relación Tau-equivalente. Este tipo impone una condición adicional; además de compartir la varianza del factor, la comparten en el mismo grado. En este caso se podría afirmar que ninguno de los aspectos contemplados por los indicadores representa mejor, es una característica más esencial en el concepto implicado por la dimensión.
- C. Relación paralela. Se trata de la restricción más severa. Los indicadores no sólo comparten la misma cantidad de varianza verdadera sino que sus residuales, la variabilidad no explicada por cada uno de ellos, es la misma. En este caso estaríamos hablando de aspectos prácticamente intercambiables desde el punto de vista métrico que se darían en cualquiera de los cuatro ítems que construyen el índice Pulse.

Para comprobar el tipo de relación que mantienen los cuatro indicadores con el índice Pulse, utilizaremos la técnica del anidamiento propuesta por Rex Kline para estos casos (1998). A partir del modelo co-genérico, tomado como modelo original, se imponen restricciones adicionales dando lugar a modelos anidados. En primer lugar se verificará la hipótesis de indicadores en relación Tau-equivalente al restringir las cargas a la igualdad. Si la pérdida de ajuste no resulta significativa, se procedería a imponer la igualdad en las varianzas de los residuos para comprobar la relación paralela.

Antes de continuar conviene referirnos brevemente al problema de la equivalencia de modelos.

### 3.5. EL PROBLEMA DE LOS MODELOS SEM: LA EQUIVALENCIA.

Una de las mayores dificultades en la interpretación de los modelos SEM procede del problema de la equivalencia. Se dice que dos “modelos (son) equivalentes (cuando) producen las mismas correlaciones o covarianzas predichas, pero lo hacen con una configuración diferente de *paths* (camino de análisis) entre las mismas variables”<sup>68</sup> (Kline, 1998, 139).

La equivalencia significa que la metodología SEM conlleva una indeterminación importante en la conclusión de los modelos ajustados. Entre modelos sencillos y complejos e incluso entre modelos alternativos, estos métodos pueden llegar a identificar la configuración que mejor se adapta a los datos, pero existe una clase de modelos en los que se muestra completamente ciega, es en los modelos equivalentes. Los métodos SEM son incapaces de determinar, entre modelos equivalentes, cuál de ellos es el más empírico el que mejor se ajusta a la realidad, sencillamente porque presentan el mismo nivel de ajuste. En consecuencia, la decisión de cuál es el modelo correcto siempre tendrá que fundamentarse sobre bases teóricas, nunca empíricas.

Según Rex Kline, cuando un investigador ha corroborado un modelo, debe considerar sus modelos equivalentes para seleccionar el más adecuado desde un punto de vista empírico (Kline, 1998, 139).

En nuestro caso, el índice Pulse tiene la forma de un modelo factorial confirmatorio uni-dimensional. Para identificar los modelos equivalentes, se aplica la “regla del indicador invertido”<sup>69</sup> que consiste en comprobar el sentido teórico de cada uno de los cuatro indicadores que reflejan el constructo, como si fueran la causa de los tres restantes (Kline, 1998).

En la figura siguiente se identifican cuatro modelos equivalentes que darían la misma calidad de ajuste y no podrían ser seleccionados solamente en base a criterios puramente empíricos. Esto nos lleva a revisar la teoría y como hemos comentado en la

---

<sup>68</sup> Traducción directa: “Equivalent models yield the same predicted correlations or covariances, but they do so with a different configuration of paths among the same variables”

<sup>69</sup> Traducción directa: “reversed indicator rule”.

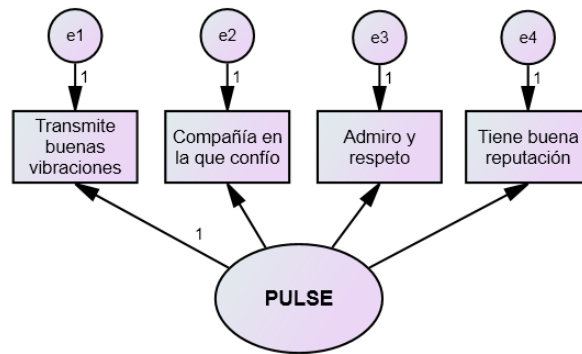
revisión científica del capítulo uno de marco teórico, el índice Pulse nace de los sentimientos asociados directamente a la idea global de reputación<sup>70</sup>.

Los modelos equivalentes plantean las hipótesis alternativas de que alguno de los sentimientos aflorados actúe como un sentimiento anterior capaz de generar el resto de evaluaciones. En el esquema para Pulse se distinguen cuatro modelos alternativos:

- A. El sentimiento “despertar buenas impresiones” sea causa de la idea global de reputación.
- B. El sentimiento “confianza” sea causa de un sentimiento global de reputación.
- C. El sentimiento “admiración y respeto” sea causa de la idea global de reputación.
- D. El sentimiento “buena reputación” sea causa de la idea global de reputación.

---

<sup>70</sup> Véase el apartado “El índice Pulse de Ponzi, Fombrun y Gardberg (2011)”.



## MODELOS EQUIVALENTES

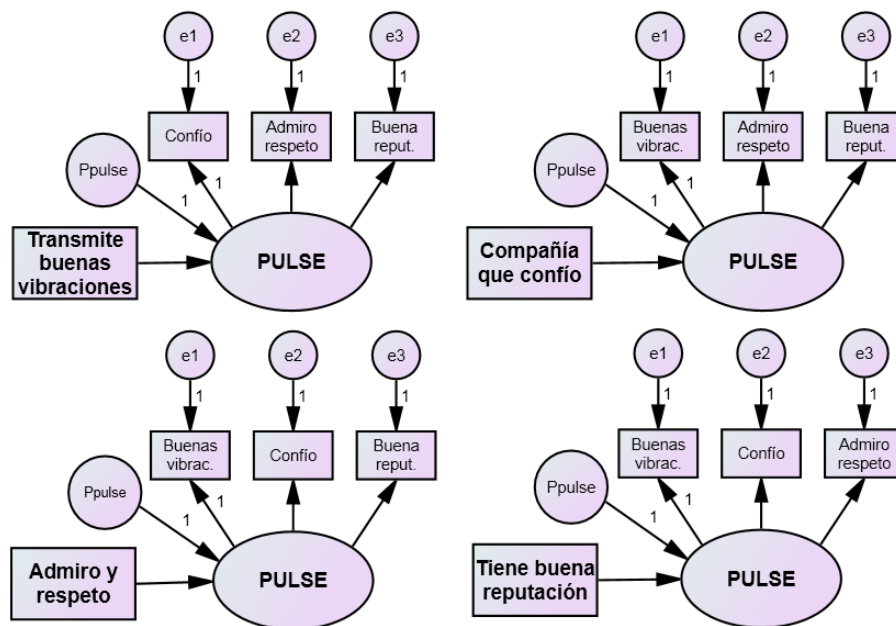


Figura nº 8. Cuatro modelos alternativos para Pulse.

De las cuatro hipótesis que hemos mencionado, la única que podría tener alguna base teórica sería la de confianza. Las tres restantes, *"transmitir buenas impresiones"*, *"reputación"* y *"admiración y respeto"* son consustanciales al contenido socialmente aceptado de reputación tal y como es definido en los diccionarios español, inglés y alemán (Carreras et al.,2013).

Sin embargo, algunas investigaciones han considerado que la confianza es un sentimiento anterior necesario para generar una relación estable (Sung y Yang, 2009; Walsh et al., 2009). Dejamos esta cuestión abierta a la espera de que se produzca un mayor desarrollo de investigación teórica sobre el vínculo de sentimientos asociados al concepto genérico de la reputación.

### **3.6. TÉCNICAS PARA COMPROBAR LA UNI-DIMENSIONALIDAD DEL ÍNDICE PULSE.**

Ya hemos comentado que las pruebas requeridas para comprobar las propiedades métricas del índice Pulse son la uni-dimensionalidad de sus indicadores y la validez de constructo (convergente, discriminante y nomológica). Ahora presentamos las técnicas concretas que nos permitirán su demostración o refutarán esta hipótesis.

La uni-dimensionalidad del índice Pulse “H1.1” quedará demostrada con la cantidad de varianza de los ítems del Pulse que queda extractada por el primer factor, en un análisis de componentes principales.

Cuando la varianza extraída es superior al 80% se considera que en la variación conjunta de los cuatro ítems subyace una sola dimensión por tanto quedaría demostrada la uni-dimensionalidad de los cuatro ítems que componen el índice Pulse.

### **3.7. TÉCNICAS PARA COMPROBAR LA VALIDEZ CONVERGENTE.**

La “validez convergente” mide el grado en que los distintos indicadores que representan un mismo constructo convergen o están altamente correlacionados entre sí. En nuestro caso mediríamos hasta qué punto los cuatro ítems del índice Pulse están o no están altamente correlacionados entre sí. Los autores han propuesto tres pruebas: (a) La consistencia interna, (b) la fiabilidad del componente y (c) la varianza extractada.

- (a) La consistencia interna (H1.2.2) de los indicadores, mide el grado de sistematicidad en las respuestas de diferentes personas (De Vellis, 2003; Kline, 1998; Nunnally y Berstein, 1994). Cuanto mayor consistencia en las respuestas de más personas, mayor inter-correlación de los indicadores, mayor estabilidad de la medida y menores fluctuaciones debidas al azar. La medida recomendada de consistencia interna de las respuestas a los cuatro indicadores Pulse es el llamado índice “*Alfa de Chronbach*”. (Churchill, 1979, 68):

$$\alpha = \left[ \frac{p}{p-1} \right] * \left[ 1 - \frac{\sum_{i=1}^p S_{p_i}^2}{S_{CL}^2} \right]$$

Siendo:

- “p” El número de indicadores
- “S<sub>pi</sub><sup>2</sup>” La varianza del indicador “p<sub>i</sub>”
- “S<sub>CL</sub><sup>2</sup>” La varianza de la combinación lineal de las variables

Interpretación: El rango de valores, entre cero y uno. Por encima del 0,7 se considera que la combinación lineal de los indicadores es fiable (De Vellis, 2003).

- (b) La fiabilidad del componente (H1.2.4). Se trata de una medida de fiabilidad del componente, menos tradicional y más adecuada para los métodos SEM (Bagozzi y Yi 2012). La fiabilidad es la cantidad de varianza total que comparten los indicadores.

$$\rho_{Pulse} = \frac{(\sum \lambda_i)^2 Var T}{(\sum \lambda_i)^2 Var T + \sum Var \theta_i}$$

Siendo:

- “λ<sub>i</sub>” La carga factorial que enlaza el indicador “i” con su latente
- “Var T” La varianza la variable latente
- “Var θ<sub>ii</sub>” La varianza de los residuos de los indicadores

Interpretación: El rango de valores, entre cero y uno. Por encima del 0,7 se considera que el componente estimado es fiable (Bagozzi y Yi, 2012,17).

- (c) La varianza extractada del componente (H1.2.3.). El índice anterior mide la cantidad de varianza de la variable latente, pero no mide “la cantidad de varianza captada por el constructo en relación con la cantidad de varianza debida al error de medida”<sup>71</sup> (Fornell y Larcker, 1981, 45). Claes Fornell y David Larcker desarrollaron una medida más fina que la fiabilidad y que

---

<sup>71</sup> Traducción directa: “the amount of variance that is captured by the construct in relation to the amount of variance due to measurement error”



denominaron AVE<sup>72</sup> o promedio de varianza extractada por el factor en los indicadores que mide la cantidad de varianza debida a constructo respecto de la varianza total:

$$AVE = \frac{\sum \lambda_i^2 Var T}{\sum \lambda_i^2 Var T + \sum Var \theta_i}$$

Interpretación. Rango entre cero u uno. Los autores consideraron que el constructo, como mínimo la varianza de los errores no debería superar la varianza captada en el conjunto de indicadores, por tanto para tener suficiente varianza extractada debería ser superior a 0,5 (Fornell y Larcker, 1981, 46).

La relación de los ítems con su constructo, la reputación global, se verificará mediante el test de anidamiento, restringiendo las cargas a la igualdad.

---

<sup>72</sup>“AVE” Acrónimo de “Average Variance Extracted”.

### 3.8. EL ANÁLISIS DE INVARIANZA POBLACIONAL.

El análisis de la invarianza es fundamental para determinar si las variaciones en el índice reflejan cambios auténticos en la actitud de las personas bajo observación (Rensvold y Cheung, 2001).

Un índice como Pulse violaría la invarianza de medida si al medir un peso idéntico en dos objetos, diera de forma sistemática distintas puntuaciones simplemente porque su forma fuera diferente. La escala como instrumento de medida debe ser una función aleatoria, condicionada exclusivamente por el atributo medido. De forma similar, un instrumento diseñado para medir un atributo psicológico como la depresión o la reputación, debe revelar medidas diferentes en individuos si en realidad difieren en los niveles de dicho atributo. Si al aplicar una escala para medir la depresión en dos grupos de hombres y mujeres que tienen el mismo promedio real de depresión, rindiera resultados diferentes, diríamos que dicha escala viola la invarianza de medida en relación al sexo.

El problema se circunscribe a la frase “grupos... con el mismo nivel de reputación” porque desconocemos el nivel real del individuo respecto de atributo. Cuando el objeto de medida es psicológico y la medida en los grupos es distinta, se cae en una incertidumbre fundamental que nos impide concluir si la diferencia registrada se debe a un error de sesgo porque en realidad la medida es la misma, o bien porque ha reflejado una diferencia real en dicho atributo (Millsap, 2011).

*“El modelo de variable latente expresa la relación probabilística entre las medidas observadas y una o más variables latentes”* (Millsap 2011,13). Dado que en el modelo factorial confirmatorio se establece que las variables observadas son función de las latentes (modo reflexivo), en una escala unidimensional como la que nos ocupa, sabemos que el vector de las observadas es una función de las cargas y del error:

$$X = \sum a_i W + e_i$$

Por tanto, si las cargas “ $a_i$ ” que relacionan un indicador con la latente (la variabilidad del indicador debida a la latente) son las mismas en diferentes grupos y el error del

mismo indicador (la variabilidad del indicador no explicada por la latente) es el mismo en diferentes grupos, los cambios en el rasgo latente serían evidenciados con la misma intensidad en las variables observadas de diferentes grupos y por tanto, la de la variable latente en dichos grupos sería la misma.

Testar la invarianza de una escala equivale a “*determinar si los ítems usados en instrumentos como las encuestas significan las mismas cosas para miembros de grupos diferentes*”<sup>73</sup> (Cheung y Rensvold, 2002, 233).

Por tanto, si disponemos “k” grupos, darían lugar a “k” estructuras de covarianzas, que en el modelo confirmatorio adoptaría la forma siguiente:

$$\Sigma_k = \Lambda_k \phi \Lambda'_k + \Theta_k$$

Siendo:

- “ $\Sigma_k$ ” La matriz de variables observadas del grupo “k”
- “ $\Lambda_x$ ” La matriz de cargas de las observadas del grupo “k”
- “ $\phi$ ” La matriz de varianzas y covarianzas de las latentes del grupo “k”
- “ $\Theta_e$ ” La matriz de varianzas y covarianzas residuales en el grupo “k”

Por tanto, el modelo puede ser estimado para cada subgrupo y valorar su ajuste mediante un test clásico de la  $\chi^2$  y dado que el valor del  $\chi^2$  es sumativo, el valor global del multi-grupo iguala la suma de los  $\chi^2$  de los modelos ajustados por separado, sin restricciones (Byrne y Stewart, 2006).

Además, los modelos pueden ser anidados si restringimos los parámetros a la igualdad o invarianza, El test chi-cuadrado de la diferencia  $\chi^2_{\text{dif}} = \chi^2_{\text{anidado}} - \chi^2_{\text{original}}$ , con grados de libertad,  $gl_{\text{dif}} = gl_{\text{anidado}} - gl_{\text{original}}$ ) prueba si las restricciones de igualdad en el modelo anidado son plausibles, es decir si podemos demostrar la invarianza en los parámetros para los diferentes grupos.

---

<sup>73</sup> Traducción directa: “... determining if items used in survey-type instruments mean the same things to members of different groups”.

El análisis de invarianza consiste en una secuencia de test anidados, dispuestos lógicamente de forma progresivamente restrictiva, capaz de identificar los parámetros para los que se cumple la invarianza hasta llegar a concluir o no la invarianza total de la escala.

En consecuencia, mediante un análisis de invarianza podríamos probar el grado en que la medida latente del Pulse resulta insensible a las variaciones poblacionales de sexo, edad y nivel de estudios, con lo que se probaría que en grupos muy diversos de población el índice refleja con exactitud las variaciones reales de reputación.

Hipótesis principal de nuestra investigación que formulamos en esos mismos términos:

***Pulse mide la reputación de una empresa u organización con independencia del sexo, edad o clase social de la población que sea entrevistada.***

### **3.8.1. LA SECUENCIA DE HIPÓTESIS DE INVARIANZA POBLACIONAL.**

Conviene distinguir dos tipos de invarianza en las escalas, la que tiene que ver con las propiedades psicométricas o invarianza débil y la invarianza fuerte, que tiene que ver con la diferencia de las medidas latentes.

Como es lógico, ambos tipos de análisis están secuencialmente condicionados. Para concluir que el nivel de reputación medio de un grupo es mayor o menor que el de otro grupo, es necesario verificar previamente la invarianza psicométrica de la escala. Esta invarianza garantiza que los valores alcanzados tienen el mismo valor en las distintas poblaciones y por tanto pueden ser comparados.

En nuestro caso, estamos interesados en demostrar la invarianza psicométrica de la escala por grupos de sexo y edad.

La invarianza psicométrica del Pulse se despliega en una secuencia de tres hipótesis restrictivas<sup>74</sup>:

- La invarianza configural que demuestra que los diferentes grupos utilizan el mismo dominio cognitivo.

$$Modelo^{(1)} = Modelo^{(2)} = \dots = Modelo^{(k)}$$

Se propone el mismo modelo para todos los grupos, dejando libres los parámetros, de forma que la aceptación de los modelos indicaría que los diferentes grupos usan el mismo conjunto de ítems.

El modelo configural representa la línea base, a partir de la cual se desarrollan los modelos anidados, secuencialmente restringidos, de forma que la evaluación de la pérdida de ajuste se realizará respecto del ajuste inicial del modelo configural.

- La invarianza de cargas que verifica si el Pulse tiene el mismo significado para los distintos grupos poblacionales.

$$\Lambda^{(1)} = \Lambda^{(2)} = \dots = \Lambda^{(k)}$$

Siendo “ $\Lambda^k$ ” La matriz de cargas de los ítems que los relacionan con el Pulse en el grupo “k”

El modelo de invarianza de cargas está anidado respecto del configural porque se le impone la restricción adicional de igualdad de cargas en los distintos grupos, aunque la fuerza entre los ítems dentro del grupo pueda diferir. Esto significa que la estructura de cargas es similar entre distintos grupos, que están entendiendo el índice de la misma forma, manteniendo la misma estructura de importancia en los mismos ítems.

Dado que este modelo está anidado respecto del configural, un test de anidamiento indicará si las restricciones han provocado una pérdida significativa del ajuste y por

---

<sup>74</sup> En los análisis de invarianza completos, se plantean cinco hipótesis secuencialmente restrictivas: (1) La invarianza configural, (2) La invarianza métrica al nivel de constructo (3) la invarianza métrica al nivel del ítem (4) la invarianza de los residuales y la invarianza métrica de los interceptores (Cheung y Rensvold 2002: 235 y ss). En nuestro caso, sólo requerimos 3 pruebas, la configural, la de constructo y la residual, debido a que nuestro constructo es unidimensional.

tanto no es invariante o bien, que la pérdida es debida al azar y por tanto los grupos le dan el mismo significado métrico a las respuestas del Pulse.

- La invarianza de la varianza del factor.

Esta prueba verifica si la variabilidad extraída por el factor es la misma en cada grupo.

$$\phi^{(1)} = \phi^{(2)} = \dots = \phi^{(k)}$$

En el caso de corroborarse la invarianza de la variabilidad, significaría que el Pulse es homocedástico en los distintos grupos, por lo tanto, los grupos muestran el mismo rango de variabilidad en los sentimientos que lo conforman.

- La invarianza residual.

Por último, el test más restrictivo es el que soporta la invarianza de los residuales. Esta prueba determina si la cantidad de varianza del ítem no asociada al constructo es la misma y por tanto contienen el mismo grado de error de medida.

$$\theta^{(1)} = \theta^{(2)} = \dots = \theta^{(k)}$$

Siendo “ $\Lambda_k$ ” La matriz de varianzas y covarianzas de los residuales de los ítems, en el grupo “k”

Matemáticamente, sobre el modelo anterior se impone una restricción adicional de igualdad de los residuos y se somete una vez más a un nuevo test de anidamiento.

### **3.8.2. LA DETERMINACIÓN DE LA VARIABLE DE REFERENCIA.**

Antes de pasar a la validación de la invarianza, conviene hacer mención al problema de la “variable de referencia” (Millsap, 2011). Para fijar la escala de la variable latente, normalmente se fija una de las cargas que enlaza el indicador con la latente a la unidad <sup>75</sup>. Sin embargo, esta operación puede producir estimaciones inexactas en el resto de cargas.

Para fijar a uno la carga, debemos de seleccionar los ítems que realmente son invariantes entre los grupos analizados porque de otro modo, las estimaciones resultantes podrían llevar a concluir erróneamente la desigualdad en otras cargas que en realidad fueran iguales. Antes de plantear el test conviene averiguar qué ítems podrían actuar como “marcadores de referencia” para la restricción de igualdad. Rensvold y Cheung (2001) han desarrollado un procedimiento iterativo en el que prueban todos los subconjuntos de variables posibles que pueden ser considerados como invariantes de cargas. Yoon y Millsap (2007) han propuesto un método alternativo basado en la eliminación sucesiva hacia atrás mediante la estimación sucesiva de modelos parcialmente invariantes. Nosotros seguimos el procedimiento de Yonn y Millsap.

### **3.8.3. LA VALIDACIÓN DEL ANÁLISIS DE INVARIANZA: LOS TEST DE ANIDAMIENTO.**

Ya hemos comentado que el análisis de invarianza consiste en un conjunto de pruebas consecutivas que corroboran una serie de hipótesis cada vez más restrictivas y exigentes y que se han dado en llamar, test de anidamiento.

Hay tres pruebas para evaluar la secuencia de hipótesis anidadas en un análisis de invarianza: los índices de ajuste global de cada modelo correspondiente a cada hipótesis de invarianza, la estimación de la pérdida de ajuste entre el modelo padre, configura y los sucesivos anidados y la diferencia entre los índices CFI y RMSEA.

---

<sup>75</sup> Véase el apartado: “El sistema completo: Los subsistemas de medida y estructural”.

Los índices de ajuste global han sido comentados en los apartados anteriores, en los siguientes apartados comentaremos los métodos de evaluación de la pérdida de ajuste y las diferencias entre los índices globales.

#### **3.8.4. . LA EVALUACIÓN DE LA PÉRDIDA DE AJUSTE POR EL ESTADÍSTICO “ $\chi^2_{DIF}$ ”.**

Como hemos comentado antes, el test de la pérdida de ajuste se aplica para verificar la significación de determinados *paths* (caminos de análisis) individuales dentro de un modelo determinado.

El planteamiento del test parte de una lógica es sencilla. Si disponemos de dos modelos completamente iguales y en el segundo imponemos una restricción adicional, esa restricción es el único elemento que lo diferencia del modelo anterior. La prueba consiste en estimar ambos modelos, el original, sin la restricción y el anidado, con la restricción, y comparar la diferencia en la función de ajuste entre ambos modelos.

La comparación nos introduce una disyuntiva: bien el segundo modelo pierde ajuste respecto del primero, o bien no lo pierde. En el caso de que pierda ajuste, esa pérdida sería debida exclusivamente al parámetro restringido y significaría que éste es empíricamente necesario porque el modelo restringido no logra reproducir toda la variación observada. En el caso de que el modelo anidado o restringido, presente el mismo nivel de ajuste, significa necesariamente que dicha restricción es empírica y consigue reproducir la variación observada original.

De esta forma la viabilidad de un parámetro (o subconjunto de ellos) puede ser testado empíricamente analizando la pérdida de ajuste debida exclusivamente a su restricción.

Se dice que un modelo “A” está anidado o dentro de otro modelo “B”, cuando el conjunto de parámetros libres del primero, es un subconjunto de los libres del segundo (Anderson y Gerbin, 1988).

El test de invarianza que hemos presentado en el apartado anterior consiste en una serie de modelos anidados respecto de un modelo base o “configural”, de modo que si



al final de la secuencia de pruebas, se demuestra que no había pérdida significativa, esto supondría que las restricciones de invarianza son empíricas y son por tanto conformes a los datos.

En los test de anidamiento se distinguen dos tipos de restricciones:

- Restricciones de igualdad:

En el que se obliga a que el valor de un parámetro en el modelo anidado sea igual al valor de otro parámetro, por ejemplo, podríamos imponer la restricción de que las cargas que relacionan una variable latente a sus respectivos indicadores, sean las mismas. En el análisis de invarianza esta es tipo de restricción que se aplica. Por ejemplo, se asume que el valor del coeficiente que relaciona el ítem primero con el Pulse, adopte el mismo valor en todos los grupos, sea el que sea y sin requerir la igualdad con los coeficientes del resto de ítems, dentro del mismo grupo:

### Tipo de restricción a la igualdad en el análisis de invarianza

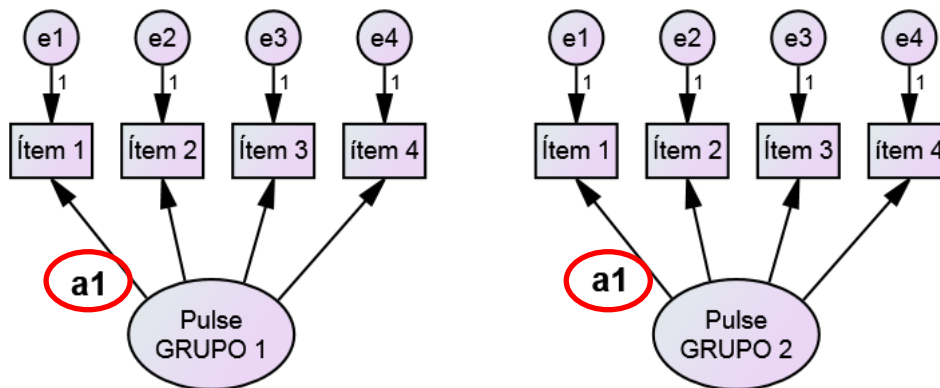


Figura nº 9. Tipo de restricción a la igualdad en el análisis de invarianza Pulse.

- Restricciones a constante:

Por las que obligamos a que el parámetro en cuestión adopte un valor concreto y no pueda ser estimado, el caso más común es la restricción a cero, por la cual afirmamos

que el parámetro restringido no opera en el modelo anidado. En los test de anidamiento, se utilizan exclusivamente las restricciones de igualdad.

Una vez más debemos tener presente que la falta de ajuste podría tener un origen muestral y no poblacional. Necesitamos evaluar la pérdida de ajuste mediante un nuevo contraste.

Existen varios procedimientos para determinar la pérdida significativa del ajuste<sup>76</sup>, aquí sólo veremos el test de anidamiento debido a la diferencia  $\chi^2$ .

Ambos modelo, el anidador o base y el anidado presentan sendas medidas de discrepancia entre sus respectivas matrices de covarianzas. Sabemos que la diferencia entre dos estadísticos  $\chi^2$  es otra variable aleatoria que se distribuye asintóticamente como otra chi-cuadrado, con el valor, diferencia de valores y con grados de libertad, diferencia entre los grados de libertad de las  $\chi^2$  comparadas. Por tanto, la pérdida de ajuste entre los modelos base y anidado puede evaluarse con un test de la diferencia de los chi-cuadrados de ambos modelos (Bollen, 1989, 292):

$$\chi^2_{dif} = (\chi^2_{anidado} - \chi^2_{base}); g \cdot l_{anidado-base}$$

El test de la diferencias prueba la hipótesis nula de la pérdida de ajuste no significativa ( $H_0: \chi^2_{anidado} - \chi^2_{original} = 0$ ) al que se le asocia un “p-valor” para aceptarla o rechazarla.

---

<sup>76</sup> Otros procedimientos se basan en el multiplicador lagrangiano y el test “W” de Wald (Bollen 1989)

### 3.8.5. LA EVALUACIÓN DE LA PÉRDIDA DE AJUSTE POR COMPARACIÓN DE ÍNDICES GLOBALES DE AJUSTE.

La aplicación del test del anidamiento en el análisis de invarianza de grupos, presenta el problema ya conocido del tamaño de la muestra. “Si el tamaño de muestra es grande, sin embargo, incluso pequeñas diferencias entre  $\chi^2$  restringidos y no restringidos puede resultar en un valor significativo de “ $\Delta\chi^2$ ”, indicando que la hipótesis nula de la no diferencia debería ser rechazada incluso cuando la diferencia fuera trivial” <sup>77</sup>. (Cheung y Rensvold , 2002, 239).

En estos casos, se propone completar la evaluación de la pérdida de ajuste, mediante un análisis adicionales con los índices índice incrementales del CFI de Bentler “ $\Delta\text{CFI}$ ”, Gamma, “ $\Delta\text{Gamma Hat}$ ”<sup>78</sup>, y el índice NCI de McDonald “ $\Delta\text{NCI}$ ”<sup>79</sup> (Cheung y Rensvold, 2002, 250). Fang Chen añade el índice incremental del RMSEA “ $\Delta\text{RMSEA}$ ” (Cheng, 2011, 501). En los siguientes casos no debería rechazarse la hipótesis nula de pérdida de ajuste:

- $\Delta\text{CFI} \leq -0,01$
- $\Delta\text{Gamma Hat} \leq -0,001$
- $\Delta\text{NCI} \leq -0.02$
- $\Delta\text{RMSEA} \leq 0,01$

En nuestro caso utilizaremos los índices CFI y RMSEA aportados por el programa AMOS.

---

<sup>77</sup> Traducción directa: “If the sample sizes are large, however, even a small difference between (Fmin ) constrained and (Fmin ) unconstrained may result in a significant value of  $\Delta\chi^2$ , indicating that the null hypothesis of no difference should be rejected even when the difference is trivial”.

<sup>78</sup>La fórmula del índice Gamma Hat de Steiger (1989):

$$\text{Gamma} - \text{hat} = \frac{p}{p + 2 \left[ \frac{\chi^2_{propuesto} - gl_{propuesto}}{gl_{propuesto}(n - 1)} \right]}$$

<sup>79</sup> La fórmula del índice NCI de no centralidad de McDonald (1989):

$$\text{NCI} = e^{\frac{1}{2} \left[ \frac{\chi^2_{propuesto} - gl_{propuesto}}{n-1} \right]}$$

Las pruebas sucesivas irán determinando el grado en que el índice Pulse va demostrando su invarianza poblacional. Si los resultados de los test verifican consecutivamente la no pérdida de ajuste, significaría que las medidas del índice Pulse, en los distintos grupos demográficos homogeneizados por sexo, edad o nivel de estudios, tienen el mismo significado, son comparables y que las variaciones de valor entre dichos grupos reflejan cambios reales en la reputación global en estos colectivos.

En caso de que las pruebas no se confirmaran, un análisis en mayor detalle llegaría a demostrar dónde no se mantiene la invarianza y de esta forma disponer de una interpretación más exacta de las diferencias registradas en las medidas de Pulse.

### 3.9. LA INVARIANZA PSICOMÉTRICA E INVARIANZA DE LAS MEDIAS.

Los autores han distinguido entre dos tipos de análisis de invarianza, la psicométrica de la escala, también denominada invarianza débil y la invarianza fuerte que sirve para demostrar la diferencia de los niveles medios de las latentes entre grupos (Byrne y Stewart, 2006; Little, 1997). Según los autores, esta diferencia depende de dónde ponga el foco de este análisis, en probar la equivalencia psicométrica de la escala entre varios grupos como sostiene la relación de hipótesis a verificar en esta tesis, o bien en comparar los niveles medios de las latentes alcanzados por estos mismos grupos.

La invarianza métrica de propiedades es un requisito previo para la comparación de segundo tipo (Little, 1997; Cheung y Rensvold, 2002). Dicho de otro modo, si conseguimos probar las secuencias de hipótesis sobre la invarianza métrica del Pulse, podemos proceder a comparar los niveles de reputación en los diferentes grupos de población analizados.

Al análisis de la comparación de medias latentes entre grupos se le denomina MACS80 o estructuras de medias y covarianzas. El análisis es ligeramente diferente del que se aplica para probar las propiedades psicométricas. Requiere, además de la invarianza de cargas entre grupos, la de los interceptores de los indicadores o invarianza escalar (Byrne y Stewart, 2006; Little, 1997).

- Invarianza de cargas. La carga de un indicador es la misma en los diferentes grupos. Cuando la carga es la misma, la cantidad de rasgo latente en el indicador es la misma en todos los grupos. De esta forma, se verificaría que la unidad de medida en la variable latente es la misma en los diferentes grupos.
- Invarianza de los interceptores. Establecida la unidad de medida, se desea saber si las puntuaciones tienen el mismo origen en los diferentes grupos. Los interceptores son valores de cada ítem que se corresponden con el valor cero el constructo latente (Byrne y Stewart, 2006). Si fuera

---

<sup>80</sup> Acrónimo en inglés “Means and Covariance Structures”.

diferente para cada grupo de comparación, ante un cambio de nivel de la latente estaríamos ante una incertidumbre porque no se sabría si afecta a la rasgo o al diferente origen de la latente. Para conseguir que el origen sea el mismo, se fuerza a que las medias de los residuos sean cero, de forma que el nivel corresponda completamente al ítem, tal y como se percibe en el esquema de la figura siguiente:

### Tipo de restricción a la igualdad en el análisis de invarianza de medias

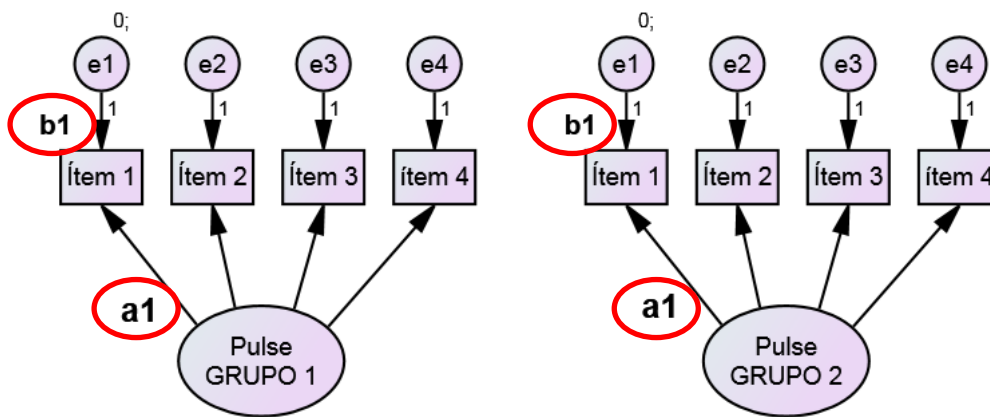


Figura nº 10. Tipo de restricción a la igualdad en el análisis de invarianza de medias Pulse.

La lógica del análisis de invarianza radica en la comparación de modelos anidados. En el análisis MACS, se comparan dos modelos, el de igualdad de cargas e interceptores, que garantiza la invarianza escalar, es decir que los indicadores de cada grupo poseen el rasgo latente en el mismo grado y con el mismo origen y el modelo que impone la igualdad de medias latentes entre los grupos. Roger Millsap ha sintetizado matemáticamente el problema de la invarianza escalar de cargas e interceptores (Millsap, 2011, 73 y ss.):

$$X_j = \tau_{jk} + \sum_{m=1}^r \lambda_{jmk} W_m + U_j$$

Siendo:

- $\tau_{jk}$  el interceptor de la j-ésima variable medida en la población “k”
- $\lambda_{jmk}$  el coeficiente de pautas factoriales de la la j-ésima variable del factor “m” en grupo “k”
- $W_m$  es el valor de la variable latente
- $U_j$  es el valor unicidad de la variable “j”.

De la ecuación fácilmente se deduce que la diferencia de nivel en la latente, exige que tanto los interceptores como las cargas sean iguales entre grupos<sup>81</sup>, porque sólo así las variaciones en la cantidad del rasgo latente en cada grupo podrán ser reflejadas en las medidas alcanzadas por dichos grupos.

---

<sup>81</sup> Debemos recordar que en las ecuaciones, las esperanzas de los errores o unicidades han sido prefijadas a cero y por tanto, no afectan al origen.

## **CAPÍTULO 4. RESULTADOS Y CONTRASTE DE LAS PRUEBAS MATEMÁTICAS DE VERIFICACIÓN DE LAS HIPÓTESIS.**

En el capítulo anterior hemos descrito las técnicas matemáticas y estadísticas concretas que hemos seleccionado como las más adecuadas gracias a la experiencia empírica de Alloza y en base a los referentes académicos y científicos consultados, para poner a prueba las propiedades métricas de Pulse con el objeto de demostrar o refutar las hipótesis de nuestra investigación.

En este capítulo veremos cuáles han sido los resultados de evaluación de las propiedades métricas de Pulse, obtenidos con la aplicación de estas técnicas.

### **4.1. LA ESTIMACIÓN DE ADECUACIÓN MÉTRICA DEL ÍNDICE PULSE.**

El índice Pulse ha sido medido mediante 4 ítems o atributos que son la respuesta a 4 preguntas sobre los sentimientos globales del entrevistado asociados con la reputación de la empresa evaluada como hemos visto en el cuestionario aplicado en la investigación de base. Las categorías de respuesta forman una escala numérica de siete posiciones con anclas semánticas extremas sobre el grado de conformidad que mantiene con cada declaración, siendo “1” *completamente en desacuerdo* y “7” *completamente de acuerdo* (ver cuestionario Pulse).

Para las estimaciones de adecuación métrica se ha aplicado en análisis factorial confirmatorio con un único factor, el índice Pulse en forma reflexiva.

Utilizamos el programa de análisis, AMOS 20.

Comprobamos la adecuación métrica del Pulse en cada país, con todos los datos disponibles.

#### **4.1.1. ESTIMACIÓN DEL PULSE PARA ESPAÑA: COMPROBACIÓN DE LA ADECUACIÓN MÉTRICA DEL ÍNDICE.**

Los datos para España parecen cuestionar los supuestos de normalidad multivariada. El test de Mardia de Kurtosis presenta un valor de 125,911, con un ratio coeficiente del  $129,094 > 1,96$ , altamente significativo. Los efectos de violación de la kurtosis



multivariada son fundamentalmente dos y afectan sobrestimando el valor de la función de ajuste “ $\chi^2$ ”, lo que llevaría a rechazar impropriamente la hipótesis nula del ajuste de matrices de covarianzas poblacionales, además de sesgar los errores estándar de los estadísticos estimados. En estos casos se recomienda aplicar el procedimiento *Bootstrap* de Bollen-Stein, disponible en AMOS. (Carreras y González, 2014). En este caso comprobaremos el sesgo provocado por la falta de normalidad multivariante.

El modelo factorial confirmatorio del modelo PULSE no se acomoda exactamente a los datos de España. Para evaluar el ajuste global de datos se han utilizados los criterios Richard Bagozzi y Youhae Yi (2012): CFI  $\geq 0,95$ ; NFI  $\geq 0,92$  y RMSEA  $\leq 0,06$ . Carmines y McIver (1981) proponen el “ $\chi^2/g.l$ ”  $< 3$ .

### Estimación del índice PULSE en la BBDD de España, 2012

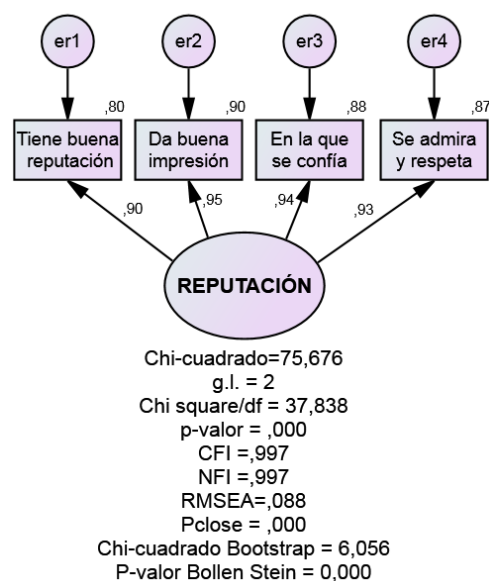


Figura nº11. Estimación de los datos España.

Para España, el modelo PULSE no se ajusta suficientemente bien a los datos. De los cuatro índices globales, sólo el índice de ajuste comparativo CFI y el normado NFI parece superar los umbrales de aceptación, sin embargo, los índices que evalúan mejor la discrepancia entre matrices, “ $\chi^2/g.l. = 37,838$ ” y el RMSEA = 0,088 se alejan demasiado de los niveles aceptables. El análisis de la matriz residual y los índices de

modificación señalan que las variables implicadas en la falta de ajuste son las dos primeras, “*tiene buena reputación global*” y “*da buena impresión*”.

Hemos calculado una covarianza entre los residuales de ambas variables y de esta forma, el modelo sí se ajusta plenamente a los datos.

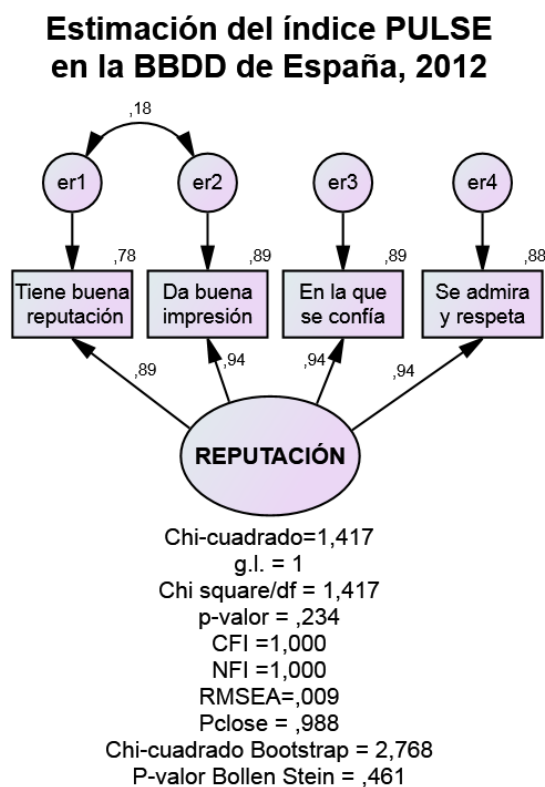


Figura nº 12. Modelo base para los datos de España.

De esta forma, estableciendo una covarianza entre los residuales de ambas variables, el modelo se ajusta totalmente a los datos. Todos los índices superan los umbrales de aceptación. Aunque la probabilidad exacta de rechazar la hipótesis nula de la igualdad de covarianzas poblacionales observadas e implicadas por el modelo sale no significativa,  $p\text{-valor} = ,234$ , el P-valor de la corrección Bollen y Stein vuelve a confirmar este supuesto  $p\text{-valor}_{B-S} = 0,461$ .

Los sesgos que provoca la falta de normalidad en las estimaciones de los coeficientes estandarizados beta y en las varianzas de los residuales son prácticamente despreciables como podemos comprobar en la tabla siguiente:

Variable	Beta	Beta <sub>Bootstrap</sub>	Sesgo	S <sup>2</sup> <sub>res</sub>	S <sup>2</sup> <sub>Bootstrap</sub>	Sesgo	p-valor <sub>sesgo</sub>
Tiene buena reputación	0,885	0,885	0,000	2,496	2,497	0,001	0,004
Da buena impresión	0,942	0,942	0,000	0,688	0,689	0	0,002
En la que se confía	0,944	0,944	0,000	0,376	0,377	0	0,002
Se admira y respeta	0,937	0,937	0,000	0,372	0,372	0	0,002
				0,413	0,415	0,002	0,002

Tabla nº 11. Sesgos en las estimaciones de España debidas a la falta de normalidad multivariante .

La corrección que ha sido requerida para que el modelo sea aceptado por los datos se refiere a la correlación entre los errores de las dos primeras variables antes citadas. El hecho de que los residuos de los indicadores estén correlacionados significa que parte de la varianza de dichos indicadores, que debería de ser aleatoria, está explicada por otra fuente sistemática de variación. Dicho de otro modo, el significado de los atributos “*tiene una buena reputación*” y “*es una empresa sobre la que tengo buena impresión*” para el entrevistado en España posiblemente despierte cierto matiz de interpretación semántica diferente que no coincide con la idea global de reputación.

La recomendación que se desprende de este resultado es que se podría someter a futuras investigaciones a testar la traducción del atributo “*es una empresa sobre la que tengo buena impresión*”, (“*Is a company I have a good feeling about*” en el original en inglés).

#### 4.1.2. ESTIMACIÓN DEL PULSE PARA ESTADOS UNIDOS: COMPROBACIÓN DE LA ADECUACIÓN MÉTRICA DEL ÍNDICE.

En el caso de EEUU, los datos también parecen violar los supuestos de normalidad multivariada. El test de Mardia de Kurtosis multivariada presenta un valor de 25,495, con un ratio coeficiente del  $130,325 > 1,96$ , altamente significativo. Hemos controlado los sesgos mediante el procedimiento Bootstrap de Bollen-Stein.

#### Estimación del índice PULSE en la BBDD de Estados Unidos, 2012

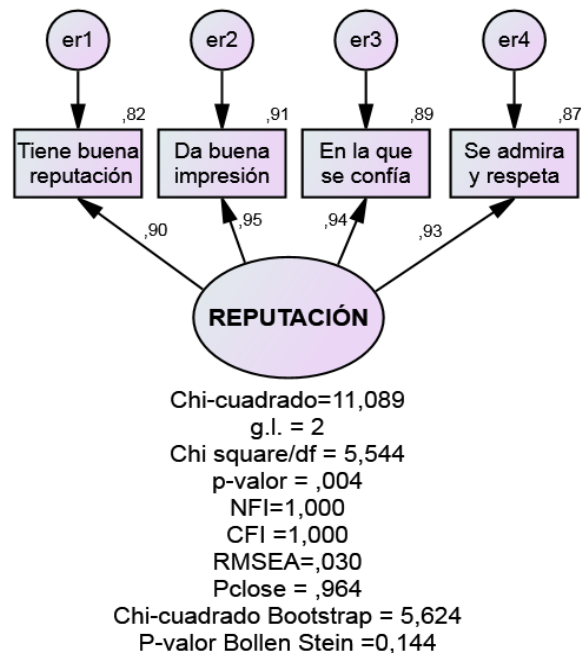


Figura nº 13. Estimación de Pulse en la base de datos de EEUU.

Los datos confirman plenamente el modelo PULSE para Estados Unidos. Aunque la función de ajuste “ $\chi^2_2 = 11,089$ ” sale significativa con un p-valor = 0,004, cuando controlamos la falta de normalidad multivariante, mediante el test Bollen Stein, el p-valor  $B-S = 0,144$ , nos indica que existe un pleno ajuste entre ambas matrices poblacionales de covarianzas, implicada y observada.

El resto de índices globales vuelven a confirmar el ajuste excelente para este país: CFI = 1,00 > de 0,95 y RMSEA = 0,03 < 0,05.

Los sesgos, en las estimaciones de los parámetros, aunque significativos son triviales tanto en las estimaciones de los coeficientes beta como en las varianzas de los errores y en la de la variable latentes PULSE como podemos comprobar en la tabla siguiente.

Variable	Beta	Beta <sub>Bootstrap</sub>	Sesgo	S <sup>2</sup> <sub>res</sub>	S <sup>2</sup> <sub>Bootstrap</sub>	Sesgo	p-valor <sub>sesgo</sub>
Tiene buena reputación	0,885	0,885	0,000	2,496	2,497	0,001	0,004
Da buena impresión	0,942	0,942	0,000	0,688	0,689	0	0,002
En la que se confía	0,944	0,944	0,000	0,376	0,377	0	0,002
Se admira y respeta	0,937	0,937	0,000	0,372	0,372	0	0,002
				0,413	0,415	0,002	0,002

Tabla nº 12. Sesgos en las estimaciones de EEUU debidas a la falta de normalidad multivariante

#### 4.1.3. ESTIMACIÓN DEL PULSE PARA CHINA: COMPROBACIÓN DE LA ADECUACIÓN MÉTRICA DEL ÍNDICE.

La estimación del modelo en China, también parece incumplir los supuestos de normalidad multivariante. El test de Mardia produce un estadístico 9,495, con un ratio de coeficiente  $20,920 > 1,96$ . También aquí verificamos que los sesgos en la estimación de los impactos son también mínimos en el contexto chino. Únicamente, la carga estándar del indicador “buena reputación” aparece con un sesgo del -0,001, prácticamente trivial. En el caso de las varianzas de los residuos, aparece con mínimo sesgo, -0,001 “buena reputación” y -0,002 “es una empresa en la que confío”.

## Estimación del índice PULSE en la BBDD de China, 2012

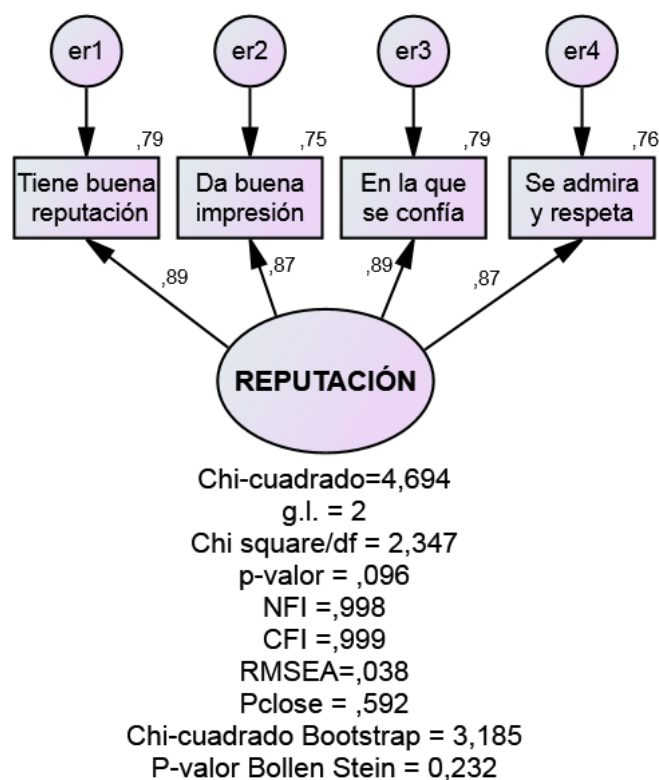


Figura nº 14. Estimación de Pulse para China.

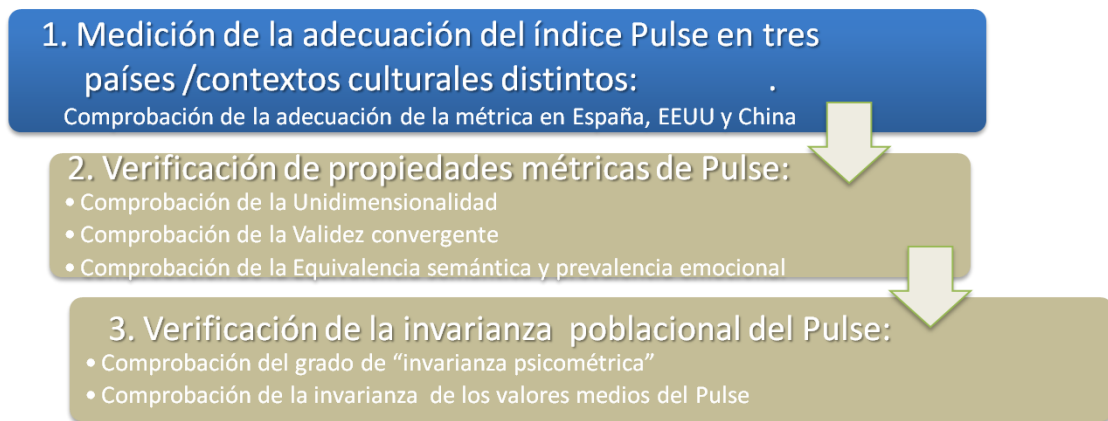
Variable	Beta			S <sup>2</sup> <sub>res.</sub>	S <sup>2</sup> <sub>Bootstrap</sub>	Sesgo	p-valor <sub>sesgo</sub>
	Beta	Boot	Sesgo				
Tiene buena reputación	0,888	0,887	-0,001	0,810	0,809	-0,001	0,004
Da buena impresión	0,868	0,868	0,000	0,218	0,218	0	0,001
En la que se confía	0,891	0,891	0,000	0,276	0,274	-0,002	0,001
Se admira y respeta	0,873	0,873	0,000	0,240	0,24	0	0,002
				0,312	0,312	0	0,001

Tabla nº 13. Sesgos en las estimaciones de China debidas a la falta de normalidad multivariante.

Debido a la falta de normalidad, deberíamos haber seleccionado teóricamente el método de distribución libre asintótica ADF de Browne ya que no requiere supuestos distribucionales. Sin embargo este método produce sólo estimaciones robustas por

encima de los 2500 casos (Ullman, 2006). En los análisis globales, las muestras de España y EEUU serían adecuadas por su tamaño pero en los cálculos de análisis de invarianza por subgrupos de sexo, edad y nivel de estudios, , que constituye la prueba principal en esta tesis, los tamaños de muestra son mucho menores. Por esa razón hemos seleccionado el método de máximo rigor estadístico, que produce las estimaciones más eficientes (Bollen, 1989), corregido con el método *Bootstrap* para tratar las desviaciones no-normales.

A continuación ofrecemos una representación gráfica de los resultados que hemos recogido en este apartado.



### Relación de hipótesis

El Pulse es un índice con adecuación métrica en tres países

H0

Gráfico nº 10 Formulación de la medición de la adecuación métrica del índice Pulse en tres países.

## 4.2. VERIFICACIÓN DE LAS PROPIEDADES MÉTRICAS DEL ÍNDICE PULSE.

Recordemos que las propiedades métricas/psicométricas básicas tal y como han sido definidas en el capítulo cuatro del marco analítico matemático y estadístico y que forman parte de las Hipótesis H1.1, H1.2. y H1.3 de esta investigación.

Veremos en las páginas siguientes de este epígrafe si conseguimos demostrar o refutar estas 3 hipótesis cuando hemos puesto a prueba las propiedades psicométricas del Índice Pulse.

- (1) Uni-dimensionalidad del índice Pulse.
- (2) Convergencia o capacidad de los 4 atributos/preguntas de Pulse para representar adecuadamente la reputación global.
- (3) El tipo de relación que mantienen los 4 atributos/preguntas con el constructo que representan.

### 4.2.1. LA UNI-DIMENSIONALIDAD DEL ÍNDICE PULSE.

Para verificar la uni-dimensionalidad del índice Pulse se ha procedido a estimar el auto-valor más alto en un análisis de componentes principales, con el programa de software *SPSS*, aplicado a las tres bases de datos.

Análisis de componentes principales aplicado a los ítems del PULSE, en los tres contextos			
	España	Estados Unidos	China
<b>KMO</b>	0,874	0,879	0,864
<b>significación del Test de Barlett</b>	0,000000	0,000000	0,000000
<b>Auto valor más alto</b>	3,59	3,61	3,32
<b>% de varianza explicada por un factor</b>	0,898	0,903	0,831

Tabla nº 14. Resultados de los tres análisis de componentes principales aplicados a las tres BBDD.

Los resultados de la tabla siguiente parecen confirmar la uni-dimensionalidad del índice PULSE en los tres contextos, español, estadounidense y chino.



La prueba de Kaiser Mayer Olkin (KMO) con valores cercanos al 0,90, indica que los cuatro indicadores comparten suficiente variación como para extraer factores comunes de covariación. El test de esfericidad de Barlett rechaza la hipótesis de que los ítems estén incorrelacionados al nivel poblacional.

El auto-valor determina el grado de uni-dimensionalidad que ha encontrado la covariación de los cuatro ítems. En España y EEUU, una única dimensión es capaz de reproducir el 90% de la variación total de la suma de los cuatro indicadores ( $3,59/4 = 0,898$  y  $3,61/4 = 0,903$ ) y el caso Chino, un poco menos, el 83,1% ( $3,32/4=0.831$ ). En consecuencia parece que se aporta evidencia empírica a favor de la uni-dimensionalidad del índice, dicho de otro modo, en la combinación lineal de los ítems en un solo índice subyace prácticamente una única dimensión (la reputación). Por tanto los cálculos nos llevarían a la aceptación de la hipótesis H1.1 de uni-dimensionalidad del índice Pulse.

### 4.3. LA VALIDEZ CONVERGENTE DEL ÍNDICE PULSE Y SU FIABILIDAD.

El índice de fiabilidad del compuesto lineal de los 4 ítems. Mediante el programa SPSS, se ha estimado el valor “Alfa de Cronbach” para comprobar el nivel de consistencia interna del índice en los tres países.

Ítems	España		Estados Unidos		China	
	$\alpha$ Cronbach	$\alpha$ si se elimina	$\alpha$ Cronbach	$\alpha$ si se elimina	$\alpha$ Cronbach	$\alpha$ si se elimina
Tiene buena reputación		0,959		0,960		0,909
Da buena impresión	<b>0,962</b>	0,944	<b>0,964</b>	0,948	<b>0,931</b>	0,913
En la que se confía		0,948		0,950		0,907
Se admira y respeta		0,950		0,953		0,913

Tabla nº 15. Índice de fiabilidad del compuesto lineal de los 4 ítems.

En los tres países, el Pulse ha demostrado tener una alta fiabilidad, por encima del umbral de aceptación del 0,7 (De Vellis 2003). Además todos los ítems parecen necesarios para mantener el alto nivel obtenido. En cada país, la columna de la derecha ha reproducido el valor del alfa si el ítem correspondiente fuera eliminado. En todos los casos, se verifica una pérdida de fiabilidad, que significa que los 4 son necesarios para medir la variable latente de reputación global.

En los tres países, la combinación lineal equiponderada de los 4 atributos del Pulse constituye una medida fiable y unidimensional. Dado que la combinación lineal no es equiponderada, Richard Bagozzi y Youhae Yi proponen comprobar esta medida con el índice de fiabilidad del compuesto de Dillon-Golstein (Bagozzi y Yi, 2012). La validez convergente quedaría corroborada con el índice “AVE” o promedio de varianza extractada por el factor (Fornell y Larcker, 1981) que representa la cantidad de varianza que en promedio explica el Pulse en sus 4 atributos.

Ítems del Pulse	Dillon-Goldstein			AVE			Fiabilidad del ítem		
	España	EEU	China	España	EEU	China	España	EEU	China
Tiene buena reputación	0,96	0,96	0,93	0,86	0,87	0,77	0,783	0,817	0,787
Da buena impresión							0,887	0,908	0,753
En la que se confía							0,891	0,889	0,794
Se admira y respeta							0,878	0,870	0,762

Tabla nº 16. Índice de fiabilidad del compuesto de Dillon-Golstein.

Tal y como comentamos en el apartado metodológico, también conviene comprobar si cada ítem en particular aporta suficiente información a la variable latente Pulse, mediante el índice de fiabilidad del ítem.

Debido a la no normalidad de los datos, los coeficientes estandarizados han sido estimados mediante procedimiento *bootstrap*. Los cálculos están incluidos en el anexo con los cálculos estadísticos.

Según se desprende de la tabla, la fiabilidad del compuesto en los tres países, supera ampliamente los umbrales de aceptación ( $> 0,6$ ), los promedios de varianza extractada por el Pulse, están por encima del 0,5 requerido y la aportación de cada ítem es alta, por encima del 0,582. Por tanto los cálculos matemáticos y estadísticos apuntan a la aceptación de la hipótesis H1.2 de convergencia de los indicadores.

---

<sup>82</sup>Para justificar os umbrales de aceptación véase el apartado “técnicas para comprobar las hipótesis de medida”

#### 4.4. EL TIPO DE RELACIÓN ENTRE LOS CUATRO ÍTEMS Y EL PULSE.

Como ya se ha comentado, en el capítulo cuatro del marco analítico estadístico y matemático anterior, hay tres modos en que los cuatro ítems/variables se relacionan con el Pulse, cuyo conocimiento permite interpretar mejor el significado del constructo que representan.

En la relación co-genérica, el peso de los ítems es diferente, indicando que algunos ítems representan mejor la idea global. En la relación tau-equivalente, se verifica la equivalencia de los pesos, lo que nivela la contribución equivalente de los contenidos aportados por los cuatro ítems. En la relación “paralela” los indicadores no sólo contribuirían del mismo modo al contenido del Pulse sino que sus residuales, la variabilidad no explicada, sería la misma.

El problema con el que nos enfrentamos radica en el tamaño excesivo de muestra, sobre todo en España con 4766 casos y EEUU con 5017. En este sentido, los test resultan demasiado sensibles al tamaño de muestra inflando los errores tipo I (Ullman 2006) lo que llevaría impropriamente a rechazar el modelo tau-equivalente.

Robert MacCallum, Michael Browne y Hazuki Sugawara, han examinado la necesidad de controlar el tamaño de muestra en función del poder de discriminación. El poder de discriminación ha sido definido como la probabilidad de rechazar un modelo que es falso. Los autores proponen que para un modelo con 4 grados de libertad y un poder de discriminación del 80%, el tamaño de muestra requerido debería de ser de 1807 casos (MacCallum et al., 1996, 144).

Para cumplir con este requisito, hemos procedido a extraer sub-muestras aleatorias más reducidas para España y Estados Unidos (en China no es necesario puesto que la muestra es inferior a los 1000 casos).

El análisis practicado con una sub-muestra aleatoria reducida a 1807 casos en cada país, ha rendido las siguientes estimaciones para España y para Estados Unidos.

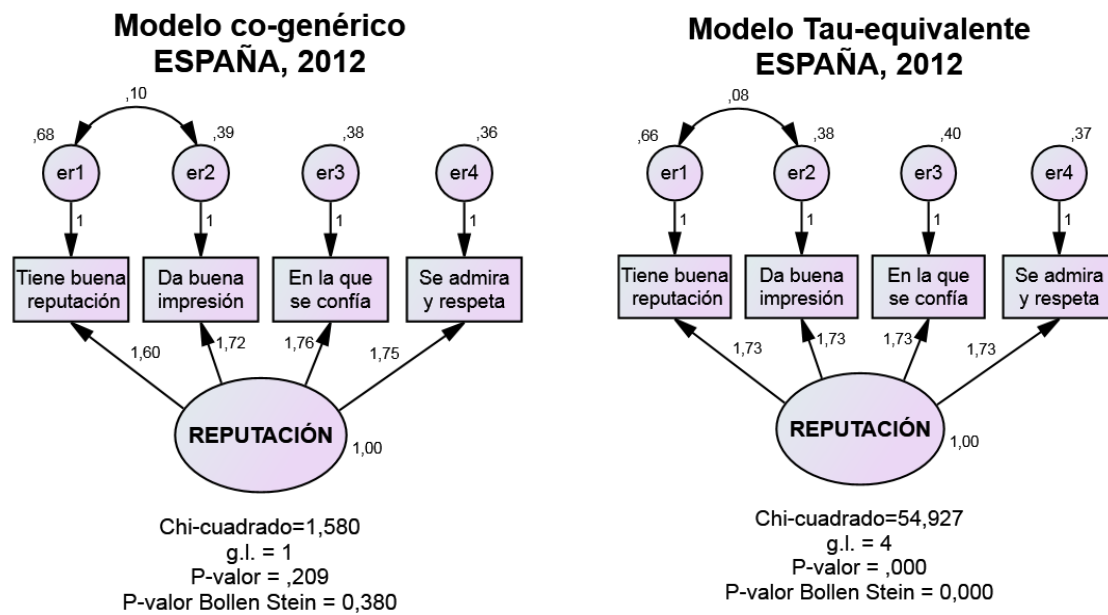


Figura nº 17. Modelo co-genérico y modelo Tau-equivalente España

El modelo tau-equivalente ha sido rechazado por los datos. El  $\chi^2_4 = 54,927$  sale no significativo (p-valor = 0.000), pero para evitar concluir erróneamente debido a la no normalidad multivariante de los datos, hemos procesado el p-valor de Bollen Stein, con este análisis complementario se vuelve a confirmar la hipótesis considerada en esta investigación.

Por otra parte, el ajuste del modelo co-genérico es excelente, El  $\chi^2_1 = 1,580$  presenta una probabilidad exacta para probar la igualdad de las matrices de covarianzas observada de implicada, del 0,209, muy por encima del umbral de rechazo del 0,05.

El p-valor obtenido mediante Bootstrap, estimado por Bollen Stein vuelve a confirmar la bondad del ajuste (p-valor  $B-S = 0.380 > 0,05$ ).

En la tabla siguiente, hemos procedido a estimar por el método *Bootstrap*, los intervalos de confianza de las cargas estandarizadas, mediante el método del percentil.

España			
Ítems	Beta <sub>Bootstrap</sub>	Lim inf	Lím sup
Tiene buena reputación	0,889	0,87	0,90
Da buena impresión	0,94	0,93	0,95
En la que se confía	0,944	0,93	0,95
Se admira y respeta	0,946	0,94	0,95

Tabla nº 17. Estimaciones bootstrap de los intervalos de confianza del modelo co-genérico para España.

En los datos para España, el indicador “*tiene buena reputación*” contribuye sistemáticamente menos que el resto de indicadores, cuyos intervalos de confianza se solapan. En la formación del juicio global de reputación, los sentimientos de “*buena impresión*”, “*confianza*” y “*admiración*” pesan significativamente más que la idea más genérica y cognitiva de “reputación”. Se confirma para España que el índice Pulse tiene una carga emocional importante.

Los resultados de EEUU son similares a los de España.

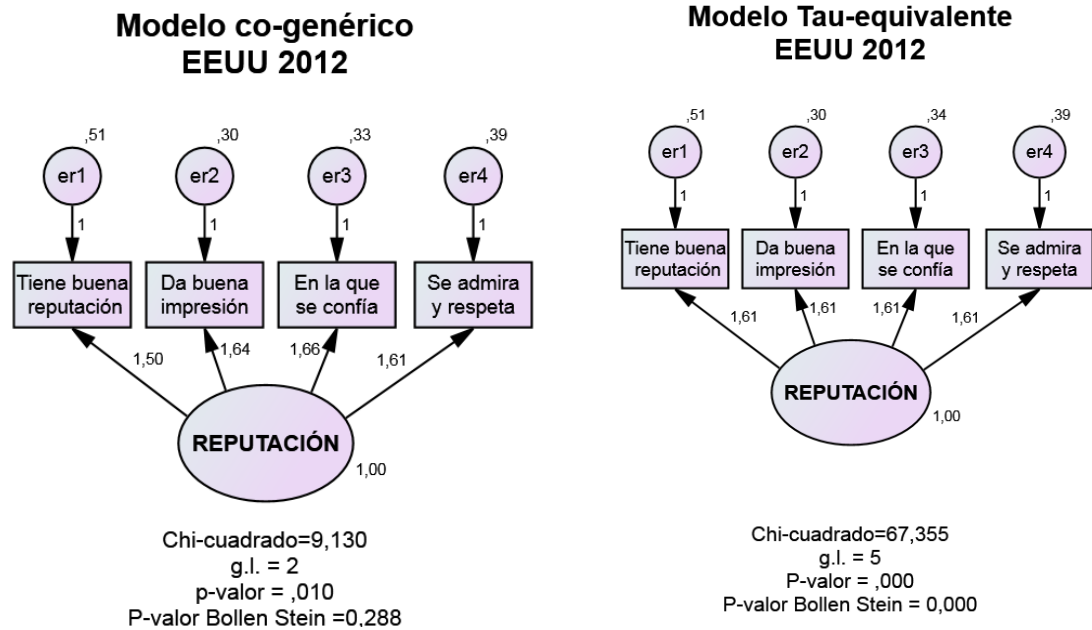


Figura nº 18. Comparación de los modelos co-genérico y tau-equivalente para EEUU.

El modelo co-genérico que deja libres las estimaciones de las cargas de los ítems, se ajusta razonablemente bien a los datos. Aunque el  $\chi^2_2 = 9,130$  sale significativo (p-valor = 0,010), la corrección de Bollen y Stein para eliminar los efectos de la no normalidad confirman que el modelo es empírico (p-valor  $B-S = 0,288 > 0,05$ ). Por el contrario, el modelo que propone la igualdad de cargas (tau-equivalente), ha sido completamente rechazado por los datos:  $\chi^2_5 = 67,355$  (p-valor = 0,000). Siendo el p-valor de Bollen y Stein también altamente significativo (0,000).

El análisis de los intervalos de confianza de los pesos estandarizados vuelve a ofrecernos una situación similar en EEUU. Los tres ítems más afectivos o emocionales, como la "buena impresión, la confianza y la admiración", parecen tener la misma contribución a la formación del sentimiento global de reputación, cuyos intervalos de confianza se solapan. Por el contrario el límite superior del intervalo de confianza correspondiente al ítem más cognitivo o racional de "buena reputación" está por debajo de los límites inferiores de los otros tres, señalando que la intensidad con la que se relaciona con el sentimiento global de reputación es significativamente menor.

Estados Unidos			
Ítems	Beta <sub>Bootstrap</sub>	Lim inf	Lím sup
Tiene buena reputación	0,903	0,89	0,91
Da buena impresión	0,949	0,94	0,96
En la que se confía	0,945	0,94	0,95
Se admira y respeta	0,933	0,92	0,94

Tabla nº 18. Estimaciones *bootstrap* intervalos de confianza para EEUU.

### Resultados para China.

En el caso de la población china de internautas, encontramos una pauta similar, aunque no ha sido plenamente confirmada por los datos.

El modelo co-genérico se ajusta perfectamente a los datos:  $\chi^2_2 = 4,694$ , con una probabilidad exacta no significativa (p-valor = 0,096) por encima del umbral de rechazo. La prueba Bollen Stein confirma su adecuación (p-valor<sub>B-S</sub> = 0,232).

Por el contrario, el modelo tau-equivalente ha sido empíricamente rechazado, el  $\chi^2_5 = 22,801$  (p-valor = 0,000 < 0,05 y eliminando el efecto de la sobrestimación de la función del ajuste mediante el cálculo de p-valor de Bollen y Stein, tampoco parece soportar el umbral de aceptación (p-valor<sub>R-S</sub> = 0,004 < 0,05).

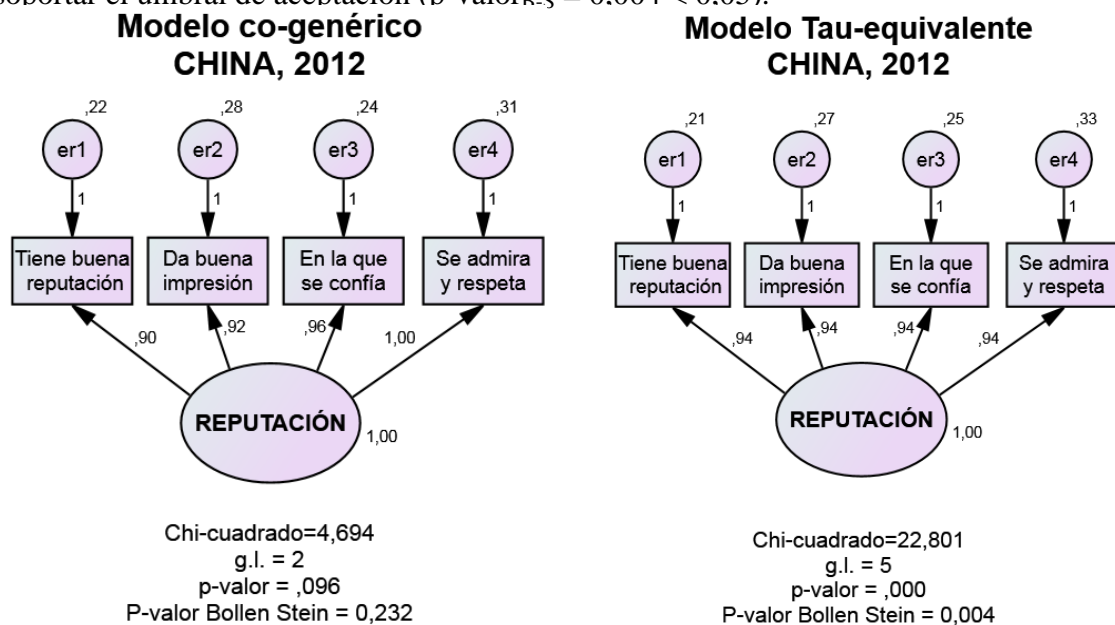


Figura nº 19. Comparación de los modelos co-genérico y tau-equivalente para China.



Sin embargo, el análisis de los intervalos de confianza de los estimadores de las cargas no llega a reflejar las diferencias detectadas en el análisis global. En el caso chino, los intervalos de confianza de los cuatro ítems se solapan entre sí.

China			
Ítems	Beta <sub>Bootstrap</sub>	Lim inf	Lím sup
Tiene buena reputación	0,888	0,87	0,91
Da buena impresión	0,868	0,84	0,89
En la que se confía	0,891	0,87	0,91
Se admira y respeta	0,873	0,85	0,89

Tabla nº 19. Estimaciones *bootstrap* de los intervalos de confianza del modelo co-genérico para China.

Los resultados presentan cierta ambigüedad y no debemos concluir que en el caso de la población de internautas en China, el primer atributo contribuya con una intensidad menor a la formación del resto de ítems. Existen varias razones por las cuales podemos concluir que los resultados en China, sí corroboran también los alcanzados en España y EEUU.

En primer lugar y la evidencia más sólida radica en que el modelo tau-equivalente ha sido rechazado por los datos, mientras que el co-genérico ha sido confirmado.

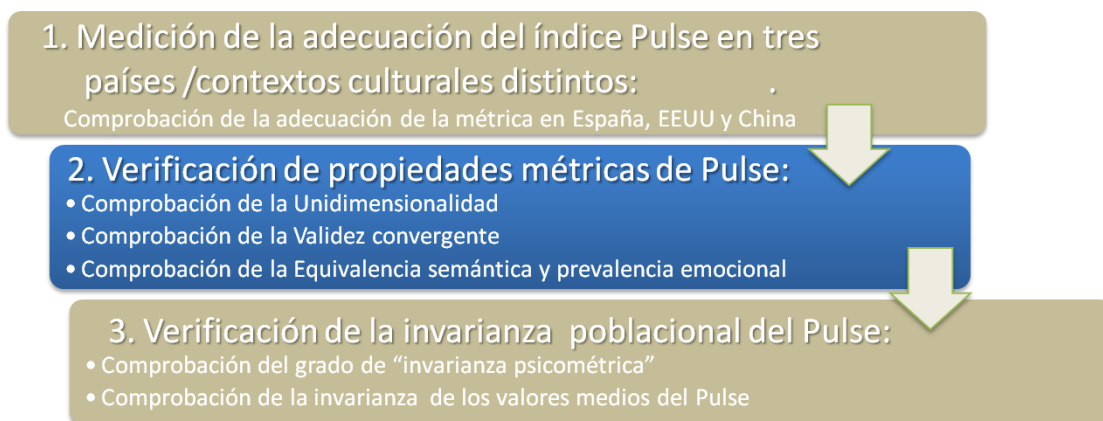
En segundo lugar, el tamaño de muestra en las BBDD (932 casos) no alcanza, en términos de Robert MacCallum y colegas la cantidad suficiente para garantizar un poder de discriminación del 0,80. Desde la perspectiva de la falta de poder de discriminación, el ajuste detectado en el modelo co-genérico no queda suficientemente fundamentado porque la hipótesis nula parte del principio de que el modelo co-genérico es verdadero, pero si fuera falso en la realidad, la prueba no tendría suficiente fuerza para detectar que era falso (Carreras y González, 2013).

Por otro lado, en el caso de China, sólo disponemos de un solo sector de actividad, el bancario, no disponemos de la misma variedad que en los otros dos países, sin embargo, la razón de mayor peso, a nuestro juicio, es que la prueba de los intervalos de confianza no alcanza la sensibilidad de otras pruebas como el cálculo de chi-cuadrado escalado de Bentler, no disponible en el programa de software utilizado

AMOS (Ullman, 2006). El modelo en China sería co-genérico, pero no tenemos seguridad suficiente para afirmar que los tres indicadores emocionales pesan más que el primero más cognitivo ("*Tiene buena reputación*").

Aunque la relación co-genérica ha sido plenamente corroborada en los tres países, el mayor peso relativo de los tres ítems emocionales se ha probado en España y EEUU, pero no en China. Por tanto los cálculos matemáticos y estadísticos parecen verificar la hipótesis la hipótesis H1.3 de relación co-genérica entre los indicadores en distintos países.

En los gráficos siguientes ofrecemos una representación visual sintética de los resultados que hemos presentado en este apartado.



Relación de hipótesis	
El Pulse es un índice unidimensional	H1.1
El Pulse es una métrica con validez convergente	H1.2
El Pulse tiene equivalencia semántica y mayor peso emocional	H1.3

Gráfico nº 11. Formulación de la verificación de las propiedades métricas del índice Pulse.

#### **4.5. EL GRADO DE INVARIANZA POBLACIONAL DEL ÍNDICE PULSE.**

Como ya hemos comentado en el capítulo cuatro de marco analítico matemático y estadístico, el análisis de invarianza poblacional (hipótesis principal de la investigación) se calcula utilizando tres pruebas: los índices globales del ajuste, los test de anidamiento de la secuencia de hipótesis concatenadas y las diferencias entre los índices incrementales “ $\Delta CFI$ ” y “ $\Delta RMSEA$ ”.

Sin embargo, la comprobación de la falta de normalidad multivariante en los datos nos obliga a dar mayor prioridad a la prueba del ajuste global corregido por Bollen y Stein y a la prueba de las diferencias entre los índices CFI y RMSEA, que a la pérdida de ajuste basado en la diferencia de los  $\chi^2_{\text{dif}}$ .

##### **4.5.1. LA IDENTIFICACIÓN DEL ÍTEM DE REFERENCIA EN LOS TRES PAÍSES.**

El análisis de invarianza poblacional requiere que se identifique previamente el ítem que puede ser fijado en condiciones de igualdad para los distintos grupos poblacionales (sexo, edad y nivel de estudios), de modo que las estimaciones del resto de cargas no resulten sesgadas.

Como comentamos en el capítulo del análisis de invarianza, el método utilizado para hallar el ítem de referencia trabaja restringiendo a la igualdad todos los ítems a la vez y fijando la varianza de la latente a la unidad. Si esta primera prueba queda rechazada, se procede a liberar sucesivamente las cargas de los diferentes ítems hasta alcanzar el último ítem fijado que representa el ítem con mayor probabilidad de presentar igualdad de cargas y por tanto sesgar menos el resto de los coeficientes estimados.

Los cálculos se han efectuado por cada una de las variables que originan los grupos de comparación, sexo, edad, nivel de estudios, sector de actividad e invarianza conjunta.

En España, el modelo base de la invarianza de cargas ha sido rechazado en todas las variables analizadas, sexo, edad y nivel de estudios, por lo que no conviene seleccionar al azar un ítem determinado como el ítem de referencia.

### ESPAÑA SEXO

Modelos	$\chi^2$	g.l.	p-valor	p-valor <sub>B-S</sub>		
Base: invarianza cargas y Latente a 1	20,3	6	0,002	0,043	Libera	Fija
Libera carga del 1	9,8	5	0,081	0,249	ítem 1 (4,08)	ítems 1,2 y 4
Libera carga del 1 y del 3	8,6	4	0,072	0,269	ítem 3 (0,601)	ítems 2 y 4
Libera carga del 1, 3 y 4	7,3	3	0,063	0,284	ítem 4 (0,520)	ítem 2
<b>Ítem invariante</b>	<b>Ítem 4</b>					

### ESPAÑA EDAD

Modelos	$\chi^2$	g.l.	p-valor	p-valor <sub>B-S</sub>		
Base: invarianza cargas y Latente a 1	9,2	6	0,16	0,37	Libera	Fija
Libera carga del 2	4,1	5	0,528	0,722	ítem 2(3,204)	ítems 1, 2 y 3
Libera carga del 2 y del 3	3,4	4	0,493	0,721	ítem 1 (0,943)	ítems 2 y 4
Libera carga del 2, 3 y 4	3,0	3	0,386	0,663	ítem 4 (0,890)	ítem 2
<b>Ítem invariante</b>	<b>Ítem 4</b>					

### ESPAÑA NIVEL DE ESTUDIOS

Modelos	$\chi^2$	g.l.	p-valor	p-valor <sub>B-S</sub>		
Base: invarianza cargas y Latente a 1	12,7	6	0,048	0,185	Libera	Fija
Libera carga del 2	8,4	5	0,136	0,352	ítem 2(2,100)	ítems 1, 3 y 4
Libera carga del 2 y del 3	8,4	4	0,078	0,303	ítem 3 (0,543)	ítems 2 y 4
Libera carga del 2, 3 y 1	7,0	3	0,073	0,321	ítem 1 (0,541)	ítem 4
<b>Ítem invariante</b>	<b>Ítem 4</b>					

### ESPAÑA SECTOR

Modelos	$\chi^2$	g.l.	p-valor	p-valor <sub>B-S</sub>		
Base: invarianza cargas y Latente a 1	8,5	6	0,202	0,418	Libera	Fija
Libera carga del 1	4,5	5	0,484	0,658	ítem 1(1,041)	ítems 2, 3 y 4
Libera carga del 1 y del 3	3,7	4	0,455	0,666	ítem 3 (0,200)	ítems 2 y 4
Libera carga del 2, 3 y 1	2,3	3	0,521	0,739	ítem 2 (0,259)	ítem 4
<b>Ítem invariante</b>	<b>Ítem 4</b>					

### ESPAÑA SEXO EDAD Y ESTUDIOS

Modelos	$\chi^2$	g.l.	p-valor	p-valor <sub>B-S</sub>		
Base: invarianza cargas y Latente a 1	69,5	36	0,001	0,035	Libera	Fija
Libera carga del 1	53,3	29	0,004	0,08	ítem 1(3,717)	ítems 2, 3 y 4
Libera carga del 1 y del 3	47,1	22	0,001	0,075	ítem 3 (0,910)	ítems 2 y 4
<b>Ítem invariante</b>	<b>Ítem 4</b>					

Tabla nº 20. Pruebas para la identificación de los ítems de referencia en España (entre paréntesis el índice de modificación).

En España, los ítems que resisten mejor las pruebas en todas las variables son el ítem 2 “buena impresión” y el 4 “admiración y respeto”. Debido a que, en los datos españoles se ha estimado una covarianza en los residuos de los ítems 1 y 2, hemos seleccionado el ítem 4 “admiración y respeto”, como el mejor candidato para restringir a la unidad los casos en los distintos análisis. De esta forma, la covarianza no afectará al posible sesgo por variable de referencia.

Para el caso de España, el ítem seleccionado como referencia será el cuatro: “admiración y respeto”.

En el caso de EEUU, en la página siguiente, tampoco conviene seleccionar al azar cualquier ítem como ítem de referencia. En todas las estimaciones realizadas se ha verificado que la igualdad de cargas entre grupos no es empíricamente plausible ( $p\text{-valor}_{BS} = 0,002 > 0,05$ ).

En la variable sexo y en la variable conjunta de grupos homogéneos por sexo, edad y estudios, en realidad ningún ítem consigue ajustar un buen modelo, mostrando p-valores de Bollen Stein significativos, ( $p\text{-valor}_{B-S} < 0,05$ ). En estos casos, el ítem fijado que presenta mejor los mejores ajustes es el ítem 3. Este mismo ítem es el que sale aceptable en el resto de variables.

En consecuencia, en el caso de EEUU, el ítem de referencia será el ítem 3: “confianza”.

En el caso de China parece confirmarse el modelo de invarianza de cargas, con lo que en principio, cualquier indicador hubiera servido para fijar la unidad en los grupos considerados.

Dado que el índice que mejor soporta las pruebas de liberación de cargas es el número cuatro “admiración y respeto”, seleccionamos éste como indicador de referencia.

En China, el índice de referencia retenido será el cuatro: “admiración y respeto”.

**EEUU SEXO**

Modelos	$\chi^2$	g.l.	p-valor	p-valor <sub>B-S</sub>		
Base: invarianza cargas y Latente a 1	59,4	8	0,000	<b>0,002</b>	Libera	Fija
Libera carga del 4	51,9	7	0,000	<b>0,004</b>	ítem 4 (3,989)	ítems 1,2 y 3
Libera carga del 4 y 1	48,7	6	0,000	<b>0,006</b>	ítem 1 (1,435)	ítems 2 y 4
Libera carga del 4, 1 y 2	47,6	5	0,000	<b>0,007</b>	ítem 4 (0,520)	ítem 2
<b>Ítem invariante</b>	<b>Ítem 3</b>					

**EEUU EDAD**

Modelos	$\chi^2$	g.l.	p-valor	p-valor <sub>B-S</sub>		
Base: invarianza cargas y Latente a 1	25,9	8	0,001	<b>0,098</b>	Libera	Fija
Libera carga del 2	24,3	7	0,001	<b>0,112</b>	ítem 2(0,953)	ítems 1, 3 y 4
Libera carga del 2 y 1	23,3	6	0,001	<b>0,105</b>	ítem 1 (0,594)	ítems 3 y 4
Libera carga del 2, 1 y 4	22,3	5	0,000	<b>0,106</b>	ítem 4 (0,524)	ítem 3
<b>Ítem invariante</b>	<b>Ítem 3</b>					

**EEUU NIVEL DE ESTUDIOS**

Modelos	$\chi^2$	g.l.	p-valor	p-valor <sub>B-S</sub>		
Base: invarianza cargas y Latente a 1	30,9	8	0,000	<b>0,052</b>	Libera	Fija
Libera carga del 2	28,9	7	0,000	<b>0,058</b>	ítem 2(1,665)	ítems 1, 3 y 4
Libera carga del 2 y del 1	26,7	6	0,000	<b>0,062</b>	ítem 1 (1,726)	ítems 3 y 4
Libera carga del 2, 1 y 4	23,9	5	0,000	<b>0,084</b>	ítem 1 (1,833)	ítem 3
<b>Ítem invariante</b>	<b>Ítem 3</b>					

**EEUU SECTOR**

Modelos	$\chi^2$	g.l.	p-valor	p-valor <sub>B-S</sub>		
Base: invarianza cargas y Latente a 1	23,3	8	0,003	<b>0,123</b>	Libera	Fija
Libera carga del 2	18,2	7	0,011	<b>0,206</b>	ítem 2(1,408)	ítems 1, 3 y 4
Libera carga del 2 y 3	16,0	6	0,014	<b>0,241</b>	ítem 3 (0,479)	ítems 1 y 4
Libera carga del 2, 3 y 4	14,0	5	0,016	<b>0,261</b>	ítem 4 (0,318)	ítem 1
<b>Ítem invariante</b>	<b>Ítem 3</b>					

**EEUU SEXO EDAD Y ESTUDIOS**

Modelos	$\chi^2$	g.l.	p-valor	p-valor <sub>B-S</sub>		
Base: invarianza cargas y Latente a 1	132,7	44	0,000	<b>0,003</b>	Libera	Fija
Libera carga del 4	122,5	37	0,000	<b>0,002</b>	ítem4(0,973)	ítems 1, 2 y 3
Libera carga del 4 y 1	106,3	30	0,000	<b>0,005</b>	ítem 1(0,423)	ítems 2 y 3
Libera carga del 4, 3 y 2	90,4	23	0,000	<b>0,015</b>	ítem 2(1,82)	ítem 3
<b>Ítem invariante</b>	<b>Ítem 3</b>					

Tabla nº 21. Pruebas para la identificación de los ítems de referencia en Estados Unidos (entre paréntesis el índice de modificación).

CHINA SEXO						
Modelos	$\chi^2$	g.l.	p-valor	p-valor <sub>B-S</sub>		
Base: invarianza cargas y Latente a 1	17,5	6	0,008	<b>0,059</b>	Libera	Fija
Libera carga del 3	12,7	5	0,026	<b>0,113</b>	ítem 3 (1,703)	ítems 1,2 y 4
Libera 3 y 4	11,7	4	0,02	<b>0,104</b>	ítem 4 (0,315)	ítems 1 y 2
Libera 1 3 y 4	10,3	3	0,016	<b>0,089</b>	ítem 1 (0,399)	ítems 2
<b>Ítem invariante</b>	<b>Ítem 4</b>					

CHINA EDAD						
Modelos	$\chi^2$	g.l.	p-valor	p-valor <sub>B-S</sub>		
Base: invarianza cargas y Latente a 1	14,3	6	0,026	<b>0,139</b>	Libera	Fija
Libera carga del 1	12,8	5	0,026	<b>0,137</b>	ítem 1 (0,479)	ítems 2, 3 y 4
Libera carga del 1 y 2	9,9	4	0,042	<b>0,174</b>	ítem 2 (0,727)	ítems 3 y 4
<b>Ítem invariante</b>	<b>Ítem 4</b>					

Tabla nº 22. Pruebas para la identificación de los ítems de referencia en China (entre paréntesis el índice de Modificación)

Una vez que hemos identificados los ítems de referencia en cada país para fijar las cargas, se procede a realizar la secuencia de diferentes análisis de invarianza.

#### 4.5.2. INVARIANZA DEL PULSE POR SEXO, EN ESPAÑA.

En el caso español, el modelo mantiene la especificidad de una covarianza necesaria entre los residuales de los ítems uno y dos. Por esa razón en este caso el análisis de invarianza mantiene cinco secuencias de hipótesis:

- El modelo configural: Ajusta a la vez el modelo en los dos grupos por sexo, sin restricciones de valor.
- El modelo de cargas iguales: Se restringen a la igualdad las cargas respectivas de ambos grupos:
  - Carga del ítem 1 de varones = carga del ítem 1 de mujeres
  - Carga del ítem 2 de varones = carga del ítem 2 de mujeres
  - Carga del ítem 3 de varones = carga del ítem 3 de mujeres
  - El ítem 4 actúa para España como ítem de referencia, por tanto se fija a 1.

- El modelo de invarianza de la varianza del Pulse. Además de la igualdad de cargas, se fija la igualdad de la varianza por grupos.
- El modelo de invarianza de la covarianza de los errores. Además de las anteriores, el modelo restringe a la igualdad la covarianza entre los residuales de los ítems 1 y 2. Indicando que la parte no explicada por el Pulse en éstos ítems es igual por sexos.
- El modelo de invarianza estricta, que además impone la restricción de las varianzas de los residuales.

En la tabla siguiente ofrecemos los resultados de los ajuste de los modelos:

ESPAÑA: INVARIANZA POR SEXO					
Modelos	$\chi^2$	g.l.	P-valor B-S	CFI	RMSEA
Modelo configural	7,2	2	<b>0,255</b>	1,000	0,023
Igualdad de cargas	19,0	5	<b>0,050</b>	0,999	0,024
Igualdad varianza PULSE	20,3	6	<b>0,043</b>	0,999	0,022
Igualdad de covarianza error	20,4	7	<b>0,005</b>	0,999	0,020
Invarianza estricta	126,1	11	<b>0,000</b>	0,995	0,047

Tabla nº 23. Estimación de los modelos anidados.

Según el p-valor ajustado por Bolen-Stein, el modelo verdadero sería el de igualdad de cargas. El peso de los ítems para formar la opinión global de reputación es el mismo en ambos grupos.

A partir de este modelo valoramos la pérdida de ajuste en la tabla siguiente:

Valoración de la pérdida de ajuste	$\Delta\chi^2$	$\Delta$ g.l.	p-valor	$\Delta$ CFI	$\Delta$ RMSEA
Igualdad de varianza PULSE	1,2	1	<b>0,265</b>	0,001	0,001
Igualdad de covarianza error	1,4	2	<b>0,500</b>	0,001	0,003
invarianza estricta	107,1	6	<b>0,000</b>	0,005	<b>-0,024</b>

Tabla nº 24 . Valoración de la pérdida de ajuste.



Cuando restringimos la variabilidad del Pulse y la covarianza de los errores a la igualdad, la pérdida de ajuste no resulta significativa. Así lo confirman las pérdidas en los índices CFI y RMSEA, cuyas diferencias son menores de 0,01.

La igualdad de los residuales no se sostiene. Presentan una pérdida significativa en la función de ajuste y una diferencia del RMSEA es superior al 0,01.

En España, el índice Pulse ha demostrado ser invariante por sexo. Tanto varones como mujeres entienden del mismo modo la métrica, presentan rangos de variación similares y la covarianza de los residuos es la misma en ambos grupos.

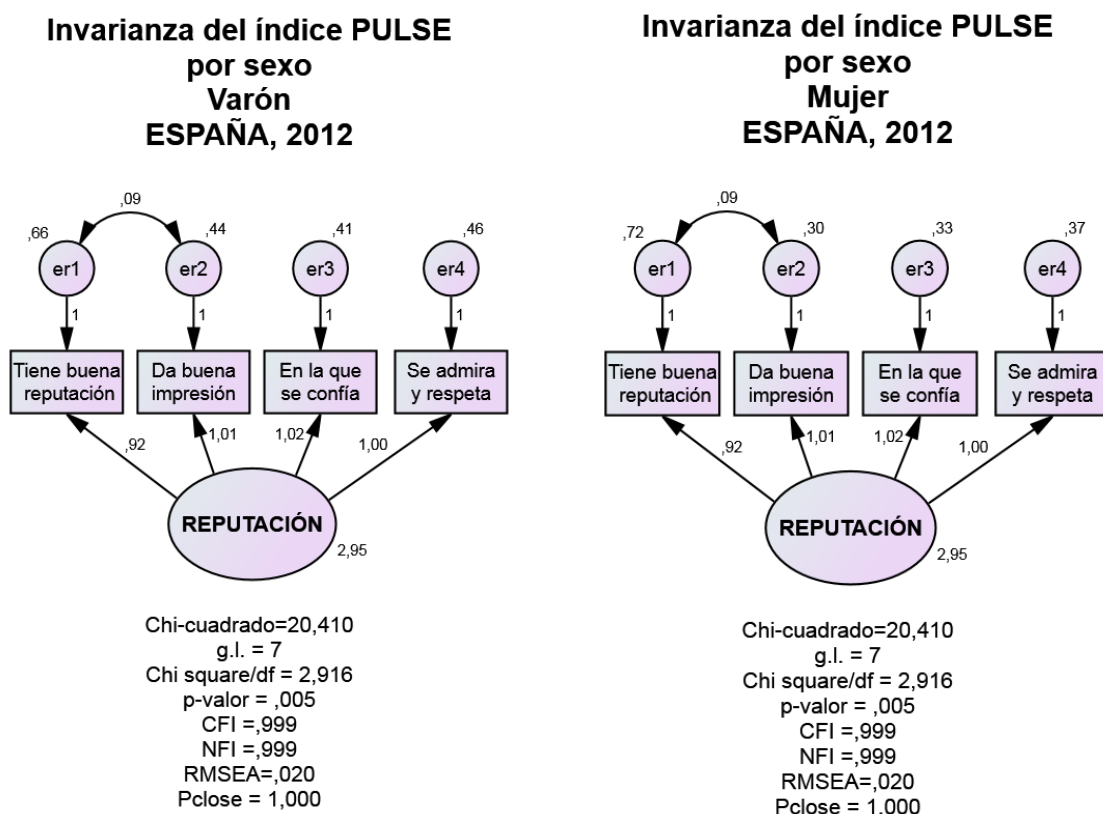


Figura nº 20. Modelo confirmado por los datos de la invarianza por sexo.

#### 4.5.3. INVARIANZA DEL PULSE POR GRUPOS DE EDAD, EN ESPAÑA.

Con el objeto de obtener suficientes casos por grupo homogéneo, en especial por el caso chino, hemos agrupado las categorías de edad en dos grupos: los jóvenes, hasta 35 años y lo adultos, mayores de 35 años.

Por edad, los resultados de los ajustes han sido los siguientes:

**ESPAÑA: INVARIANZA POR EDAD**

Modelos	$\chi^2$	g.l.	P-valor B-S	CFI	RMSEA
Modelo configural	1,4	2	<b>0,774</b>	1,000	0,000
Igualdad de cargas	5,5	5	<b>0,624</b>	1,000	0,004
Igualdad varianza PULSE	9,2	6	<b>0,370</b>	1,000	0,011
Igualdad de covarianza error	13,2	7	<b>0,343</b>	1,000	0,014
Invarianza estricta	219,4	11	<b>0,000</b>	0,991	0,063

Tabla nº 25. Estimación de los modelos anidados.

En este caso, el modelo verdadero conforme con los datos es el que restringe a la igualdad las cargas, la varianza del Pulse y la covarianza de los dos primeros residuales. La probabilidad exacta de cometer error al rechazar la hipótesis nula de igualdad de las matrices de covarianzas observadas e implicadas en ambos grupos, corregido por la no normalidad, ha salido muy alta, p-valor  $Bollen-Stein = 0,343$ , muy por encima del umbral de significación, 0,05, por tanto concluimos, que a este nivel, el modelo de invarianza es conforme a datos.

La única comparación a comprobar es la pérdida de ajuste respecto de la invarianza estricta, cuando los residuales de medida son iguales.

Valoración de la pérdida de ajuste	$\Delta\chi^2$	$\Delta g.l.$	p-valor	$\Delta CFI$	$\Delta RMSEA$
invarianza estricta	218,0	9	<b>0,000</b>	0,009	-0,063

Tabla nº 26. Valoración de la pérdida de ajuste.

Como en el caso del sexo, para España, la igualdad de residuos de los ítems observados ha sido rechazada, tanto porque la pérdida de ajuste es altamente significativa (p-valor = 0,000) como por las diferencias de ambos índices que cumplen o superan el 0,01 de diferencia, en especial el RMSEA.

En consecuencia, el índice Pulse ha demostrado ser invariante por edad. Las personas mayores de 35 años, como las menores entienden y usan métrica de forma similar, la dispersión de datos del Pulse en ambos grupos es la misma y la cantidad de información no explicada por el Pulse, de los ítems “buena reputación” y “tengo buena impresión” es la misma en ambos grupos.

El modelo de invarianza es el siguiente:

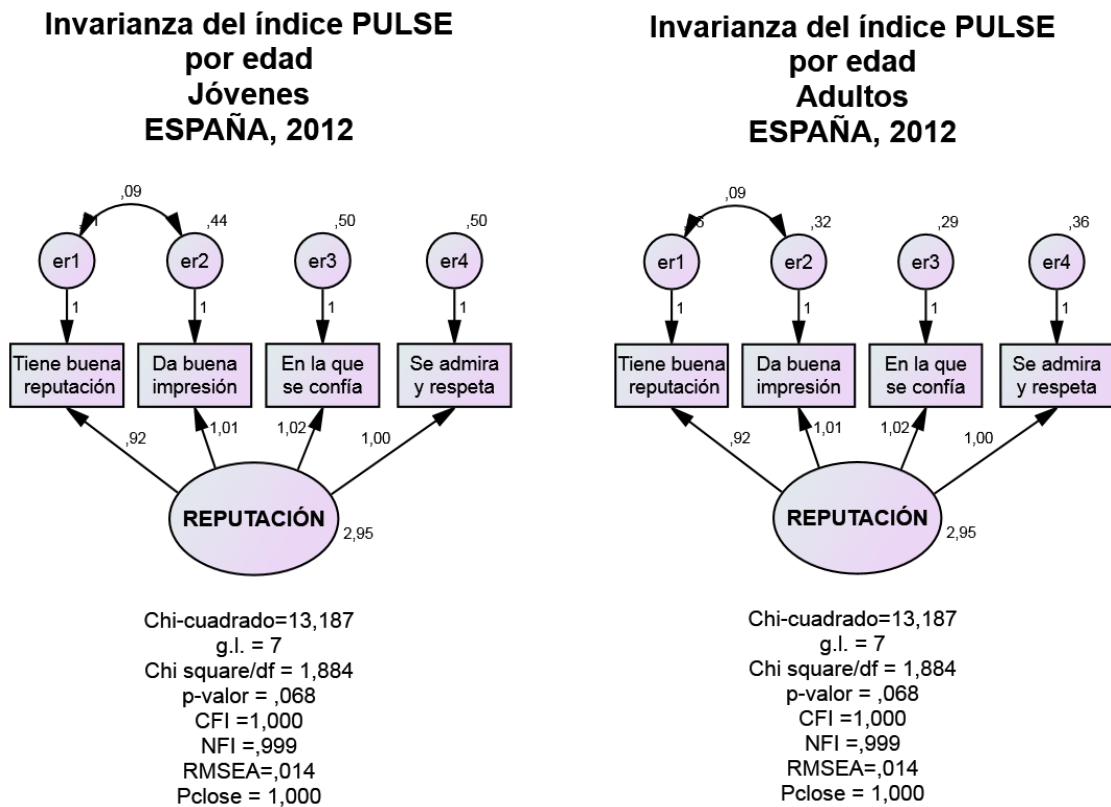


Figura nº 21. Modelo confirmado por los datos de la invarianza por edad.

Como podemos comprobar, las estimaciones de cargas, varianza y covarianza errores son las mismas.

#### 4.5.4. INVARIANZA DEL PULSE POR NIVEL DE ESTUDIOS, EN ESPAÑA.

Para asegurarnos suficiencia de casos en los subgrupos de comparación, hemos agrupado los casos de las categorías del nivel de estudios en dos grupos: el colectivo con estudios bajos, que agrega contingentes sin estudios, primarios y secundarios y el grupo con estudios altos, que agrupa personas con estudios superiores.

Por nivel de estudios, los resultados han sido los siguientes:

### ESPAÑA: INVARIANZA POR NIVEL DE ESTUDIOS

Modelos	$\chi^2$	g.l.	P-valor B-S	CFI	RMSEA
Modelo configural	2,9	2	<b>0,590</b>	1,000	0,009
Igualdad de cargas	6,5	5	<b>0,525</b>	1,000	0,008
Igualdad varianza PULSE	12,7	6	<b>0,185</b>	1,000	0,015
Igualdad de covarianza error	18,3	7	<b>0,163</b>	1,000	0,018
Invarianza estricta	61,4	11	<b>0,017</b>	0,998	0,031

Tabla nº 27. Estimación de los modelos anidados.

También en el caso del nivel de estudios, el modelo que se ajusta a los datos es el que asume la igualdad de las cargas, la varianza del Pulse y la igualdad de la covarianza entre grupos. El p-valor del procedimiento de Bollen Stein, es 0,163, indicando que la matriz poblacional de covarianzas reproducidas por este modelo reproduce fielmente la matriz poblacional de covarianzas observadas.

Según se aprecia en la tabla siguiente, tampoco en el caso del nivel de estudios se verifica la igualdad de los residuos de los ítems entre niveles de estudios diferentes.

Valoración de la pérdida de ajuste	$\Delta\chi^2$	$\Delta$ g.l.	p-valor	$\Delta$ CFI	$\Delta$ RMSEA
invarianza estricta	58,5	9	<b>0,000</b>	0,002	-0,022

Tabla nº 28. Valoración de la pérdida de ajuste.

La diferencia en chi-cuadrado del modelo de invarianza estricta, respecto del anterior ( $\Delta\chi^2 = 58,5$ ) es altamente significativa, indicando que los datos no lo verifican, además la pérdida de RMSEA supera el umbral del 0,01 especificado por Fang Chen (2007).

Parece confirmarse la invarianza del índice Pulse por nivel de estudios. Las personas con estudios altos, otorgan la misma importancia a los ítems que componen el Pulse, muestran la misma dispersión de datos y presentan la misma covariante no explicada entre los ítems primero y segundo.

Normalmente, la invarianza por nivel de estudios no se suele conseguir porque las personas con estudios más altos suelen estar habituadas a procesar una mayor cantidad

de información. Precisamente en el caso de nuestro índice, al ser de naturaleza más emocional, esta circunstancia no le haya afectado y así, la gente con estudios altos tienden a sentir de la misma manera que la gente con un nivel menos cualificado.

En la figura siguiente podemos comprobar el grado de invarianza conseguida por el índice Pulse entre diferentes niveles de estudios.

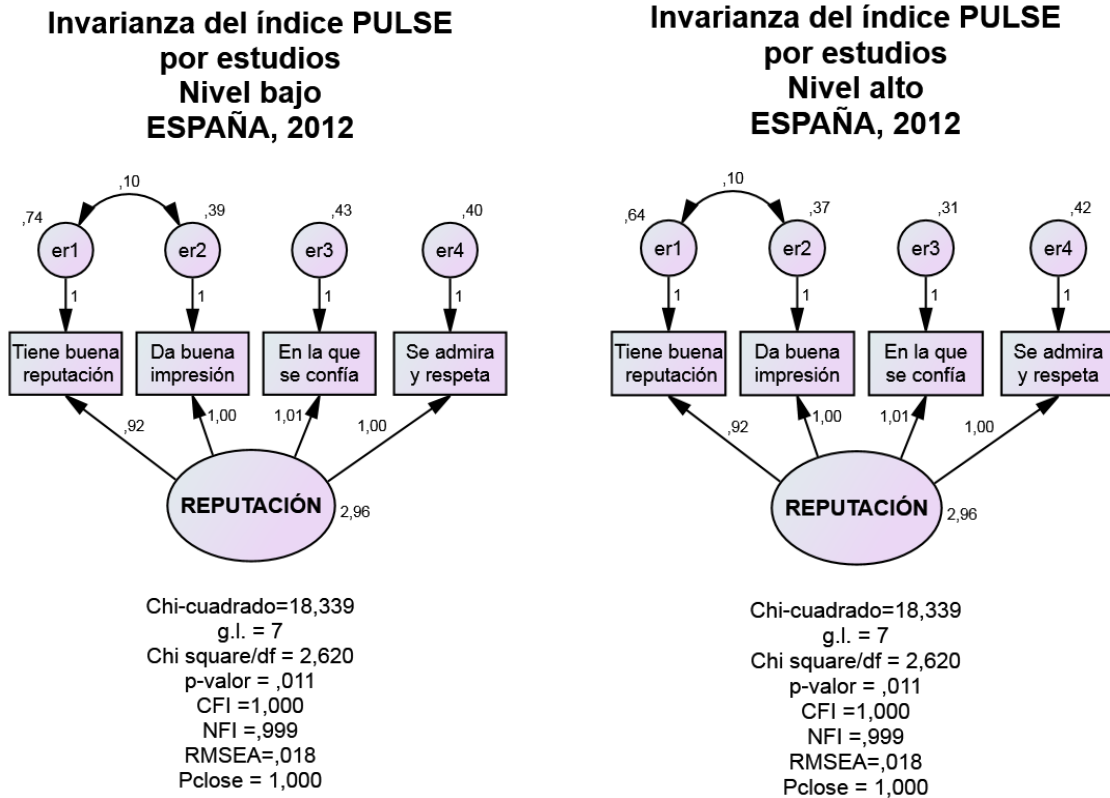


Figura nº 22. Modelo confirmado por los datos de la invarianza por nivel de estudios.

#### 4.5.5. INVARIANZA DEL PULSE POR SECTOR DE ACTIVIDAD EN ESPAÑA.

Las muestras analizadas contienen dos sectores de actividad, la banca y el sector de las telecomunicaciones<sup>83</sup>. Esta estructura nos permite verificar si el índice Pulse se muestra invariante por sector de actividad.

En realidad, el sector de actividad no es una variable demográfica, más bien, se trata de una variable de ámbito o alcance. Sin embargo, hemos considerado someterla a prueba para comprobar si el peso de los ítems es el mismo en las citadas áreas productivas.

La tabla siguiente presenta los resultados del análisis de invarianza por sector de actividad.

ESPAÑA: INVARIANZA POR SECTOR DE ACTIVIDAD			
Modelos	$\chi^2$	g.l.	P-valor B-S
Modelo configural	1,4	2	<b>0,778</b>
Igualdad de cargas	6,9	5	<b>0,492</b>
Igualdad varianza PULSE	8,5	6	<b>0,418</b>
Igualdad de covarianza error	22,5	9	<b>0,396</b>
Invarianza estricta	37,7	11	<b>0,135</b>

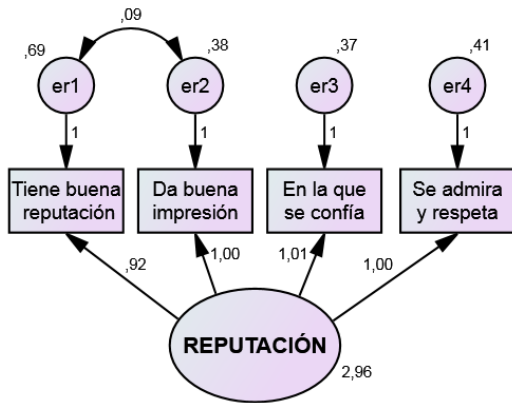
Tabla nº 29. Estimación de los modelos anidados.

En el caso español, la invarianza es completa. Los pesos de los ítems son los mismos por sector, la dispersión de las puntuaciones Pulse también es la misma, la covarianza de los errores se verifica en ambos sectores y los residuales, es decir, la varianza no explicada por cada indicador, la misma. Estamos ante una situación de invarianza métrica plena de la escala. La invarianza estricta por sector es la siguiente:

---

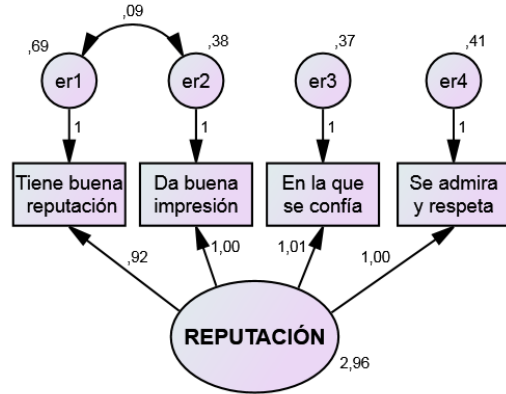
<sup>83</sup> Véase el apartado “Descripción de las muestras utilizadas”

**Invarianza del índice PULSE  
por sector de actividad  
Bancario  
ESPAÑA, 2012**



Chi-cuadrado=37,715  
g.l. = 11  
Chi square/df = 3,429  
p-valor = ,000  
CFI = ,999  
NFI = ,998  
RMSEA=,023  
Pclose = 1,000

**Invarianza del índice PULSE  
por sector de actividad  
Telecomunicaciones  
ESPAÑA, 2012**



Chi-cuadrado=37,715  
g.l. = 11  
Chi square/df = 3,429  
p-valor = ,000  
CFI = ,999  
NFI = ,998  
RMSEA=,023  
Pclose = 1,000

Figura nº 23. Modelo confirmado por los datos de la invarianza por sector de actividad.

#### **4.5.6. CORROBORACIÓN DE LA INVARIANZA DEL PULSE EN LOS CONTEXTOS AMERICANO Y CHINO.**

Conviene ahora analizar si el grado de invarianza detectado en España, es extrapolable a los contextos culturales norteamericano y chino.

##### **4.5.6.1. LA INVARIANZA POR SEXO EN EEUU Y CHINA.**

La especificidad española de la covarianza entre los residuales de los ítems 1 “buena reputación” y 2 “buena impresión” no se verifica en los casos estadounidense y chino. Por lo tanto, los residuales, la variabilidad de cada uno de los ítems que no está explicada por el Pulse se comporta de modo aleatorio y no apunta a otros efectos sistemáticos anteriores, además del Pulse.

En estos casos el análisis de invarianza queda simplificado porque se reduce en un modelo. Así probaremos cuatro modelos de invarianza:

- La invarianza configural, que significa que la estructura factorial de un factor con cuatro índices es adecuada para todos los grupos, aunque varíen en la importancia de cada ítem. Representa el estado de invarianza mínima (Schmitt y Kuljanin, 2008).
- La invarianza de cargas, los grupos otorgan la misma importancia a los ítems. Los ítems significan lo mismo para ambos ítems.
- La invarianza de la varianza del Pulse. La dispersión de valores de la variable latente es la misma en ambos grupos
- La invarianza estricta que asume las invarianzas anteriores más la de los residuos de los ítems.



Los resultados del análisis de invarianza por sexo para EEUU y China son los siguientes:

**EEU: INVARIANZA POR SEXO**

Modelos	$\chi^2$	g.l.	P-valor B-S	CFI	RMSEA
Modelo configural	43,0	4	<b>0,009</b>	0,998	0,044
Igualdad de cargas	53,4	7	<b>0,003</b>	0,998	0,036
Igualdad varianza PULSE	59,4	8	<b>0,002</b>	0,998	0,036
Invarianza estricta	126,9	12	<b>0,000</b>	0,995	0,044

Valoración de la pérdida de ajuste	$\Delta\chi^2$	$\Delta$ g.l.	p-valor	$\Delta$ CFI	$\Delta$ RMSEA
Igualdad de cargas	10,4	3	<b>0,015</b>	0,000	0,008
Igualdad de varianza PULSE	16,4	4	<b>0,003</b>	0,000	0,008
invarianza estricta	83,9	8	<b>0,000</b>	0,003	0,000

**CHINA: INVARIANZA POR SEXO**

Modelos	$\chi^2$	g.l.	P-valor B-S	CFI	RMSEA
Modelo configural	9,3	4,0	<b>0,201</b>	0,998	0,038
Igualdad de cargas	15,8	7,0	<b>0,133</b>	0,997	0,037
Igualdad varianza PULSE	19,3	8,0	<b>0,105</b>	0,996	0,039
Invarianza estricta	26,4	12,0	<b>0,340</b>	0,996	0,027

Tabla nº 30. Estimación de los modelos anidados por sexo, para EEUU y China

En el caso americano el modelo configural sale significativo, incluso con la corrección de la no normalidad ( $p\text{-valor}_{\text{Bollen-Stein}} = 0,009$ ). Gordon Cheung y Roger Rensvold sostienen que el test de la diferencia es sensible al tamaño de muestra y en este sentido recomiendan comprobar los valores de los índices de ajuste CFI y RMSEA.

En nuestro caso, los valores del CFI  $0,998 > 0,95$  y del RMSEA  $= 0,04 < 0,05$ , indican que el modelo es adecuado a los datos. Jodie Ullman recuerda que cuando el ajuste no está asegurado, conviene examinar las matrices de las covarianzas residuales estandarizadas (Ullman, 2006).

Cuando el residual estandarizado alcanza o supera el valor de 1,96, significa que el modelo no ha conseguido reproducir la varianza o covarianza de los ítems implicados y por tanto es un indicador de modelo desajustado. La tabla siguiente muestra la cantidad de residuales estandarizados encontrados en los modelos de invarianza por sexo en EEU

Los modelos configural e igualdad de cargas no presentan residuales estandarizados significativos tanto en varones como en mujeres, indicando que el índice Pulse descrito en estos modelos llega a reproducir razonablemente bien las variaciones de los ítems. Sin embargo los modelos que proponen la igualdad de varianza Pulse y la invarianza restringida, serían candidatos a ser rechazados porque contienen 2 violaciones en los 10 puntos muestrales (4 varianzas y 6 covarianzas).

Cantidad de residuales estandarizados significativos > 1,96 según modelos ajustados		
Modelos	Varones	Mujeres
Configural	0	0
Igualdad de cargas	0	0
Igualdad de varianza Pulse	2	2
invarianza irrestricta	2	2

Tabla nº 3. Covarianzas residuales de las matrices observada e implicada divididas por su error estándar <sup>84</sup>.

Por lo tanto, proponemos que en el caso americano el índice Pulse se ha verificado empíricamente la igualdad de cargas. Hombres y mujeres otorgan la misma importancia a los ítems cuando forman el valor global del Pulse. La invarianza de la dispersión del Pulse no ha sido corroborada. La tabla siguiente somete a prueba el nivel de significación de la diferencia de varianzas Pulse entre mujeres y varones.

	Mujeres	Varones	diferencia	Prueba F	g.l. mujeres	g.l. varones	Significación
Varianza Pulse	2,802	2,549	0,253	1,09925	2404	2611	<b>0,009</b>

Tabla nº 30 . Determinación de significación de la diferencia en varianzas en hombres y mujeres EEUU.

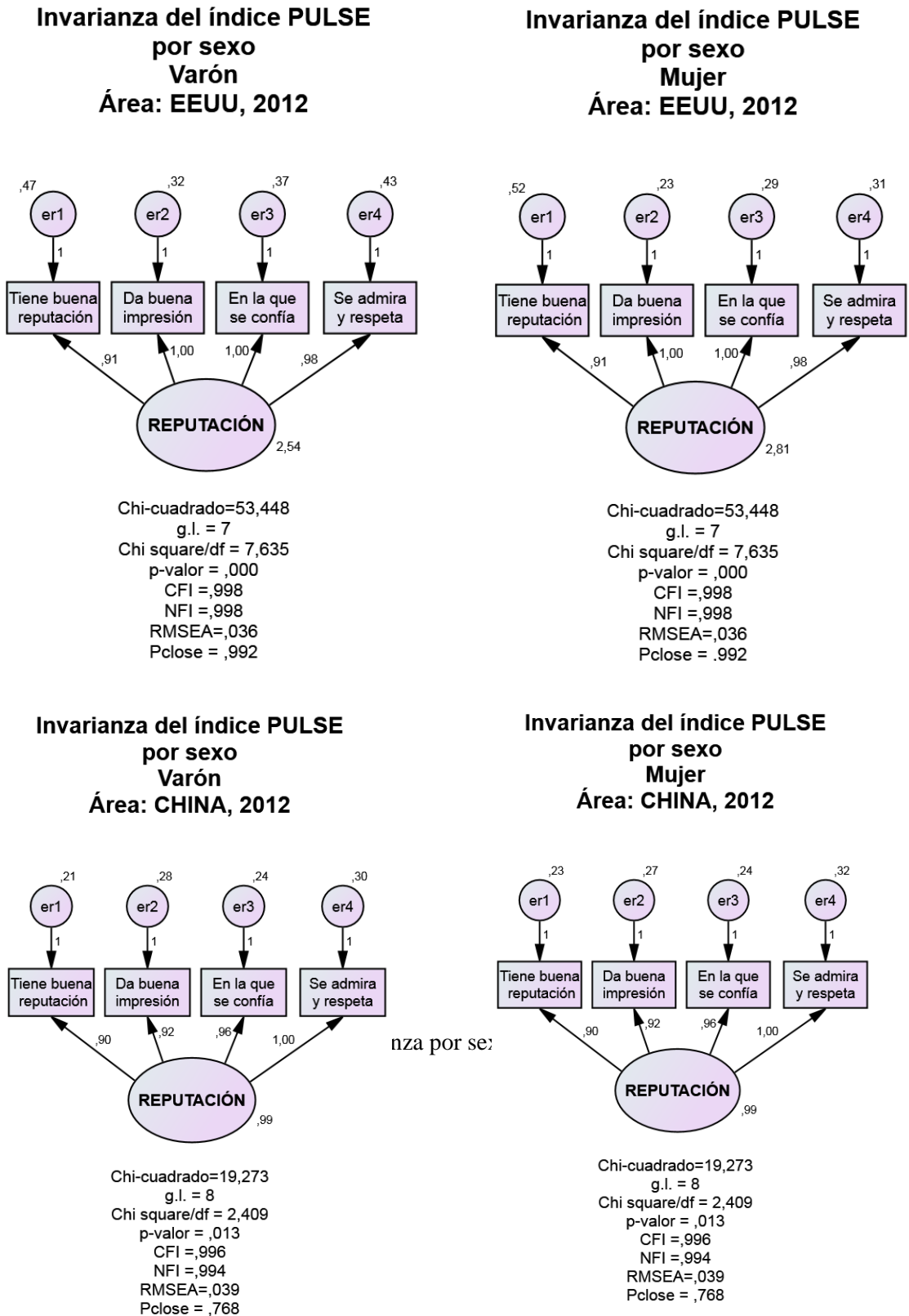
En términos de desviaciones típicas la variabilidad de las mujeres sería superior en 0,08 puntos a la de los varones <sup>85</sup>. La diferencia es significativa debido al tamaño de muestra pero las diferencias son triviales.

<sup>84</sup> Las matrices de varianzas y covarianzas residuales estandarizadas para los modelos de la invarianza por sexo en EEUU están disponibles en el anexo 10.2.6.

<sup>85</sup> “ $S_{mujer} = 1,674$ ”, frente a “ $S_{varón} = 1,597$ ”

En China, la invarianza del Pulse por sexo es clara. La probabilidad exacta de cometer error al rechazar la igualdad de matrices de covarianzas poblacionales supera el umbral del 0,05 ( $p\text{-valor}_{\text{Bollen-Stein}} = 0,136$ ).

Los modelos de invarianza pulse por sexo en EEUU y China son: (figura nº 24):



Los resultados de las pruebas estadísticas para España parecen confirmarse también en EEUU y China. El índice Pulse habría mostrado un grado de invarianza importante por sexo que alcanzaría la invarianza de cargas y la de dispersión del índice<sup>86</sup>.

#### 4.5.6.2. LA INVARIANZA POR EDAD EN EEUU Y CHINA.

Esta serie de pruebas estadísticas han sido aplicadas para tratar de demostrar o refutar 4 Hipótesis de la tesis de no diferencia por edad en el Índice Pulse: H3.1 invarianza configural, H3.2 invarianza de cargas, H3.3 invarianza de la varianza del índice y H.3.4. invarianza estricta.

Vamos a ver a continuación los resultados de las pruebas estadísticas y matemáticas realizadas para la verificación o rechazo de estas hipótesis.

En los contextos chino y estadounidense el Pulse no se ve afectado por posibles cambios en la estructura de edad de la población.

##### EEUU: INVARIANZA POR EDAD

Modelos	$\chi^2$	g.l.	P-valor B-S	CFI	RMSEA
Modelo configural	17,8	4	<b>0,176</b>	0,999	0,026
Igualdad de cargas	18,7	7	<b>0,229</b>	1,000	0,018
Igualdad varianza PULSE	25,9	8	<b>0,098</b>	0,999	0,021
Invarianza estricta	323,1	12	<b>0,000</b>	0,988	0,072

Valoración de la pérdida de ajuste	$\Delta\chi^2$	$\Delta$ g.l.	p-valor	$\Delta$ CFI	$\Delta$ RMSEA
invarianza estricta	305,3	8	<b>0,000</b>	0,01	-0,046

##### CHINA: INVARIANZA POR EDAD

Modelos	$\chi^2$	g.l.	P-valor B-S	CFI	RMSEA
Modelo configural	10,6	4	<b>0,159</b>	0,998	0,042
Igualdad de cargas	14,5	7	<b>0,182</b>	0,998	0,034
Igualdad varianza PULSE	15,7	8	<b>0,218</b>	0,997	0,032
Invarianza estricta	24,2	12	<b>0,171</b>	0,996	0,027

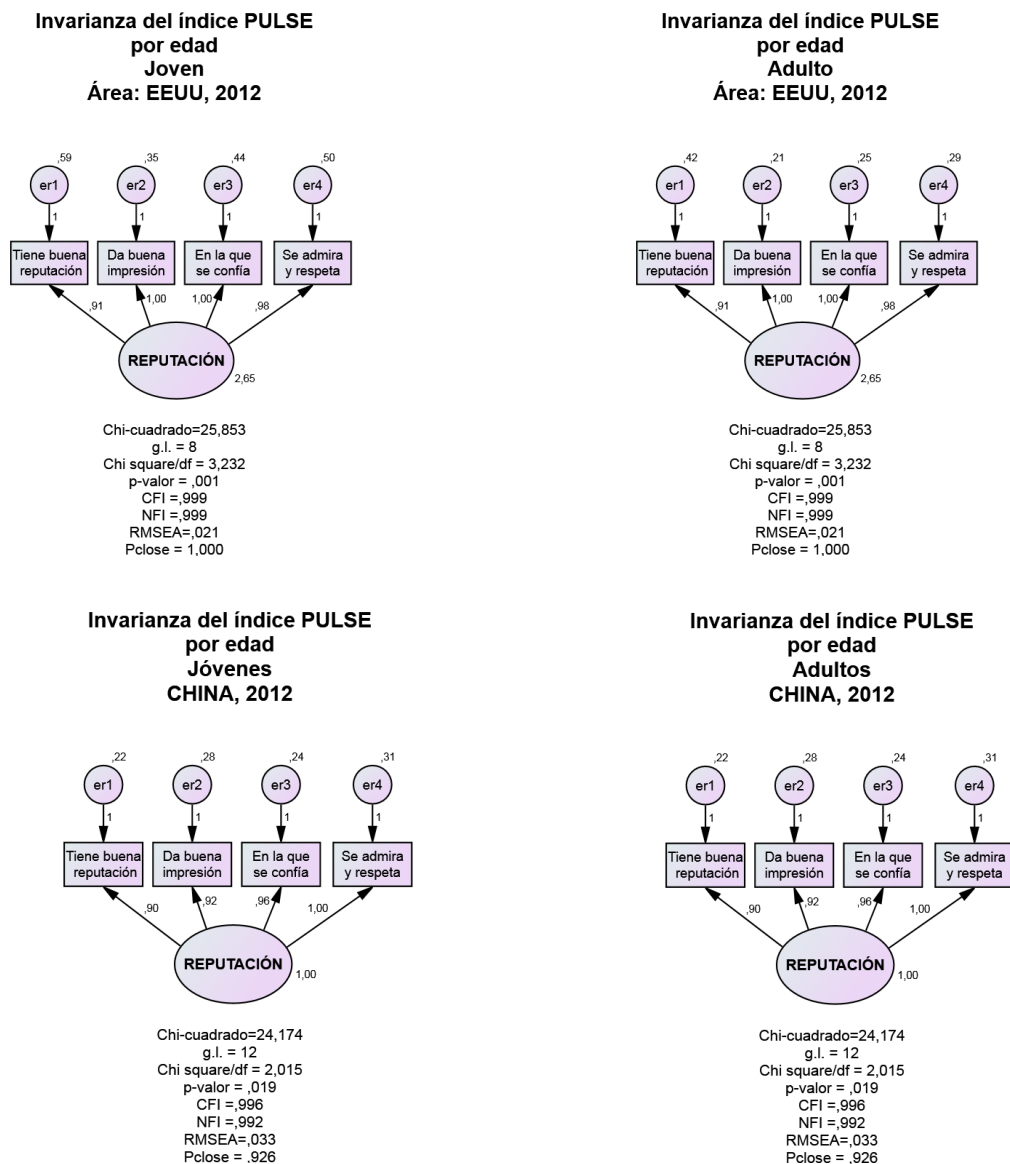
Tabla nº 33. Estimación de los modelos anidados por edad para EEUU y China

<sup>86</sup> La única excepción del caso americano sólo impugnaría la hipótesis de invarianza por varianza latente, pero hemos visto que en ese caso es más un cuestión de la sensibilidad al excesivo tamaño de muestra que al efecto real, que en todo caso sería mínimo.

En EEUU, el modelo verdadero es el que asume la invarianza de cargas y la dispersión de las puntuaciones Pulse en jóvenes y adultos ( $p\text{-valor}_{B-S} = 0,098 > 0,05$ ). Sin embargo la invarianza de los residuales de los ítems no estamos seguros de verificarla según el test de pérdida de ajuste ( $\Delta\chi^2_8 = 305,3$ , con un  $p\text{-valor} = 0,000$ ). La diferencia en el RMSEA también indicaría que el modelo de invarianza estricta es insostenible porque la pérdida en la función es ajuste es excesiva ( $\Delta\text{RMSEA} = |0,046| > 0,01$ ).

Para China, la invarianza estricta del índice Pulse por edad ha quedado corroborada ( $p\text{-valor}_{B-S} = 0,171$ ).

En la figura siguiente nº 25 representamos los dos modelos invariantes en estas poblaciones.



#### 4.5.6.3. LA INVARIANZA POR NIVEL DE ESTUDIOS EN EEUU.

Esta serie de pruebas estadísticas han sido aplicadas para refutar o demostrar 4 Hipótesis de la tesis de no diferencia por nivel de estudios en el Índice. Los resultados de las pruebas estadísticas apuntan a una verificación de estas hipótesis.

Los resultados y explicación de las pruebas de verificación de las hipótesis son los que describimos a continuación.

La invarianza del Pulse por nivel de estudios también ha quedado corroborada, además de España, en Estados Unidos. El modelo corroborado empíricamente se corresponde con el de invarianza de la varianza extractada por el Pulse ( $p\text{-valor} = 0,052 > 0,05$ ). Por otro lado, los valores del CFI ( $0,999 > 0,95$ ) y del RMSEA ( $0,027 < 0,05$ ) indican un ajuste de las matrices de covarianzas excelente. Tanto los coeficientes que unen los ítems con el índice como la varianza del Pulse en ambos grupos es idéntica.

**EEUU: INVARIANZA POR NIVEL DE ESTUDIOS**

Modelos	$\chi^2$	g.l.	P-valor B-S	CFI	RMSEA
Modelo configural	12,8	4	<b>0,332</b>	1,000	0,021
Igualdad de cargas	14,3	7	<b>0,397</b>	1,000	0,014
Igualdad varianza PULSE	30,9	8	<b>0,052</b>	0,999	0,024
Invarianza estricta	54,4	12	<b>0,040</b>	0,998	0,027

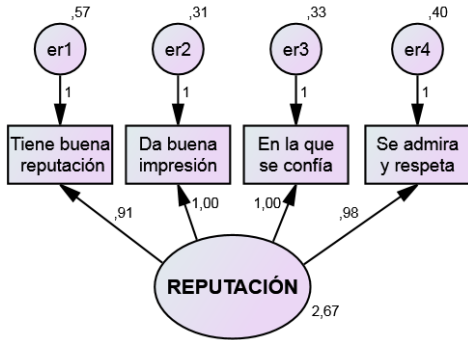
Valoración de la pérdida de ajuste	$\Delta\chi^2$	$\Delta g.l.$	p-valor	$\Delta CFI$	$\Delta RMSEA$
invarianza estricta	23,5	4	<b>0,000</b>	0	-0,006

Tabla nº 34. Estimación de los modelos anidados por nivel de estudio en EEUU.

El gráfico siguiente se puede apreciar que los coeficientes no estandarizados (0,91, 1,00, 1,00 y 0,98) la varianza del índice 2,67 son las mismas en ambos grupos. Sólo difieren en los residuales de las cargas.

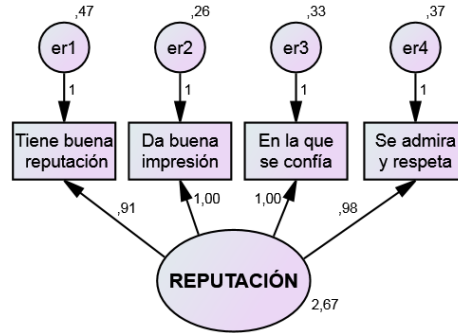
Modelos de invarianza Pulse en EEUU.

**Invarianza del índice PULSE  
por nivel de estudios  
Bajos  
Área: EEUU, 2012**



Chi-cuadrado=30,896  
g.l. = 8  
Chi square/df = 3,862  
p-valor = ,000  
CFI = ,999  
NFI = ,999  
RMSEA=,024  
Pclose = 1,000

**Invarianza del índice PULSE  
por nivel de estudios  
Alto  
Área: EEUU, 2012**



Chi-cuadrado=30,896  
g.l. = 8  
Chi square/df = 3,862  
p-valor = ,000  
CFI = ,999  
NFI = ,999  
RMSEA=,024  
Pclose = 1,000

Figura nº 26. Modelos de Invarianza Pulse en EEUU por nivel de estudios.

#### 4.5.6.4. LA INVARIANZA POR SECTOR DE ACTIVIDAD EN EEUU

Al igual que en España el índice Pulse se muestra invariante por sector de actividad. El grado de invarianza es completo y alcanza el nivel estricto de igualdad de residuales de los ítems, a pesar del ingente volumen de casos de la base de datos estadounidense.

**EEUU. INVARIANZA POR SECTOR DE ACTIVIDAD**

Modelos	$\chi^2$	g.l.	P-valor B-S	CFI	RMSEA
Modelo configural	12,5	4	<b>0,219</b>	0,9990	0,030
Igualdad de cargas	15,4	7	<b>0,249</b>	0,9990	0,022
Igualdad varianza PULSE	17,6	8	<b>0,200</b>	0,9990	0,022
Invarianza estricta	36,8	12	<b>0,226</b>	0,998	0,029

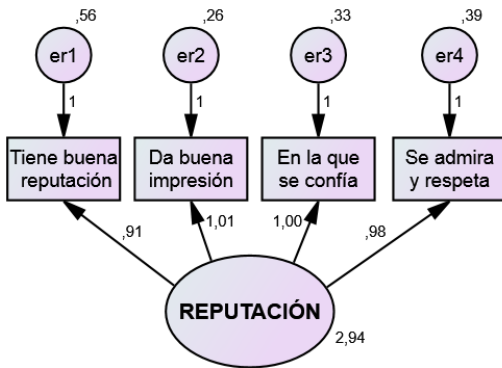
Tabla nº 35. Invarianza de los modelos anidados por sector de actividad en EEUU.

El modelo de la invarianza estricta aparece como el modelo verdadero que consigue reproducir la matriz de covarianzas observadas con un p-valor de  $0,182 > 0,05$ .

La confirmación de estos resultados con los de España, permite plantear la última prueba de invarianza poblacional sobre grupos homogéneos por sexo, edad y nivel de estudios, en los dos sectores conjuntamente, que ha quedado despejado que el sector de actividad sea fuente de no invarianza.

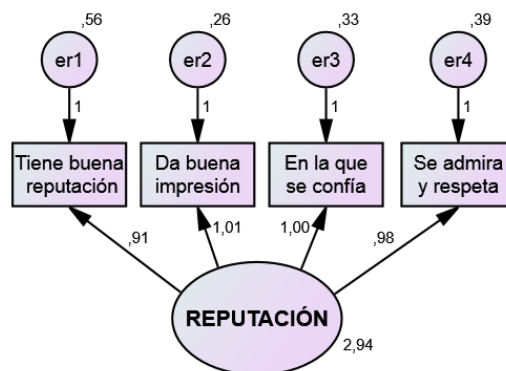
El modelo verdadero de invarianza por sector es el siguiente: Figura nº 27.

**Invarianza del índice PULSE  
por sector de actividad  
Bancario  
Área: EEUU, 2012**



Chi-cuadrado=36,354  
g.l. = 12  
Chi square/df = 3,029  
p-valor = ,000  
CFI = ,998  
NFI = ,997  
RMSEA = ,029  
Pclose = ,999

**Invarianza del índice PULSE  
por sector de actividad  
Telecomunicaciones  
Área: EEUU, 2012**



Chi-cuadrado=36,354  
g.l. = 12  
Chi square/df = 3,029  
p-valor = ,000  
CFI = ,998  
NFI = ,997  
RMSEA = ,029  
Pclose = ,999



Tanto las cargas no estandarizadas (0,91, 1,01, 1,00 y 0,98), como la varianza del índice de Reputación (2,94) como los residuales (0,56, 0,26, 0,33 y 0,39) son iguales para cada grupo.

#### 4.5.6.5. LA INVARIANZA CONJUNTA POR SEXO, EDAD Y ESTUDIOS EN LOS TRES CONTEXTOS.

Por último sometemos al Pulse a una prueba fuerte de invarianza demográfica, homogeneizando los grupos por sexo, edad y nivel de estudios en España y EEUU y por sexo y edad en China<sup>87</sup>.

Según se refleja en la tabla siguiente, el cruce de las tres variables ha dado lugar a la creación de ocho grupos homogéneos.

Grupos homogéneos por sexo, edad y estudios	España	EEUU
Varón, menos de 35 años nivel de estudios bajos	497	273
Varón, 35 años o más nivel de estudios bajos	809	297
Varón, menos de 35 años nivel de estudios altos	511	852
Varón, 35 años o más nivel de estudios altos	639	1190
Mujer, menos de 35 años, nivel de estudios bajos	364	266
Mujer, 35 años o más nivel de estudios bajos	729	300
Mujer, menos de 35 años nivel de estudios altos	514	796
Mujer, 35 años o más nivel de estudios altos	703	1043
	Suma	Suma
	4766	5017

Tabla nº 36. Distribución muestral de los ocho grupos demográficos homogeneizados en España y EEUU.

Como ya comentamos más arriba, la base de datos de China ha facilitado la identificación de 4 grupos igualados por sexo y edad que representamos en la tabla siguiente.

---

<sup>87</sup> Recordemos que en la BBDD de China la variable nivel de estudios no presentaba una distribución suficientemente variada como para someterla a prueba.

Grupos homogéneos por sexo y edad	China
Varón, menos de 35 años	354
Varón, 35 o más años	181
Mujer, menos de 45 años	289
Mujer, 35 años o más	108
	Suma
	932

Tabla nº 37. Distribución muestral de los cuatro grupos demográficos en China.

#### 4.5.6.6. INVARIANZA CONJUNTA EN ESPAÑA.

En España, el grado de invarianza del índice Pulse por grupos homogeneizados por las principales variables demográficas es muy buena. Los resultados de la tabla siguiente dejan claro que el índice cumple la de la igualdad de cargas ( $p$ -valor  $B-S = 0,079$ ) y el test de pérdida de ajuste del modelo que restringe la varianza del Pulse no es significativa ( $\Delta\chi^2_7 = 11,9$ ;  $p$ -valor =  $0,103$ ) confirmando a ese nivel el grado alcanzado.

#### ESPAÑA: INVARIANZA CONJUNTA POR SEXO EDAD Y NIVEL DE ESTUDIOS

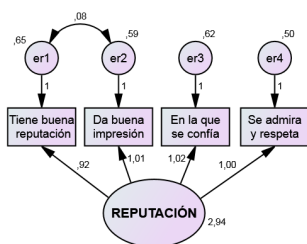
Modelos	$\chi^2$	g.l.	P-valor B-S	CFI	RMSEA
Modelo configural	27,3	8	<b>0,216</b>	0,999	0,022
Igualdad de cargas	57,6	29	<b>0,079</b>	0,999	0,014
Igualdad varianza PULSE	69,5	36	<b>0,035</b>	0,999	0,014
Igualdad de covarianza error	85,4	43	<b>0,105</b>	0,994	0,024
Invarianza estricta	476,1	71	<b>0,000</b>	0,983	0,035

Valoración de la pérdida de ajuste	$\Delta\chi^2$	$\Delta$ g.l.	p-valor	$\Delta$ CFI	$\Delta$ RMSEA
Igualdad de varianza PULSE	11,9	7	<b>0,103</b>	0	-0,008
Igualdad de covarianza error	27,8	14	<b>0,015</b>	-0,005	0,002
invarianza estricta	418,6	42	<b>0,000</b>	-0,016	0,013

Tabla nº 38. Invarianza de los modelos anidados por grupos homogéneos de sexo, edad y estudios. España.

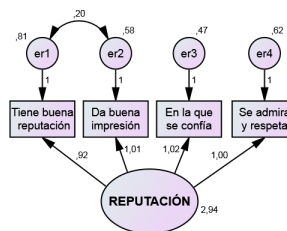
A partir de aquí los resultados son menos congruentes porque el modelo que restringe la igualdad de los errores, aunque parece cumplir con los datos ( $p$ -valor  $B-S = 0,105$ ), el test de pérdida de ajuste resulta significativo ( $p$ -valor =  $0,015 < 0,05$ ). En consecuencia daremos por válido el modelo de invarianza de la varianza del Pulse. En la figura siguiente representamos este modelo por grupos homogeneizados.

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón, joven, nivel bajo  
ESPAÑA, 2012**



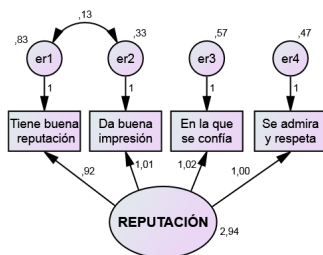
Chi-cuadrado=69,482  
g.l. = 36  
Chi square/df = 1,930  
p-valor = ,001  
CFI = ,999  
NFI = ,997  
RMSEA=.014  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón, joven, nivel alto  
ESPAÑA, 2012**



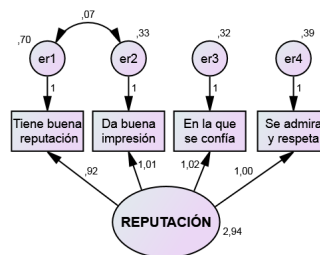
Chi-cuadrado=69,482  
g.l. = 36  
Chi square/df = 1,930  
p-valor = ,001  
CFI = ,999  
NFI = ,997  
RMSEA=.014  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel bajo  
ESPAÑA, 2012**



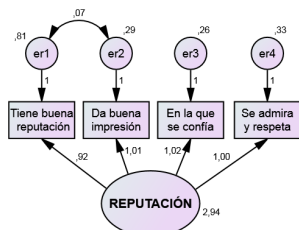
Chi-cuadrado=69,482  
g.l. = 36  
Chi square/df = 1,930  
p-valor = ,001  
CFI = ,999  
NFI = ,997  
RMSEA=.014

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel alto  
ESPAÑA, 2012**



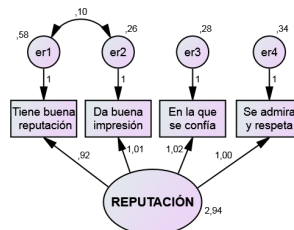
Chi-cuadrado=69,482  
g.l. = 36  
Chi square/df = 1,930  
p-valor = ,001  
CFI = ,999  
NFI = ,997  
RMSEA=.014

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel bajo  
ESPAÑA, 2012**



Chi-cuadrado=69,482  
g.l. = 36  
Chi square/df = 1,930  
p-valor = ,001  
CFI = ,999  
NFI = ,997  
RMSEA=.014  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel alto  
ESPAÑA, 2012**



Chi-cuadrado=69,482  
g.l. = 36  
Chi square/df = 1,930  
p-valor = ,001  
CFI = ,999  
NFI = ,997  
RMSEA=.014  
Pclose = 1,000

Figura nº 28. Modelos confirmados por los datos de la invarianza conjunta en España.

En los ocho modelos se verifica que los diferentes grupos mantienen igualdad de cargas, es decir otorgan la misma importancia a los ítems a la hora de formarse el juicio global de calidad y además producen la misma dispersión de valores Pulse en los diferentes grupos (invarianza de varianza del índice).

#### 4.5.6.7. INVARIANZA CONJUNTA EN EEUU.

Los resultados de EEUU corroboran los alcanzados en España, pero al nivel de la invarianza de cargas del Pulse.

En la tabla siguiente apreciamos que el modelo configural es corroborado con la corrección de la no normalidad ( $p\text{-valor}_{\text{Bollen-Stein}} = 0,084$ ) pero el de igualdad de cargas ya sale significativo ( $p\text{-valor}_{\text{Bollen-Stein}} = 0,034$ ).

En consecuencia, para EEUU, los resultados parecen apoyar el supuesto de que el modelo configural es verdadero y por tanto debemos aceptar que los diferentes grupos homogéneos por sexo, edad y nivel de estudios utilizan el mismo conjunto de indicadores para valorar la reputación global.

Cuando imponemos la restricción adicional de la igualdad de cargas sobre el modelo configural (parte inferior de la tabla siguiente) los resultados confirman que dicha restricción es empírica, con un  $p\text{-valor}$  no significativo ( $p\text{-valor}_{\text{cargas}} = 0,070$ ).

#### EEUU: INVARIANZA CONJUNTA POR SEXO EDAD Y NIVEL DE ESTUDIOS

Modelos	$\chi^2$	g.l.	P-valor B-S	CFI	RMSEA
Modelo configural	64,458	16	<b>0,084</b>	0,998	0,025
Igualdad de cargas	95,694	37	<b>0,034</b>	0,998	0,018
Igualdad varianza PULSE	132,681	44	<b>0,003</b>	0,996	0,02
Invarianza estricta	803,289	72	<b>0,000</b>	0,971	0,045

Valoración de la pérdida de ajuste	$\Delta\chi^2$	$\Delta\text{g.l.}$	p-valor	$\Delta\text{CFI}$	$\Delta\text{RMSEA}$
Igualdad de cargas	31,2	21	<b>0,070</b>	0,000	-0,007
Igualdad de varianza PULSE	68,2	28	<b>0,000</b>	-0,002	-0,005
invarianza estricta	738,8	56	<b>0,000</b>	-0,027	0,020

Tabla nº 39. Invarianza conjunta de los modelos anidados por grupos homogéneos de sexo, edad y estudios. EEUU.

Este resultado viene corroborado al comprobar que las variaciones incrementales de los índices “ $\Delta CFI$ ” y “ $\Delta RMSEA$ ” están por debajo de los límites críticos señalados por Gordon Cheung y Roger Rensvold (2002), antes comentados.

En consecuencia aceptamos la igualdad de cargas de los indicadores Pulse entre los colectivos homogeneizados por sexo, edad y nivel de estudios.

Ya hemos indicado anteriormente que los autores avisan de la sensibilidad de estos test de ajuste a los tamaños grandes de muestra y que conviene completar la interpretación de resultados con el análisis de residuales.

La matriz de covarianzas residuales estandarizadas confirma que los modelos configural e invarianza de cargas son adecuados a los datos<sup>88</sup> (tabla siguiente). La ausencia de covarianzas residuales significativas indica que el modelo reproduce razonablemente bien la relación entre los indicadores en los ocho colectivos.

Cantidad de residuales estandarizados significativos > 1,96									
MODELO	Varones jóvenes bajos	Varones jóvenes altos	Varones adultos bajos	Varones adultos altos	Mujeres jóvenes bajos	Mujeres jóvenes altos	Mujeres adultas bajas	Mujeres adultas altas	Total residuales significativos
Configural	0	0	0	0	0	0	0	0	0
Igualdad de cargas	0	0	0	0	0	0	0	0	0
Igualdad de varianza	0	9	9	0	9	1	7	0	35
invarianza irrestricta	1	9	7	0	10	0	4	0	31

Tabla nº 40. Covarianzas residuales de las matrices observada e implicada divididas por su error estándar<sup>89</sup>.

Ahora bien, el modelo de invarianza de la varianza del Pulse no se sostiene. De los ocho grupos, en cinco, el modelo no consigue reproducir la estructura de varianzas y covarianzas observadas<sup>90</sup>:

<sup>88</sup> Recordemos que cuando el residual estandarizado de algún grupo alcanza o supera el valor de 1,96, significa que el modelo, para dicho grupo, no ha conseguido reproducir la varianza o covarianza de los ítems implicados y por tanto es un indicador de modelo poco aceptable.

<sup>89</sup> Las matrices de varianzas y covarianzas residuales estandarizadas para los modelos de la invarianza por sexo en EEUU están disponibles en el anexo 10.2.7.

En cuatro grupos el desajuste es casi total. En tres de ellos, “varones jóvenes” (con estudios altos y bajos) y “mujeres jóvenes con estudios bajos”, los residuales significativos son 9, de un total de 10 (4 varianzas y 6 covarianzas) y en el grupo de “mujeres adultas con estudios bajos” los residuales significativos son 7. En el grupo de mujeres jóvenes con estudios altos registra un residual significativo: varianza del ítem 1.

A la luz de estos resultados parece que no es conveniente aceptar la igualdad de dispersión de las medidas del Pulse. En consecuencia valoraremos el grado en que las varianzas de los ocho grupos difieren mediante una serie de test de contraste<sup>91</sup>, siendo la hipótesis nula la igualdad de las varianzas entre grupos: “ $H_0: S^2_{\text{grupo A}} - S^2_{\text{grupo B}} = 0$ ”. Se trata de una prueba complementaria al test de invarianza de la “varianza” del Pulse, que nos permite visualizar mejor el alcance y los colectivos que muestran un mayor grado de dispersión en las puntuaciones de la variable latente Pulse.

Para estimar los valores de la varianza del Pulse en los ocho colectivos demográficos, hemos utilizado el modelo más válido de la igualdad de cargas, siendo sus valores los de la tabla siguiente.

	Varones jóvenes bajos	Varones jóvenes altos	Varones adultos bajos	Varones adultos altos	Mujeres jóvenes bajos	Mujeres jóvenes altos	Mujeres adultas bajos	Mujeres adultas altos
<b>Varianza del Pulse</b>	2,73	2,13	3,27	2,54	3,54	2,50	3,13	2,80
<b>n-1</b>	272	296	851	1189	265	299	795	1042

Tabla nº 41. Varianzas del Pulse y grados de libertad de los ocho grupos de comparación.

Al ser 8 grupos, da lugar a 28 comparaciones que serán presentadas en 3 matrices consecutivas, en la primera, se presentan las diferencias de las varianzas, en la segunda, el valor del estadístico de contraste “F” y en la tercera el nivel de significación. con los grados de libertad asociados.

<sup>90</sup> El apartado 10.2.7 contiene las ocho matrices por cada modelo.

<sup>91</sup> El test de contraste de igualdad de varianzas se basa en la distribución de probabilidad F de Snedecor. Siendo el estadístico de contraste:

$$F = \frac{S^2_{\text{mayor}}}{S^2_{\text{menor}}}, \text{ con g.l. } (n_{\text{mayor}} - 1), \text{ g.l. } (n_{\text{menor}} - 1) \approx F \text{ de Snedecor}$$

<b>Diferencias entre las varianzas</b>							
	Varones jóvenes altos	Varones adultos bajos	Varones adultos altos	Mujeres jóvenes bajos	Mujeres jóvenes altos	Mujeres adultas bajos	Mujeres adultas altos
Varones jóvenes bajos	0,60	0,54	0,19	0,81	0,23	0,40	0,07
Varones jóvenes altos		1,14	0,41	1,41	0,37	1	0,67
Varones adultos bajos			0,73	0,27	0,77	0,14	0,47
Varones adultos altos				1	0,04	0,59	0,26
Mujeres jóvenes bajos					1,04	0,41	0,74
Mujeres jóvenes altos						0,63	0,30
Mujeres adultas bajos							0,33
<b>Valor de contraste F</b>							
	Varones jóvenes altos	Varones adultos bajos	Varones adultos altos	Mujeres jóvenes bajos	Mujeres jóvenes altos	Mujeres adultas bajos	Mujeres adultas altos
Varones jóvenes bajos	1,282	1,198	1,075	1,297	1,092	1,147	1,026
Varones jóvenes altos		1,535	1,192	1,662	1,174	1,469	1,315
Varones adultos bajos			1,287	1,083	1,308	1,045	1,168
Varones adultos altos				1,394	1,016	1,232	1,102
Mujeres jóvenes bajos					1,416	1,131	1,264
Mujeres jóvenes altos						1,120	1,118
Mujeres adultas bajos							1,118
<b>Nivel de significación</b>							
	Varones jóvenes altos	Varones adultos bajos	Varones adultos altos	Mujeres jóvenes bajos	Mujeres jóvenes altos	Mujeres adultas bajos	Mujeres adultas altos
Varones jóvenes bajos	<b>0,02</b>	<b>0,04</b>	<b>0,22</b>	<b>0,02</b>	<b>0,23</b>	<b>0,09</b>	<b>0,39</b>
Varones jóvenes altos		<b>0,00</b>	<b>0,03</b>	<b>0,00</b>	<b>0,08</b>	<b>0,00</b>	<b>0,00</b>
Varones adultos bajos			<b>0,00</b>	<b>0,22</b>	<b>0,00</b>	<b>0,27</b>	<b>0,01</b>
Varones adultos altos				<b>0,00</b>	<b>0,44</b>	<b>0,00</b>	<b>0,05</b>
Mujeres jóvenes bajos					<b>0,00</b>	<b>0,12</b>	<b>0,01</b>
Mujeres jóvenes altos						<b>0,12</b>	<b>0,11</b>
Mujeres adultas bajos							<b>0,05</b>

Tabla nº 42. Serie de contrastes de la igualdad de varianzas entre los grupos homogéneos por sexo edad y nivel de estudios.

Como era de esperar, el 61% de las 28 comparaciones, en concreto 17, han resultado significativas. Los resultados son congruentes con el test de invarianza, sin embargo debemos recordar que la cantidad de muestra vuelve el test muy sensible, llevando inapropiadamente a rechazar la hipótesis nula de igualdad de varianzas. Por ello conviene revisar el tamaño de las diferencias en la dispersión.

La tabla siguiente contiene las diferencias en las desviaciones típicas del Pulse en los diferentes colectivos.

Colectivo	Desv.	Colectivo	Desv.	DIF (abs)
Varones jóvenes bajos	1,65	Varones jóvenes altos	1,46	0,19
Varones jóvenes bajos	1,65	Varones adultos bajos	1,81	0,16
Varones jóvenes bajos	1,65	Mujeres jóvenes bajos	1,88	0,23
Varones jóvenes altos	1,46	Varones adultos bajos	1,81	0,35
Varones jóvenes altos	1,46	Varones adultos altos	1,59	0,13
Varones jóvenes altos	1,46	Mujeres jóvenes bajos	1,88	0,42
Varones jóvenes altos	1,46	Mujeres adultas bajos	1,77	0,31
Varones jóvenes altos	1,46	Mujeres adultas altos	1,67	0,21
Varones adultos bajos	1,81	Varones adultos altos	1,59	0,21
Varones adultos bajos	1,81	Mujeres jóvenes altos	1,58	0,23
Varones adultos bajos	1,81	Mujeres adultas altos	1,58	0,23
Mujeres jóvenes bajos	1,88	Mujeres jóvenes altos	1,58	0,30
Mujeres jóvenes bajos	1,88	Mujeres adultas altos	1,67	0,21
Mujeres adultas bajos	1,77	Mujeres adultas altos	1,67	0,10

Tabla nº 43. Diferencias en valores absolutos entre las desviaciones típicas de los grupos con diferencias significativas.

La mayoría de las diferencias entre las dispersiones resultan menores, con una diferencia menor de 0,23 puntos, pero aparecen cuatro comparaciones con diferencias más acusadas, superando el 0,30 puntos sobre rangos de desviación que no superan los 2 puntos:

- Las mujeres jóvenes con estudios bajos mantienen una dispersión Pulse mayor que el colectivo de varones jóvenes con estudios más altos ( $1,88 - 1,46 = |0,42|$ ).
- Los varones adultos con menos estudios también mantienen mayor variabilidad en las puntuaciones Pulse que los varones jóvenes con estudios altos ( $1,81 - 1,46 = |0,35|$ ).
- Las mujeres jóvenes con estudios bajos varían más en sus puntuaciones que el mismo grupo con estudios altos ( $1,88 - 1,58 = |0,30|$ ).

Parece que las personas con menos estudios tienden a percibir más diferencias en la reputación que los colectivos con estudios más altos que tienden a uniformizar más sus puntuaciones.

Resumiendo las evidencias obtenidas sobre el grado de invarianza de los grupos homogéneos por sexo, edad y nivel de estudios en EEUU, tenemos que la invarianza



configural y de cargas han sido plenamente verificadas. No así en el caso de la dispersión de los valores del índice entre los grupos.

En el caso e EEUU, el índice Pulse ha mostrado tener invarianza de cargas. Podemos concluir que los distintos colectivos demográficos utilizan los mismos indicadores para formarse una idea del índice global de reputación Pulse y le dan el mismo significado a la aportación relativa de cada uno de ellos (invarianza de cargas). Ambos requisitos son imprescindibles para garantizar la comparabilidad plena de las medidas alcanzadas en dichos colectivos. En consecuencia, para EEUU, las medidas Pulse de estos colectivos son plenamente comparables.

Probablemente, para el caso norteamericano, resultaría conveniente ampliar el análisis de invarianza demográfica con la variable “raza” dado el carácter multiétnico de este país.

En los gráficos siguientes representamos los modelos de los ocho colectivos demográficos homogéneos para EEUU.

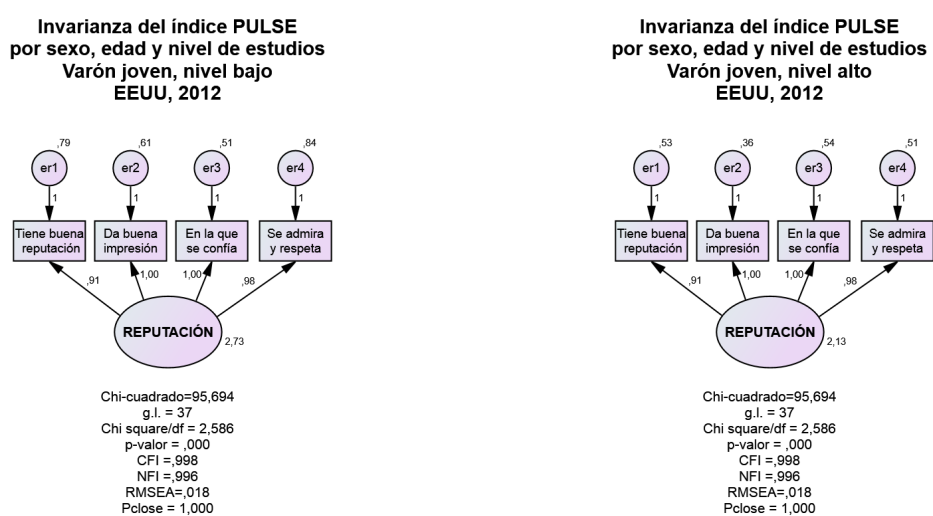
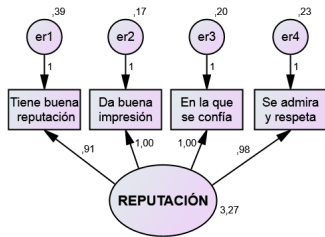


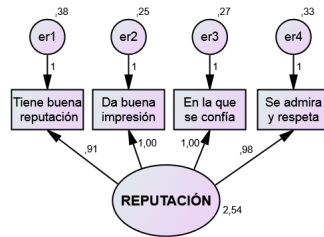
Figura nº 29. Modelos confirmados por los datos de la invarianza conjunta en EEUU.

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón adulto, nivel bajo  
EEUU, 2012**



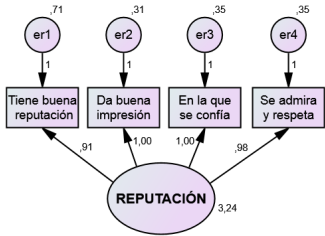
Chi-cuadrado=95,694  
g.l. = 37  
Chi square/df = 2,586  
p-valor = ,000  
CFI = ,998  
NFI = ,996  
RMSEA=.018  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón adulto, nivel alto  
EEUU, 2012**

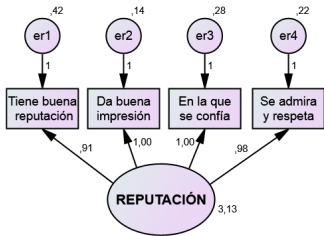


Chi-cuadrado=95,694  
g.l. = 37  
Chi square/df = 2,586  
p-valor = ,000  
CFI = ,998  
NFI = ,996  
RMSEA=.018  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel bajo  
EEUU, 2012**

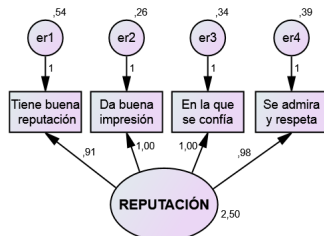


**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel bajo  
EEUU, 2012**

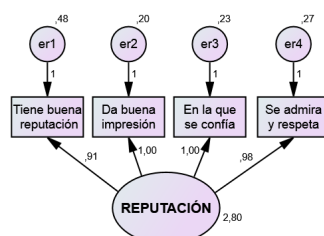


Chi-cuadrado=95,694  
g.l. = 37  
Chi square/df = 2,586  
p-valor = ,000  
CFI = ,998  
NFI = ,996  
RMSEA=.018  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel alto  
EEUU, 2012**

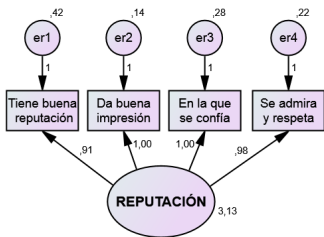


**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel alto  
EEUU, 2012**



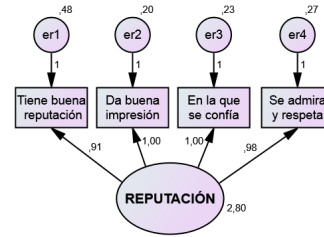
Chi-cuadrado=95,694  
g.l. = 37  
Chi square/df = 2,586  
p-valor = ,000  
CFI = ,998  
NFI = ,996  
RMSEA=.018  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel bajo  
EEUU, 2012**



Chi-cuadrado=95,694  
g.l. = 37  
Chi square/df = 2,586  
p-valor = ,000  
CFI = ,998  
NFI = ,996  
RMSEA=.018

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel alto  
EEUU, 2012**



Chi-cuadrado=95,694  
g.l. = 37  
Chi square/df = 2,586  
p-valor = ,000  
CFI = ,998  
NFI = ,996  
RMSEA=.018

#### 4.5.6.8. INVARIANZA CONJUNTA EN CHINA.

En el caso de China, el Pulse cumple con las propiedades de invarianza de varianza por grupos homogéneos por sexo y edad.

Para las audiencias de las empresas evaluadas en China, los ítems pesan los mismo para los diferentes grupos de sexo y edad, sus dispersiones son las mismas y la varianza no explicada en los 4 ítems también es la misma, indicando que los sentimientos asociados al Pulse mantienen una comunalidad similar.

**CHINA: INVARIANZA CONJUNTA POR SEXO Y EDAD**

Modelos	$\chi^2$	g.l.	P-valor B-S	CFI	RMSEA
Modelo configural	60,99	8	0,010	0,995	0,052
Igualdad de cargas	68,50	17	0,015	0,996	0,035
Igualdad varianza PULSE	74,54	20	0,005	0,995	0,033
Invarianza estricta	257,16	32	0,005	0,981	0,054

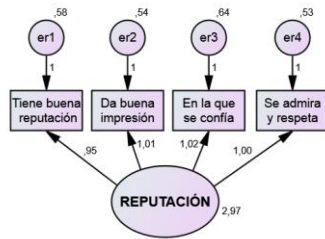
Valoración de la pérdida de ajuste	$\Delta\chi^2$	$\Delta$ g.l.	p-valor	$\Delta$ CFI	$\Delta$ RMSEA
Igualdad de cargas	7,5	9	0,584	0,001	-0,017
Igualdad de varianza PULSE	13,5	12	0,331	0,000	-0,019
invarianza estricta	196,2	24	0,000	-0,014	0,002

Tabla nº de los modelos anidados por grupos homogéneos de sexo, edad y estudios. China.

Aunque el modelo configural sale significativo, con un  $P\text{-valor}_{B-S} = 0,01$ , los índice  $CFI = 0,995$  y  $RMSEA = 0,052$ , cumplen con el umbral de aceptación. A partir de este modelo, los modelos anidados de igualdad de cargas y de igualdad de varianza no pierden ajuste significativo. Sólo el modelo de varianza estricta reúne evidencia de que es rechazado por los datos. En consecuencia, el índice Pulse por grupos homogéneos de sexo y edad en China, ha demostrado tener suficiente equivalencia métrica.

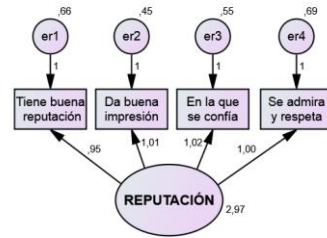
En la página siguiente representamos los diferentes modelos de invarianza de varianza Pulse por los cuatro grupos homogéneos.

**Invarianza del índice PULSE  
por sexo y edad  
Varón joven  
CHINA, 2012**



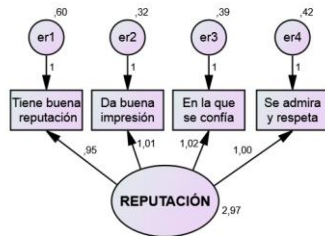
Chi-cuadrado=74,539  
g.l. = 20  
Chi square/df = 3,727  
p-valor = ,000  
CFI = ,995  
NFI = ,994  
RMSEA=,033  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo y edad  
Varón adulto  
CHINA, 2012**

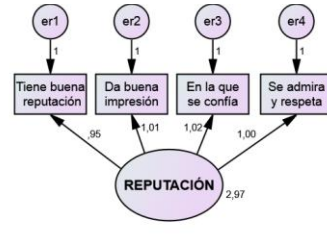


Chi-cuadrado=74,539  
g.l. = 20  
Chi square/df = 3,727  
p-valor = ,000  
CFI = ,995  
NFI = ,994  
RMSEA=,033  
Pclose = 1,000

**Mujer joven  
CHINA, 2012**



Chi-cuadrado=74,539  
g.l. = 20  
Chi square/df = 3,727  
p-valor = ,000  
CFI = ,995  
NFI = ,994  
RMSEA=,033



Chi-cuadrado=74,539  
g.l. = 20  
Chi square/df = 3,727  
p-valor = ,000  
CFI = ,995  
NFI = ,994  
RMSEA=,033  
Pclose = 1,000

Figura nº 30. Modelos confirmados por los datos de la invarianza conjunta en China.

**4.5.6.9. LA COMPARACIÓN DEL NIVEL DEL PULSE POR SEXO, EDAD Y NIVEL DE ESTUDIOS.**

Hasta ahora nos hemos focalizado en las propiedades métricas del índice Pulse. Los resultados anteriores han demostrado que el índice Pulse, formado reflexivamente por los cuatro ítems, cumple las propiedades psicométricas y ha demostrado ser una medida equivalente entre los diferentes grupos de sexo, edad y nivel de estudios. En consecuencia podemos decir que producirá medidas plenamente comparables a nivel demográfico. Dicho de otro modo, las variaciones del índice reflejarán variaciones reales del rasgo “reputación”, sea cual sea la composición poblacional analizada.

Establecida la invarianza poblacional de la métrica podemos pasar a comparar los niveles alcanzados por estos grupos. Este tipo de comparaciones resulta extremadamente útil porque señala la cantidad de reputación que llegan a percibir los distintos grupos de audiencias. Así podemos determinar si los jóvenes perciben más reputación que los mayores o bien si las mujeres tienden a percibir mejor reputación que los hombres. Sin embargo estas comparaciones no tendrían fundamento si la equivalencia anterior no hubiera sido verificada.

En el programa AMOS, la comparación de las medias latentes entre grupos se obtiene tomando un grupo como referencia que adopta el valor cero y sobre él se comparan las medias de los otros grupos. Así, en la comparación del sexo, los varones actuarán como referencia con el valor cero y se representará la media del grupo de mujeres. Si alcanza valores positivos significa que valoran mejor la reputación que los varones, por el contrario, las puntuaciones negativas indicarían si presenta una puntuación menor que la de los varones. Dado que la variable latente está asociada a un indicador por el 1, significa que la escala de la latente es la misma que la de dicho indicador, en nuestro caso del 1 al 7.

Como dijimos en el apartado metodológico correspondiente<sup>92</sup>, para comparar los niveles medios de cada grupo necesitamos comprobar la validez del modelo igualdad de cargas e igualdad de interceptores. A partir de él se somete a prueba la igualdad de las medias.

---

<sup>92</sup> “La invarianza de medias y la invarianza psicométrica”

#### 4.5.6.10. LOS NIVELES DEL PULSE POR SEXO Y PAÍS.

Para determinar si las mujeres valoran la reputación de las empresas de forma diferente, debemos determinar que ambos grupos son invariantes en cargas e interceptores en todos los países (tabla siguiente).

España: SEXO					
Modelos	$\chi^2$	g.l.	P-valor B-S	CFI	RMSEA
Invarianza de cargas e interceptores	28,1	9	<b>0,055</b>	0,999	0,022
Invarianza de medias	33,6	10	<b>0,029</b>	0,999	0,014
Invarianza estricta	140,5	15	<b>0,000</b>	0,995	<b>0,042</b>
Valoración de la pérdida de ajuste					
	$\Delta\chi^2$	$\Delta$ g.l.	p-valor	$\Delta$ CFI	$\Delta$ RMSEA
Invarianza de medias	5,4	1,0	<b>0,020</b>	0,000	<b>0,02</b>
EEUU: SEXO					
Modelos	$\chi^2$	g.l.	P-valor B-S	CFI	RMSEA
Invarianza de cargas e interceptores	66,3	10	<b>0,001</b>	0,998	1,000
Invarianza de medias	70,3	11	<b>0,001</b>	0,998	1,000
Invarianza estricta	190,5	15	<b>0,000</b>	0,993	0,663
Valoración de la pérdida de ajuste					
	$\Delta\chi^2$	$\Delta$ g.l.	p-valor	$\Delta$ CFI	$\Delta$ RMSEA
Invarianza de medias	3,9	1,0	<b>0,047</b>	0,000	<b>0,000</b>
China: SEXO					
Modelos	$\chi^2$	g.l.	P-valor B-S	CFI	RMSEA
Invarianza de cargas e interceptores	18,3	10	<b>0,181</b>	0,997	0,030
Invarianza de medias	27,4	11	<b>0,030</b>	0,995	0,040
Invarianza estricta	31,3	16	<b>0,134</b>	0,995	0,032
Valoración de la pérdida de ajuste					
	$\Delta\chi^2$	$\Delta$ g.l.	p-valor	$\Delta$ CFI	$\Delta$ RMSEA
Invarianza de medias	9,1	1,0	<b>0,003</b>	-0,002	<b>0,010</b>

Tabla nº 43. Comparación de modelos para valorar las medias por sexo.

En los tres países se verifica que el modelo de invarianza de cargas e interceptores es apropiado a los datos y que la restricción adicional de igualdad de las medias Pulse en ambos grupos es rechazada por los datos: España  $\Delta\chi^2_1 = 5,4$ ; p-valor = 0,020; EEUU  $\Delta\chi^2_1 = 3,9$ ; p-valor = 0,047 y China  $\Delta\chi^2_1 = 9,1$ ; p-valor = 0,003.

El gráfico siguiente representa la reputación global que despiertan las empresas del sector bancario y de las telecomunicaciones en los diferentes países, en los hombres y mujeres.

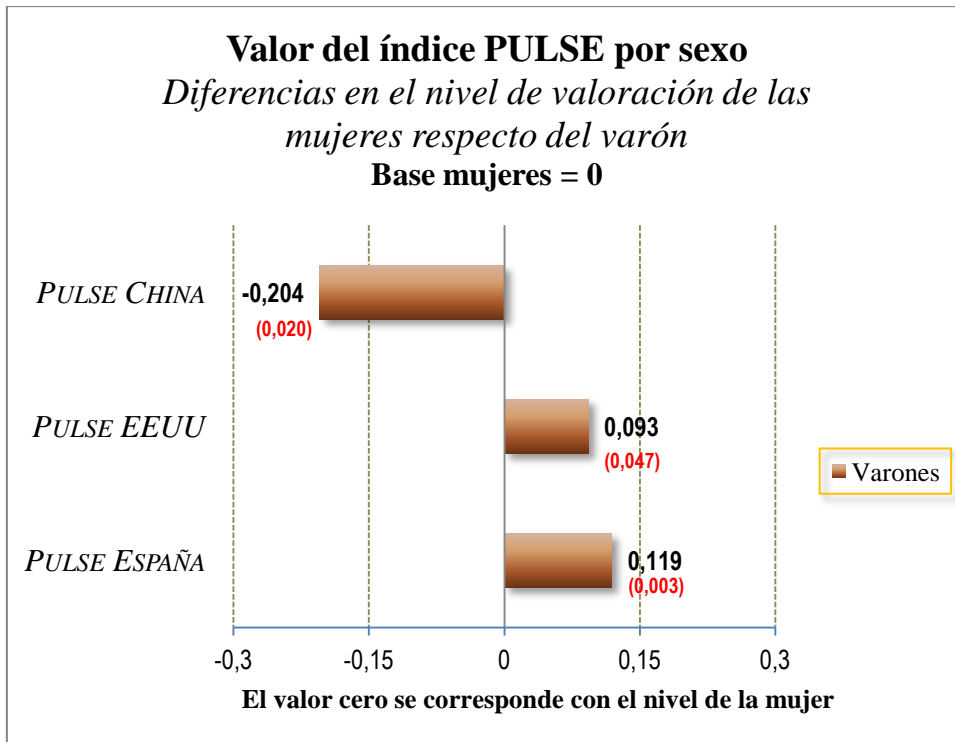


Figura nº 31. Valor del índice Pulse por sexo.

	Diferencia	E.S. dif	C.R.	P-valor
Pulse España	<b>0,119</b>	0,051	2,335	<b>0,020</b>
Pulse EEUU	<b>0,093</b>	0,047	1,985	<b>0,047</b>
Pulse China	<b>-0,204</b>	0,068	-3,02	<b>0,003</b>

Tabla nº 44. Test de las diferencias entre el Pulse por sexo.

En España y Estados Unidos las mujeres son ligeramente más críticas que los varones. Los hombres tienden a percibir una mejor reputación en las empresas que las mujeres. Sobre una escala de siete posiciones, aproximadamente unos 0,10 puntos más. Sin embargo en China, la situación es la contraria, la mujer china valora mucho más positivamente la reputación de las empresas que los varones.

Según el test del contraste de diferencia de medias, asociado al gráfico, las puntuaciones Pulse de las mujeres son significativamente diferentes a las de los hombres (p-valor < 0,05) y por tanto las diferencias son extrapolables a dichas poblaciones.

En el gráfico y tabla siguientes podemos apreciar cómo han afectado las percepciones diferenciales al nivel de los indicadores.

En España, las mujeres tienden a fiarse un poco menos ( $p$ -valor = 0,046), mantienen una peor impresión (0,045) y sobre todo admiran y respetan menos a las empresas evaluadas (0,002).

En EEUU, las mujeres mantienen una opinión más crítica en “buena impresión “ (0,014) y “admiración y respeto” (0,011). Por el contrario en China, las mujeres valoran mejor a las empresas en los cuatro indicadores, con diferencias significativas en todos ellos.

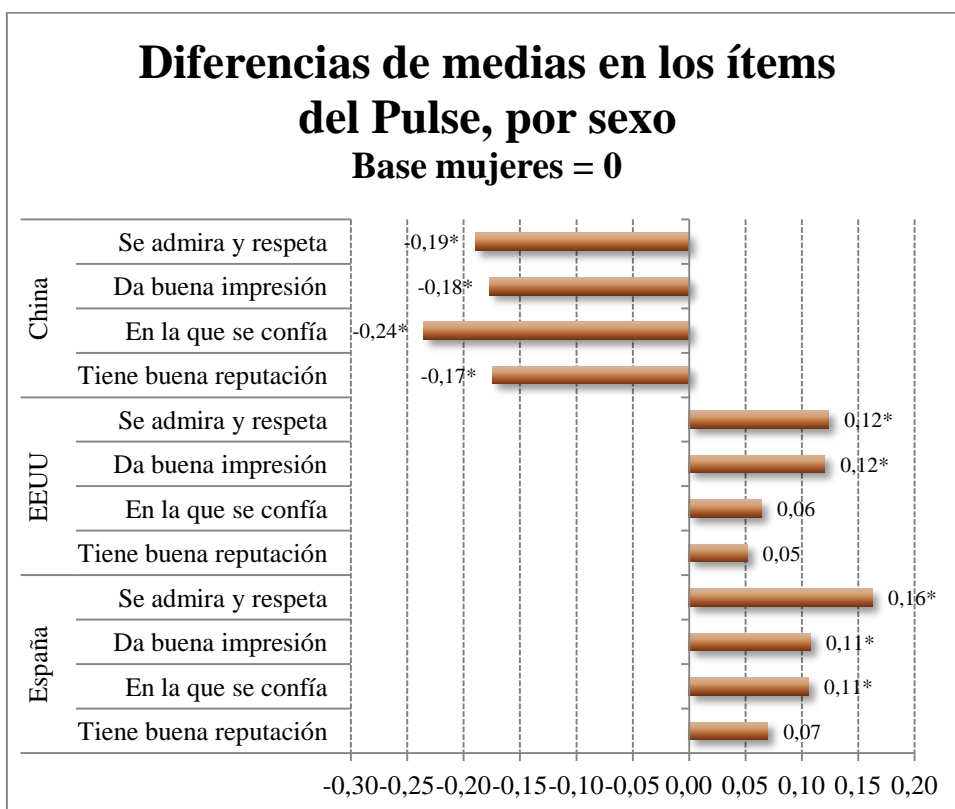


Figura nº 32 (\*) . Diferencia de medias por sexo. El asterisco indica diferencia significativas

Sexo	España				EEUU				China			
	Tiene buena reputación	En la que se confía	Da buena impresión	Se admira y respeta	Tiene buena reputación	En la que se confía	Da buena impresión	Se admira y respeta	Tiene buena reputación	En la que se confía	Da buena impresión	Se admira y respeta
Hombres	4,50	4,37	4,27	4,24	5,06	4,86	4,87	4,80	5,88	5,80	5,85	5,70
Mujeres	4,43	4,26	4,16	4,08	5,00	4,79	4,75	4,68	6,05	6,04	6,03	5,89
Diferencia	0,07	0,11	0,11	0,16	0,05	0,06	0,12	0,12	-0,17	-0,24	-0,18	-0,19
Sig.	0,176	0,046	0,045	0,002	0,265	0,183	0,014	0,011	0,009	0,001	0,013	0,012

Tabla nº 45 Test de contraste de las diferencias de medias en los indicadores.



#### 4.5.6.11. LOS NIVELES DEL PULSE POR EDAD Y PAÍS.

En los tres países, el modelo de invarianza de cargas e interceptores, se demuestra conforme a datos (España, p-valor  $B-S = 0,433$ ; EEUU, p-valor  $B-S = 0,082$ ; China , p-valor  $B-S = 0,240$ ). A partir de él examinamos la igualdad de las medias mediante el cálculo del nivel de significación debido a la pérdida de ajuste al forzar la igualdad de las medias.

En los casos de España y China no se han detectado diferencias en el nivel de reputación por grupos de edad. En España, la pérdida no es significativa ( $\Delta\chi^2_1 = 3,3$ ; p-valor =  $0,068 > 0,05$ ), pero además el test de ajuste global del modelo de igualdad de las medias nos estaba indicando que era un supuesto conforme a datos (p-valor  $B-S = 0,329$ ). Lo mismo se ha registrado en el caso de China, con una pérdida de ajuste no significativa ( $\Delta\chi^2_1 = 2,3$ ; p-valor =  $0,133$ ) y con el modelo de igualdad de medias empíricamente corroborado (p-valor  $B-S = 0,201$ ).

En España y China, los menores de 35 años han percibido el mismo nivel de reputación en las empresas evaluadas en ambos sectores que los más mayores. La edad no ha producido una diferencia en el nivel medio de reputación.

No es el caso de EEUU. En este país, la igualdad de las medias por grupos de edad ha sido rechazada por los datos. La pérdida de ajuste ha resultado significativa ( $\Delta\chi^2_1 = 61,4$ ; p-valor =  $0,002$ ). Además, en el análisis previo, el modelo de igualdad de medias ya estaba siendo rechazado por los datos (p-valor  $B-S = 0,002$ ) indicando que cuando forzamos la igualdad de las medias grupales por edad, no conseguimos reproducir la realidad de las matrices de varianzas, covarianzas y medias de los indicadores.

**España: EDAD**

Modelos	$\chi^2$	g.l.	P-valor B-S	CFI	RMSEA
Invarianza de cargas e interceptores	13,9	9	<b>0,433</b>	1,000	0,011
Invarianza de medias	17,2	10	<b>0,329</b>	1,000	0,012
Invarianza estricta	227,3	15	<b>0,000</b>	0,991	0,055

Valoración de la pérdida de ajuste	$\Delta\chi^2$	$\Delta$ g.l.	p-valor	$\Delta$ CFI	$\Delta$ RMSEA
Invarianza de medias	3,3	1,0	<b>0,068</b>	0,000	0,001

**EEUU: EDAD**

Modelos	$\chi^2$	g.l.	P-valor B-S	CFI	RMSEA
Invarianza de cargas e interceptores	30,1	10	<b>0,082</b>	0,999	0,020
Invarianza de medias	61,4	11	<b>0,002</b>	0,998	0,030
Invarianza estricta	487,2	15	<b>0,000</b>	0,981	0,079

Valoración de la pérdida de ajuste	$\Delta\chi^2$	$\Delta$ g.l.	p-valor	$\Delta$ CFI	$\Delta$ RMSEA
Invarianza de medias	31,3	1,0	<b>0,000</b>	-0,001	<b>0,010</b>

**China: EDAD**

Modelos	$\chi^2$	g.l.	P-valor B-S	CFI	RMSEA
Invarianza de cargas e interceptores	16,7	10	<b>0,240</b>	0,998	0,027
Invarianza de medias	19,0	11	<b>0,201</b>	0,997	0,028
Invarianza estricta	28,8	16	<b>0,181</b>	0,996	0,029

Valoración de la pérdida de ajuste	$\Delta\chi^2$	$\Delta$ g.l.	p-valor	$\Delta$ CFI	$\Delta$ RMSEA
Invarianza de medias	2,3	1,0	<b>0,133</b>	-0,001	<b>0,001</b>

Tabla nº 46. Comparación de modelos para valorar las medias por edad.

En el gráfico siguiente se perciben mejor estos resultados.

Las diferencias de nivel por edad en China y España no han resultado suficientemente destacadas como para detectar diferencias significativas, por lo que debemos aceptar que las diferencias en el gráfico para España de 0,094 puntos y para China 0,112 por lo que concluimos que las diferencias del gráfico siguiente para estos países no reflejan diferencias al nivel poblacional y debemos interpretarlas como un simple un error de muestreo.

En la figura siguiente muestra una situación distinta para EEUU. En este caso, el colectivo de jóvenes mantiene un nivel medio de reputación de las empresas superior al de los varones en 0,26 puntos en una escala de siete posiciones, lo que representa una diferencia altamente significativa (p-valor = 0,000).

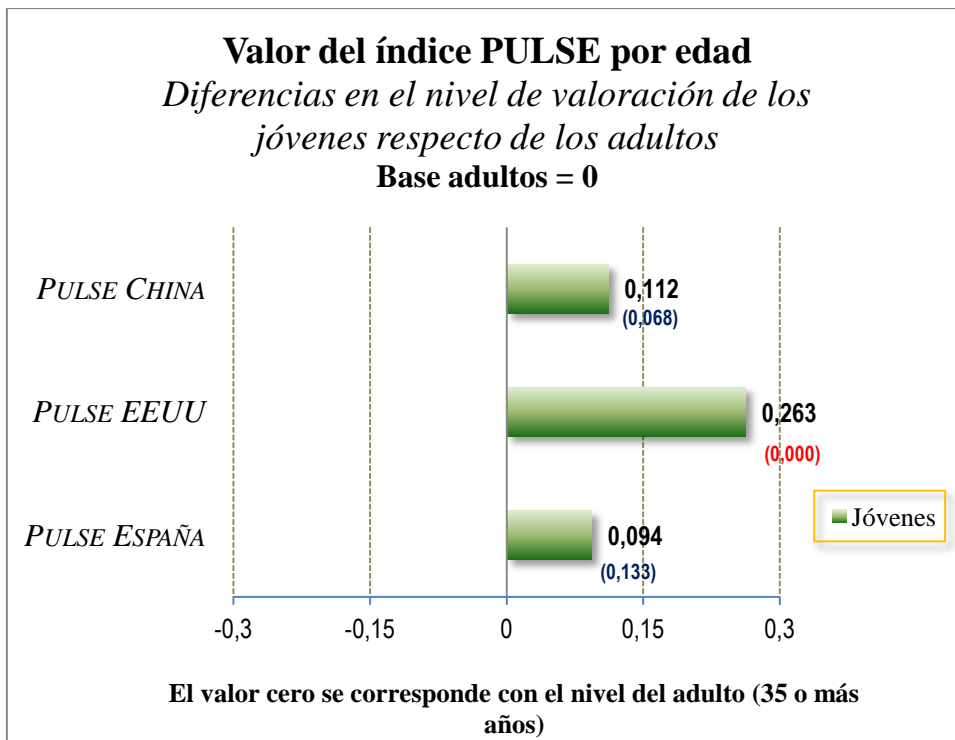


Figura nº 33. Valor del índice Pulse por edad.

El test de la diferencia de medias indica que mientras que en los casos de China y España, no se aprecian diferencias significativas en los niveles medios de jóvenes y adultos (p-valor diferencia  $_{China} = 0,068$  y p-valor diferencia  $_{España} = 0,133$ ), en el caso estadounidense los jóvenes se muestran significativamente más favorables que sus conciudadanos más adultos (p-valor diferencia  $_{EEUU} = 0,000$ ).

	Diferencia	E.S.dif	C.R.	P-valor
Pulse España	<b>0,094</b>	0,052	1,822	<b>0,068</b>
Pulse EEUU	<b>0,263</b>	0,047	5,601	<b>0,000</b>
Pulse China	<b>0,112</b>	0,075	1,503	<b>0,133</b>

Tabla nº 47. Test de las diferencias entre el Pulse por edad.

Las diferencias al nivel de los indicadores permiten comprender mejor los resultados alcanzados.

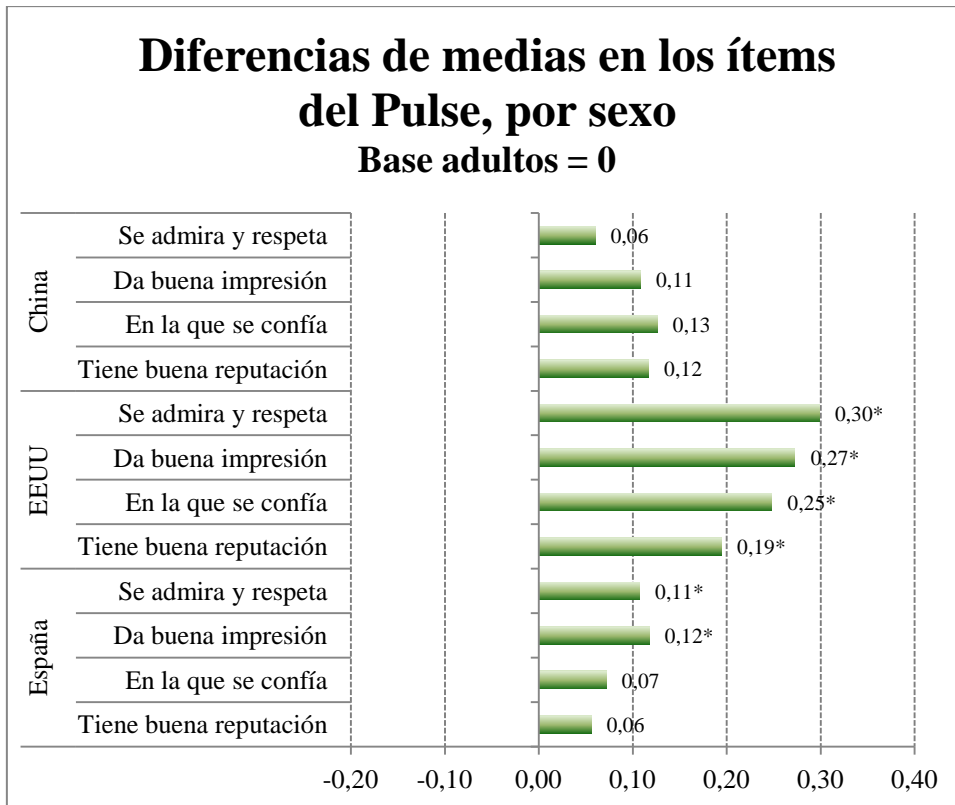


Figura nº 34. Diferencias de medias por sexo.

Edad	España				EEUU				China			
	Tiene buena reputación	En la que se confía	Da buena impresión	Se admira y respeta	Tiene buena reputación	En la que se confía	Da buena impresión	Se admira y respeta	Tiene buena reputación	En la que se confía	Da buena impresión	Se admira y respeta
Jóvenes	4,50	4,36	4,29	4,23	5,14	4,96	4,97	4,91	5,99	5,94	5,96	5,80
Adultos	4,44	4,29	4,17	4,12	4,95	4,72	4,69	4,61	5,87	5,81	5,85	5,74
Diferencia	0,06	0,07	0,12	0,11	0,19	0,25	0,27	0,30	0,12	0,13	0,11	0,06
Sig.	0,291	0,185	0,032	0,049	0,000	0,000	0,000	0,000	0,10	0,092	0,157	0,459

Tabla nº 48. Test de contraste de las diferencias de medias en los indicadores.

En el caso chino, se percibe con claridad que en los cuatro indicadores las medias entre colectivos e edad son prácticamente las mismas y en todo caso las diferencias no son significativas. Para España, los resultados indican diferencias significativas en dos indicadores, “buena impresión” (p-valor = 0,032) y “se admira y respeta” (p-valor = 0,049). En estos aspectos concretos, los más jóvenes mantienen una ligera mejor percepción que los más adultos, sin embargo, estas diferencias no se han hecho sentir en el sentimiento latente de la reputación (la diferencia de 0,09 a favor de los jóvenes se asocia a una probabilidad de error del 0,068, por encima del umbral de significación

del 0,05). Debemos concluir que buena parte de lo que vuelve significativos estos resultados se correspondería con la varianza del indicador que representa parte única o error, en todo caso, está fuera de la varianza común del sentimiento Pulse, que sabemos que no ha experimentado variación real.

Este resultado aboga a favor de la utilización de índices multi-indicador, para determinar con la mayor certeza variaciones reales en el fenómeno psicológico que se está observando y no en las errores de observación.

En el caso de EEUU, los cuatro indicadores han presentado diferencias significativas. Los jóvenes americanos perciben una reputación superior en los cuatro aspectos. Resulta de interés señalar que los tres indicadores más emocionales, como “admiración y respeto”, “buena impresión” y “confianza” presentan diferencias más acusadas. En este caso, el sentimiento global de reputación en este grupo es mejor.

#### 4.5.6.12. LOS NIVELES DEL PULSE POR ESTUDIOS Y PAÍS.

Recordemos que en el caso de estudios, el análisis se limita a España y EEUU, donde las distribuciones tienen suficiente tamaño como para facilitar la comparación.

España: NIVEL DE ESTUDIOS					
Modelos	$\chi^2$	g.l.	P-valor B-S	CFI	RMSEA
Invarianza de cargas e interceptores	13,4	9	<b>0,453</b>	1,000	0,0100
Invarianza de medias	17,5	10	<b>0,302</b>	1,000	0,013
Invarianza estricta	66,6	15	<b>0,017</b>	0,998	0,027
Valoración de la pérdida de ajuste					
	$\Delta\chi^2$	$\Delta$ g.l.	p-valor	$\Delta$ CFI	$\Delta$ RMSEA
Invarianza de medias	4,1	1,0	<b>0,043</b>	0,000	<b>-0,003</b>
EEUU: NIVEL DE ESTUDIOS					
Modelos	$\chi^2$	g.l.	P-valor B-S	CFI	RMSEA
Invarianza de cargas e interceptores	14,7	10	<b>0,522</b>	1,000	0,010
Invarianza de medias	14,8	11	<b>0,562</b>	1,000	0,008
Invarianza estricta	53,2	16	<b>0,104</b>	0,998	<b>0,022</b>
Valoración de la pérdida de ajuste					
	$\Delta\chi^2$	$\Delta$ g.l.	p-valor	$\Delta$ CFI	$\Delta$ RMSEA
Invarianza de medias	0,1	1,0	<b>0,727</b>	0,000	-0,002

Tabla nº 49. Comparación de modelos para valorar las medias por nivel de estudios.

En ambos casos, los modelos de invarianza de cargas e interceptores son adecuados a los datos (España, p-valor  $B-S = 0,453$  y en EEUU, p-valor  $B-S = 0,522$ ) y por tanto permite fundamentar el análisis comparativo de las medias de reputación por nivel de estudio. Según se desprende de la tabla anterior, las diferencias son significativas para España (p-valor = 0,017), pero para EEUU, cuyo test de pérdida de ajuste no resulta significativo (p-valor = 0,727) y además su modelo de igualdad de medias está respaldado por los datos (p-valor  $B-S = 0,562$ ). Los resultados son más fáciles de detectar en el gráfico siguiente.

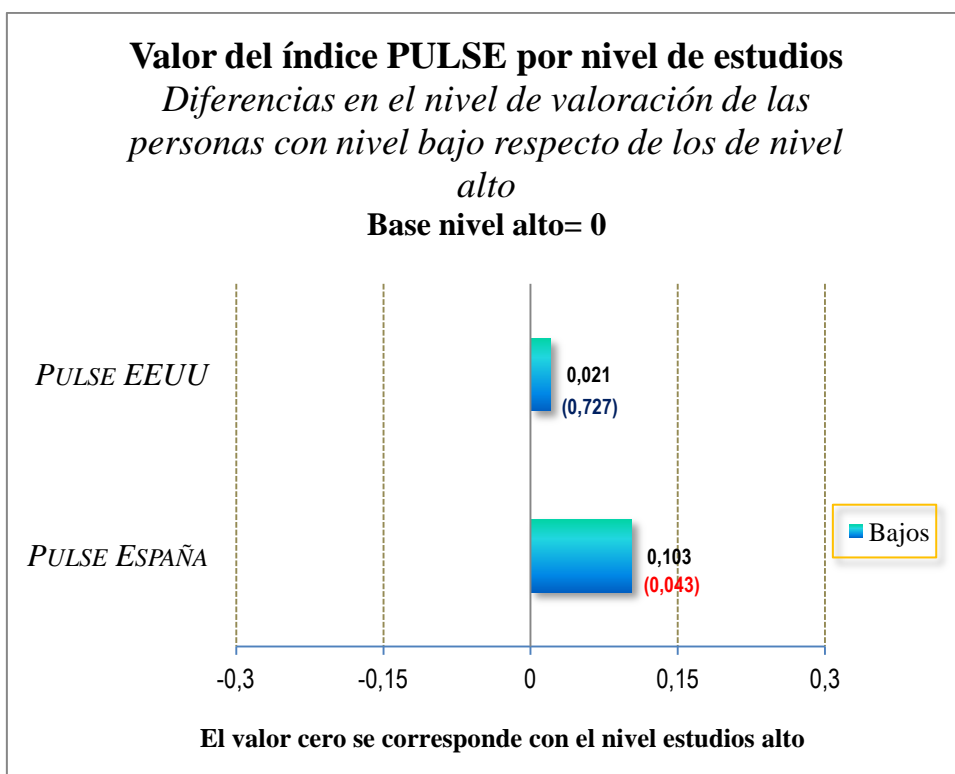


Figura nº 35. Diferencias por nivel de estudios.

	Diferencia	E.S.dif	C.R.	P-valor
Pulse España	0,103	0,051	2,022	<b>0,043</b>
Pulse EEUU	0,021	0,059	0,35	<b>0,727</b>

Tabla nº 50. Test de las diferencias entre el Pulse por nivel de estudios.

Mientras que el estadounidense con estudios altos mantiene el mismo nivel reputacional que sus conciudadanos con estudios inferiores, sobre las empresas valoradas en los sectores bancario y de telecomunicación (p-valor 0,727), en España,

las personas con estudios más altos tienden a ser globalmente más críticas con las empresas en ambos sectores ( $p\text{-valor} = 0,043 < 0,05$ ).

Al nivel de cada indicador, comprobamos que en el caso de EEUU, las diferencias no resultan significativas en ningún caso. Las diferencias de nivel en los cuatro ítems no resultan significativas.

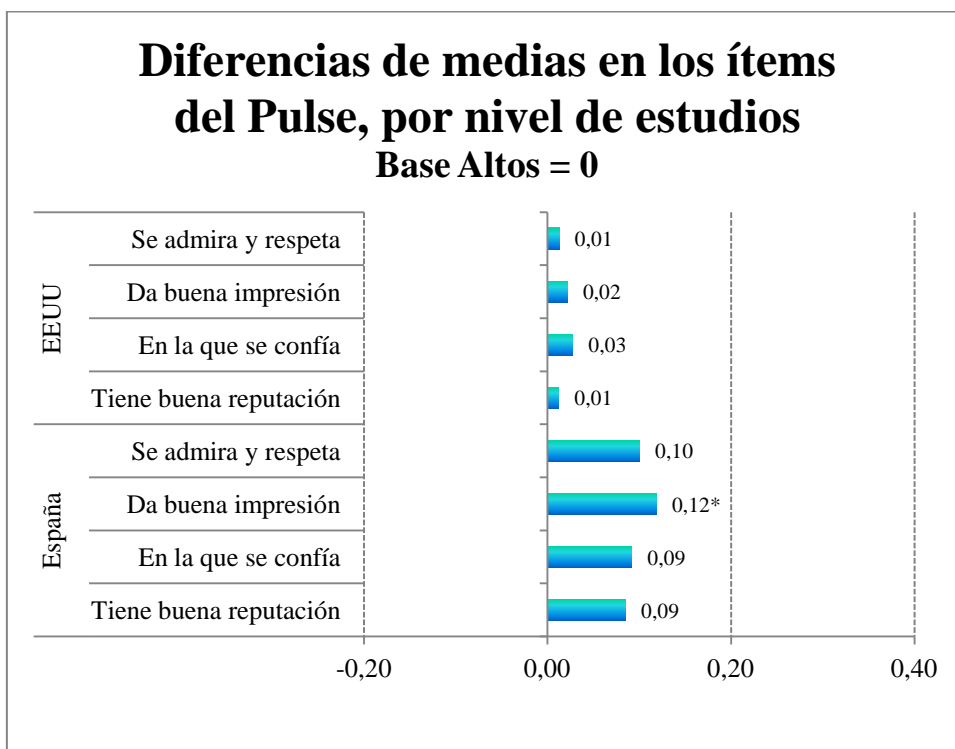


Figura nº 36. Diferencias en las medias por nivel de estudios.

Nivel de estudios	España				EEUU			
	Tiene buena reputación	En la que se confía	Da buena impresión	Se admira y respeta	Tiene buena reputación	En la que se confía	Da buena impresión	Se admira y respeta
Bajos	4,50	4,36	4,28	4,21	5,04	4,85	4,83	4,75
Altos	4,42	4,27	4,16	4,11	5,03	4,82	4,81	4,74
Diferencia	0,09	0,09	0,12	0,10	0,01	0,03	0,02	0,01
Sig.	0,099	0,082	0,027	0,059	0,819	0,629	0,709	0,821

Tabla nº 51. Test de contraste de las diferencias de medias en los indicadores.

En España, aunque el índice global sí resulta significativo, al nivel de los ítems, sólo el de “buena impresión” ha superado el umbral de significación ( $p$ -valor = 0,027). Una vez más se comprueba que las variaciones registradas en los ítems sólo reflejan parcialmente las variaciones reales en el sentimiento subyacente. Los resultados parecen indicar que la varianza de la unicidad ha tendido a homogeneizar sus diferencias, mientras que la varianza común parece ser que ha obrado la diferencia latente. Debemos concluir por tanto que en España, las personas con estudios más altos se muestran globalmente más críticas con las empresas de ambos sectores que aquellas con estudios inferiores.

En el gráfico siguiente ofrecemos una representación visual sintética de los resultados que hemos presentado en este apartado.



Invarianza poblacional (hipótesis principal de la investigación)	
Invarianza por sexo	H2.1
Invarianza por grupos de edad	H3.1
Invarianza por nivel de estudios	H4.1
Invarianza por grupos homogéneos por sexo, edad y nivel de estudios	H5.1

Gráfico nº 12. Formulación de la verificación de la invarianza poblacional del Pulse.



Los resultados de las pruebas de verificación estadísticas y matemáticas nos permitan ahora afirmar que se han verificado favorablemente la hipótesis principal así como las 2 sub-hipótesis que han sido la base de esta investigación:

Hipótesis principal:

*" Pulse mide la reputación global de una empresa u organización con independencia del sexo, edad o clase social de la población que sea entrevistada sobre esta cuestión" (invarianza poblacional).*

Sub-hipótesis:

- 1. "La puntuación resultante de Pulse refleja fielmente la fortaleza de esa reputación" (equivalencia y fortaleza métricas).*
- 2. "Los individuos de una población entienden lo mismo por reputación" (equivalencia semántica y contenido emocional).*

En el capítulo siguiente veremos la síntesis de las pruebas matemáticas y estadísticas utilizadas para la verificación de las hipótesis así como el apartado concreto de contraste de hipótesis que relaciona cada hipótesis con las pruebas de verificación empleadas en esta investigación.

## CAPÍTULO 5. SÍNTESIS DE LAS PRUEBAS MATEMÁTICAS Y ESTADÍSTICAS Y CONTRASTE DE LAS HIPÓTESIS

### 5.1. SÍNTESIS DE LAS PRUEBAS DE VERIFICACIÓN MATEMÁTICAS Y ESTADÍSTICAS.

Las pruebas de verificación matemáticas y estadísticas principales que enunciamos al principio de este trabajo de investigación sobre el índice Pulse han quedado corroboradas. Hemos comprobado:

- *La hipótesis matemática H1.1 de uni-dimensionalidad del índice Pulse ha sido corroborada en los tres contextos*
- *La hipótesis matemática H1.2 de convergencia de los indicadores ha sido corroborada (así como las sub-hipótesis H1.2.1; H1.2.2; H1.2.3.;H1.2.4) en los tres contextos*
- *La hipótesis matemática H1.3 de relación co-genérica entre los indicadores y prevalencia emocional ha quedado demostrada en los tres países.*

Contraste de hipótesis matemáticas y estadísticas en detalle:

**Hipótesis matemática H0 de adecuación métrica** del índice Pulse a tres países/contextos culturales distintos: se verifica totalmente la hipótesis para China y EEUU. En España se verifica parcialmente puesto que se detecta un matiz de interpretación semántica algo diferente para uno de los 4 ítems del Pulse. Esta diferencia de interpretación no invalida la hipótesis de adecuación métrica para España.

Hipótesis matemática H1 de las propiedades métricas de Pulse:

- Hipótesis matemática H1.1 de fiabilidad y validez convergente: corroborada
- Hipótesis matemática H1.2. de unidimensionalidad: corroborada
- Hipótesis matemática H1.3. de equivalencia semántica y prevalencia emocional: se rechaza la hipótesis de tau-equivalencia (los 4 ítems reflejan de forma idéntica la variable latente de reputación) y se acepta la hipótesis co-genérica (algunos ítems reflejan mejor esta variable latente). Los ítems que reflejan mejor la reputación son los 3 ítems de contenido más emocional (admiración, confianza y buenas impresiones) y menos en el ítem más racional (buena

reputación). Por tanto la equivalencia semántica y la prevalencia emocional del índice Pulse quedan corroboradas.

Hipótesis H2 matemática de invarianza poblacional del Pulse:

- Hipótesis matemática H2.1. Se ha demostrado la invarianza por sexo.
- Hipótesis matemática H3.1. Se ha contrastado positivamente la invarianza por edad.
- Hipótesis matemática H4.1. Se ha comprobado la invarianza por nivel de estudios.
- Hipótesis matemática H5.1. Ha quedado demostrada la plena invarianza poblacional de Pulse, ha superado la prueba fuerte de invarianza por colectivos homogéneos de sexo, edad y nivel de estudios en los tres países.

De esta forma se ha verificado la hipótesis principal de la investigación sobre la propiedad de invarianza poblacional que posee el índice Pulse.

Enumeramos a continuación los detalles de los resultados de contraste de cada una de las pruebas matemáticas y estadísticas que hemos utilizado para verificar o refutar la invarianza poblacional de Pulse. Cabe destacar que aunque algunas pruebas específicas han sido rechazadas, estos resultados no afectan a la verificación y contraste positivo de la hipótesis principal sobre la invarianza poblacional de Pulse.

Los detalles de contraste de todas las pruebas matemáticas y estadísticas son los siguientes:

- *H2.1. Invarianza configural por sexo: comprobada.*
- *H2.2. Invarianza de cargas por sexo corroborada.*
- *H2.3. Invarianza de la varianza extractada por el factor Pulse se comprueba en China y España pero queda rechazada para EEUU.*
- *H2.4. Invarianza de los residuos de los ítems por sexo: comprobada.*
- *H3.1. Invarianza configural por edad: comprobada.*
- *H3.2. Invarianza de cargas por edad: validada.*
- *H3.3. Invarianza de la varianza extractada por el factor Pulse: corroborada.*
- *H3.4. Invarianza de los residuos de los ítems por edad (o invarianza estricta), ha quedado rechazada para España y EEUU y comprobada en China.*
- *H4.1. Invarianza configural por nivel de estudios: comprobada.*

- *H4.2 . Invarianza de cargas: comprobada.*
- *H4.3. Invarianza de la varianza extractada por el factor Pulse: validada.*
- *H4.4 La invarianza de los residuos de los ítems por nivel de estudios(o invarianza estricta) ha quedado rechazada.*
- *H5.1. La invarianza configural por grupos homogéneos de sexo, edad y nivel de estudios ha sido corroborada.*
- *H5.2. La invarianza de cargas por grupos homogéneos ha sido aceptada.*
- *H5.3. La invarianza de la varianza del factor Pulse por grupos homogéneos de sexo, edad y estudios ha sido corroborada.*
- *H5.4. La invarianza de los residuos de los indicadores Pulse por grupos homogéneos ha sido rechazada.*

## 5.2. CONTRASTE DE LAS HIPÓTESIS.

Como hemos visto, los resultados de las pruebas de verificación estadísticas y matemáticas nos permitan ahora afirmar que se han verificado favorablemente la hipótesis principal así como las 2 sub-hipótesis que han sido la base de esta investigación:

*" Pulse mide la reputación global de una empresa u organización con independencia del sexo, edad o clase social de la población que sea entrevistada sobre esta cuestión" (invarianza poblacional).*

Sub-hipótesis:

*1. "La puntuación resultante de Pulse refleja fielmente la fortaleza de esa reputación" (equivalencia y fortaleza métricas).*

*2. "Los individuos de una población entienden lo mismo por reputación" (equivalencia semántica y contenido emocional).*

En concreto la hipótesis principal se sometió a las pruebas matemáticas y estadísticas que denominamos

La hipótesis principal: *" Pulse mide la reputación global de una empresa u organización con independencia del sexo, edad o clase social de la población que sea entrevistada sobre esta cuestión" (invarianza poblacional).*

Se corroboró esta hipótesis principal de nuestra investigación mediante las pruebas matemáticas y estadísticas a las que dimos el nombre de *hipótesis matemáticas H.2.1; H.3.1; H4.1; H.5.1*. Estos análisis verificaron la invarianza poblacional de Pulse por sexo, edad y clase social/nivel de estudios en los tres países.

La sub-hipótesis 1: *"La puntuación resultante de Pulse refleja fielmente la fortaleza de esa reputación" (equivalencia y fortaleza métricas).*

Se verificó esta sub-hipótesis mediante las pruebas matemáticas y estadísticas que denominamos *Hipótesis matemática H0* que midió la adecuación métrica de Pulse en tres países/contextos culturales diferentes.

La sub-hipótesis 2: "*Los individuos de una población entienden lo mismo por reputación*" (*equivalencia semántica y contenido emocional*).

Se verificó esta sub-hipótesis sobre las propiedades métricas de Pulse gracias a los resultados positivos los análisis matemáticos y estadísticos que denominamos *hipótesis matemáticas H1.1; H1.2, H1.3*. Las tres hipótesis matemáticas verificaron que Pulse es un índice unidimensional (H1.1.), tiene validez convergente (H1.2.) y tiene equivalencia semántica y mayor peso emocional (H.1.3.).

### **5.3. LIMITACIONES DE LOS RESULTADOS DEL CONTRASTE DE LAS HIPÓTESIS.**

En general, tanto las hipótesis principales y sub-hipótesis de la investigación han superado los contrastes matemáticos y estadísticos, sin embargo conviene señalar algunas limitaciones en los resultados debidas a aspectos técnicos e inherentes también al diseño empírico de la investigación y a la disponibilidad limitada de las bases de datos (países y sectores de actividad) utilizadas.

El estudio está limitado a grandes empresas privadas pertenecientes a dos sectores de actividad, el sector bancario y de las telecomunicaciones. Convendría comprobar si las propiedades de invarianza se verifican en otros sectores, en empresas del sector público y de menor dimensión.

Los tamaños de muestra han sido tal vez excesivamente elevados, sobre todo para España y para EEUU. Esto ha permitido fundamentar mejor el grado de invarianza, pero ha sembrado algunas dudas sobre la importancia práctica de estas diferencias.

La no invarianza en reputación detectada por sexo en el caso norteamericano, puede que sea real pero prácticamente despreciable (el elevado tamaño de muestra hace saltar los test de significación). En el caso chino, ocurre un efecto similar en los colectivos demográficos homogeneizados.

El estudio se ha basado en el índice Pulse, convendría replicarlo con otros índices globales como el de Highhouse y colegas (2009) o el de nuestros colegas españoles Casaló, Flaviá y Guinaliú (2008) para demostrar la invarianza poblacional se mantiene más allá de los ítems del Pulse.

### 6.1. CONCLUSIONES PRINCIPALES.

Los resultados positivos de esta investigación son relevantes para el mundo de la gestión de la reputación en el que resulta esencial que las métricas, en este caso el índice de reputación global Pulse reflejen lo más fielmente posible las variaciones reales del fenómeno que se está evaluando, máxime cuando se trata de un recurso intangible no directamente observable como es el caso de la reputación corporativa.

El gestor necesita de un indicador global de reputación como el que hemos puesto a prueba en esta investigación para adoptar las iniciativas correctas que le lleven tomar decisiones empresariales relevantes para la protección o el fortalecimiento de la reputación de su organización. Un indicador agregado como el que hemos investigado que le permita establecer un diagnóstico del estado de la reputación (obtener una puntuación en una escala numérica) en un momento concreto del tiempo y la trayectoria de su evolución a lo largo del tiempo y siempre con la posibilidad de comparar los resultados de su organización frente a los de la competencia utilizando la misma métrica.

La reputación global fue definida por Ponzi y colegas (2011) como un indicador emocional que expresa los sentimientos de admiración, respeto, confianza y buenas sensaciones que despiertan las empresas en un sujeto.

El índice global de reputación es una métrica que expresa los sentimientos y evaluaciones que los distintos *stakeholders* hacen sobre una empresa, institución o país/ciudad. Este indicador global de reputación debería ser capaz predecir la actitud intencional, que es la responsable de las conductas de apoyo o de rechazo hacia el objeto evaluado, que es la organización reputada.

Este indicador global, hemos visto que posee un carácter sumario o sintético que, gracias al contraste de las hipótesis de esta investigación, lo hace especialmente válido para su uso en la gestión organizativa porque representa un disparador de considerable impacto sobre de las conductas de valor o de destrucción de valor hacia dichas



empresas (Carreras et al. 2013, 286-288), en este sentido se podría afirmar que existe una relación directa entre el nivel que alcanza esta reputación global y el negocio/la generación de valor.

También hemos hablado de que este doble carácter, de métrica sintética y estratégica que lo convierte en un indicador fundamental de un recurso intangible, como la reputación, que podría ser incorporado a los cuadros de mando de las organizaciones, complementando a los indicadores financieros clave que ya se utilizan en todos los cuadros de mando para controlar la estrategia de la organización y establecer los mecanismos y baremos de seguimiento, control y compensación salarial derivados del grado de cumplimiento de dicha estrategia.

En este sentido, el indicador de la reputación global que hemos puesto a prueba ha comprobado su potencialidad para convertirse en una métrica adecuada para ayudar a implementar una gestión más equilibrada de las organizaciones, ya que toma en cuenta tanto los indicadores financieros ya existentes como este nuevo indicador no financiero que juntos permiten realizar un seguimiento del valor agregado de la empresa (valor tangibles y valor intangible), en el sentido que le dan al cuadro de mando estratégico Kaplan y Norton, en la medida en que su uso por parte de la alta dirección facilita el alineamiento de una parte significativa de los procesos de la organización para conseguir sus objetivos de negocio, de creación de valor compartido y de distribución más equilibrada de ese valor entre los distintos *stakeholders* de la organización (Kaplan y Norton, 1996, 85).

En el cuadro siguiente presentamos un caso real de aplicación a una entidad financiera española (no es posible desvelar el nombre del banco) del cuadro de mando con indicadores de intangibles clave, desarrollado conceptualmente por Ángel Alloza y el equipo de investigación de Corporate Excellence y cuyos resultados matemáticos y efectos económicos han sido estimado por el equipo de investigación de "*Conento Advanced Analytics*" dirigido por Macarena Estévez. Este cuadro de mando demuestra el impacto positivo del indicador de reputación Pulse que, en este caso concreto, explica en un 32.2% la formación de las actitudes favorables que desencadenan comportamientos positivos hacia el banco y que se traducen finalmente en un incremento en el negocio (medido aquí como resultado de los recursos de explotación

del banco). Estos resultados empíricos que han evidenciado el efecto de Pulse en el incremento del negocio, han sido previos a la realización de la investigación de esta tesis doctoral.

En la figura siguiente presentamos a título puramente ilustrativo un cuadro de mando en el que se incorporó el índice Pulse como indicador global de reputación del banco. Los resultados de la investigación empírica mostraron la relación causal entre Pulse y las actitudes clave de los consumidores y el impacto directo de estas actitudes en el negocio del banco (medido como incremento de los recursos de explotación)

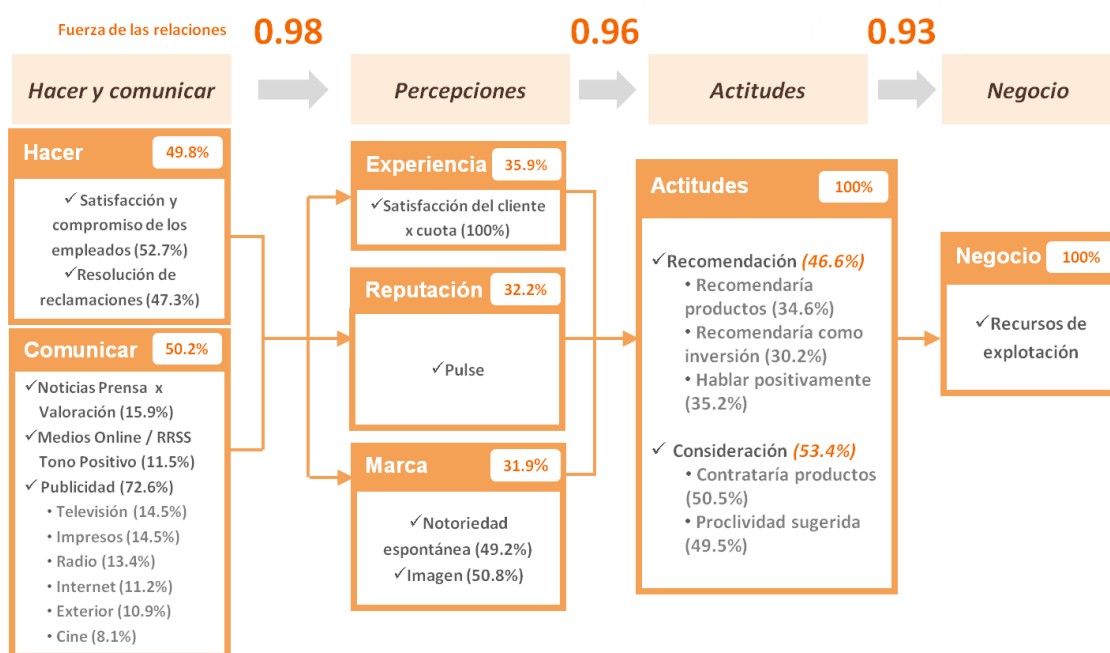


Figura nº 37. Cuadro de mando conceptualizado por Ángel Alloza y el equipo de investigación de Corporate Excellence. Fuente: (Estévez, M. et al. 2014).

De esta forma el indicador global de reputación Pulse impactaría directamente en la lealtad psicológica de los clientes (actitudes clave) que desencadenaría comportamientos favorables hacia la empresa como podemos ver en el modelo propuesto por Carreras y Alloza que presentamos en el gráfico siguiente (Carreras, E. et al. , 2013):

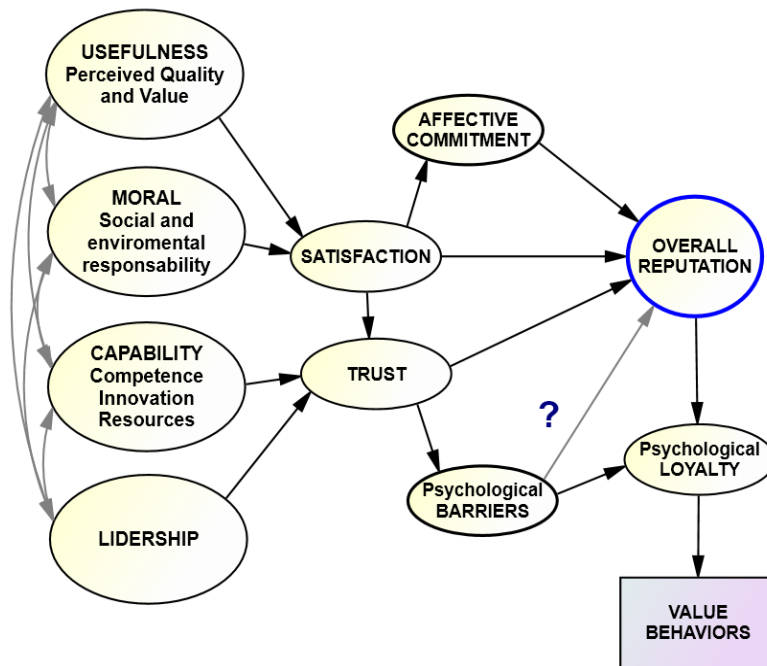


Figura nº 38. Fuente: Procede de "Reputación Corporativa", Carreras, Alloza, y Carreras, 2013,320.

Contar con un indicador global de reputación supone para las organizaciones que deciden adoptarlo para la gestión de la estrategia empresarial introduce una perspectiva que (1) concilia la consecución de objetivos a corto plazo sin comprometer los de medio y largo plazo, (2) permite pasar de una perspectiva "mono-stakeholder" (foco exclusivo en la maximización del valor para el accionista) a una "multi-stakeholder" (Retolaza, et al. 2014), (3) transforma a la organización para crear valor compartido (Porter y Kramer 2011) y (4) sirve de forma equilibrada las expectativas y necesidades de todos sus grupos de interés (Bartikowski y Walsh 2011; Brown y Perry 1994; Casaló et al. 2008; Fombrun 1996; Fuente-Sabaté y Quevedo-Puente 2003; Flanagan et al. 2011; Retolaza y San José, 2011).

Por el contrario las organizaciones que fijan la estrategia, su seguimiento y los sistemas de compensación únicamente con indicadores financieros generan un sesgo claro hacía la consecución de objetivos de corto plazo y favorecen los problemas de *agency* que implican a los intereses de los directivos responsables de la ejecución de la estrategia no están alineados con los del resto de *stakeholders*.

La incorporación de un indicador no financiero como el de la reputación global Pulse podría ayudar a conciliar y alinear los intereses de los directivos, propietarios y resto de *stakeholders* puesto que crecer en ese indicador de reputación supone establecer una estrategia de largo plazo, relaciones de confianza con los grupos de interés y por tanto un foco *multi-stakeholder* que no poseen las empresas y organizaciones regidas por indicadores exclusivamente financieros (Barton, 2011).

**LA CONSECUENCIA DE HABER DEMOSTRADO LA INVARIANZA POBLACIONAL DEL ÍNDICE PULSE PARA LA GESTIÓN DE LA REPUTACIÓN EN LAS ORGANIZACIONES.**

En terminología técnica y matemática la hipótesis principal y las sub-hipótesis de esta investigación se pueden formular de la siguiente forma:

La investigación ha intentar demostrado que Pulse cumple con la hipótesis de **"invarianza poblacional "**.

Debido a que el índice Pulse adopta la forma de una variable latente evaluada por la combinación lineal de esos cuatro atributos, dispuestos en forma reflexiva <sup>93</sup>, hemos conseguido probar que Pulse cumple con dos requisitos matemáticos:

- La escala de Pulse cumple con la condición de uni-dimensionalidad interna.
- La escala de Pulse tiene validez convergente

Estos resultados de nuestra investigación complementan los presentados recientemente Leonard Ponzi que demostraban la validez del índice Pulse en contextos culturales muy diferentes (Ponzi et al 2011). Sin embargo, como señalan en un texto reciente, Enrique Carreras, Ángel Alloza y Ana Carreras, la invarianza poblacional del índice no había quedado demostrada (Carreras et al. 2013). En efecto, de poco sirve validar un índice en una encuesta en un momento concreto del tiempo si no demostramos que,

---

<sup>93</sup> En los apartados “las propiedades psicométricas del índice reflexivo Pulse” y “El sistema completo: Los sub-sistemas de medida y estructural”, desarrollamos este concepto.

en las sucesivas oleadas de tomas de información, las variaciones en las puntuaciones del índice reflejarán lo más fielmente posible variaciones en la actitud global que expresa el Pulse y no variaciones debidas a cambios en la composición de la muestra por sexo, edad o clase social.

Este fue uno de los motivos que llevaron a Alloza a poner a prueba el índice Pulse mediante el análisis de **invarianza poblacional** que se ha empleado para despejar esta incógnita, con el último fin de garantizar al gestor de la reputación que las variaciones en el indicador global de reputación están razonablemente libres de ambigüedades procedentes de posibles interpretaciones diferentes por parte de distintos segmentos de la población que podrían dar más importancia a alguno de los cuatro atributos del Pulse que a otros.

Si no se hubiera verificado la hipótesis de invarianza poblacional eso hubiera implicado una limitación importante en el índice global de reputación. Significaría que estaríamos midiendo sentimientos y valoraciones diferentes, según el colectivo socio demográfico implicado y en ese caso no sabríamos determinar si la gestión reputacional está realmente modificando la actitud de los *stakeholders* o bien es un reflejo de sensibilidades distintas.

Como han señalado autores de reconocido prestigio en el campo de la invarianza<sup>94</sup>, “la ambigüedad de tales (medidas) deriva de la falta de seguridad en saber si las diferencias son debidas a diferencias actitudinales verdaderas o más bien a diferencias psicométricas relacionadas con las respuestas a los atributos”. Despejar esta fuente de incertidumbre es una necesidad prioritaria desde la perspectiva de la gestión práctica de la reputación en las organizaciones.

---

<sup>94</sup> Cita que recogen Byrne y Stewart, del artículo de Cheung y Rensvold (2002) ambos en la bibliografía: The ambiguity of such interpretations arises from the lack of definitiveness in knowing whether the differences are due to true attitudinal differences or, rather, to psychometric differences related to the item responses (Byrne y Stewart 2006: 288)

**LA IMPLICACIÓN DE HABER DEMOSTRADO LA HIPÓTESIS SOBRE LA VALIDEZ DEL ÍNDICE GLOBAL EN DIFERENTES PAÍSES Y SECTORES DE ACTIVIDAD.**

Además del análisis de invarianza, la tesis ha querido abordar otra de las cuestiones clave para la gestión de la reputación en empresas e instituciones de diferentes sectores de actividad y con presencia internacional en numerosos países.

En este sentido la investigación ha servido para demostrar las hipótesis que nos hicieron someter a prueba al índice Pulse para evaluar hasta qué punto cumplía con:

- Las propiedades psicométricas del Pulse, en concreto la validación de la implantación del índice en los tres contextos culturales (España, EEUU y China).
- La validez convergente de los cuatro atributos
- La fiabilidad de la medida
- El tipo de relación entre los cuatro atributos
- El índice global (Pulse) se construye en modo reflexivo a partir de los cuatro atributos.

Nuestra investigación corrobora y amplía los resultados alcanzados por Leonard Ponzi, Charles Fombrun y Naomi Gardberg unos tres años atrás, cuándo demostraron que el índice Pulse es válido para contextos culturales distintos, con índices de fiabilidad y validez demostrados en diferentes sectores productivos como el transporte, el farmacéutico y el energético (Ponzi et al. 2011).

Nuestro estudio incorpora dos nuevos países entre los 38 citados por Ponzi, como son España y China en los sectores de empresas de telecomunicación y banca. Los índices de fiabilidad basados en la consistencia interna como el alfa de Cronbach y el de la fiabilidad del constructo de Dillon-Goldstein, oscilan prácticamente en la misma banda, entre 0,93 y 0,96 para España, EEUU y China y en el rango de 0,93 al 0,98 encontrados en el estudio de Ponzi y colegas.

En nuestra investigación hemos añadido el “AVE” o promedio de varianza extractada por el índice en cada país, oscilando entre el 0,77 para China y el 0,86 para EEUU, en todo caso superando ampliamente el umbral del 0,5 propuesto por sus mentores

(Fornell y Lacker 1981). También hemos verificado la fiabilidad de cada ítem identificada en los términos de Bagozzi y Yi (1988; 2012) como la cantidad de varianza del ítem que refleja el Pulse. En los tres contextos superaron el 0,78.

**EL INTERÉS DE HABER DEMOSTRADO LA HIPÓTESIS SOBRE LA VALIDEZ DEL ÍNDICE GLOBAL COMO INDICADOR SINTÉTICO DE REPUTACIÓN EN LOS CUADROS DE MANDO.**

Para completar el análisis de convergencia, controlamos el grado en que el Pulse se puede considerar como una métrica uni-dimensional. En los tres países, ha quedado demostrado que el Pulse es un índice uni-dimensional, con una varianza extractada por el primer factor superior al 0,86. Aunque los autores del Pulse (Ponzi et al., 2011) no citan esta propiedad a nosotros nos parece importante porque, desde el punto de vista de la gestión de la reputación en las organizaciones, resulta fundamental conocer cuántas dimensiones subyacen en un índice para evaluar con mayor precisión sus variaciones y utilizar este índice global de reputación como el KPI clave y único para incorporar la reputación corporativa en los cuadros de mando de las organizaciones.

Ahora sabemos por los resultados de nuestra investigación que Pulse expresa y sintetiza realmente la reputación global de una organización, institución, país, etc.

**LA CONSECUENCIA DE HABER DEMOSTRADO LA HIPÓTESIS SOBRE LA RELACIÓN ENTRE EL ÍNDICE GLOBAL Y LOS COMPORTAMIENTOS FAVORABLES DE LOS *STAKEHOLDERS* (RELACIÓN ENTRE REPUTACIÓN Y CREACIÓN DE RIQUEZA).**

Nos ha parecido necesario, de cara a la gestión empresarial, verificar también el tipo de relación que mantiene cada uno de los cuatro atributos con el índice global porque así podemos poner a prueba el contenido semántico del índice. Hemos utilizado el método del modelo co-genérico que supone una contribución diferenciada de los cuatro atributos del Pulse con el tau-equivalente, que propone la misma contribución, lo que significaría que cada uno de los atributos representa el Pulse del mismo modo.

Este último supuesto ha sido rechazado por los datos. La contribución del atributo “buena reputación” es menor que la de los otros tres indicadores.

Este resultado confirma la naturaleza emocional del Pulse propuesta por las investigaciones previas de otros autores (Ponzi et al 2011: 16). Esta naturaleza emocional del índice es clave y necesaria para desencadenar comportamientos favorables y de "lealtad" por parte de los grupos de interés (Carreras, et al. 2013).

Vemos así que el atributo "buena reputación" tiene una formulación más cognitiva y más racional (y menos emocional) que los tres atributos restantes. En efecto, los tres atributos restantes: una buena impresión<sup>95</sup> y los sentimientos de "confianza" y "admiración", son de naturaleza más afectiva. En los tres casos, la contribución al constructo de reputación global de estos atributos ha sido significativamente mayor que la del primero ("buena reputación").

Este resultado tiene efectos importantes a la hora de estimar las puntuaciones globales del índice. La estrategia más adecuada para obtener las puntuaciones sería la media ponderada, donde los tres ítems emocionales pesaran entre sí lo mismo y más que el cuarto ítem "buena reputación".

En consecuencia, las hipótesis sobre la calidad psicométrica del Pulse han sido corroboradas, el índice ha demostrado tener suficiente fiabilidad y validez convergente<sup>96</sup>. La reputación global representada por el Pulse es válida, fiable, en ella subyace una sola dimensión y contiene un sentido emocional acusado que lo asocia con fuerza a los comportamientos positivos o negativos de los *stakeholders*.

Estos resultados se enmarcan en la teoría general que ha dado lugar al desarrollo del modelo explicativo de lealtad del cliente que ha recibido el nombre de la teoría de la acción razonada<sup>97</sup> (TAR) de Ajzen y Fishbein. El interés para el gestor de la reputación es que se trata de una teoría general muy consolidada, que se ha verificado ya en numerosos campos de la conducta humana, explicando con éxito los

---

<sup>95</sup> El término en inglés es "feeling".

<sup>96</sup> La investigación no contempla la validez discriminante porque es menos relevante para el análisis de invarianza. Por otro lado, la investigación de Ponzi y colegas ya demuestra este tipo de validez.

<sup>97</sup> El texto de Carreras, Alloza y Carreras (2013) cita la teoría de la acción planeada, que es un desarrollo posterior de Izeck Ajzen en 1991. En aras de la simplificación, aquí sólo mencionamos la teoría de la acción razonada.



mecanismos de toma de decisión (desde las decisiones de practicar deporte o dejar de fumar hasta las decisiones de compra en el ámbito de consumo).

La teoría viene a decir que una persona se decide a realizar una conducta porque la considera atractiva o beneficiosa (componente de la utilidad), siente que debe hacerlo (componente normativo o moral) y se ha formado la intención de llevarla a cabo pero esta evaluación racional no es suficiente, de ahí la importancia de haber demostrado la presencia de un fuerte componente emocional en el índice Pulse. En este sentido, Richard Bagozzi (1992) incorporó el elemento emocional como el mediador necesario entre la actitud y la intención de pasar a la acción. Hemos representado el modelo de forma sintética en el esquema siguiente que expresa el modelo general de la toma de decisiones.

### MODELO GENERAL DE EXPLICACIÓN DE UNA DECISIÓN

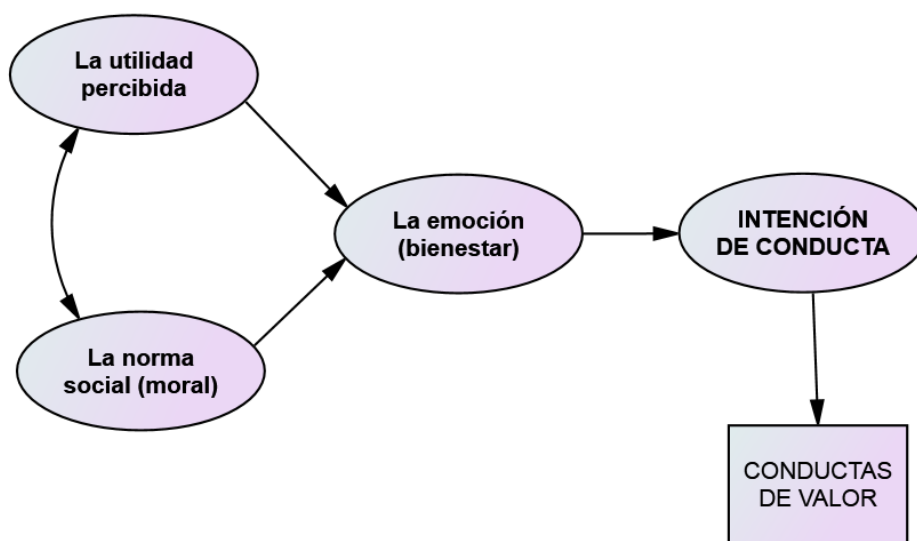


Figura nº 39. Modelo de explicación de la decisión basado en las teorías de la acción razonada (Ajzen y Fishbein) y la de la emoción (Lazarus).

A pesar de las buenas propiedades que hemos comprobado para Pulse, conviene no olvidar que en el caso español, la formulación idiomática del índice ha requerido una covarianza adicional entre los residuales de “buena impresión” y “buena reputación”. Aunque es mínima, en términos de correlación, el 0,18, que arroja la prueba resulta significativa y su ausencia empobrece el análisis. Este resultado en la prueba de covarianza nos avisa de que tal vez el término elegido en la traducción del inglés, aunque correcto, presenta un cierto sesgo cultural de naturaleza más cognitiva/racional (y menos emocional que es lo que se pretende con la versión empleada en inglés). Obviamente, en el análisis para España hemos considerado la invarianza de esta correlación y hemos comprobado que no afecta a la confirmación de la hipótesis de la investigación que afirma que el constructo Pulse es válido, fiable, en él subyace una sola dimensión y contiene un sentido emocional acusado (validez y fiabilidad convergentes).

#### **LA IMPORTANCIA DE HABER DEMOSTRADO LA HIPÓTESIS SOBRE LA INVARIANZA POBLACIONAL DEL PULSE.**

El resultado más importante de esta investigación, desde el punto de vista académico y en especial para la consolidación de la gestión de la reputación en cualquier organización haya sido la demostración de la invarianza poblacional del Pulse.

En los tres países, las hipótesis de invarianza configural, la de cargas, la de la varianza del Pulse han quedado demostradas y en algunos casos como en China se ha llegado hasta confirmar las hipótesis de invarianza métrica.

La invarianza configural y de cargas, resulta esencial para deshacer la ambigüedad de los factores reflexivos que nos planteaban Barbara Byrne y Sunita Stewart (2006). Sólo si se confirma la hipótesis de la invarianza configural entonces podemos afirmar que las medidas obtenidas tendrán un cierto grado de comparabilidad porque significa que los diferentes segmentos de población (sexo, edad y nivel de estudios) utilizan los mismos indicadores cuando tienen en mente esta métrica/índice global de reputación Pulse.

Obviamente, si la hipótesis de invarianza configural hubiera sido rechazada, significaría que cada segmento de población necesitaría utilizar atributos diferentes

para construir el constructo o índice de reputación global. Bajo estas condiciones, las medidas reflexivas de dichos segmentos de población no serían comparables y por tanto no tendríamos la posibilidad de utilizar un mismo índice global de reputación para toda la población u opinión pública. En ese sentido por ejemplo, las mujeres tendrían una representación distinta de reputación que la de los varones, los jóvenes de los mayores, etc. No nos referimos a niveles de reputación diferentes en una escala numérica, sino a formas diferentes de entender la reputación de ahí la importancia de haber podido verificar esta hipótesis.

La corroboración de la hipótesis sobre la invarianza de cargas en los tres países, consolida la posibilidad de comparar el índice global entre diferentes países porque no sólo reconoce que los segmentos de población de cada país comparten los mismos conceptos sino que los utilizan del mismo modo. De esta forma, la naturaleza emocional del índice, que otorga mayor relevancia a los tres atributos con carga afectiva se verifica en todos los grupos demográficos y también por países.

Con la verificación de la invarianza configural y de cargas, Pulse ya podría considerarse como un índice con capacidad plenamente comparativa. Sin embargo, hemos querido dar un paso en las pruebas a las que hemos sometido al índice con el fin de afianzar la robustez de Pulse: de esta forma hemos seguido los criterios de mayor exigencia planteados por algunos autores que proponen comprobar si el índice cumple o no con otro tipo de invarianza más estricta como es la de los residuales. Los resultados de esta prueba estadística no confirman esta hipótesis para España ni para EEUU. Significaría que la cantidad de varianza no explicada por el Pulse en cada ítem sería la misma, sin embargo hay que relativizar este resultado aparentemente negativos puesto que estamos midiendo sólo la varianza residual y en todo caso esa varianza residual no contribuye de forma significativa a la formación del constructo (Byrne y Stewart 2006, 299). Por tanto el rechazo de hipótesis de invarianza residual de Pulse no debe considerarse como un resultado relevante que pudiera afectar a la comparabilidad de Pulse entre países diferentes.

La hipótesis de invarianza de la varianza del Pulse tampoco resulta relevante para la comparación de las medidas. Indica que la dispersión de las puntuaciones entre los diferentes segmentos de población es la misma.

El análisis de invarianza se suele aplicar a poblaciones que puedan apuntar modos de vida diferentes, con diferentes formas de interpretar el mundo. Por esa razón, los análisis suelen construir sus grupos en función de las variables demográficas básicas como el sexo, la edad o el nivel de estudios (como aproximación a la clase social). Las mujeres suelen tener ópticas diferentes a la de los varones, los jóvenes suelen tener apreciaciones distintas a la de los mayores y las personas con diferente grado de estudios, pueden optar por estilos de vida distintos, con distintas preferencias.

Por esta razón la investigación de nuestra tesis defiende que la demostración de la invarianza se realice atendiendo a estas tres variables demográficas. Como ya se ha dicho en el capítulo 4 (Marco Analítico Estadístico y Matemático), el análisis matemático y estadístico que hemos realizado para verificar o refutar las hipótesis de la investigación ha sido muy exigente, no sólo se ha planteado tomando las tres variables por separado, sino la combinación de ocho grupos que podrían representar estilos de vida muy diversos, desde las mujeres jóvenes con estudios bajos hasta los varones adultos con estudios altos. Se trata de una hipótesis exigente de invarianza porque incorpora a grupos socio-demográficos muy diversos. En estas condiciones de especial rigor científico, como se ha venido demostrando las páginas anteriores, los resultados han sido satisfactorios y corroboran las hipótesis de partida de la tesis.

Podemos afirmar por tanto, tras las pruebas realizadas, **que el índice Pulse ha demostrado la equivalencia métrica<sup>98</sup> por sexo, edad y nivel de estudios, pero sobre todo, ha quedado probado que existe este tipo de invarianza en los ocho grupos de segmentos de población considerados<sup>99</sup>.**

La prueba de invarianza poblacional no significa que los colectivos estudiados mantengan el mismo nivel de reputación, únicamente que esos niveles alcanzados reflejan magnitudes distintas de la actitud global subyacente. Para diferenciar

---

<sup>98</sup> La prueba fuerte de la equivalencia métrica de una escala es el análisis de invarianza, probando esta, se prueba la primera (Carreras et al. 2013).

<sup>99</sup> En realidad los ochos grupos se han formado para España y Estados Unidos. En el caso Chino, sólo se han comparado cuatro casos, definidos por el cruce de sexo y edad. En estos cuatro colectivos las hipótesis de invarianza métrica han quedado corroboradas.

plenamente ambos aspectos, no hemos querido finalizar la tesis sin comprobar los niveles absolutos de reputación alcanzados por cada uno de estos colectivos.

De esta forma hemos comprobado que para la mujer china, el nivel medio de reputación de las empresas en el sector bancario es superior al del varón, y que un efecto contrario ha sido detectado en España y EEUU, donde las mujeres son más críticas que los varones. En cuestiones de edad, sólo los jóvenes norteamericanos se han mostrado más favorables a las empresas que los más mayores y en lo que respecta al nivel de estudios, sólo en España, el colectivo con estudios más altos resulta más crítico.

La existencia de estas diferencias de nivel en el índice global de reputación está ahora corroborada por las hipótesis de invarianza psicométrica que han quedado demostradas.

La relevancia de haber podido demostrar la invarianza poblacional de Pulse radica en que podemos utilizar este índice para medir la reputación global de una organización entre sus diferentes *stakeholders* con la posibilidad de comparar la reputación que tenemos entre nuestros empleados, clientes, accionistas, proveedores, etc. Dado que la definición socio demográfica de cada colectivo que denominamos *stakeholder* es, en último caso una combinación específica de las 3 variables que hemos puesto a prueba en esta investigación: sexo, edad y nivel de estudios (proxy de la clase o estatus social).

Disponemos por tanto de un indicador único que permite medir la reputación global entre los diferentes *stakeholders* sabiendo que los resultados de esta métrica (expresados en una escala de 1 a 100 puntos) son directamente comparables entre diferentes *stakeholders*, diferentes países y sectores de actividad.

En consecuencia podemos afirmar que el índice **Pulse se puede considerar como una métrica plenamente interpretable, cuyas variaciones reflejan cambios reales en las actitudes subyacentes de los *stakeholders*, sin importar la estructura poblacional que en cada caso le sea de aplicación.** El índice Pulse se nos desvela como una herramienta con propiedades óptimas para el gestor de la reputación corporativa.

## 6.2. CONCLUSIONES MÁS RELEVANTES.

La invarianza poblacional de Pulse, su fortaleza métrica y la equivalencia semántica y el contenido emocional de este índice, son las propiedades que han quedado verificadas en esta investigación. La consecuencia para la gestión de la reputación es que Pulse permite medir la reputación global de una organización, su evolución a lo largo tiempo, y la comparación de esta reputación global de la organización con otras organizaciones, entre diferentes segmentos de población, sectores de actividad y países.

El gestor de la reputación podría utilizar este índice como indicador clave o KPI de la reputación global de la organización sabiendo que se trata de un indicador para el que se han verificado las propiedades métricas que hemos puesto a prueba en las hipótesis de esta investigación y que por tanto se trata de un indicador sólido y válido para medir la reputación global de una organización.

Este KPI se podría por tanto incluir en el cuadro de mando de alto nivel, junto al resto de indicadores clave tanto financieros como no financieros, que definen y miden el grado de cumplimiento de la estrategia de la organización.

Esta conclusión es de especial relevancia para los gestores de la reputación (también para los investigadores académicos) puesto que el Reptrak TM Pulse, es un indicador que fue hecho público en el año 2006 (Prado, 2008), a partir de ese momento y hasta el momento actual, más de un centenar de organizaciones (empresas y otras instituciones) en todo el mundo utilizan este indicador para gestionar su reputación (Reputation Institute, 2014).

Algunas de estas organizaciones y empresas han incorporado Pulse a sus cuadros de mando y también sabemos que se ha utilizado Pulse para diversas investigaciones y pruebas empíricas para medir el impacto que tienen los incrementos o descensos en Pulse en otras variables actitudinales y comportamentales en los *stakeholders* y en último término en la generación de ingresos económicos y creación de valor para las empresas.

El índice Pulse en este sentido ha sido objeto de diversas investigaciones no publicadas puesto que se ha tratado de investigación aplicada destinada a demostrar la relación positiva entre los incrementos en Pulse y su impacto positivo en las actitudes clave de los consumidores o su impacto directo en el incremento de negocio de algunas de las empresas que forman parte de la fundación Corporate Excellence así como de otras organizaciones y empresas en todo el mundo.

Estas investigaciones, algunas de ellas, en curso actual en el marco de la fundación Corporate Excellence, son coordinadas por Alloza y realizadas por el equipo de colaboradores de la fundación<sup>100</sup>. Los resultados no se han hecho públicos puesto que los resultados son confidenciales y propiedad de las empresas miembro de la fundación.

Estos resultados empíricos han permitido demostrar el impacto positivo de Pulse como veremos en los ejemplos reales siguientes (se ha eliminado el nombre de la empresa a la que corresponden estos resultados en el sector bancario y en el sector seguros). Los resultados de esta tesis podrían permitir argumentar con mayor solvencia la idoneidad de utilizar Pulse para la gestión de la reputación en las organizaciones.

Así la fundamentación teórica y experimental de la tesis que ha permitido demostrar la hipótesis de la invarianza poblacional de Pulse podría contribuir a este objetivo operativo.

A modo de ilustración de este tipo de actividades, en el gráfico siguiente vemos la correlación positiva entre el incremento de Pulse y su impacto en dos variables clave: las actitudes positivas hacia la empresa y el incremento del negocio. Es un ejemplo real de una empresa en el sector bancario en España. En palabras del equipo de investigación de Conento (2014): "... hay una relación directa clara entre reputación (Pulse) y actitudes y entre reputación y negocio...".

---

<sup>100</sup> Investigaciones realizadas por el equipo de matemáticos de la consultora "Conento Consultoría Analítica de Marketing" bajo la dirección de su fundadora Macarena Estévez.

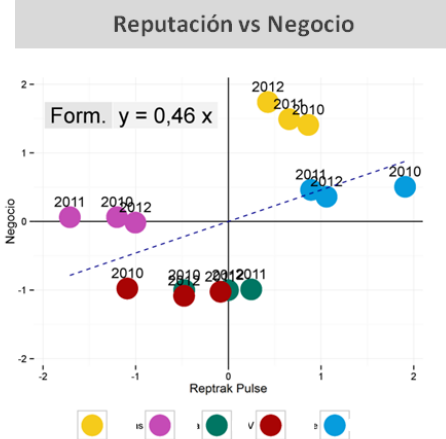
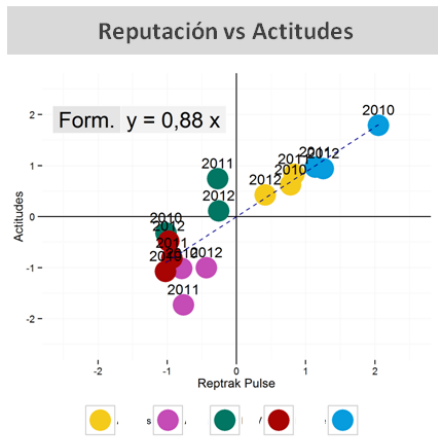


Gráfico nº 13. Relación entre Pulse y actitud favorable a recomendar el banco.

En el gráfico 14, vemos el impacto positivo entre Pulse y la actitud favorable de los clientes de un banco en España para recomendar ese banco "a sus amigos y familiares. Se trata de un ejemplo real de una empresa en el sector bancario en España.

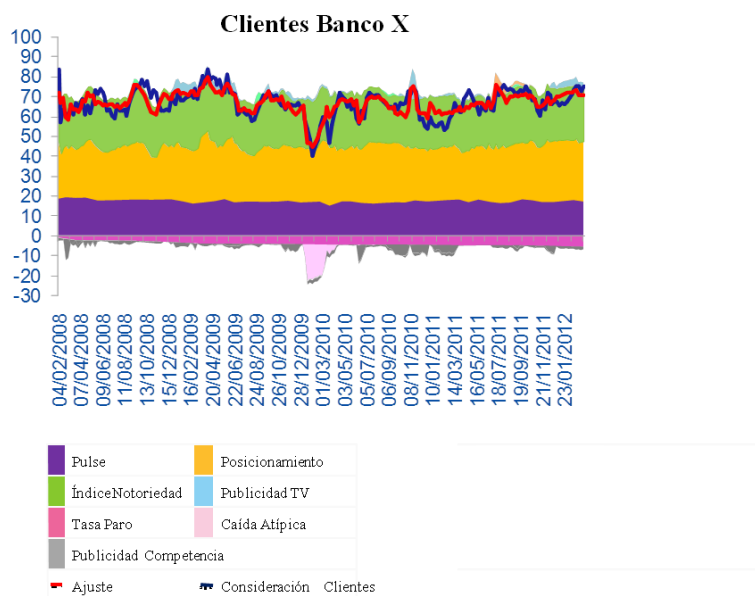
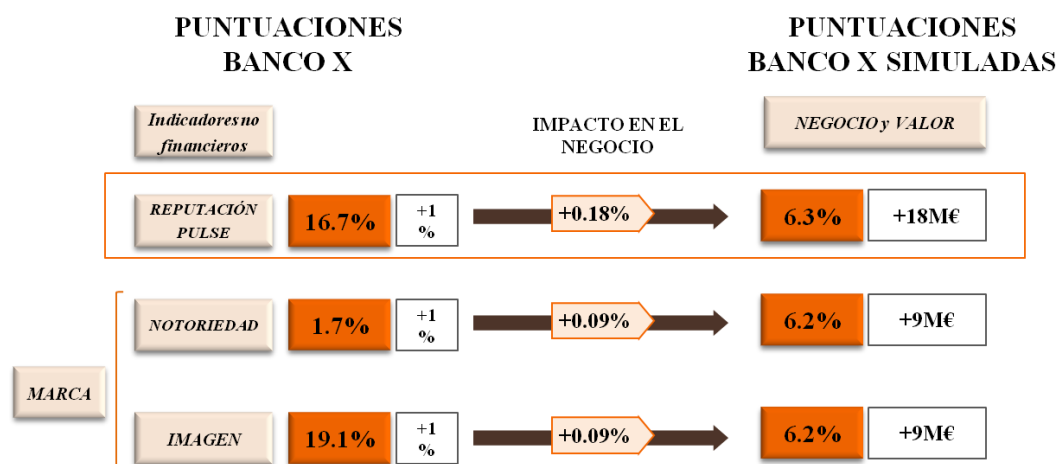


Gráfico nº 14. Relación entre Pulse y el impacto en el negocio en Euros.

Finalmente en el gráfico siguiente vemos la relación entre el incremento en Pulse en un banco en España y su impacto en negocio cuantificado en Euros. Así un incremento de un punto en el índice Pulse reportaría un negocio adicional valorado en 1.8 millones de Euros. Se compara en esta investigación el incremento de negocio que



supone un aumento de 1 punto porcentual en la notoriedad de la marca (conocimiento espontaneo del banco) y el incremento en atributos de imagen de posicionamiento de la marca bancaria y su impacto en negocio. En este ejercicio el mayor impacto en negocio viene dado por la reputación (Pulse) frente a las dos variables antes mencionadas (datos correspondientes al año 2012).



Un incremento del 1% en Reputación supone un aumento del 0,18% en Negocio (+18M€).

Gráfico nº 15. Relación entre Pulse y el impacto en el negocio en Euros.

Esta línea de investigación empírica sobre la relación entre Pulse y el negocio seguirá avanzando en otros sectores de actividad además del bancario y sector seguros que se han citado, en concreto se están iniciando los trabajos en los sectores de gran consumo, distribución, energía, petróleo y telecomunicaciones.

### 6.3. SÍNTESIS DE LAS CONCLUSIONES.

El índice Pulse se presenta como una métrica estratégica para la gestión:

Estudios anteriores han demostrado efectos significativos en las conductas de valor de los *stakeholders*.

Desde el punto de vista métrico contiene las características fundamentales para su gestión:

- Es una medida fiable (los cambios del nivel representan cambios reales en el sentimiento global de reputación)
- Es una medida unidimensional por tanto, el sentimiento global no es complejo (facilita el diagnóstico, la toma de decisiones y la evaluación de los resultados alcanzados)
- Es una medida no ambigua (cambios demográficos en los *stakeholder* no interfieren en la interpretación del nivel de reputación real)
- Los cambios detectados en la métrica reflejan cambios reales en la actitud evaluativa subyacente en las poblaciones hacia esa empresa, con lo que los efectos de ésta en el valor podrán ser detectados con mayor precisión.
- El carácter intercultural del estudio refuerza la generalidad de los resultados

Desde el punto de vista de la gestión de la reputación la conclusión más relevante radica en que la verificación de las hipótesis de esta investigación sobre las propiedades métricas de Pulse, determinan que este índice podría ser incorporado como indicador clave de reputación en el cuadro de mando de alto nivel, junto al resto de indicadores clave tanto financieros como no financieros, que definen y miden el grado de cumplimiento de la estrategia de una organización.

Finalmente, los trabajos empíricos de los gestores de la reputación de numerosas organizaciones y empresas en todo el mundo que han venido utilizando Pulse en investigaciones internas y no publicadas que han mostrado el efecto positivo de los incrementos de Pulse en la creación de valor para estas organizaciones, ven refrendado, gracias a los resultados de nuestra investigación, la elección de Pulse

como *KPI* que además de producir efectos empíricos en la creación o destrucción de valor de sus empresas, ahora se puede afirmar que se trata de un índice con validez y robustez métrica, argumento que lo fortalece para ser considerado como un KPI importante y sólido para las organizaciones.

### **LIMITACIONES DE LOS RESULTADOS OBTENIDOS.**

Conviene también repetir aquí las limitaciones que consideramos que contiene este trabajo de investigación y que ya citamos en el capítulo de contraste de hipótesis.

El estudio está limitado a grandes empresas privadas pertenecientes a dos sectores de actividad, el sector bancario y de las telecomunicaciones. Convendría comprobar si las propiedades de invarianza se verifican en otros sectores, en empresas del sector público y de menor dimensión.

Los tamaños de muestra han sido tal vez excesivamente elevados, sobre todo para España y para EEUU. Esto ha permitido fundamentar mejor el grado de invarianza, pero ha sembrado algunas dudas sobre la importancia práctica de estas diferencias.

La no invarianza en reputación detectada por sexo en el caso norteamericano, puede que sea real pero prácticamente despreciable (el elevado tamaño de muestra hace saltar los test de significación)

En el caso chino, ocurre un efecto similar en los colectivos demográficos homogeneizados.

El estudio se ha basado en el índice Pulse, convendría replicarlo con otros índices globales como el de Highhouse y colegas (2009) o el de nuestros colegas españoles Casaló, Flaviá y Guinaliú (2008) para demostrar la invarianza poblacional se mantiene más allá de los ítems del Pulse.

## **CAPITULO 7. LÍNEAS FUTURAS DE INVESTIGACIÓN.**

### **7.1. PRINCIPALES LÍNEAS DE INVESTIGACIÓN.**

#### **VERIFICACIÓN DE LAS HIPÓTESIS CON OTROS ÍNDICES GLOBALES DE REPUTACIÓN.**

Una posible línea de investigación futura consistiría en verificar las hipótesis de invarianza poblacional a las que hemos sometido al índice Pulse aplicándolas a otros índices globales de reputación que hemos recogido en la literatura académica.

Sería conveniente replicar el estudio con otros índices globales como el de Highhouse y colegas o el de nuestros colegas españoles Casaló, Flavia y Guinalí para demostrar que esta tendencia a la invarianza en poblaciones diferentes de los sentimientos globales se mantiene más allá de los ítems del Pulse.

#### **VERIFICACIÓN DE LAS HIPÓTESIS EN OTROS SECTORES DE ACTIVIDAD.**

Por otro lado, el estudio está obviamente limitado a dos sectores de actividad, el sector bancario y de las telecomunicaciones. Convendría comprobar si la tesis de invarianza poblacional se verifica en otros sectores y poner a prueba la hipótesis de una invarianza sectorial.

#### **VERIFICACIÓN DE LAS HIPÓTESIS CON MUESTRAS DE TAMAÑO MÁS ADECUADO.**

La tercera cuestión es la relativa a los tamaños de muestra. Los datos disponibles han sido tal vez excesivamente elevados. Esto no ha permitido fundamentar mejor el grado de invarianza, pero ha sembrado algunas dudas sobre el nivel de no invarianza detectado. Por ejemplo, las diferencias en reputación detectadas en el nivel de estudios en el caso español o las diferencias por sexo en el norteamericano, puede que sean, aunque reales, prácticamente despreciables, simplemente porque la lupa para detectar estas discrepancias era excesivamente sensible. Convendría replantearse tamaños de

muestra más equilibrados, ahora que las evidencias sobre la invarianza psicométrica ha quedado despejada.

### **VERIFICACIÓN DE LAS HIPÓTESIS CON OTROS *STAKEHOLDERS*.**

La importancia estratégica de las métricas de reputación radica en que representan actitudes o disposiciones colectivas que propician las conductas de valor por parte de los principales grupos de interés, clientes, empleados e inversores y por tanto, esenciales para el futuro de las organizaciones, “atraer clientes, talento y capital”.

Uno de los problemas del desarrollo actual de las métricas es la carencia de un índice global que represente la reputación global de la compañía para todos sus grupos de interés directos. Nuestra tesis se ha limitado a demostrar la validez del índice global de reputación en un único grupo de interés: la población general pero no ha abordado el mismo ejercicio de validación para el resto de los grupos de interés directos de las organizaciones: entre otros sus empleados, clientes, accionistas, reguladores, proveedores, etc.

Aunque la equivalencia semántica ha sido verificada para el “RQ” (Groeland 2002) o para la escala de carácter corporativo (Davies et al. 2004), estos índices de reputación racional no pueden ser integrados en un índice único porque no cumplen con la equivalencia métrica (Chun y Davies 2006; Walsh et al. 2009), las medidas que se obtienen no tienen el mismo valor para los distintos grupos de interés y por tanto tenemos únicamente una métrica de reputación para cada uno de los grupos de interés.

La situación es diferente respecto de los índices globales. Esta tesis junto con otras investigaciones recientes han verificado que estos índices como el Pulse (Reprtrak Pulse) o el de Highhouse, han demostrado tener equivalencia semántica, pero sólo se han demostrado evidencias parciales de su potencial equivalencia métrica (Highhouse et al 2009; Ponzi et al 2011), nuestra tesis ha profundizado en demostrar la equivalencia métrica en la población general.

Una línea de investigación futura buscaría establecer los procedimientos metodológicos y las técnicas más adecuadas para poder estudiar la equivalencia métrica del índice Pulse aplicando el análisis de invarianza para comprobar el grado de

equivalencia del índice entre los distintos colectivos o *stakeholders*. Se trataría entonces de poner a prueba la hipótesis de una invarianza por *stakeholders*.

#### **VALIDACIÓN EMPÍRICA DEL EFECTO DEL ÍNDICE PULSE SOBRE OTRAS VARIABLES Y SOBRE LOS COMPORTAMIENTOS.**

Una línea de investigación empírica, de aplicación práctica para las empresas y las organizaciones buscaría establecer las relaciones causales entre los incrementos y caídas en el indicador Pulse y su correlato con los incrementos o caídas en las actitudes que preceden a los comportamientos de los *stakeholders* (expresadas como la intención de comprar productos, considerarlos como posibles alternativas al actualmente consumido, intención de recomendarlos a terceros. etc.).

En este mismo campo se podría verificar o refutar el efecto que producen estos incrementos o pérdidas en el índice Pulse y los incrementos o caídas en indicadores económicos o proxies de indicadores económicos. Por ejemplo verificar o refutar los efectos causales de Pulse sobre la disminución de la pérdida/rotación de clientes, incremento de la relación con la empresa (número de productos adquiridos/contratados), incrementos o pérdidas de cuota de mercado, etc.

Esta línea de investigación nos permitiría averiguar si podemos cuantificar en Euros perdidos o Euros adicionales ganados como consecuencia de los movimientos al alza o a la baja del índice Pulse.

Esta línea de investigación podría abordar estos ejercicios en distintos sectores de actividad empresarial para verificar si existen comportamientos "universales", generalizables entre los incrementos de Pulse y su correlato con los cambios en las actitudes clave y en los indicadores económicos.

#### **VALIDACIÓN EMPÍRICA DEL EFECTO DEL ÍNDICE PULSE EN OTROS TIPO DE EMPRESAS, ORGANIZACIONES, INSTITUCIONES, CIUDADES Y PAÍSES.**

El estudio quedó limitado a grandes empresas privadas pertenecientes a dos sectores de actividad, el sector bancario y de las telecomunicaciones. Convendría comprobar si las propiedades de invarianza de Pulse se verifican en otros sectores, en empresas del

sector público y de menor dimensión. En este sentido podríamos poner a prueba la hipótesis de una invarianza por tamaño de empresa.

. En este mismo sentido se podría abordar la validación de la invarianza poblacional de Pulse para poder utilizarlo como indicador global de reputación de instituciones, ciudades y países. Siguiendo con la misma lógica verificaríamos o refutaríamos las hipótesis de una invarianza institucional, invarianza por ciudades, invarianza por países, etc.

## CAPITULO 8. GLOSARIO Y GUÍA MATEMÁTICA Y ESTADÍSTICA.

En esta investigación hemos utilizado un elevado número de pruebas matemáticas y estadísticas de considerable especialización y complejidad técnica que aconsejan incluir este capítulo a modo de glosario y soporte didáctico.

### 8. 1. PROPIEDADES PSICOMÉTRICAS DE UNA ESCALA MULTIDIMENSIONAL.

Una escala psicológica compuesta por varios ítems destinada a medir un **constructo** (concepto complejo que explica la presencia de un grupo de comportamientos), debe cumplir con dos conjuntos de cualidades o propiedades esenciales: la “**validez de contenido**” y la “**validez de constructo**”.

- La validez de contenido, de orden teórico, garantiza que la métrica contiene todos los rasgos o dimensiones relevantes del fenómeno que pretende medir o “dominio de contenido”.
- La validez de constructo mide el grado en que una medida se relaciona con otras de acuerdo al conjunto de hipótesis definidas en el constructo. La validez de constructo requiere tres tipos de validez más concretas.
  - La “**validez convergente**”. Las conductas observables a través de los ítems de la escala convergen de forma fiable en el constructo que representan.
  - La “**validez discriminante**”. Los ítems de una dimensión en un constructo divergen de los ítems que indican a otras dimensiones de forma fiable (cuando la escala contiene más de una dimensión, los ítems correspondientes discriminan entre ellas).
  - La “**validez nomológica**” o teórica. El constructo, con sus dimensiones y conductas asociadas, se relaciona con las conductas de otros constructos en un sentido congruente con la teoría que les sirve de referencia.

Las propiedades psicométricas garantizan que las medidas de la escala son teóricamente adecuadas y producen valores fiables y válidos del fenómeno considerado.



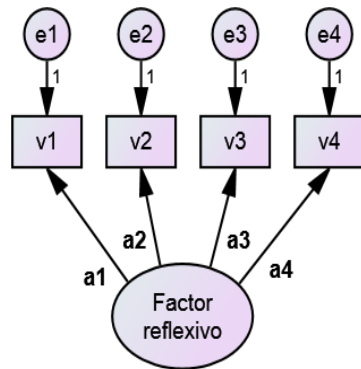
El correlato empírico del constructo y sus dimensiones son las variables latentes o factores.

## 8.2. MODOS DE OPERATIVIZAR UNA VARIABLE LATENTE: REFLEXIVO Y FORMATIVO.

Los ítems pueden relacionarse con una variable latente (factor no observado) de dos modos:

- El **modo reflexivo**. Los ítems son un reflejo de la variable latente no observada. En este caso, el rasgo latente sería el factor común que explica el comportamiento de los ítems que lo representan. Por tanto, la variación de cada ítem se descompone a la variación debida al rasgo latente y una variación residual o única que representa el “error” debido a que cada ítem en cuestión supone una aproximación imperfecta del rasgo latente.
- El **modo formativo**. Los ítems causan o “forman” la variable latente. El contenido o extensión empírica del rasgo latente es reflejado exactamente por la combinación de los ítems.

Modo reflexivo



Modo formativo

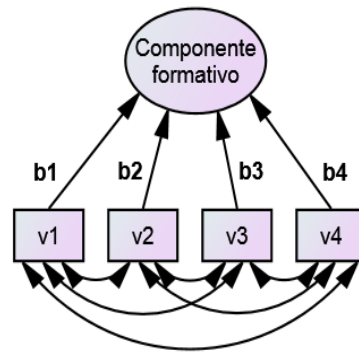


Figura nº 40. Modo reflexivo y modo formativo para hacer operativa una variable latente.

El modo reflexivo sería más adecuado para aproximar constructos psicológicos, como en el caso de la reputación global (Carreras et al. 2013; Highhouse et al. 2009; Ponzi et al. 2011; Nessel y Helgesen 2009; Sung y Yang 2009; Walsh et al. 2009).

### **8.3. ANÁLISIS DE INVARIANZA MULTIGRUPO.**

El análisis de invarianza multigrupo es una técnica concreta de las ecuaciones estructurales destinada a comparar variables latentes aproximadas en modo reflexivo en colectivos distintos. Debido a que las variables latentes son constructos no observables si no es a través de los indicadores o ítems que los representan, antes de comparar los niveles alcanzados en dichos colectivos, se requiere comprobar la invarianza o grado en que estos grupos usan y entienden la métrica del mismo modo.

Neal Schmitt y Goran Kuljanin han expresado esta idea de forma sencilla: una escala viola la invarianza cuando dos individuos de poblaciones diferentes con idéntica posición en el constructo, puntúan de forma distinta los ítems que lo aproximan (Schmitt y Kuljanin 2008: 211). Dicho de otro modo, la invarianza de escala entre grupos sólo puede sostenerse si las puntuaciones en los diversos ítems reflejan la misma posición en el rasgo latente que miden y esto sólo es posible si ambos colectivos “usan e interpretan” de la misma forma dichos ítems.

El análisis de invarianza procede en dos momentos:

#### **1. El establecimiento del grado de “invarianza métrica” de la escala.**

El grado de **invarianza métrica**, también denominada psicométrica o invarianza de la medida, pretende establecer el grado en que los diferentes colectivos entienden y usan los ítems el constructo del mismo modo. El análisis procede mediante una secuencia de hipótesis jerárquicas y concatenadas que describen un continuum en el grado de invarianza, desde un mínimo, la “invarianza configural” hasta un máximo denominado “invarianza estricta”.

El grado de **invarianza** mínima o “**configural**” se alcanza cuando las poblaciones usan los mismos ítems para evaluar el rasgo latente. En ocasiones, poblaciones con distinto nivel de sensibilidad o conocimiento, consideran aspectos diferentes del

mismo fenómeno considerado. La primera prueba de invarianza consiste en determinar si los colectivos implicados en la comparación usan los mismos ítems para expresar su posición en el raso latente. Este nivel es el mínimo requerido para proceder a la comparación.

La “**invarianza de cargas**” supone un nivel de invarianza superior cuando se demuestra que las poblaciones interpretan los ítems del mismo modo. En este caso, el test consiste en verificar la igualdad de las cargas o pendientes que enlazan los ítems con sus latentes en los diferentes grupos. La verificación de esta hipótesis significa que los distintos grupos no sólo utilizarían los mismos ítems sino que les otorgan la misma importancia cuando los puntúan.

En el nivel más alto correspondería a la “invarianza de los residuales” que William Meredith denominó “**invarianza estricta**” (Meredith 1993). En este caso, las poblaciones no sólo serían comparables al nivel de rasgo, sino al nivel de los ítems porque habrían sido explicados por el rasgo de la misma forma en las distintas poblaciones. Este último nivel se considera innecesario para proceder a la comparación entre los parámetros latentes (Byrne y Stewart 206; Cheung Rensoldt 2002; Little 1997; Schmitt y Kuljanin 2008).

Debido a que el grado de invarianza de las escalas puede ser diferente, los autores han distinguido entre niveles de “**invarianza débil**”, como por ejemplo la invarianza configural, que representa el grado mínimo de comparación y la “**invarianza fuerte**” que representaría el nivel pleno de comparabilidad, por ejemplo la invarianza de los residuales. Si una escala sólo mostrara un grado débil, la comparación de los parámetros debería explicitar los supuestos incumplidos.

## **2. La invarianza estructural: “medias latentes” y “relaciones estructurales”.**

Una vez establecido el grado de “invarianza métrica” se puede proceder a la comparación de los parámetros estructurales entre los diferentes grupos. Debemos diferenciar entre dos tipos de parámetros, los de relación y los de nivel:

## 2.1. Comparación entre parámetros de “relaciones estructurales” ó relaciones entre latentes:

- Los **coeficientes de impacto** entre variables latentes. La invarianza probaría que la influencia de una latente sobre otra sería la misma en distintos colectivos. Esta prueba, para ser fuerte, debe completarse con hipótesis más restrictivas como la varianza de los perturbadores o la covarianza entre éstos perturbadores de las variables latentes endógenas.
- Las **covarianzas** entre las latentes exógenas. La invarianza establecería que el nivel de correlación entre latentes explicativas es el mismo en diferentes grupos.
- La **varianza**<sup>101</sup> **de la variable latente** o grado en que el rasgo latente se dispersa o varía entre los individuos de la población. En nuestro caso, la invarianza del Pulse establece que la reputación global, en cuanto rasgo latente, se dispersa en el mismo grado entre los distintos colectivos.

En nuestro caso y dado que disponemos de una única variable latente, sólo procede verificar la varianza de la variable latente.

## 2.2. La comparación entre los niveles medios de los grupos.

Otro tipo de comparación consiste en detectar la igualdad o desigualdad entre los niveles medios del rasgo latente en diferentes grupos. Algunos autores denominan este tipo de invarianza de medias, la **invarianza escalar**<sup>101</sup>.

Para testar la igualdad o desigualdad de las medias globales de distintos grupos, además de la invarianza métrica (configural y de cargas) se requiere una prueba de invarianza adicional, la equivalencia inter-grupo de los interceptores. Cuando se restringe a la igualdad los interceptores entre los diferentes grupos, los orígenes de los valores latentes y su unidad de medida son equivalentes y pueden ser comparados.

---

<sup>101</sup> Incluimos a la varianza como un parámetro relacional porque desde un punto de vista matemático, la varianza sería la covarianza de una variable consigo misma.

## 8.4. NOTAS SOBRE LA IMPLANTACIÓN DE UN ANÁLISIS DE INVARIANZA.

- **La determinación de la variable de referencia**

La lógica del análisis de invarianza descansa en la valoración de la pérdida de ajuste al restringir los parámetros a la igualdad en la estimación simultánea de los distintos grupos.

Este supuesto entra en conflicto con el problema de la identificación que exige restringir a la unidad<sup>102</sup> el coeficiente que enlaza un indicador con la latente con el objeto de dotarla de una escala. Obviamente cuando se opera un análisis de invarianza, la restricción a la unidad se impone en todos los grupos, asumiendo sin evidencia alguna que dicho indicador es en realidad invariante.

Si el indicador restringido a la igualdad en realidad es no invariante, la estimación del resto de cargas resulta sesgada (Millsap 2011; Rensvold y Cheung 2001) Por esa razón se recomienda para verificar cuál de los indicadores disponibles resultaría óptimo para asumir esta restricción a la unidad. En el estudio hemos seguido el método propuesto por Roger Millsap (2011) para averiguar qué ítem puede actuar como variable de referencia cuya carga sobre la latente será restringida a la unidad en todos los grupos.

### 8.4.1. RESTRICCIONES DE INVARIANZA.

En ecuaciones estructurales, los modelos que configuran las relaciones entre variables se definen imponiendo un conjunto de restricciones a partir del “**modelo saturado**” que contiene todas las relaciones posibles entre las variables.

Por ejemplo, un modelo con 3 variables “A”, “B” “C”, estaría saturado si contiene 3 varianzas,  $Var(A)$ ,  $Var(B)$  y  $Var(C)$  y 3 covarianzas que las enlazan,  $Cov(A, B)$ ,  $Cov(A, C)$  y  $Cov(B, C)$ . Supongamos que formulamos un modelo en el que establecemos

---

<sup>102</sup> En realidad se puede restringir a cualquier valor. El más utilizado es la unidad porque así, la escala de la variable latente coincide con la del indicador.

que la relación entre “A” y “C” es cero y que su relación está mediada a través de “B”. En dicha propuesta, la influencia directa o pendiente que va de “A” a “C” se restringiría a cero, dejando que toda la covarianza entre ambas fluya a través de los enlaces de “A” con “B” y “B” con “C”. Las restricciones sobre los parámetros perfilan los modelos más sencillos que sometemos a prueba empírica.

Existe dos tipos de restricciones:

- **Restricción a constante.** La restricción a una constante admite dos modalidades, cuando el valor del parámetro está forzado a cero, indicando ausencia de relación o cuando es distinto de cero, que en cuyo caso se obliga a que el parámetro asuma un valor determinado en el nuevo modelo.
- **Restricciones a igualdad.** La restricción a la igualdad afecta a un conjunto de parámetros, no se impone que asuman un valor determinado, pero se obliga a que asuma el mismo valor.

El análisis de invarianza se basa fundamentalmente en la restricción a igualdad. La hipótesis de invarianza fuerza que el valor de un parámetro sea el mismo en todas las poblaciones implicadas.

La ilustración de la figura siguiente (nº 44), presenta un esquema de dos restricciones para verificar la equivalencia métrica del factor en dos colectivos “A” y “B”. En él se obliga a que la carga que relaciona la variable V1” con el factor latente sea la misma en cada grupo, dejando para la estimación su valor concreto; además, la variable V2 actúa como variable de referencia con el valor “1” para escalar el factor, a la vez que cumple la restricción de igualdad impuesta por la equivalencia de cargas en ambos grupos (Byrne y Stewart 2006; Rensvold y Cheung 2002).

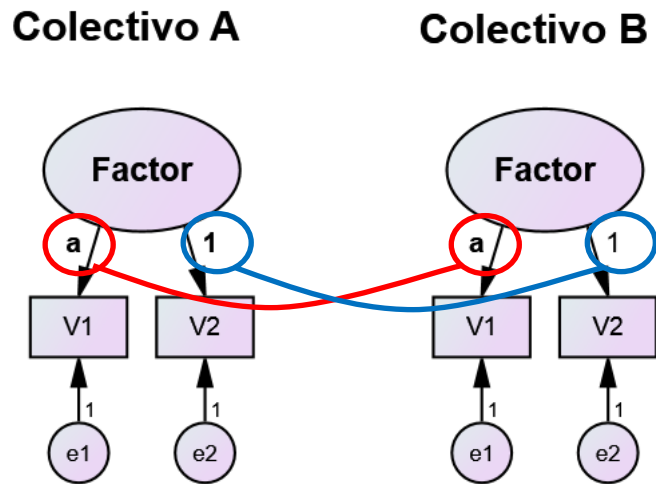


Figura nº 41. Esquema de dos restricciones para verificar la equivalencia métrica del factor.

#### 8.4.2. EQUIVALENCIA DE ESCALA.

La equivalencia de una escala indica el nivel de invarianza alcanzado entre grupos y queda determinado por la corroboración de la hipótesis de mayor rango jerárquico en la serie concatenada. Los niveles de equivalencia ordenados desde la mínima a la máxima invarianza son:

- La **equivalencia configural** de la escala queda establecida cuando el análisis verifica la hipótesis configural, pero rechaza hipótesis de rango superior.
- La **equivalencia métrica** se verifica cuando el test ha superado la configural y la métrica, pero no alcanza a verificar la estricta
- La **equivalencia estricta** se afirma cuando se supera además la invarianza de los residuales de los ítems en los distintos grupos.

En relación con las pruebas sobre la igualdad de los parámetros latentes, tenemos dos:

- La **equivalencia de parámetros** verifica la igualdad de dichos parámetros entre los diferentes grupos, en nuestro caso, la varianza de la variable latente Pulse.

- La **equivalencia escalar** verifica la igualdad de niveles medios de los grupos. Como ya hemos comentado, en esta secuencia, la equivalencia métrica debe asumir, además de la equivalencia métrica, la igualdad de interceptores entre grupos.

Este tipo de equivalencia estructural requiere como mínimo la equivalencia configural.

#### **8.4.3. PRUEBAS PARA ESTABLECER EL NIVEL DE INVARIANZA: ANIDAMIENTO E ÍNDICE INCREMENTAL.**

La secuencia de hipótesis jerárquicas forma un conjunto de modelos progresivamente “anidados”. Así la hipótesis de equivalencia métrica supondría restringir a igualdad las cargas correspondientes a los grupos, un modelo anidado al de equivalencia configural. De esta forma, el nivel de invarianza puede verificarse mediante un test de pérdida de ajuste detectado en el modelo superior, más restrictivo, respecto del modelo base o jerárquicamente más relajado.

El **modelo anidado** y **base** son exactamente iguales con excepción de la restricción impuesta sobre el modelo anidado. Dicha restricción hace que el modelo anidado, al alejarse más del modelo saturado, aumente su valor de discrepancia (“ $\chi^2_{\text{ANIDADO}}$ ” > “ $\chi^2_{\text{BASE}}$ ”) y dado que no tiene que estimar el parámetro restringido, este modelo anidado dispondrá de más grados de libertad (“G.L. ANIDADO” > “G.L. BASE”).

Es factible establecer un contraste de hipótesis sobre la pérdida de ajuste entre el modelo base y anidado. Como la función  $\chi^2$  es sumativa, la pérdida de ajuste se distribuirá como otra chi-cuadrado, diferencia de los chi-cuadrados de ambos modelos ( $\chi^2_{\text{dif}} = \chi^2_{\text{ANIDADO}} - \chi^2_{\text{BASE}}$ ), con grados de libertad, la diferencia de sus correspondientes grados de libertad (G.L.<sub>ANIDADO</sub> – G.L.<sub>BASE</sub>). El test de contraste se formula en los siguientes términos:

- $H_0: \chi^2_{\text{dif}} = 0$
- $H_1: \chi^2_{\text{dif}} > 0$



Si la hipótesis nula se acepta, la pérdida de ajuste se debe a un error muestral y por tanto, el modelo anidado es verdadero en la población, la hipótesis más restrictiva ha quedado corroborada, por el contrario, si la hipótesis nula queda rechazada, la pérdida de ajuste es real en la población por lo que el modelo anidado es empíricamente insostenible.

Aplicado al análisis de invarianza, una vez se establece que el modelo configural es verdadero en la población, se propone como “modelo base” y se anota el valor obtenido en la función de ajuste “ $\chi^2_{CONF}$ ”. Sobre él se introducen las restricciones de igualdad de cargas, dando lugar a un modelo de equivalencia de cargas, anidado al anterior. Se estima el nuevo modelo calculando la nueva función de ajuste “ $\chi^2_{CARGAS}$ ”. Se calcula la pérdida registrada con el anterior ( $\chi^2_{dif} = \chi^2_{CARGAS} - \chi^2_{CONF}$ ), con sus grados de libertad ( $G.L._{CARGAS} - G.L._{CONF}$ ). Se plantea el contraste de hipótesis y se calcula el p-valor para decidir si la pérdida es debida al azar, en cuyo caso se demuestra la equivalencia de cargas o bien la pérdida es real en la población, con lo que volvería empíricamente insostenible la equivalencia de cargas, aceptando la equivalencia configural.

El test de la pérdida de ajuste basado en la “ $\chi^2_{dif}$ ” es sensible al tamaño de muestra, lo que lleva rechazar la hipótesis nula ante pérdidas triviales. Por esta razón algunos autores proponen **valorar la pérdida** significativa de ajuste mediante **los índices incrementales** como el “ $\Delta CFI$ ” que es la diferencia entre el CFI del modelo anidado, respecto del correspondiente al modelo base, si la diferencia es menor a una centésima, la pérdida de ajuste se podría considerar despreciable y proceder a aceptar el modelo más restringido (Bagozzi y Yi 2012); Byrne y Stewart 2006; Cheung y Rensvold 2002).

#### **8.4.4. DIMENSIONALIDAD DE LA VARIABLE LATENTE.**

La variable no observada o latente nace de la combinación de varios ítems o indicadores que reflejan el rasgo psicológico que se pretende medir. La presencia de indicadores distintos en la misma métrica abre la posibilidad de que la variación en

dicho rasgo latente no sea unívoca y presente varias fuentes de variación ó dimensiones.

Desde un punto de vista gerencial, la determinación del número de dimensiones subyacentes resulta fundamental para interpretar con mayor precisión la evolución del índice compuesto. Si un índice compuesto de 4 indicadores contuviera dos dimensiones, éstas podrían progresar de forma divergente compensándose entre sí y mostrando una evolución estable que no obedece a la realidad. El gestor necesita saber cómo evolucionan las dos dimensiones para calibrar mejor la intervención subsiguiente.

Encontramos dos procedimientos para determinar la cantidad de dimensiones subyacentes en una medida compuesta por varios indicadores: el análisis factorial y el índice de Dillon-Goldstein.

La familia de técnicas factoriales tiene la propiedad de extraer las dimensiones subyacentes a un conjunto de variables observadas. Dentro de esta familia, el método “**componentes principales**” identifica las dimensiones utilizando toda la información que incorporan los indicadores y no sólo la información común como hace el resto de técnicas factoriales.

En realidad, los análisis factoriales producen tantas dimensiones como variables observadas. Su valor reside en la forma en que organiza y extrae la información de las variables. El primer componente extrae el máximo de información de la variación conjunta y los componentes sucesivos van extrayendo la información residual hasta agotar la información original. De esta forma, se puede determinar el porcentaje de la información extraída por el primer factor. Si la varianza extractada es superior al 80%, se puede suponer que los indicadores prácticamente están implicados en una única fuente de variación y por tanto podrían considerarse uni-dimensionales<sup>103</sup>.

---

<sup>103</sup> Matemáticamente, la estimación de la matriz diagonal de los autovalores “ $\lambda$ ” obtenida a partir del algoritmo de Jacoby sobre la matriz de intercorrelación, identifica el porcentaje de varianza extractada por cada factor.

El **índice “Dillon-Goldstein”** también ha sido considerado como un buen método para evaluar la uni-dimensionalidad de una variable latente, siempre que las correlaciones entre los ítems y sus latentes sean positivas (Chin 1998). El componente resulta uni-dimensional si las cargas que relacionan los indicadores con la latente son grandes y sus residuales pequeños. Wynne Chin considera que valores en el “ $\rho$ ” superiores a 0,7 garantizan uni-dimensionalidad. (Chin 1998: 320).

Así como el índice Dillon-Goldstein evalúa al componente en su totalidad, también se puede evaluar la aportación de cada indicador al componente mediante el índice de **fiabilidad del indicador** (Bagozzi y Yi 2012). El índice varía entre “0” y “1”, sin embargo los autores no reportan un umbral de aceptación porque la selección del indicador debe regirse por criterios teóricos y no empíricos. En ocasiones un aspecto podría resultar esencial desde el punto de vista teórico e incorporarse al componente, aunque su unicidad sea grande y su contribución pobre, por debajo del 0,6.

## 8.5. TÉCNICAS DE MODELADO DE ECUACIONES ESTRUCTURALES (SEM).

Los métodos de ecuaciones estructurales parten del supuesto de que la relación total de un conjunto de variables observables puede ser reproducida mediante la formulación de un **modelo de relaciones simplificado** que evidencia la forma en que influyen realmente unas variables sobre otras. Estos modelos explicativos pueden incorporar variables latentes que organizan y simplifican la covariación de dichas variables observables.

El modelo simplificado de relaciones se construye introduciendo restricciones (a cero, a constante o a igualdad) en el modelo total o saturado de relaciones y es definido por una serie de parámetros que deberán ser estimados como las cargas factoriales de los ítems sobre su latente, los coeficientes de unas latentes sobre otras, las varianzas residuales, etc. Estos parámetros expresan las relaciones de influencia entre las variables tal y como lo propone el investigador<sup>104</sup>. Mediante un álgebra de ecuaciones, cualquier modelo propuesto será transformado en la matriz de covarianzas implicada por dicho modelo.

La relación total de relaciones de las variables iniciales se expresa en la **matriz de covarianzas observada** “ $\Sigma$ ” y representa la cantidad de información total que vehiculan dichas variables. El investigador propone un modelo de relaciones más simple, que se supone capaz de explicar la variación total anterior. Mediante un sistema general de ecuaciones, cualquier modelo simplificado puede ser traducido en un matriz de covarianzas, que se denomina “**matriz de covarianzas implicada**” por los parámetros “ $\theta$ ” que configuran el modelo “ $\Sigma(\theta)$ ”.

La lógica de la metodología SEM es sencilla, si la matriz de covarianzas implicada por el modelo simplificado reproduce con exactitud la matriz observada significa que las restricciones introducidas son conformes a la realidad y el modelo es por tanto empíricamente corroborado, por el contrario, si la matriz implicada no consiguiera

---

<sup>104</sup> Conviene recordar que estos modelos son confirmatorios. El investigador propone un modelo de impactos con ayuda de la teoría y la realidad lo corrobora o lo rechaza.

reproducir la matriz observada significaría que el modelo, tal y como ha sido formulado, no responde a la realidad.

El sistema de ecuaciones que modela la relación entre variables se divide en dos subsistemas: subsistema de medida y subsistema estructural:

- El **subsistema de medida** contiene todas las ecuaciones que relacionan los ítems observables con sus variables latentes
- El **sistema estructural** contiene las ecuaciones que enlazan las variables latentes entre sí.

Los parámetros de ambos sistemas de ecuaciones son estimados simultáneamente. Para encontrar una solución óptima, los sistemas tienen que estar “sobre-identificados”, es decir disponer de más ecuaciones que parámetros a estimar (los grados de libertad) de modo que se puedan iterar posibles soluciones para retener la óptima. Hay varios **métodos de estimación**, pero todos trabajan buscando los posibles valores de los parámetros a estimar de forma que se minimicen la distancia entre la matriz de covarianzas observada con la implicada por el modelo. Entre los métodos de estimación de los parámetros destacan el de “máxima verosimilitud” (ML), “mínimos cuadrados generalizados (GLS) y la distribución libre asintótica (ADF). El método de máxima verosimilitud estima los valores más probables dado que se ha obtenido esta muestra por azar. Asume los supuestos de no normalidad multivariante y cuando los datos no la cumplen, se debe valorar el ajuste mediante procedimientos específicos como el p-valor de Bollen-Stein<sup>105</sup>.

Estimados los valores de los parámetros, se evalúa el grado en que la matriz de covarianzas implicada por el modelo reproduce con exactitud la matriz observada mediante un test contraste de la discrepancia entre las matrices poblacionales. Dado que los datos proceden de una muestra aleatoria, pudiera ser que la discrepancia observada en la matriz de residuales (diferencia entre la matriz observada y la implicada) se deba a un error de muestreo.

---

<sup>105</sup> Este es el procedimiento que incorpora el programa AMOS. Peter Bentler ha propuesto el estadístico “ $\chi^2_{\text{corregida}}$ ” y lo ofrece en su programa EQS.

La hipótesis básica de contraste:

- $H_0: \Sigma = \Sigma(\theta)$
- $H_1: \Sigma \neq \Sigma(\theta)$

La hipótesis nula plantea la posibilidad de que las matrices de covarianza observada “ $\Sigma$ ” e implicada “ $\Sigma(\theta)$ ” poblacionales, coincidan y que la discrepancia detectada sea un error muestral. El contraste se evalúa mediante el p-valor o probabilidad exacta de rechazar la hipótesis nula siendo verdadera. En nuestro caso buscamos aceptar la hipótesis nula porque significaría que el modelo propuesto es conforme a datos (p-valor  $> 0,05$ ).

Debido a que el test de contraste es muy sensible al tamaño de muestra, los autores recomiendan completar la conclusión del p-valor con una serie de índices globales de ajuste que se detallan en el capítulo correspondiente.

## **8.6. PROGRAMA DE ANÁLISIS, AMOS VERSIÓN 20.**

El programa AMOS es un software diseñado por James L. Arbuckle para el procesamiento de ecuaciones estructurales basados en estructuras de covarianzas. En la actualidad es un módulo del programa estadístico SPSS.

## CAPÍTULO 9. BIBLIOGRAFÍA.

### 9.1. BIBLIOGRAFÍA CONSULTADA.

Aaker, D. (2008): *Spanning Silos*. Harvard Business School Press, Boston., Capítulo 2, (pp. 38-40).

Ajzen, I.(1991): "*The Theory of Planned Behavior*", en *Organizational Behavior and Human Decision Processes*, vol. 50, (pp. 179-211).

Ajzen, I. (2001): «*Nature and Operation of Attitudes*», en *Annual Review Psychology*, vol. 52, (pp. 27-58).

Ajzen, I. y Fishbein, M. (2005): "*The Influence of Attitudes on Behavior*". En Dolores Albarracín , Blair Johnson y Mark Zanna (eds): *The Handbook of attitudes*. New Jersey. Mahwah Erlbaum, (pp 51-56).

Ajzen, I. (2011): «*The Theory of Planned Behavior: Reactions and Reflections*», en *Psychology & Health*, vol. 26 (9), (pp. 1113-1127).

Alameda, García, D. (2006): *Una nueva realidad publicitaria: la generación de valores corporativos en publicidad*. Ediciones del Laberinto, 2006.

Álvarez Villanueva, C. (2012a): *Activos intangibles: catalogación de métodos de valoración*. Madrid: Fragua (Colección Fragua Comunicación, n.º 119).

Álvarez Villanueva, C. (2012b): *Manual de valoración de activos intangibles*. Santa Cruz de Tenerife: Sociedad Latina de Comunicación Social (Colección Cuadernos Artesanos Latina, n.º 27).

Álvarez, C, Benlloch, M., Olivares, F. (2012): "*Evolución de la gestión estratégica de la marca. Reflexiones sobre la aportación del branding a la competitividad empresarial*", *Brandtrends Journal of Communication and Branding*, nº 2.

Alloza, A. et al. (2000):*¿Cómo funciona la publicidad en televisión?*. La investigación en marketing, tomo II, AEDEMO y Celeste Ediciones, pp.819-857.

Alloza, A. (2001): «*La gestión estratégica de la marca*» en Villafañe, Justo (ed.), *El estado de la publicidad y el Corporate en España y Latinoamérica*, informe anual, capítulo 4, Pirámide, Madrid (pp. 207-269).

Alloza, A. (2002): "*La marca experiencia y su papel estratégico en el sector servicios. Las marcas renombradas españolas : un activo estratégico para la internacionalización de España*", *Foro de Marcas Renombradas* , (pp. 169-182).

Alloza, A; Conley, S; Prado, F; Farlán, J. and Espantaleón, R. (2004): "*Creating the BBVA Experience: Beyond Traditional Brand Management*"; *Corporate Reputation Review* Vol. 7 nº 1.

Alloza, A.(2005):“*La reputación corporativa, la eficacia de la comunicación, la marca y la creación de valor*”. Revista Investigación y Marketing de la Asociación de Empresas de Estudios de Opinión y Marketing (AEDEMO), nº 86, (pp.6-15).

Alloza, A. (2006): “*Brand engagement y marca experiencia, homenaje a P. Drucker*”; Revista Capital Intelectual nº 2.

Alloza, A., y Martínez, L.C. (2007): «*La medición y el fortalecimiento de la reputación corporativa*», en Hernández Robledo, M. A.; Losada Vázquez, A.; y Macías Castillo, A. (coords.). Estrategia y conducta social de la organización, Universidad Pontificia de Salamanca, Salamanca, (pp.113-128).

Alloza, A., García-Perrote, J.M., Panadero, G. (2007): “*Hacia una nueva dirección de comunicación: de la orquesta a la banda de jazz*”. La comunicación empresarial y la gestión de los intangibles en España y Latinoamérica. Pearson Prentice Hall.

Alloza, A., (2008): “*Brand Engagement and Brand Experience at BBVA, the transformation of a 150 years Old Company*”. Corporate Reputation Review Vol. 11. pp. 371-378.

Alloza, A., (2009): “*La reputación de España en el mundo*”. Investigación y marketing, AEDEMO, Nº. 103, pág. 22

Alloza, A., (2010): «*De Maquiavelo a la neurociencia: las marcas que queremos*», en Villafañe, Justo (ed.), La comunicación empresarial y la gestión de los intangibles en España y Latinoamérica, informe anual, capítulo 5, pp. 131-172, Prentice Hall, Madrid.

Alloza, A., (2010): “*Gestión del Riesgo Reputacional gestión proactiva de intangibles*”. Auditoría interna: publicación periódica del Instituto de Auditores Internos de España , págs. 72-74

Alloza, A., (2011): «*La economía de la reputación: un nuevo modelo de gestión empresarial*», en Harvard Deusto Business Review, vol. 207 (diciembre), pp. 22-32.

Alloza, A., (2012a): «*Reflexiones sobre la reputación necesaria*», adComunica. Revista Científica de Estrategias, Tendencias e Innovación en Comunicación, vol. 3, Asociación para el Desarrollo de la Comunicación. Universidad Complutense de Madrid y Universitat Jaume I, Castellón, (pp. 27-47).

Alloza Losana, A., (2012b): «*La evolución de la gestión de los intangibles en España*», en Harvard Deusto Business Review (noviembre), pp. 26-38.

Alloza, A., (2012c): “*La emergencia de la economía de la reputación corporativa*”. Uno: publicación de Llorente y Cuenca, 2011, vol. 6. (pp.13- 16).



Alloza, A., (2012d): "*Métricas de Reputación Corporativa. Fundamentos conceptuales, metodologías y aplicaciones para la gestión en el marco de la función de Dirección de Comunicación Corporativa de las organizaciones*". Asignatura SAG003- Observatorio de Nuevas Tendencias y Procesos de innovación en Comunicación. Trabajo Fin de Máster: Trabajo de investigación. Dirección de la Dra. Magdalena Mut Camacho. (Inédito), Universidad Jaume I.

Alloza, A., (2013). "*Construyendo creencias: un nuevo modelo para fomentar el carácter corporativo y la influencia auténtica*". *Prospectivas y tendencias para la comunicación en el siglo XXI*, (pp. 453-472).

Alloza, A., (2014): "*El nuevo paradigma de la marca y la reputación corporativas. Está naciendo un nuevo ciclo económico y un nuevo modelo de hacer empresa*". *Revista Investigación y Marketing*, editada por AEDEMO (Asociación española de estudios de marketing y opinión pública), nº especial diciembre, (pp. 50-55).

Alloza, A., y Noya, F. J. (2004): «*Capital disonante: la imagen de las inversiones españolas en América Latina*», en Documento de Trabajo nº 6/2004, Real Instituto de Estudios Internacionales y Estratégicos.

Anderson, J. C. and Gerbing G.W., (1988): "*Structural Equation Modelling in Practice: A Review and Recommended Two-Step Approach*". En *Psychological Bulletin* Vol. 103 (3), (pp.411-423).

Anholt, S. (2002): «*Foreword to the special issue on place branding*», en *Journal of Brand Management*, vol. 9 (4–5), (pp. 229–239).

Argenti, P. et al. (2005): "*The strategic communication imperative*". *Mit Sloan Management Review*.

Argenti, P. (2014): "*Comunicación estratégica y su contribución a la reputación*", Ed. Lid Biblioteca Corporate Excellence, (pp. 59-72).

Arthur W. Page Society (2012): "*Building Belief: A New Model for Activating Corporate Character and Authentic Advocacy*", Nueva York.

Bagozzi, R. y Youjae Y. (1988): "*On the Evaluation of Structural Equation Models*", *Journal of the Academy of Marketing Science*, 16 (1), (pp. 74-94).

Bagozzi, R. P. (1992), «*The self-Regulation of Attitudes, Intentions, and behavior*», en *Social Psychology Quarterly*, vol. 55, nº 2, pp. 178-204.

Bagozzi, R. P. y Youjae Y. (2012): "*Specification, evaluation, and interpretation of structural equation models*". En *Journal of the Academy of Marketing Science*. Vol. 40, (pp.8-34).

Bartikowski, B., y Walsh, G. (2011): «*Investigating mediators between corporate reputation and customer citizenship behaviors*», en *Journal of Business Research*, vol. 64, (pp. 39–44).

Barton, Dominic, (2011): «*Capitalism for the long term*», en Harvard Business Review.

Batista-Foguet, J. M., Coenders, G. y Alonso, J.(2004): "*Análisis factorial confirmatorio*", Med Clin (Barc); nº122(Supl 1), (pp. 25-26).

Benavides Delgado, J. (1997): *Lenguaje publicitario*. Síntesis.

Benavides Delgado, J., (1993): Dimensiones conceptuales de la comunicación corporativa. en Revista de Ciencias de la Información 7, (pp. 79-86).

Benjumea Acevedo, J.M. et al. (2006): *Matemáticas avanzadas y estadística para ciencias e ingenierías*. Secretariado de Publicaciones de la Universidad de Sevilla: (pp.237-240).

Bentler, P. (1990): "*Comparative fit indexes in structural models*". Psychological bulletin 107 (2), (pp.238).

Benlloch Osuna, M. T. (2012): *Marca territorio, competitividad y empresa. La gestión estratégica de los intangibles comunicativos en el clúster cerámico español: un estudio de la empresa cerámica de Castellón(2005-2010)*. Castellón: tesis doctoral, Departamento de Comunicación, Universitat Jaume I(inédito).Dir. Rocío Blay , Rafael López Lita.

Benlloch, M. y Álvarez, C. (2014): "*KPI de creación de valor a largo plazo y marca corporativa para el 'management': reflexiones sobre la situación en España y propuesta de indicadores*", Harvard Deusto Business Review Vol. 3 nº1.

Bergh, D.; Ketchen, D. J. Jr.; Boyd, B. K.; y Bergh, J. (2010): «*New Frontiers of the Reputation–Performance Relationship: Insights From Multiple Theories*», en Journal of Management, vol. 36 (3), (pp. 620-632).

Bergkvist, L., y Rossiter, J.R. (2007): «*The Predictive Validity of Multiple-Item Versus Single-Item Measures of the Same Constructs*», en Journal of Marketing Research, vol. 44 (mayo), (pp. 175-184).

Bollen, K. A. (1989): *Structural Equation Models with Latent Variables*, John Wiley & Sons, New York.

Bowden, J. Lay-Hwa (2009): "*The process of customer engagement: A conceptual Framework*". En Journal of Marketing Theory and Practice. Vol. 17 (1), (pp. 63–74).

Browne, M. W. (1984): "*Asymptotically distribution-free methods for the analysis of covariance structures*". En British Journal of Mathematical and Statistical Psychology, Vol. 37, (pp.62–83).

Brown, B., y Perry, S. (1994): «*Removing the financial performance halo from Fortune's Most Admired Companies*», en Academy of Management Journal, vol. 37 (5), (pp. 1347-1359).

Brown, S. P. (1995): «*The Moderating Effects of Insupplier/Outsupplier Status on Organizational Buyer Attitudes*», en *Journal of the Academy of Marketing Sciences*, vol. 23 (3), (pp. 170-181).

Bunge, M. (2000): *La investigación científica: su estrategia y su filosofía*, Editorial Siglo XXI, (pp. 38-42).

Byrne, B. M., y Stewart, S. M. (2006): «*The MACS Approach to Testing for Multigroup Invariance of a Second-Order Structure: A Walk Through the Process*», en *Structural Equation Modeling*, vol. 13 (2), (pp. 287–321).

Carmines E.G. y J.P. McIver. (1981): “*Analyzing Models with Unobserved Variables*” en G.W. Borhnstedt y Borgatta E.F (eds): *Social Measurement: Current Issues*. Beverly Hills, CA: Sage.

Carreras, E. (2006): «*El pronóstico en modelos estructurales: Predecir la satisfacción latente*», en Lévy Mangin, Jean-Pierre (ed.), *Modelización con Estructuras de Covarianzas en Ciencias Sociales*, Netbiblo, La Coruña.

Carreras, E. (2009): "*Curso de estadística avanzada*". Colegio Oficial de Sociólogos de Madrid. Madrid.

Carreras, E., y González, B. (2013): «*El paradigma intencional en la evaluación ciudadana de la administración local: la utilidad frente al bienestar*», en *Revista Internacional de Sociología (RIS)*, manuscrito aceptado, eISSN: 1988-429x, doi: 10.3989/ris.

Carreras, E.; González, B.; y Carreras, A. (2012): «*Modelización para la gestión de la atracción social de la universidad: los mecanismos de la actitud intencional del egresado*», IX Foro Internacional sobre la Evaluación de la Calidad de la Investigación y de la Educación Superior (FECIES), Santiago de Compostela, 12-15 de junio.

Carreras, Ana (2010): "*Modelo estadístico de evaluación de los servicios públicos locales mediante la aplicación del paradigma intencional*". Tesis doctoral, leída en la Universidad CEU San Pablo de Madrid (Facultad de Económicas y Empresariales), Julio.

Carreras, E, Alloza, A. Carreras, A.(2013). "*Reputación Corporativa*". Editorial Lid Biblioteca Corporate Excellence.

Carreras, E, Alloza, A. Carreras, A.(2014). "*Corporate Reputation*". Editorial Lid Biblioteca Corporate Excellence.

Carreras, E. y González, B. (2014): “*Calidad y responsabilidad institucional como fuentes de la atracción social de la Universidad: La actitud intencional del egresado*”. Inédito.

Caruana, A., (1997): "*Corporate Reputation: Concept and Measurement*". *The Journal of Product and Brand Management*, 6(2), (pp.109-118).

Casaló, L.; Flavia, C.; y Guinalí, M., (2008): «*The role of perceived usability, reputation, satisfaction and consumer familiarity on the website loyalty formation process*», en *Computers in Human Behaviour*, vol. 24, (pp. 325–345).

Cater, B. y Vesna Z., (2008): “*Antecedents and consequences of commitment in marketing research services: The client's perspective.*” En *Industrial Marketing Management* . Vol. 30, ( pp. 1-13).

Chetthamrongchai, P. (2010): «*Revalidating Two Measures of Reputation in Thailand*», en *Corporate Reputation Review*, vol. 13 (3), (pp. 209–219).

Chen, F. (2007): “*Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance*”. En *Structural Equation Modeling*. Vol. 14(3), (pp. 464–504).

Cheung, G. W., y Rensvold, R. B., (2002): «*Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance*», en *Structural Equation Modeling*, vol. 9 (2), (pp. 233–255).

Chin, W. W. (1998): “*The partial Least Square Approach or Structural Equation Modeling*”. En George A. Marcoulides (Ed): *Modern Methods for Business Research*. Lawrence Earlbaum Associates (pp.295-336).

Chun, R., y Davies, G. (2006): «*The Influence of Corporate Character on Customers and Employees: Exploring Similarities and Differences*», en *Journal of the Academy of Marketing Science*, vol. 34 (2), (pp. 138-146).

Churchill, G. A. (1979): «*A Paradigm for Developing Better Measures of Marketing Constructs*», en *Journal of Marketing Research*, (febrero), (pp. 64-73).

Cornelissen, J.; Christensen, L. T.; y Firat, A. F. (2009): «*New tensions and challenges in integrated communications*», en *Corporate Communications: An International Journal*, vol. 14 (2), (pp. 207–219).

Corporate Excellence (2010): “*Estudio La Reputacion de España en el Mundo*”, Ediciones CE.

Corporate Excellence (2012): “*Reputación de España 2012*”, Ediciones CE.

Corporate Excellence (2010): “*Guía Normativa 2010 Informes de sostenibilidad en la Unión Europea*”, Ediciones CE.

Corporate Excellence (2010): “*Documento Ejecutivo: El reporting no financiero en Europa*”, Ediciones CE.

Corporate Excellence (2012-2014): “*BEO, Balance de Expresiones Online. Resultados empíricos*”, Ediciones CE.

Corporate Excellence (2011): “*Asuntos Públicos. Marco Conceptual y Ranking Global Issues 2011*”, Ediciones CE.

Corporate Excellence (2013): "*What makes a Chief Communications Officer Excellent?*". Ediciones CE.

Corporate Excellence (2013): "*Identificación de Rankings Relevantes 2013*", Ediciones CE.

Corporate Excellence (2011): "*Introducción a la gestión de los Riesgos Reputacionales*", Ediciones CE.

Corporate Excellence (2013): "*Modelo de Gestión del Riesgo Reputacional*", Ediciones CE.

Costa, J. (1995): *Comunicación corporativa y revolución de los servicios*. Madrid: Ediciones Ciencias Sociales.

Costa, J. (2004b): *El futuro de la comunicación en las organizaciones*. Epílogo en Losada Díaz, J. C. (coord.) en *La gestión de la comunicación en las organizaciones*, Ariel, (pp. 544-545).

Costa, J. (2012): *No prólogo*. En Olivares, F. (2012). *Rebelión en las marcas*, 11-14. Madrid: Lid.

Costa, J. (2011): *El ADN del DirCom. Origen, necesidad, expansión y futuro de la Dirección de Comunicación*. Costa Punto Com Editor, Barcelona 2011.

Costa, J. (2013): *Los 5 pilares del branding. Anatomía de la marca*. Costa Punto Com, Barcelona.

Costa, J., (2009): "*Dircom, estrategia de la complejidad. Nuevos paradigmas para la dirección de comunicación*". Publicacions Universitat de Valencia, (pp. 17-18).

Cravens, K.; Goad Oliver, E.; y Ramamoorti, S., (2003): «*The Reputation Index: Measuring and Managing Corporate Reputation*», en *European Management Journal*, vol. 21 (2), (pp. 201–212).

Cronin J.; Taylor, Jr.; Steven A. (1992): «*Measuring Service Quality: A Reexamination and Extention*», en *Journal of Marketing*, vol 56 (julio), (pp. 55-68).

Cronin, J, Jr.; Brady, M.K.; y Hult, T. M. (2000): «*Assessing the Effects of Quality, Value, and Customer Satisfaction on Consumer, Behavioral Intentions in Service Environments*», en *Journal of Retailing*, vol. 76 (2), (pp. 193–218).

Curran, P. J. (1996): "*Psychological Methods*", in the *American Psychological Association, Inc. Vol 1, No. 1, (pp.16-29)*.

Dabholkar, P. A.; Shepherd, C. D.; y Thorpe, D. I. (2000): «*A Comprehensive Framework for Service Quality: An Investigation of Critical Conceptual and Measurement Issues, Through a Longitudinal Study*», en *Journal of Retailing*, vol. 76 (2), (pp. 139-173).

Davies, G.; Chun, R.; Da Silva, R. V.; y Roper, S. (2004): «*A Corporate Character Scale to Assess Employee and Customer Views of Organization Reputation*», en *Corporate Reputation Review*, vol. 7 (2), (pp.125-146).

De Salas Nestares, M. I. (2002): *La comunicación empresarial a través de Internet*, Fundación Universitaria San Pablo-CEU.

De Salas Nestares, M. I. (2010): "*La estrategia de comunicación basada en valores ecológicos como instrumento de la RSC.*" V. Ros Diego (edit by), Dykinson, (pp. 15-32).

De Vellis, R. F. (1991): "*Scale Development: Theory and Applications*", Newbury Park, California: SAGE.

De Vellis, R., (2012), "*Scale Development: Theory and Applications*", Sage Publications, (pp. 59-72).

DiStefano, Ch., Min Z. y Diana M. (2009): "*Understanding and Using Factors Scores: Considerations for The Applied Researcher*". En *Practical Assessment, Research & Evaluation* Vol. 20 (14), (pp.1-11).

Dunteman, G.H. (1989): "*Principal Components Analysis*", Issue 69. Sage Publications.

Estévez, M. (2009): "*Dashboards inteligentes: la mejor defensa ante la crisis*", Anuncios, (pp. 18-19).

Estévez, M. (2011): "*Tendencias del Rol. Un análisis del retorno de la inversión en los modelos econométricos*", Anuncios, (pp. 22-23).

Fazio, R., (2001): «*On the automatic activation of associated evaluations: An overview*», en *Cognition and Emotion*, vol. 15 (2), (pp. 115-141).

Fazio, R., (2007): «*Attitudes as Object-Evaluation Associations of Varying Strength*», en *Social Cognition*, vol. 25 (5), (pp. 603-637).

Flanagan, D. J.; O'Shaughnessy, K.C.; y Palmer, T. B. (2011): «*Re-Assessing the Relationship between the Fortune Reputation Data and Financial Performance: Overwhelming Influence or Just a Part of the Puzzle?*», en *Corporate Reputation Review*, vol. 14 (1), (pp. 3-14).

Fombrun, Charles (1996): "*Reputation: Realizing Value from the Corporate Image*", Harvard Business School Press, Boston.

Fombrun, Ch., y Van Riel, C. (1997): «*The Reputational Landscape*», en *Corporate Reputation Review*, vol. 1 (2), (pp. 5-13).

Fombrun, Ch.; Gardberg, N. A.; y Sever, J. M. (2000): «*A Multi-stakeholder Measure of Corporate Reputation*», en *The Journal of Brand Management*, vol. 7 (4), (pp. 241-255).



Fombrun, Ch., y Van Riel, C. (2004): *Fame and Fortune: How Successful Companies Build Winning Reputations*, Pearson-Financial Times, Nueva York.

Fombrun, Ch. (2006), «*The RepTrak system*»: Presentado al 10th Anniversary Conference on Reputation, Image, Identity and Competitiveness, 25-28 mayo de 2006, Nueva York.

Fombrun, Ch., y Pan, M.(2006): «*Corporate Reputations in China: How Do Consumers Feel About Companies?*», en *Corporate Reputation Review*, vol. 9 (3), (pp. 165-170).

Farrell, A. M. y Rudd, J., (2009): “*Factor Analysis and Discriminant Validity: A Brief Review of Some Practical Issues*”. En Australian and Neazeland Marketing Academy Annual Conference (ANZMAC), Melburne, 30 de noviembre.

Fornell, C. y Larcker, D. (1981): “*Evaluation Structural Equation Models with Unobservable Variables and Measurement Error*”, *Journal of Marketing Research*, 18, (pp. 39-50).

Fornell, C., y Wernerfelt, B. (1987): «*Defensive Marketing Strategy by Customer Complaint Management: A Theoretical Analysis*», en *Journal of Marketing Research*, vol. 24, (4), (pp. 337-346).

Fornell, C.; Johnson, M. D.; Anderson, E.W.; Cha, J.; y Everitt B., (1996): «*The American Customer Satisfaction Index: Nature, Purpose, and Findings*», en *Journal of Marketing*, vol. 60, (pp. 7-18).

Freeman, E., et al. (2010): *Stakeholder Theory: The State of the Art*. Cambridge University Press, (pp. 63-66).

Fryxell, G.E. and Wang, J. (1994): «*The Fortune Corporate Reputation Index: Reputation for what?*», en *Journal of Management*, vol. 20 (1), (pp. 1–14).

Fuente Sabaté, J. M. De la, y . Quevedo Puente, E. (2003): “*In Practice Empirical Analysis of the Relationship Between Corporate Reputation and Financial Performance: A Survey of the Literature*”. *Corporate Reputation Review*. Vol. 6, No. 2, (pp. 161-177).

Garbarino, Ellen y Mark Johnson (1999): “*The Different Roles of Satisfaction, Trust, and Commitment in Customer Relationship*”. En *Journal of Marketing*. Vol. 63, (pp.70-87).

Gardberg, N. A. (2006): «*Reputatie, Reputation, Réputation, Reputazione, Ruf: A Cross-Cultural Qualitative Analysis of Construct and Instrument Equivalence*», en *Corporate Reputation Review*, vol. 9 (1), (pp. 39-61).

Goldberg, M., y H., Jon (1990): «*The Effect of Advertiser Reputation and Extremity of Advertising Claims on Advertising Effectiveness*», en *Journal of Consumer Research*, vol. 17 (2), (pp. 172-179).

Goldberg, A.I., Cohen, G. and Feigenbaum, A. (2003). *"Reputation Building: Small Business Strategies for Successful Venture Development"*. Journal of Small Business Management, 41 (2), (pp.168-86).

Gotsi, M. and Wilson, A. M., (2001): *"Corporate Reputation: Seeking a Definition"*. Corporate Communications: An International Journal, 6 (1), (pp. 24-30).

Groenland, E. A. G. (2002): *«Qualitative Research to Validate RQ Dimensions»*, en Corporate Reputation Review, vol. 4 (4), (pp. 305-315).

Gustafsson, A.; Johnson, M. D.; y Roos, I. (2005): *«The Effects of Customer Satisfaction, Relationship Commitment Dimensions, and Triggers on Customer Retention»*, en Journal of Marketing, vol. 69 (octubre), (pp. 210-218).

Hancock, G.R. (2006): *"Power Analysis in Covariance Structure Modeling"*. En Gregory R. Hancock and Ralph O. Mueller (Eds): Structural Equation Modeling: A second Course. Information Age Publishing, Connecticut.

Hannington, T., (2006): *Cómo medir y gestionar la reputación de su empresa*. Ediciones Instituto Análisis Intangibles.

Harris, L. C., y Goode, M., (2004): *«The four levels of loyalty and the pivotal role of Trust: a study of online service dynamics»*, en Journal of Retailing, vol. 80, (pp. 139-158).

Helm, S. (2005): *«Designing a Formative Measure for Corporate Reputation»*, en Corporate Reputation Review, vol. 8 (2), (pp. 95-109).

Hernández S.; Fernández - Collado, R., Baptista Lucio, Pilar (2006): *Metodología de la investigación*. 4ª Edición, Mac Graw Hill. México.

Highhouse, S.; Broadfoot, A.; y Yugo, J. E. (2009): *«Examining Corporate Reputation Judgments With Generalizability Theory»*, en Journal of Applied Psychology, vol. 94 (3), (pp. 782-789).

Hoffmeyer-Zlotnik, J. (2008): *"Harmonisation of Demographic and Socio-Economic Variables in Cross-National Survey Research"*, en Bulletin de Méthodologie Sociologique, nº 98, (pp. 5-24).

Hooper, D.; Coughlan, J, y Mullen, M.R. (2008): *"Structural Equation Modelling: Guidelines for Determining Model Fit"*. En Electronic Journal of Business Research Methods. Vol. 6 (1), (pp. 53-60).

Hosmer, L. T. (1995): *"Trust: the connecting link between organizational theory and philosophical ethics"*. En Academy of Management Review. Vol. 20 (2), (pp. 379-403).

Hu & Bentler (1999): *"Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives"*, Structural Equation Modeling, 6 (1), (pp.1-55).



IAI, Instituto de Análisis de los Intangibles (2005): *"Estado del arte de los activos intangibles"*, Ediciones Instituto Análisis Intangibles.

IAI, Instituto de Análisis de los Intangibles (2005): *"Información sobre intangibles en la nueva economía"* por Monclús, R; Torres, T. y Rodríguez, Ediciones Instituto Análisis Intangibles.

IAI, Instituto de Análisis de los Intangibles (2007): *"Valoración de empresas cotizadas"*, dirigido por Conesa, I.M. y García Meca, E., Ediciones Instituto Análisis Intangibles.

IAI, Instituto de Análisis de los Intangibles (2007): *"Estado real de la Implantación de la Responsabilidad Corporativa en las empresas españolas"*, dirigido por Beneytez, B. Ediciones Instituto Análisis Intangibles.

IAI, Instituto de Análisis de los Intangibles (2007): *"Evaluación del Renombre y la Notoriedad de Marcas"*, dirigido por Nomen, E. y Montaña, J. Ediciones Instituto Análisis Intangibles.

IAI, Instituto de Análisis de los Intangibles (2008): *"Los intangibles en la regulación contable"* dirigido por Cañibano, L.; García Meca, E., García Osma, B. y Clemente, A.G. Ediciones Instituto Análisis Intangibles.

IAI, Instituto de Análisis de los Intangibles (2008 a 2011, 4 informes), *"La Reputación de España"*, dirigido por Alloza, A. Ediciones Instituto Análisis Intangibles.

INTEGRATED REPORTING COUNCIL (2013), *"The International Integrated Reporting Framework"*. IRC Publications.

Jackson, D. L. (2003): *"Revisiting sample size and number of parameter estimates: Some support for the N: q hypothesis"*. En *Structural Equation Modeling*, Vol. 10 (1), (pp.128-141).

Michael, J.; Gustafsson, A.; Wallin A.; Line; Jaesung, C. (2001): *«The Evolution and Future of National Customer Satisfaction Index Models»*, en *Journal of Economic Psychology*, vol. 22, (pp. 217-245).

Johnson, M.D.; Herrmann, A.; y Huber, F. (2006): *«The Evolution of Loyalty Intentions»*, en *Journal of Marketing*, vol. 70 (April), (pp. 122-132).

Jones, K. (1996): *"Trust as an Affective Attitude"*. En *Ethics*. Vol. 107 (October): (pp.4-25).

Kaplan, R. S., y Norton, D.P. (1996): *"Using the Balanced Scorecard as a Strategic Management System"*. *Harvard Business Review*. Vol. 74, No. 1, (pp. 75-85).

Kim, J. y Müeller, Ch. W. (1978): *"Introduction to Factor Analysis: What is and How to Do it"*, Newbury Park, California: Sage University Paper.

Kline, R. B. (2011): *Principles and Practice of Structural Equation Modelling*. Third Edition, New York: The Guilford Press.

Kothandapani, V. (1971): "Validation of feeling, belief, and intention to act as three components of attitude and their contribution to prediction of contraceptive behavior". En *Journal of Personality and Social Psychology*. Vol. 19(3), (pp. 321-333).

Lazarus, R.S. (2006): "Emotions and Interpersonal Relationships: Toward a Person-Centered Conceptualization of Emotions and Coping". En *Journal of Personality* Vol. 74 (1), (pp.9-46).

Little, T. (1997): "Mean and Covariance Structures (MACS) Analysis of Cross-Cultural Data: Practical and Theoretical Issues". En *Multivariate Behavioral Research*. Vol. 32 (1), (pp. 53-76).

Little, R. J. A. y Rubin, D. B. (1989): "The Analysis of Social Science Data with Missing Values." *Sociological Methods and Research* 18, (pp.292-326).

Lloyd, S. (2011): «Triangulation Research to Inform Corporate Reputation Theory and Practice», en *Corporate Reputation Review*, vol. 14 (3), (pp. 221-233).

López Lita, R. (2000): *Comunicación: la clave del bienestar social*. Drac, Madrid.

MacCallum, R. C., Browne, M. y Sugawara., H. (1996): "Power Analysis and Determination of Sample Size for Covariance Structure Modelling". En *Psychological Methods*. Vol. I (2), (pp. 130-149).

Martín de Castro, G.; Navas López, J. E.; y López Sáez, P. (2006): «Business and Social Reputation: Exploring the Concept and Main Dimensions of Corporate Reputation», en *Journal of Business Ethics*, vol. 63 (4), (pp. 361-370).

Marradi, A. (1981). "Factor Analysis as an aid in the formation and refinement of empirically useful concepts". En Jakson y E.F. Borgatta. "Factor Analysis and Measurement in Sociological Research : A multidimensional Perspective". Beverly Hills, SAGE, (pp. 110-112).

Marzal, J., and Casero, A., (2007): "El desarrollo de la televisión digital en España." *Nettíblo*, La Coruña.

McDonald, R.P. (1989): "An Index of Goodness-of-Fit Based on Noncentrality. *Journal of Classification*", 6, (pp. 97-103).

Millsap, R. E. (2011): "Statistical Approaches to Measurement Invariance", Routledge, Nueva York-Londres.

Meredith, W. (1993): "Measurement invariance, factor analysis, and factorial invariance". En *Psychometrika*. Vol. 58, (pp.525–543).

- Money, K.; Rose, S.; y Hillenbrand, C. (2010): «*The impact of the corporate identity mix on corporate reputation*», en Brand Management, vol. 18 (3), (pp. 197-211).
- Montañés, P. (2011): "*Aquí quién manda. Levantando el mapa de poder en las organizaciones del siglo XXI*". Prentice Hall, (pp. 40-44).
- Morgan, R. M. y Shelby D. (1994): "*The Commitment-Trust Theory of Relationship Marketing*". En Journal of Marketing. Vol. 58 (July), (pp.20- 38).
- Mut, M. (2011): "*El director de comunicación del cambio*", Revista Internacional de relaciones públicas, nº 2, vol.1, (pp. 107-118).
- Mut, M. et al. (2013): "*Manual de Comunicación*", editado por la asociación DIRCOM.
- Nesset, E., y Helgesen, Ø. (2009): «*Modelling and Managing Student Loyalty: A Study of a Norwegian University College*», en Scandinavian Journal of Educational Research, vol. 53 (4), (pp. 327-345).
- Newell, S. J.; y Goldsmith, R. E. (2001): «*The Development of a Scale to Measure Perceived Corporate Credibility*», en Journal of Business Research, vol. 52, (pp. 235-247).
- Noar, S. M. (2003): "*The Role of Structural Equation Modelling in Scale Development*", Structural Equation Modeling: A Multidisciplinary Journal, 10 (4), (pp. 622-647).
- Nomen, E. (2005): "*España frente al cambio de las reglas contables en la UE para los intangibles*", Ediciones IAI Instituto de Análisis de los Intangibles.
- Nomen E. (2005): "*Valor razonable de los activos intangibles: el efecto mariposa en la segunda deslocalización*", Ediciones Instituto Análisis Intangibles.
- Nunnally, J.C., y Berstein, I. H. (1994): *Psychometric Theory*, McGraw Hill, Series in Psychology, Nueva York.
- Oliver, X., y Alloza, A. (2009): "*Atrapados por el consumo*" Lid Editorial, Colección Acción Empresarial, Madrid.
- O'Loughlin, C. y Coenders, G., (2002): "*Application of the European Customer Satisfaction Index to Postal Services. Structural Equation Models versus Partial Least Squares*" (<http://www.udg.edu/fcee/economia/n4.pdf>).
- Olivares, F. (2012): "*Construcción de Marca*". Brandtrends journal, 3(2), (pp. 130). Disponible en [www.brandtrendsjournal.com](http://www.brandtrendsjournal.com)
- Oliver, R. L. (1999): "*Value as Excellence in the Consumption Experience*". En Morris B. Holbrook (Ed.): Consumer value: A framework for Analysis and Research. Routledge, London, (pp. 34).

Oroval, J. M. (2011): "*El rol de las marcas en la internacionalización de las empresas españolas*", Informe interno inédito para el Observatorio de Empresas Multinacionales Españolas (OEME).

Ponzi, L.; Fombrun, Ch.; y Gardberg, N. A. (2011): «*RepTrak™ Pulse: Conceptualizing and Validating a Short-Form Measure of Corporate Reputation*», en *Corporate Reputation Review*, vol. 14 (1), (pp. 15-35).

Porter, M. y Kramer, M. (2011): "*Creating shared value*", *Harvard Business Review*, n.4, (pp. 1-13).

Prado, F. (2008): "*RepTrak als Verfahren der Reputationsmessung*", en [www.communicationcontrolling.de;\(www.communicationcontrolling.de/index.php?id=290&type=98&tx\\_ttnews\[tt\\_news\]=342\)](http://www.communicationcontrolling.de/(www.communicationcontrolling.de/index.php?id=290&type=98&tx_ttnews[tt_news]=342))

Quero-Gervilla, M. J. y Ventura Fernández R. (2011): "*El compromiso como variable mediadora para la predicción de las futuras intenciones de consumo en los servicios. Una aproximación empírica a los consumidores de artes escénicas en España*". En *Cuadernos de Gestión*. Vol. 11 (1), (pp. 15-36).

REPUTATION INSTITUTE (2014): "*Reputation Leaders Study*", Ed. Reputation Institute.

REPUTATION INSTITUTE (2012): "*Global CRO Study*", Ed. Reputation Institute.

Rensvold R. B. y Cheung, G. W. (2001): "*Testing for metric invariance using structural equation models: Solving the Standardization Model*" En C. A. Schriesheim y L. L. Neider (Eds): *Research in Management*. Vol 1 (pp. 25-50).

Retolaza, J.L. y San José, L. (2011): "*Social economy and stakeholder theory, an integrative framework for socialization of the capitalism*", CIRIEC- España. *Revista de Economía*, n 73, (pp. 192-213).

Retolaza, J.L., Ruiz-Roqueñi, M., San José, L., Barrutia, J. (2014): "*Cuantificación del valor social: propuesta metodológica y aplicación al caso de Lantegi Batuak*", en *Revista de Servicios Sociales Zerbitzuan Gobierno Vasco* (pp.17-33).

Rossiter, J. R. (2002): «*The COARSE procedure for scale development in marketing*», en *International Journal of Research in Marketing*, vol. 19, (pp. 305-335).

Rossiter, J. R. (2008): «*Content Validity of Measures of Abstract Constructs in Management and Organizational Research*», en *British Journal of Management*, vol. 19, (pp. 380-388).

Ruíz Limón, R (2006): *Historia y evolución del pensamiento científico*. México. Editado por el autor (pp. 131).

Ryan, M.J.; Buzas, Th.; y Ramaswamy, V. (1995): «*Making CSN a Power Tool: Composite indices boost the value of satisfaction measures for decision making*», en *Marketing Research*, vol. 7 (3), (pp. 10-16).

Salinas, G. (2007): *Valoración de marcas* . Ediciones Instituto Análisis Intangibles.

Shiu, E., Simon J. P., Liliana L. B., y Beatty, S (2011): “*Reflections on discriminant validity: Reexamining the Bove et al. (2009) findings*”. En *Journal of Business Research* Vol. 64, (pp.497–500).

Schafer, J.L. y Graham, J.W. (2002): “*Missing Data: Our View of the State of the Art*”. *Psychological Methods*, Vol. 7, No. 2, (pp. 147–77).

Schmitt, N. y Goran K. (2008): “*Measurement invariance: Review of practice and implications*”. En *Human Resource Management Review*. Vol 18 (4), (pp. 210–222).

Schultz, M.; Mouritsen, J.; y Gabrielsen, G., (2001): «*Sticky Reputation: Analyzing a Ranking System*», en *Corporate Reputation Review*, vol. 4 (1), (pp. 24-41).

Schwaiger, M. (2004): «*Components and parameters of Corporate Reputation: An empirical Study*», en *Schmalenbach Business Review*, vol. 56 (enero), (pp. 46-71).

Schwaiger, M.; Raithel, S.; y Schloderer, M., (2009): *Recognition or rejection—How a company’s reputation influences stakeholder behaviour*, en Klewes, Joachim, y Wreschniok, Robert (eds.), *Reputation Capital: Building and Maintaining Trust in the 21st Century*, Springer Hedelberg Dordrecht, Londres-Nueva York.

Schmitt, N. y Goran K. (2008): “*Measurement invariance: Review of practice and implications*”. En *Human Resource Management Review*. Vol 18 (4), (pp.210–222).

Sirdeshmukh, D., Jagdip S., and Barry S., (2002): “*Consumer Trust, Value, and loyalty in Relational Exchanges*”. En *Journal of Marketing* Vol. 66 (January), (pp. 15-37).

Shrum, W. (1988): «*Reputational Status of Organizations in Technical Systems*», en *American Journal of Sociology*, vol. (4), (pp. 882-912).

SpencerStuart y Weber Shandwich (2014):“*The rising CCO Study 4th edition*”.

Spratt, D.; Czellar, S.; y Spanderberg, E.(2009): «*The Importance of a General Measure of Brand Engagement on Market Behavior: Development and Validation of a Scale*», en *Journal of Marketing Research*, vol. 46 (febrero), (pp. 92-104).

Spurgin, R., y Tamarkin, M. (2005): «*Switching Investments Can be a Bad Idea when Parrondo’s Paradox Applies*», en *Journal of Behavioral Finance*, vol. 6 (1), (pp. 15-18).

Srivastava, R. K.; McInish, T. H.; Wood, R. A.; y Capraro, A. J. (1997): «*How do reputation affect corporate performance?: The Value of Corporate Reputation: Evidence from the Equity Markets*», en *Corporate Reputation Review*, vol. 1 (1), (pp. 61-68).

Steiger, J. H. (2007): "*Understanding the limitations of global fit assessment in structural equation modelling*". En *Personality and Individual Differences*. Vol. 42 (5), (pp. 893-98).

Suchman, M. (1995): "*Managing legitimacy: Strategic and Institutional Approaches*", en *Academy of Management Review*, vol. 20 (3), (pp. 571-610).

Suki, N. M. (2011): "*Assessing patient satisfaction, trust, commitment, loyalty and doctors' reputation towards doctor services*". En *Pakistan Journal of Medical Sciences*. Vol. 27 (5), (pp.1207-1210).

Sung, M. y Yang, S., (2009): «*Student–university relationships and reputation: a study of the links between key factors fostering students' supportive behavioral intentions towards their university*», en *Higher Education*, vol. 57, (pp. 787–811).

Tucker L. R. and Lewis, Ch. (1973): "*A reliability coefficient for maximum likelihood factor analysis*", *Psychometrika* , Volume 38, Issue 1, (pp. 1-10).

Ullman, J. B. (2006): "*Structural Equation Modelling: Reviewing the Basics and Moving Forward*". En *Journal of Personality Assessment*. Vol. 87(1), (pp.35–50).

Van Riel, C. M. B. (2012): *Alinear para ganar*. LID Editorial Empresarial Biblioteca Corporate Excellence, Madrid.

Van Riel, C.M.B; Alloza, A. et al. (2013): "*El Chief Communications Officer del futuro. Factores y funciones que hacen de los Chief Communications Officers (CCO) profesionales de éxito*", editado por Corporate Excellence, Centre for Reputation Leadership.

Vidaver-Cohen, D., (2007): «*Reputation Beyond the Rankings: A Conceptual Framework for Business School Research*», en *Corporate Reputation Review*, vol. 10 (4), (pp. 278-304).

Villafañe, Gallego, J. (2011): "*Las empresas y sus directivos se transfieren la reputación*". *Capital Humano* 24.254, (pp. 58-62).

Villafañe, Gallego, J. (2003): *Influencia de la comunicación en la reputación corporativa. Construir confianza: ética de la empresa en la sociedad de la información y las comunicaciones*. Trotta.

Villafañe, Gallego, J. (2002): *Imagen positiva: gestión estratégica de la imagen de las empresas*. Ediciones Pirámide.

Villafañe, Gallego, J. (2004): *La buena reputación: Claves del valor intangible de las empresas*, Pirámide, (pp.45).



Visauta Vinacua, B (1989): *Técnicas de investigación social, recogida de datos*. Promociones y publicaciones universitarias (pp.110).

Vogel, V.; Evanschitzky, H.; y Ramaseshan, B. (2008): «*Customer Equity Drivers and Future Sales*», en *Journal of Marketing*, vol. 72 (noviembre), (pp. 98-108).

Walsh, G., Dinnie, K., y Wiedmann, K. P. (2006): “*How do corporate reputation and customer satisfaction impact customer defection? A study of private energy customers in Germany*”. En *Journal of Services Marketing*. Vol. 20 (6), (pp. 412–420).

Walsh, G., y Beatty, Sh. E. (2007):«*Customer-Based Corporate Reputation of a Service Firm: Scale Development and Validation*», en *Journal of the Academy Science*, vol. 35, (pp. 127-143).

Walsh, G.; Beatty, Sh.; y Shiu, E. M. K. (2007): «*The Consumer-based corporate reputation scale: Replication and short form*», en *Journal of Business Research*, vol. 62 (10), (pp. 924-930).

Walsh, G.; Mitchell, V. W.; Jackson, P. R.; y Beatty, Sh. E. (2009): «*Examining the Antecedents and Consequences of Corporate Reputation: A Customer Perspective*», en *British Journal of Management*, vol. 20, (pp. 187-203).

West, S. G., John F. F.y Patrick J. C. (1995): “*Structural Equation Models with Non-Normal Variables: Problems and Remedies*”. En Rick H. Hoyle (Ed.): *Structural Equation Modeling: Concepts, Issues and Applications* . Sage Publication, California (pp.56-75).

Wicker, A.W. (1969): "*Attitudes versus Actions: The Relationship of Verbal and Overt Behavioral Responses to Attitude Objects*", in *Journal of Social Issues*, Volume 25, Issue 4, (pp. 41–78).

Yoon M. y Millsap, R. E. (2007): “*Detecting violations of factorial invariance using data based specification searches: A Monte Carlo Study*”. En *Structural Equation Models*. Vol. 14, (pp.435-463).

Zerfass, A.; Verčič, D.; Verhoeven, P.; Moreno, A.; y Tench, R. (2014): "*European Communication Monitor*", A study conducted by the European Public Relations Education and Research Association (EUPRERA), the European Association of Communication Directors (EACD) and Communication Director Magazine.

## ANEXOS.

### 1. CÁLCULOS ESTADÍSTICOS REALIZADOS.

#### 1. 1. COEFICIENTES DE FIABILIDAD DEL COMPUESTO, AVE Y FIABILIDAD DEL INDICADOR.

PULSE: España						
Ítems	betas <sub>i</sub>	Var(error)	betas <sup>2</sup>	( $\Sigma r_i$ ) <sup>2</sup>	$\Sigma(r_i^2)$	Fiabilidad <sub>i</sub>
Tiene buena reputación	0,885	0,22	0,78	13,75	3,44	<b>0,783</b>
Da buena impresión	0,942	0,11	0,89	$\Sigma(1-r_i^2)$		<b>0,887</b>
En la que se confía	0,944	0,11	0,89	0,56		<b>0,891</b>
Se admira y respeta	0,937	0,12	0,88			<b>0,878</b>
<b>Dillon Goldstein</b>	<b>0,961</b>					
<b>AVE</b>	<b>0,860</b>					

PULSE: EUU						
Ítems	betas <sub>i</sub>	Var(error)	betas <sup>2</sup>	( $\Sigma r_i$ ) <sup>2</sup>	$\Sigma(r_i^2)$	Fiabilidad <sub>i</sub>
Tiene buena reputación	0,904	0,18	0,82	13,94	3,49	<b>0,817</b>
Da buena impresión	0,953	0,09	0,91	$\Sigma(1-r_i^2)$		<b>0,908</b>
En la que se confía	0,943	0,11	0,89	0,51		<b>0,889</b>
Se admira y respeta	0,933	0,13	0,87			<b>0,870</b>
<b>Dillon Goldstein</b>	<b>0,964</b>					
<b>AVE</b>	<b>0,871</b>					

PULSE: China						
Ítems	betas <sub>i</sub>	Var(error)	betas <sup>2</sup>	( $\Sigma r_i$ ) <sup>2</sup>	$\Sigma(r_i^2)$	Fiabilidad <sub>i</sub>
Tiene buena reputación	0,887	0,21	0,79	12,38	3,10	<b>0,787</b>
Da buena impresión	0,868	0,25	0,75	$\Sigma(1-r_i^2)$		<b>0,753</b>
En la que se confía	0,891	0,21	0,79	0,90		<b>0,794</b>
Se admira y respeta	0,873	0,24	0,76			<b>0,762</b>
<b>Dillon Goldstein</b>	<b>0,932</b>					
<b>AVE</b>	<b>0,774</b>					

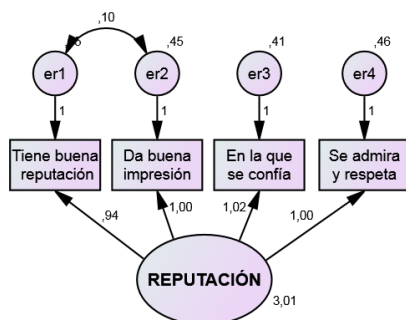


## 1.2. ANÁLISIS DE INVARIANZA EN ESPAÑA.

### 1.2.1. INVARIANZA POR SEXOS. ESPAÑA.

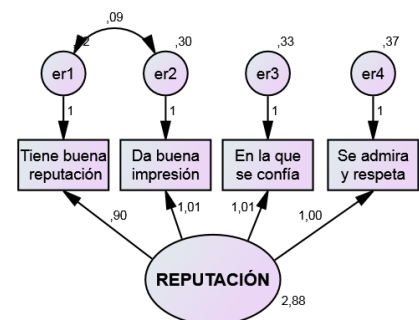
#### Modelo configural

**Invarianza del índice PULSE  
por sexo  
Varón  
ESPAÑA, 2012**



Chi-cuadrado=7,207  
g.l. = 2  
Chi square/df = 3,604  
p-valor = ,027  
CFI = 1,000  
NFI = 1,000  
RMSEA = ,023  
Pclose = ,990

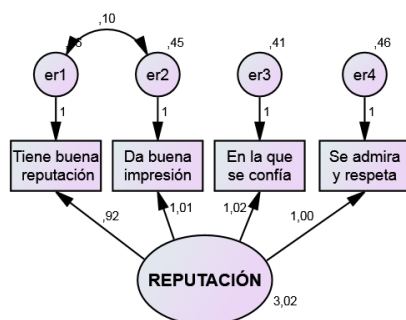
**Invarianza del índice PULSE  
por sexo  
Mujer  
ESPAÑA, 2012**



Chi-cuadrado=7,207  
g.l. = 2  
Chi square/df = 3,604  
p-valor = ,027  
CFI = 1,000  
NFI = 1,000  
RMSEA = ,023  
Pclose = ,990

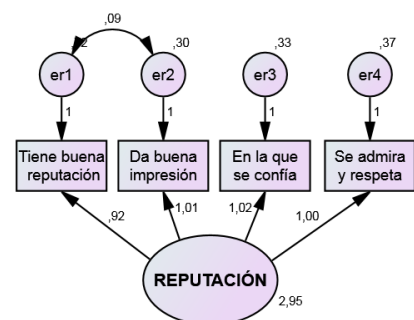
#### Modelo invarianza de cargas

**Invarianza del índice PULSE  
por sexo  
Varón  
ESPAÑA, 2012**



Chi-cuadrado=19,023  
g.l. = 5  
Chi square/df = 3,805  
p-valor = ,002  
CFI = ,999  
NFI = ,999  
RMSEA = ,024  
Pclose = 1,000

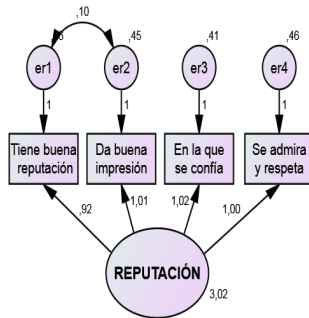
**Invarianza del índice PULSE  
por sexo  
Mujer  
ESPAÑA, 2012**



Chi-cuadrado=20,266  
g.l. = 6  
Chi square/df = 3,378  
p-valor = ,002  
CFI = ,999  
NFI = ,999  
RMSEA = ,022  
Pclose = 1,000

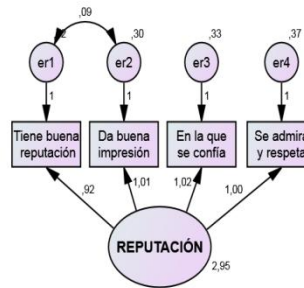
## Invarianza de varianza Pulse

Invarianza del índice PULSE  
por sexo  
Varón  
ESPAÑA, 2012



Chi-cuadrado=19,023  
g.l. = 5  
Chi square/df = 3,805  
p-valor = ,002  
CFI = ,999  
NFI = ,999  
RMSEA=,024  
Pclose = 1,000

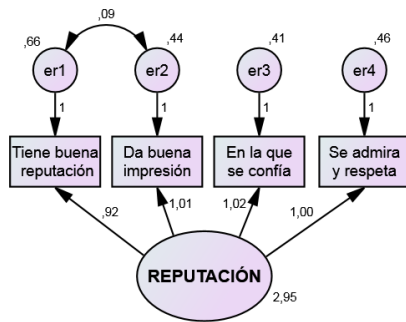
Invarianza del índice PULSE  
por sexo  
Mujer  
ESPAÑA, 2012



Chi-cuadrado=20,266  
g.l. = 6  
Chi square/df = 3,378  
p-valor = ,002  
CFI = ,999  
NFI = ,999  
RMSEA=,022  
Pclose = 1,000

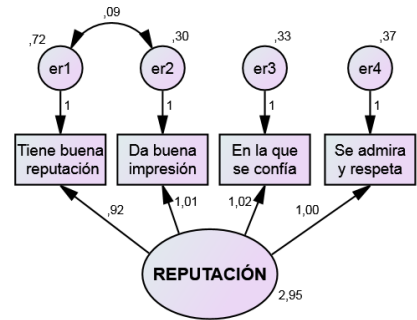
**Invarianza de covarianza errores**

**Invarianza del índice PULSE  
por sexo  
Varón  
ESPAÑA, 2012**



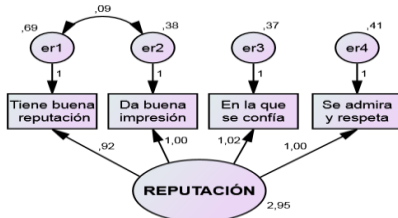
Chi-cuadrado=20,410  
g.l. = 7  
Chi square/df = 2,916  
p-valor = ,005  
CFI =,999  
RMSEA =,020  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo  
Mujer  
ESPAÑA, 2012**



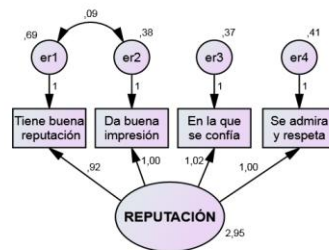
Chi-cuadrado=20,410  
g.l. = 7  
Chi square/df = 2,916  
p-valor = ,005  
CFI =,999  
NFI =,999  
RMSEA =,020  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo  
Varón  
ESPAÑA, 2012**



Chi-cuadrado=126,074  
g.l. = 11  
Chi square/df = 11,461  
p-valor = ,000  
CFI =,995  
NFI =,995  
RMSEA =,047  
Pclose = ,745

**Invarianza del índice PULSE  
por sexo  
Mujer  
ESPAÑA, 2012**



Chi-cuadrado=126,074  
g.l. = 11  
Chi square/df = 11,461  
p-valor = ,000  
CFI =,995  
NFI =,995  
RMSEA =,047  
Pclose = ,745

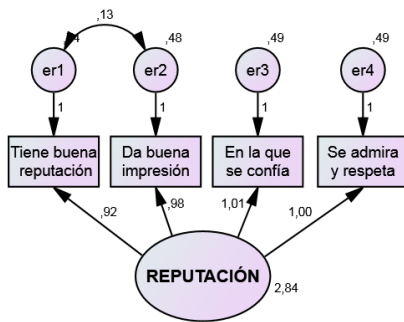
Tablas de comparación siendo verdadero el modelo de igualdad de cargas.

Assuming model cargas factoriales to be correct:			
Model	DF	CMIN	P
cargas y varianza Pulse	1	1,242	0,265
Residuales y covarianza residuales	6	107,05	0
cargas, varianza Pulse y covarianza errores	2	1,387	0,5

## 1.2.2. INVARIANZA POR EDAD. ESPAÑA

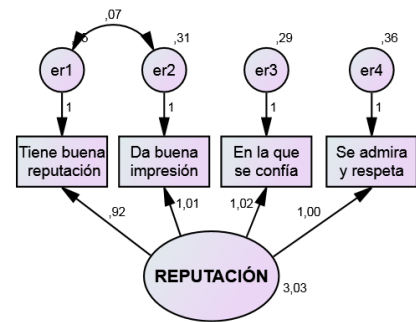
### Modelo configural

**Invarianza del índice PULSE por edad Jóvenes ESPAÑA, 2012**



Chi-cuadrado=1,403  
g.l. = 2  
Chi square/df = ,702  
p-valor = ,496  
CFI =1,000  
NFI =1,000  
RMSEA=,000  
Pclose = 1,000

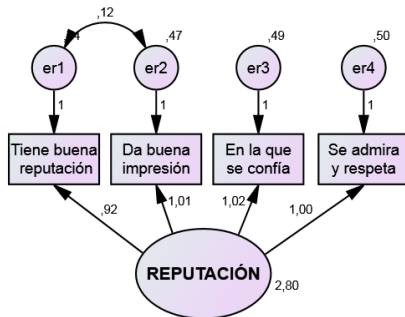
**Invarianza del índice PULSE por edad Adultos ESPAÑA, 2012**



Chi-cuadrado=1,403  
g.l. = 2  
Chi square/df = ,702  
p-valor = ,496  
CFI =1,000  
NFI =1,000  
RMSEA=,000  
Pclose = 1,000

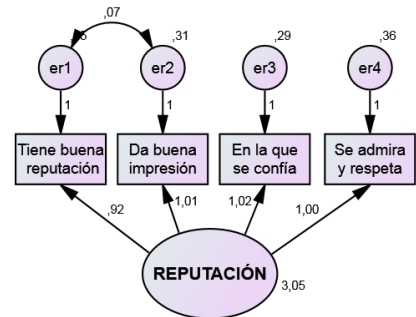
### Modelo invarianza de cargas

**Invarianza del índice PULSE por edad Jóvenes ESPAÑA, 2012**



Chi-cuadrado=5,469  
g.l. = 5  
Chi square/df = 1,094  
p-valor = ,361  
CFI =1,000  
NFI =1,000  
RMSEA=,004  
Pclose = 1,000

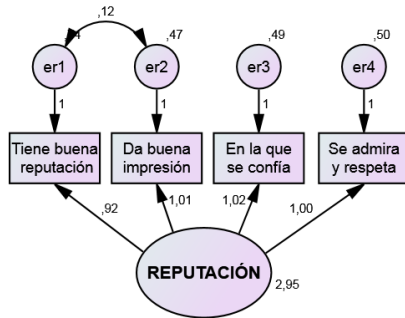
**Invarianza del índice PULSE por edad Adultos ESPAÑA, 2012**



Chi-cuadrado=5,469  
g.l. = 5  
Chi square/df = 1,094  
p-valor = ,361  
CFI =1,000  
NFI =1,000  
RMSEA=,004  
Pclose = 1,000

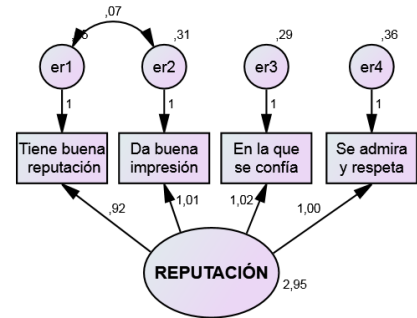
## Invarianza de varianza Pulse

### Invarianza del índice PULSE por edad Jóvenes ESPAÑA, 2012



Chi-cuadrado=9,245  
g.l. = 6  
Chi square/df = 1,541  
p-valor = ,160  
CFI = 1,000  
NFI = 1,000  
RMSEA=,011  
Pclose = 1,000

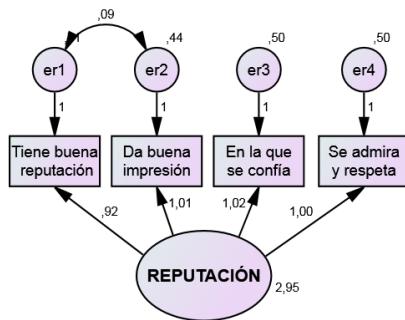
### Invarianza del índice PULSE por edad Adultos ESPAÑA, 2012



Chi-cuadrado=9,245  
g.l. = 6  
Chi square/df = 1,541  
p-valor = ,160  
CFI = 1,000  
NFI = 1,000  
RMSEA=,011  
Pclose = 1,000

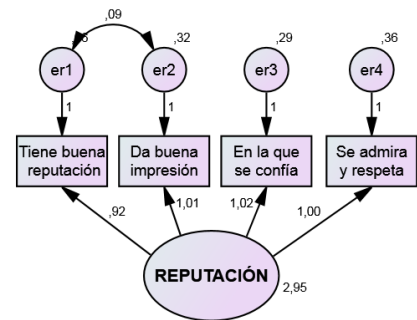
## Invarianza de covarianza errores

### Invarianza del índice PULSE por edad Jóvenes ESPAÑA, 2012



Chi-cuadrado=13,187  
g.l. = 7  
Chi square/df = 1,884  
p-valor = ,068  
CFI = 1,000  
NFI = ,999  
RMSEA=,014  
Pclose = 1,000

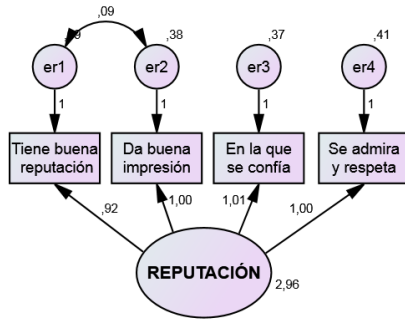
### Invarianza del índice PULSE por edad Adultos ESPAÑA, 2012



Chi-cuadrado=13,187  
g.l. = 7  
Chi square/df = 1,884  
p-valor = ,068  
CFI = 1,000  
NFI = ,999  
RMSEA=,014  
Pclose = 1,000

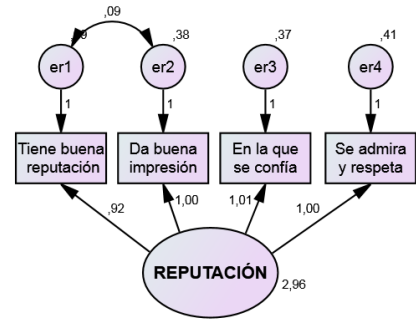
**Invarianza estricta**

**Invarianza del índice PULSE  
por edad  
Jóvenes  
ESPAÑA, 2012**



Chi-cuadrado=219,382  
g.l. = 11  
Chi square/df = 19,944  
p-valor = ,000  
CFI = ,991  
NFI = ,991  
RMSEA=,063  
Pclose = ,001

**Invarianza del índice PULSE  
por edad  
Adultos  
ESPAÑA, 2012**



Chi-cuadrado=219,382  
g.l. = 11  
Chi square/df = 19,944  
p-valor = ,000  
CFI = ,991  
NFI = ,991  
RMSEA=,063  
Pclose = ,001

Tablas de comparación siendo verdadero el modelo de igualdad de cargas.

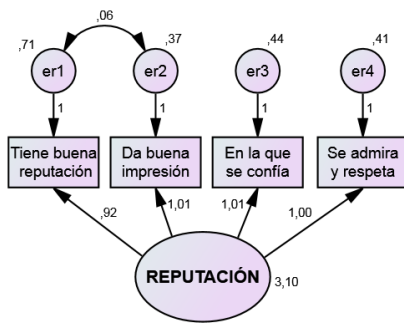
**Assuming model Invarianza covarianza errores to be correct:**

Model	DF	CMIN	P	NFI Delta-1	IFI Delta-2	RFI rho-1	TLI rho2
Invarianza estricta	4	206,195	,000	,009	,009	,009	,009

### 1.2.3. INVARIANZA POR NIVEL DE ESTUDIOS. ESPAÑA

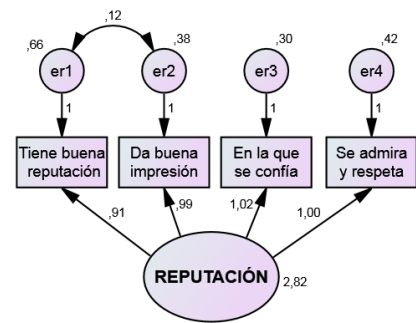
#### Modelo configural

**Invarianza del índice PULSE  
por estudios  
Nivel bajo  
ESPAÑA, 2012**



Chi-cuadrado=2,858  
g.l. = 2  
Chi square/df = 1,429  
p-valor = ,239  
CFI = 1,000  
NFI = 1,000  
RMSEA=,009  
Pclose = 1,000

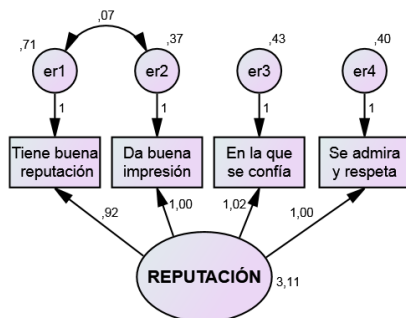
**Invarianza del índice PULSE  
por estudios  
Nivel alto  
ESPAÑA, 2012**



Chi-cuadrado=2,858  
g.l. = 2  
Chi square/df = 1,429  
p-valor = ,239  
CFI = 1,000  
NFI = 1,000  
RMSEA=,009  
Pclose = 1,000

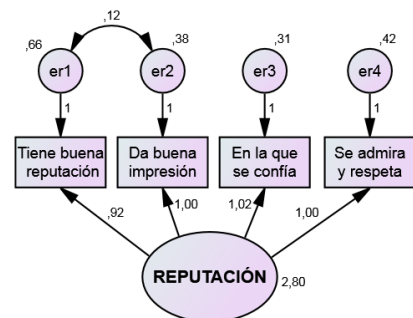
#### Modelo invarianza de cargas

**Invarianza del índice PULSE  
por estudios  
Nivel bajo  
ESPAÑA, 2012**



Chi-cuadrado=6,450  
g.l. = 5  
Chi square/df = 1,290  
p-valor = ,265  
CFI = 1,000  
NFI = 1,000  
RMSEA=,008  
Pclose = 1,000

**Invarianza del índice PULSE  
por estudios  
Nivel alto  
ESPAÑA, 2012**

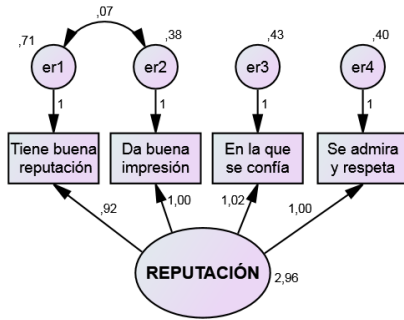


Chi-cuadrado=6,450  
g.l. = 5  
Chi square/df = 1,290  
p-valor = ,265  
CFI = 1,000  
NFI = 1,000  
RMSEA=,008  
Pclose = 1,000



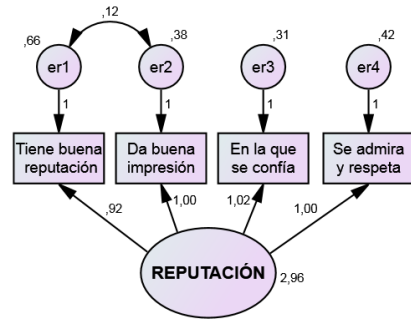
## Invarianza de varianza Pulse

**Invarianza del índice PULSE  
por estudios  
Nivel bajo  
ESPAÑA, 2012**



Chi-cuadrado=12,680  
g.l. = 6  
Chi square/df = 2,113  
p-valor = ,048  
CFI = 1,000  
NFI = ,999  
RMSEA=,015  
Pclose = 1,000

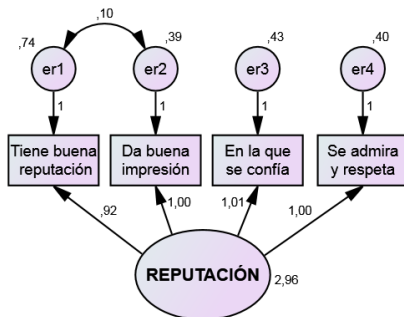
**Invarianza del índice PULSE  
por estudios  
Nivel alto  
ESPAÑA, 2012**



Chi-cuadrado=12,680  
g.l. = 6  
Chi square/df = 2,113  
p-valor = ,048  
CFI = 1,000  
NFI = ,999  
RMSEA=,015  
Pclose = 1,000

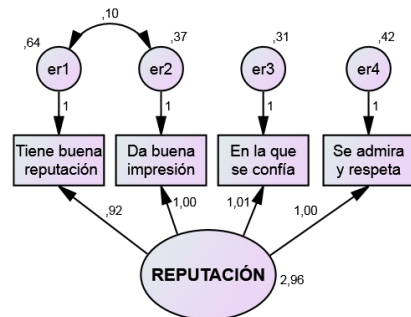
## Invarianza de covarianza errores

**Invarianza del índice PULSE  
por estudios  
Nivel bajo  
ESPAÑA, 2012**



Chi-cuadrado=18,339  
g.l. = 7  
Chi square/df = 2,620  
p-valor = ,011  
CFI = 1,000  
NFI = ,999  
RMSEA=,018  
Pclose = 1,000

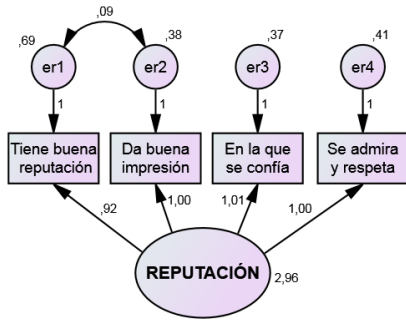
**Invarianza del índice PULSE  
por estudios  
Nivel alto  
ESPAÑA, 2012**



Chi-cuadrado=18,339  
g.l. = 7  
Chi square/df = 2,620  
p-valor = ,011  
CFI = 1,000  
NFI = ,999  
RMSEA=,018  
Pclose = 1,000

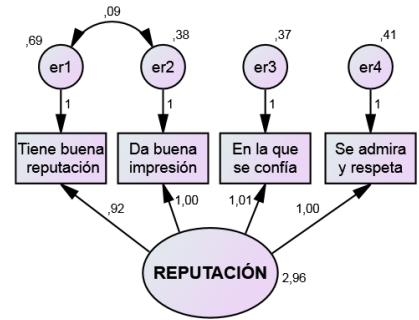
**Invarianza estricta**

**Invarianza del índice PULSE  
por estudios  
Nivel bajo  
ESPAÑA, 2012**



Chi-cuadrado=61,368  
g.l. = 11  
Chi square/df = 5,579  
p-valor = ,000  
CFI = ,998  
NFI = ,997  
RMSEA=,031  
Pclose = 1,000

**Invarianza del índice PULSE  
por estudios  
Nivel alto  
ESPAÑA, 2012**



Chi-cuadrado=61,368  
g.l. = 11  
Chi square/df = 5,579  
p-valor = ,000  
CFI = ,998  
NFI = ,997  
RMSEA=,031  
Pclose = 1,000

Tablas de comparación siendo verdadero el modelo de igualdad de cargas.

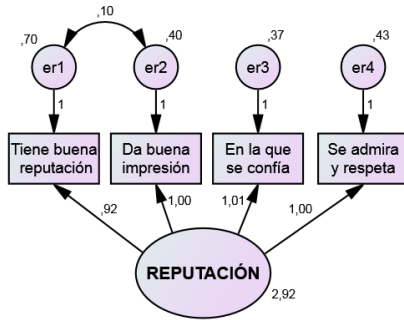
**Assuming model Invarianza covarianza errores to be correct:**

Model	DF	CMIN	P	NFI Delta-1	IFI Delta-2	RFI rho-1	TLI rho2
Invarianza estricta	4	43,030	,000	,002	,002	,002	,002

## 1.2.4. INVARIANZA POR SECTOR DE ACTIVIDAD. ESPAÑA.

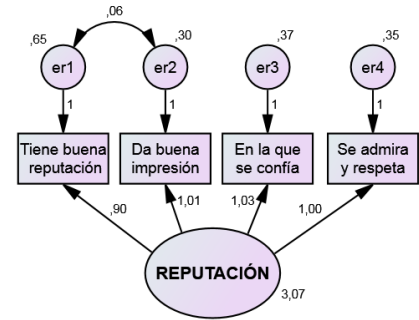
### Modelo configural

**Invarianza del índice PULSE  
por sector de actividad  
Bancario  
ESPAÑA, 2012**



Chi-cuadrado=1,370  
g.l. = 2  
Chi square/df = ,685  
p-valor = ,504  
CFI =1,000  
NFI =1,000  
RMSEA=,000  
Pclose = 1,000

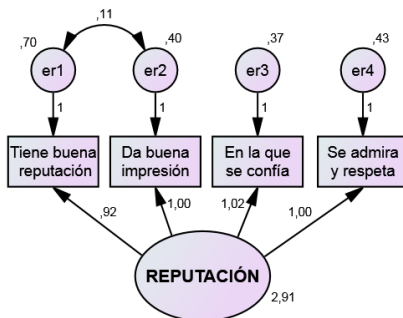
**Invarianza del índice PULSE  
por sector de actividad  
Telecomunicaciones  
ESPAÑA, 2012**



Chi-cuadrado=1,370  
g.l. = 2  
Chi square/df = ,685  
p-valor = ,504  
CFI =1,000  
NFI =1,000  
RMSEA=,000  
Pclose = 1,000

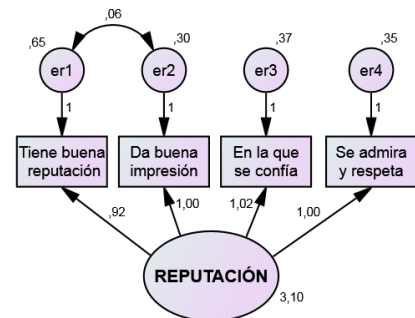
### Modelo invarianza de cargas

**Invarianza del índice PULSE  
por sector de actividad  
Bancario  
ESPAÑA, 2012**



Chi-cuadrado=6,900  
g.l. = 5  
Chi square/df = 1,380  
p-valor = ,228  
CFI =1,000  
NFI =1,000  
RMSEA=,009  
Pclose = 1,000

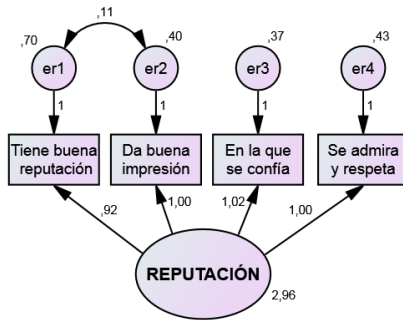
**Invarianza del índice PULSE  
por sector de actividad  
Telecomunicaciones  
ESPAÑA, 2012**



Chi-cuadrado=6,900  
g.l. = 5  
Chi square/df = 1,380  
p-valor = ,228  
CFI =1,000  
NFI =1,000  
RMSEA=,009  
Pclose = 1,000

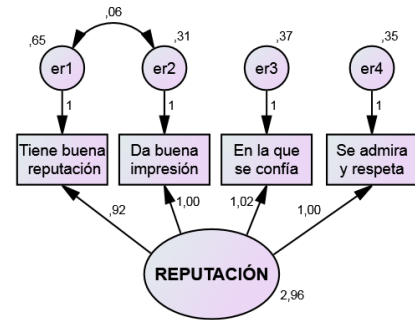
## Invarianza de varianza Pulse

**Invarianza del índice PULSE  
por sector de actividad  
Bancario  
ESPAÑA, 2012**



Chi-cuadrado=8,520  
g.l. = 6  
Chi square/df = 1,420  
p-valor = ,202  
CFI =1,000  
NFI =1,000  
RMSEA=,009  
Pclose = 1,000

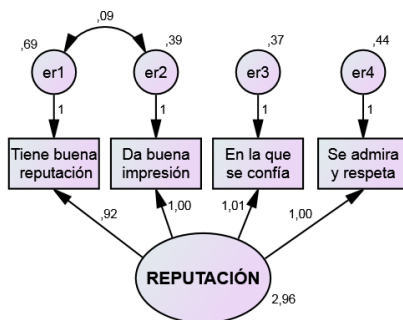
**Invarianza del índice PULSE  
por sector de actividad  
Telecomunicaciones  
ESPAÑA, 2012**



Chi-cuadrado=8,520  
g.l. = 6  
Chi square/df = 1,420  
p-valor = ,202  
CFI =1,000  
NFI =1,000  
RMSEA=,009  
Pclose = 1,000

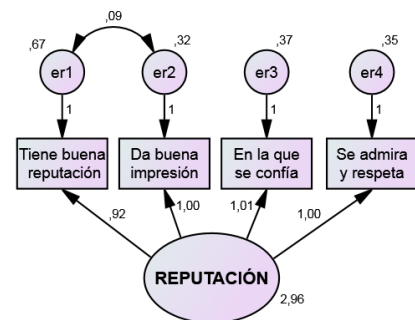
## Invarianza de covarianza errores

**Invarianza del índice PULSE  
por sector de actividad  
Bancario  
ESPAÑA, 2012**



Chi-cuadrado=11,959  
g.l. = 7  
Chi square/df = 1,708  
p-valor = ,102  
CFI =1,000  
NFI =,999  
RMSEA=,012  
Pclose = 1,000

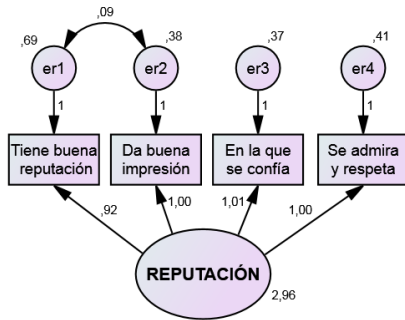
**Invarianza del índice PULSE  
por sector de actividad  
Telecomunicaciones  
ESPAÑA, 2012**



Chi-cuadrado=11,959  
g.l. = 7  
Chi square/df = 1,708  
p-valor = ,102  
CFI =1,000  
NFI =,999  
RMSEA=,012  
Pclose = 1,000

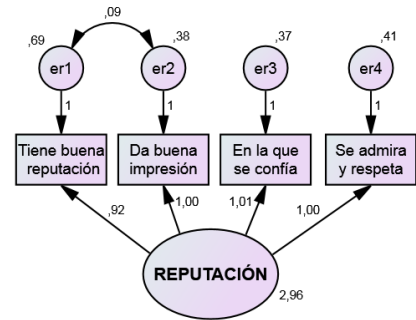
**Invarianza estricta**

**Invarianza del índice PULSE  
por sector de actividad  
Bancario  
ESPAÑA, 2012**



Chi-cuadrado=37,715  
g.l. = 11  
Chi square/df = 3,429  
p-valor = ,000  
CFI =,999  
NFI =,998  
RMSEA=,023  
Pclose = 1,000

**Invarianza del índice PULSE  
por sector de actividad  
Telecomunicaciones  
ESPAÑA, 2012**



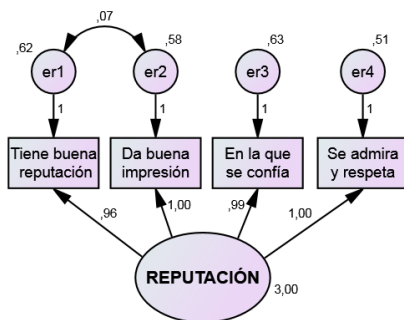
Chi-cuadrado=37,715  
g.l. = 11  
Chi square/df = 3,429  
p-valor = ,000  
CFI =,999  
NFI =,998  
RMSEA=,023  
Pclose = 1,000

El modelo verdadero es el de invarianza estricta. No hay tabla de comparación de pérdidas de ajuste.

**1.2.5. INVARIANZA CONJUNTA POR SEXO, EDAD Y NIVEL DE ESTUDIOS.  
ESPAÑA.**

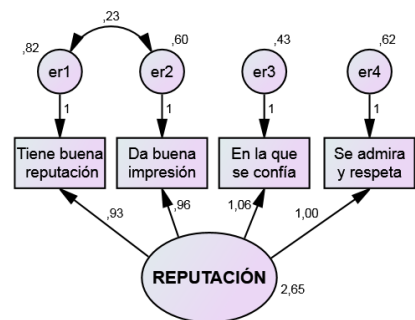
**Modelo configural**

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón, joven, nivel bajo  
ESPAÑA, 2012**



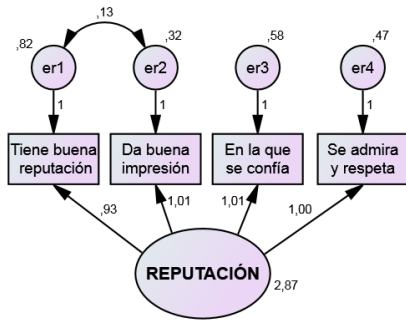
Chi-cuadrado=27,263  
g.l. = 8  
Chi square/df = 3,408  
p-valor = ,001  
CFI =,999  
NFI =,999  
RMSEA=,022  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón, joven, nivel alto  
ESPAÑA, 2012**



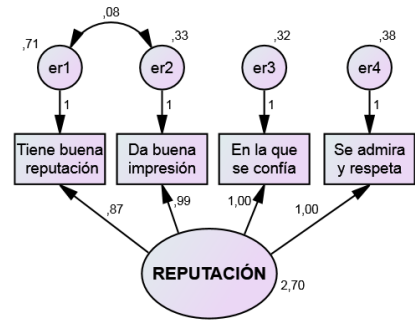
Chi-cuadrado=27,263  
g.l. = 8  
Chi square/df = 3,408  
p-valor = ,001  
CFI =,999  
NFI =,999  
RMSEA=,022  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel bajo  
ESPAÑA, 2012**



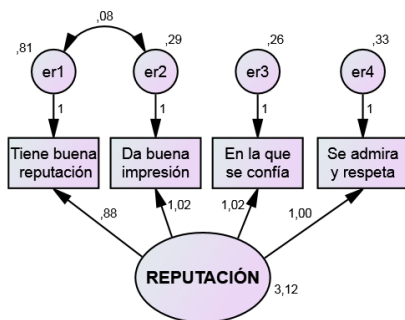
Chi-cuadrado=27,263  
g.l. = 8  
Chi square/df = 3,408  
p-valor = ,001  
CFI =,999  
NFI =,999  
RMSEA=,022  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel alto  
ESPAÑA, 2012**



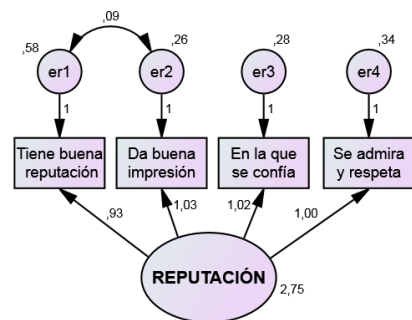
Chi-cuadrado=27,263  
g.l. = 8  
Chi square/df = 3,408  
p-valor = ,001  
CFI =,999  
NFI =,999  
RMSEA=,022

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel bajo  
ESPAÑA, 2012**



Chi-cuadrado=27,263  
g.l. = 8  
Chi square/df = 3,408  
p-valor = ,001  
CFI =,999  
NFI =,999  
RMSEA=,022  
Pclose = 1,000

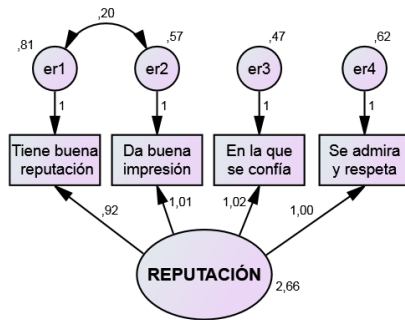
**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel alto  
ESPAÑA, 2012**



Chi-cuadrado=27,263  
g.l. = 8  
Chi square/df = 3,408  
p-valor = ,001  
CFI =,999  
NFI =,999  
RMSEA=,022  
Pclose = 1,000

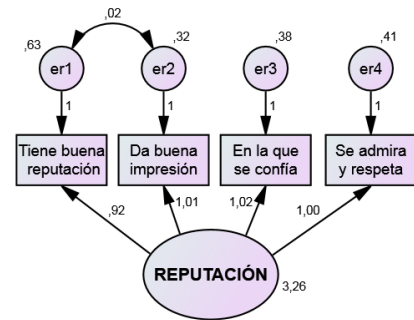
## Modelo invarianza de cargas

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón, joven, nivel alto  
ESPAÑA, 2012**



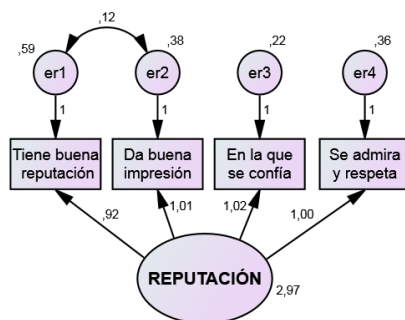
Chi-cuadrado=57,563  
g.l. = 29  
Chi square/df = 1,985  
p-valor = ,001  
CFI =,999  
NFI =,998  
RMSEA=,014  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón, adulto, nivel bajo  
ESPAÑA, 2012**



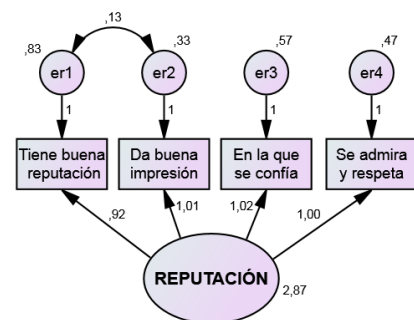
Chi-cuadrado=57,563  
g.l. = 29  
Chi square/df = 1,985  
p-valor = ,001  
CFI =,999  
NFI =,998  
RMSEA=,014  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón, adulto, nivel alto  
ESPAÑA, 2012**



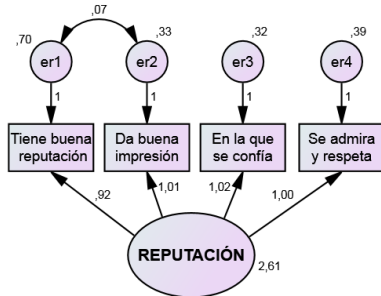
Chi-cuadrado=57,563  
g.l. = 29  
Chi square/df = 1,985  
p-valor = ,001  
CFI =,999  
NFI =,998  
RMSEA=,014  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel bajo  
ESPAÑA, 2012**



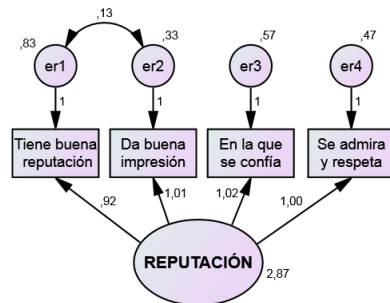
Chi-cuadrado=57,563  
g.l. = 29  
Chi square/df = 1,985  
p-valor = ,001  
CFI =,999  
NFI =,998  
RMSEA=,014  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel alto  
ESPAÑA, 2012**



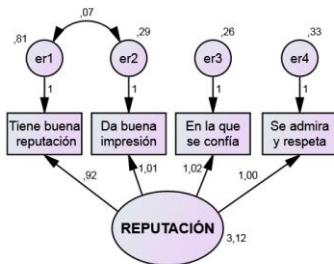
Chi-cuadrado=57,563  
g.l. = 29  
Chi square/df = 1,985  
p-valor = ,001  
CFI = ,999  
NFI = ,998  
RMSEA=,014

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel bajo  
ESPAÑA, 2012**



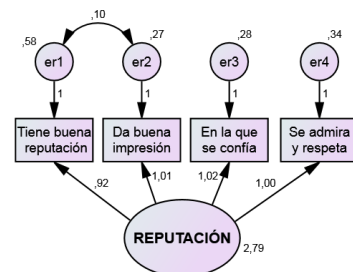
Chi-cuadrado=57,563  
g.l. = 29  
Chi square/df = 1,985  
p-valor = ,001  
CFI = ,999  
NFI = ,998  
RMSEA=,014  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel bajo  
ESPAÑA, 2012**



Chi-cuadrado=57,563  
g.l. = 29  
Chi square/df = 1,985  
p-valor = ,001  
CFI = ,999  
NFI = ,998  
RMSEA=,014  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel alto  
ESPAÑA, 2012**

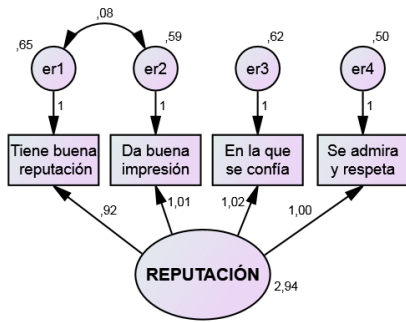


Chi-cuadrado=57,563  
g.l. = 29  
Chi square/df = 1,985  
p-valor = ,001  
CFI = ,999  
NFI = ,998  
RMSEA=,014  
Pclose = 1,000



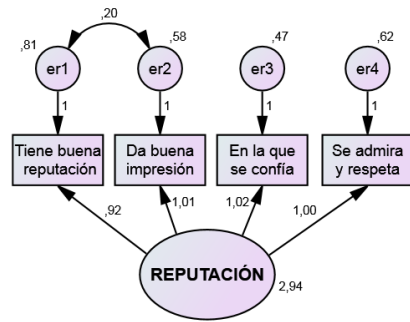
**Invarianza de varianza Pulse**

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón, joven, nivel bajo  
ESPAÑA, 2012**



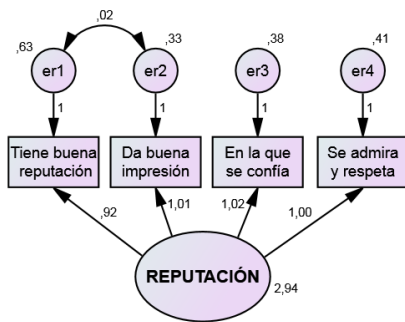
Chi-cuadrado=69,482  
g.l. = 36  
Chi square/df = 1,930  
p-valor = ,001  
CFI =,999  
NFI =,997  
RMSEA=,014  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón, joven, nivel alto  
ESPAÑA, 2012**



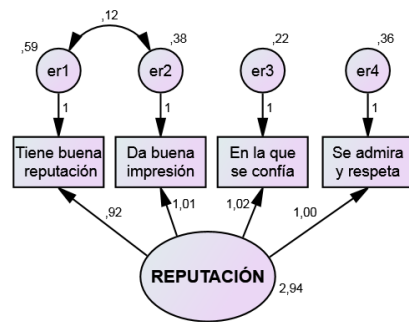
Chi-cuadrado=69,482  
g.l. = 36  
Chi square/df = 1,930  
p-valor = ,001  
CFI =,999  
NFI =,997  
RMSEA=,014  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón, adulto, nivel bajo  
ESPAÑA, 2012**



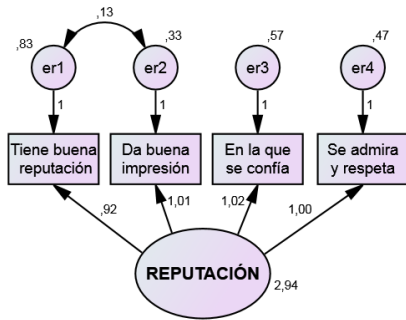
Chi-cuadrado=69,482  
g.l. = 36  
Chi square/df = 1,930  
p-valor = ,001  
CFI =,999  
NFI =,997  
RMSEA=,014  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón, adulto, nivel alto  
ESPAÑA, 2012**



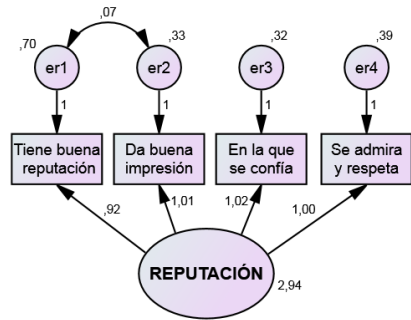
Chi-cuadrado=69,482  
g.l. = 36  
Chi square/df = 1,930  
p-valor = ,001  
CFI =,999  
NFI =,997  
RMSEA=,014  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel bajo  
ESPAÑA, 2012**



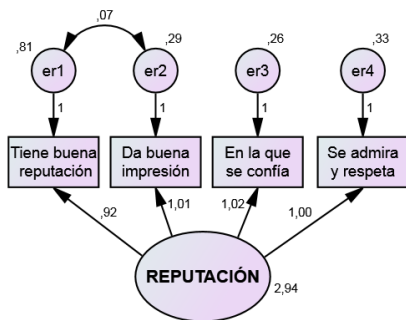
Chi-cuadrado=69,482  
g.l. = 36  
Chi square/df = 1,930  
p-valor = ,001  
CFI = ,999

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel alto  
ESPAÑA, 2012**



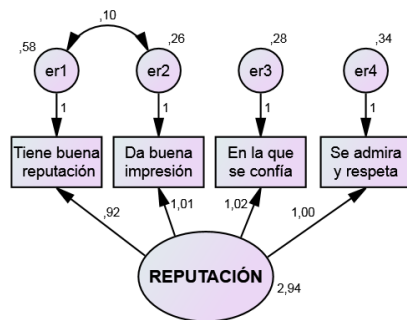
Chi-cuadrado=69,482  
g.l. = 36  
Chi square/df = 1,930  
p-valor = ,001

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel bajo  
ESPAÑA, 2012**



Chi-cuadrado=69,482  
g.l. = 36  
Chi square/df = 1,930  
p-valor = ,001  
CFI = ,999  
NFI = ,997  
RMSEA = ,014  
Pclose = 1,000

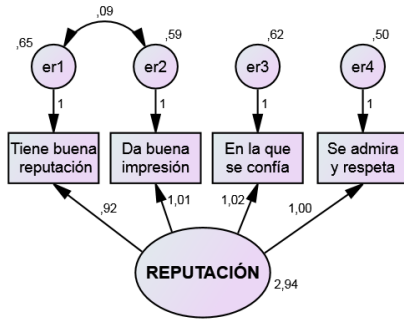
**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel alto  
ESPAÑA, 2012**



Chi-cuadrado=69,482  
g.l. = 36  
Chi square/df = 1,930  
p-valor = ,001  
CFI = ,999  
NFI = ,997  
RMSEA = ,014  
Pclose = 1,000

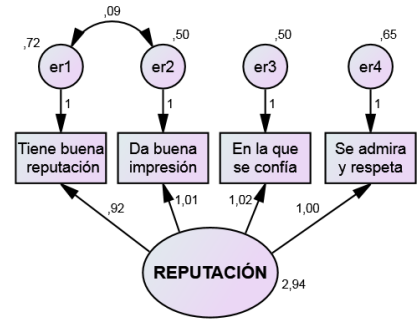
**Invarianza de covarianza errores**

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón, joven, nivel bajo  
ESPAÑA, 2012**



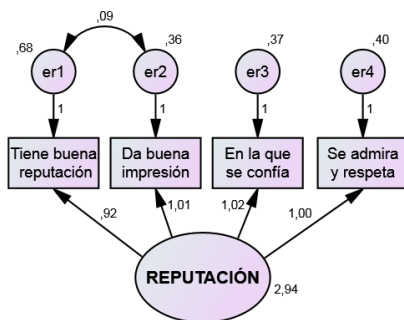
Chi-cuadrado=85,374  
g.l. = 43  
Chi square/df = 1,985  
p-valor = ,000  
CFI =,998  
NFI =,996  
RMSEA=,014  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón, joven, nivel alto  
ESPAÑA, 2012**



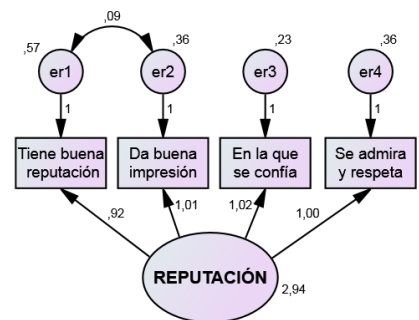
Chi-cuadrado=85,374  
g.l. = 43  
Chi square/df = 1,985  
p-valor = ,000  
CFI =,998  
NFI =,996  
RMSEA=,014  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón, adulto, nivel bajo  
ESPAÑA, 2012**



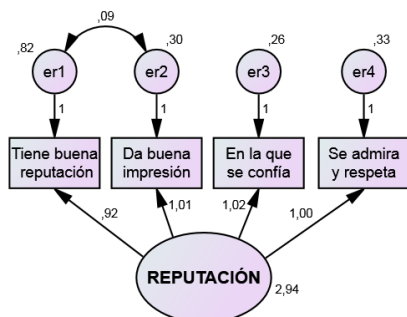
Chi-cuadrado=85,374  
g.l. = 43  
Chi square/df = 1,985  
p-valor = ,000  
CFI =,998  
NFI =,996  
RMSEA=,014  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón, adulto, nivel alto  
ESPAÑA, 2012**



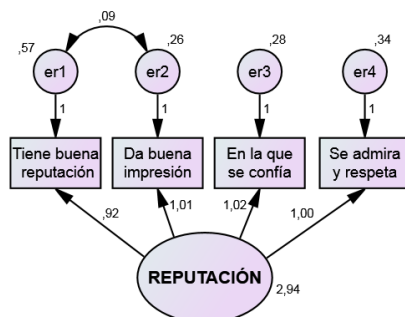
Chi-cuadrado=85,374  
g.l. = 43  
Chi square/df = 1,985  
p-valor = ,000  
CFI =,998  
NFI =,996  
RMSEA=,014  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel bajo  
ESPAÑA, 2012**



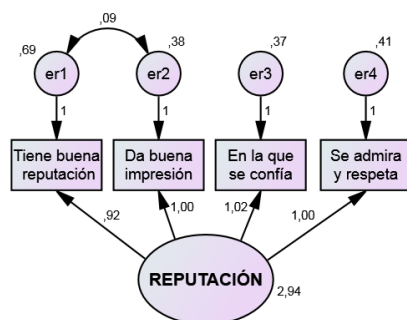
Chi-cuadrado=85,374  
g.l. = 43  
Chi square/df = 1,985  
p-valor = ,000  
CFI = ,998  
NFI = ,996  
RMSEA=,014  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel alto  
ESPAÑA, 2012**



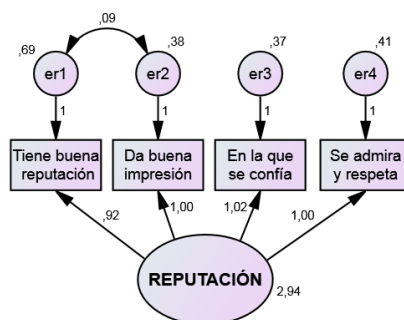
Chi-cuadrado=85,374  
g.l. = 43  
Chi square/df = 1,985  
p-valor = ,000  
CFI = ,998  
NFI = ,996  
RMSEA=,014  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón, joven, nivel bajo  
ESPAÑA, 2012**



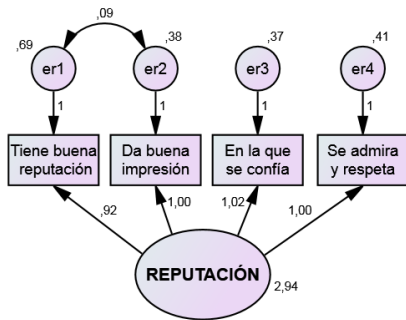
Chi-cuadrado=476,117  
g.l. = 71  
Chi square/df = 6,706  
p-valor = ,000  
CFI = ,983  
NFI = ,980  
RMSEA=,035  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón, joven, nivel alto  
ESPAÑA, 2012**



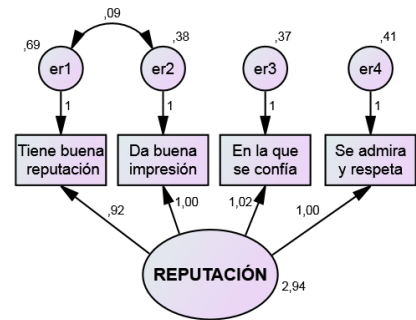
Chi-cuadrado=476,117  
g.l. = 71  
Chi square/df = 6,706  
p-valor = ,000  
CFI = ,983  
NFI = ,980  
RMSEA=,035  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón, adulto, nivel bajo  
ESPAÑA, 2012**



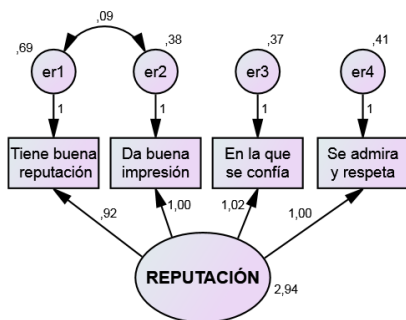
Chi-cuadrado=476,117  
g.l. = 71  
Chi square/df = 6,706  
p-valor = ,000  
CFI =,983  
NFI =,980  
RMSEA=,035  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón, adulto, nivel alto  
ESPAÑA, 2012**



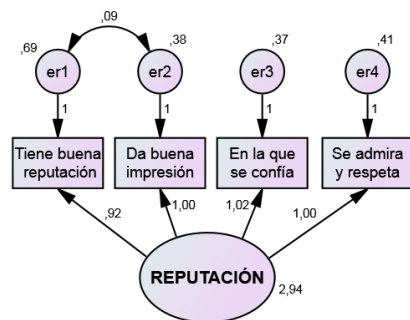
Chi-cuadrado=476,117  
g.l. = 71  
Chi square/df = 6,706  
p-valor = ,000  
CFI =,983  
NFI =,980  
RMSEA=,035  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel bajo  
ESPAÑA, 2012**



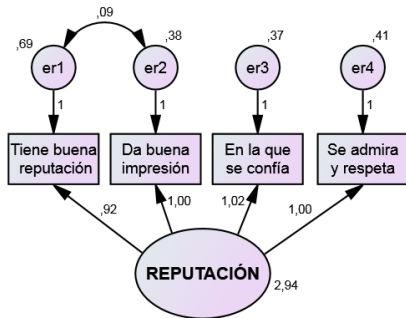
Chi-cuadrado=476,117  
g.l. = 71  
Chi square/df = 6,706  
p-valor = ,000  
CFI =,983  
NFI =,980  
RMSEA=,035  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel alto  
ESPAÑA, 2012**



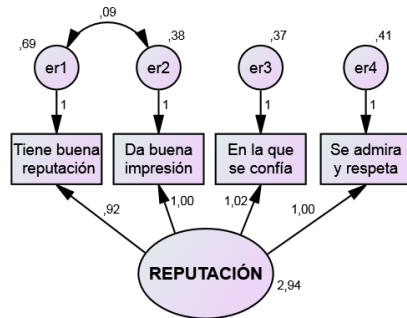
Chi-cuadrado=476,117  
g.l. = 71  
Chi square/df = 6,706  
p-valor = ,000  
CFI =,983  
NFI =,980  
RMSEA=,035

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel bajo  
ESPAÑA, 2012**



Chi-cuadrado=476,117  
g.l. = 71  
Chi square/df = 6,706  
p-valor = ,000  
CFI = ,983  
NFI = ,980  
RMSEA=,035  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel alto  
ESPAÑA, 2012**



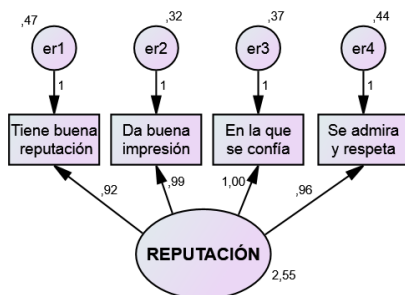
Chi-cuadrado=476,117  
g.l. = 71  
Chi square/df = 6,706  
p-valor = ,000  
CFI = ,983  
NFI = ,980  
RMSEA=,035  
Pclose = 1,000

U

**2.1. INVARIANZA POR SEXOS. EEUU.**

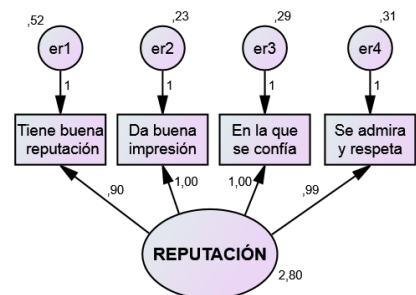
**Modelo configural**

**Invarianza del índice PULSE  
por sexo  
Varón  
Área: EEUU, 2012**



Chi-cuadrado=43,043  
g.l. = 4  
Chi square/df = 10,761  
p-valor = ,000  
CFI = ,998  
NFI = ,998  
RMSEA=,044  
Pclose = ,771

**Invarianza del índice PULSE  
por sexo  
Mujer  
Área: EEUU, 2012**

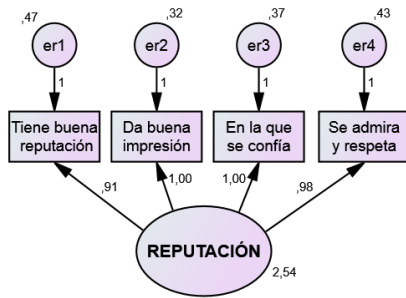


Chi-cuadrado=43,043  
g.l. = 4  
Chi square/df = 10,761  
p-valor = ,000  
CFI = ,998  
NFI = ,998  
RMSEA=,044  
Pclose = ,771



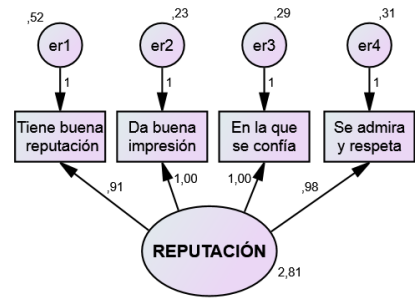
## Modelo invarianza de cargas

**Invarianza del índice PULSE  
por sexo  
Varón  
Área: EEUU, 2012**



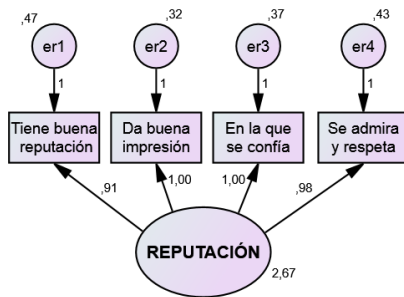
Chi-cuadrado=53,448  
g.l. = 7  
Chi square/df = 7,635  
p-valor = ,000  
CFI = ,998  
NFI = ,998  
RMSEA=,036  
Pclose = ,992

**Invarianza del índice PULSE  
por sexo  
Mujer  
Área: EEUU, 2012**



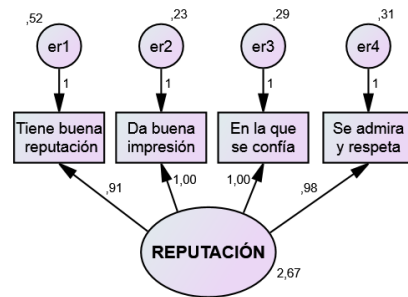
Chi-cuadrado=53,448  
g.l. = 7  
Chi square/df = 7,635  
p-valor = ,000  
CFI = ,998  
NFI = ,998  
RMSEA=,036  
Pclose = ,992

**Invarianza del índice PULSE  
por sexo  
Varón  
Área: EEUU, 2012**



Chi-cuadrado=59,418  
g.l. = 8  
Chi square/df = 7,427  
p-valor = ,000  
CFI = ,998  
NFI = ,998  
RMSEA=,036  
Pclose = ,997

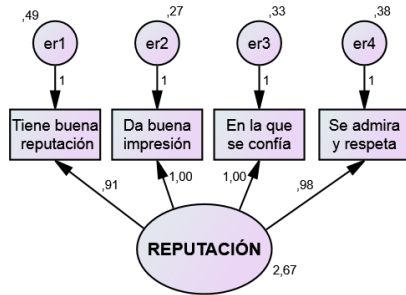
**Invarianza del índice PULSE  
por sexo  
Mujer  
Área: EEUU, 2012**



Chi-cuadrado=59,418  
g.l. = 8  
Chi square/df = 7,427  
p-valor = ,000  
CFI = ,998  
NFI = ,998  
RMSEA=,036  
Pclose = ,997

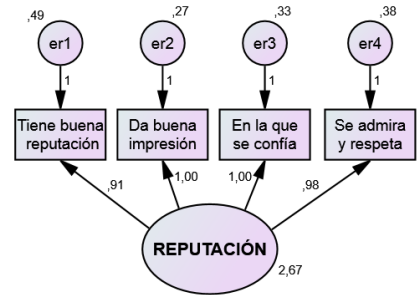
**Invarianza estricta**

**Invarianza del índice PULSE  
por sexo  
Varón  
Área: EEUU, 2012**



Chi-cuadrado=177,696  
g.l. = 12  
Chi square/df = 14,808  
p-valor = ,000  
CFI =,993  
NFI =,993  
RMSEA=,052  
Pclose = ,263

**Invarianza del índice PULSE  
por sexo  
Mujer  
Área: EEUU, 2012**



Chi-cuadrado=177,696  
g.l. = 12  
Chi square/df = 14,808  
p-valor = ,000  
CFI =,993  
NFI =,993  
RMSEA=,052  
Pclose = ,263

Test Bollen Stine asumiendo la invarianza estricta. P-valor = 0,000

Tablas de comparación siendo verdadero el modelo de igualdad de cargas.

**Assuming model Invarianza covarianza errores to be correct:**

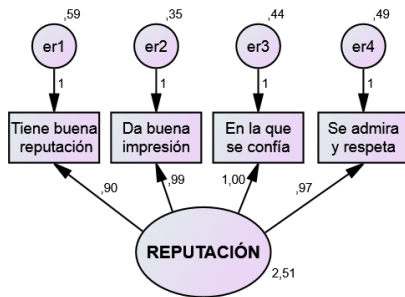
Model	DF	CMIN	P	NFI	IFI	RFI	TLI
				Delta-1	Delta-2	rho-1	rho2
invarianza estricta	4	118,279	,000	,005	,005	,004	,004



## 2.2. INVARIANZA POR EDAD. EEUU.

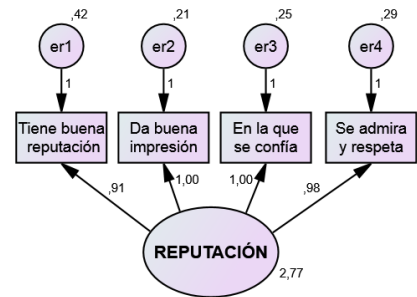
### Modelo configural

**Invarianza del índice PULSE  
por edad  
Joven  
Área: EEUU, 2012**



Chi-cuadrado=17,795  
g.l. = 4  
Chi square/df = 4,449  
p-valor = ,001  
CFI = ,999  
NFI = ,999  
RMSEA=,026  
Pclose = ,999

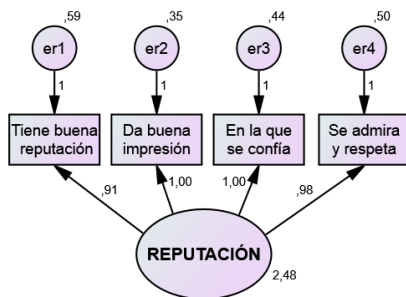
**Invarianza del índice PULSE  
por edad  
Adulto  
Área: EEUU, 2012**



Chi-cuadrado=17,795  
g.l. = 4  
Chi square/df = 4,449  
p-valor = ,001  
CFI = ,999  
NFI = ,999  
RMSEA=,026  
Pclose = ,999

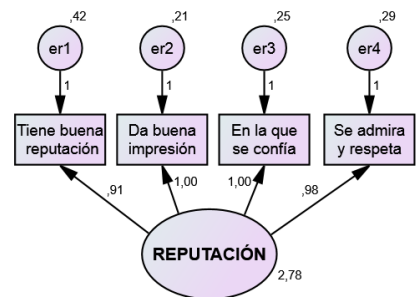
### Modelo invarianza de cargas

**Invarianza del índice PULSE  
por edad  
Joven  
Área: EEUU, 2012**



Chi-cuadrado=18,740  
g.l. = 7  
Chi square/df = 2,677  
p-valor = ,009  
CFI = 1,000  
NFI = ,999  
RMSEA=,018  
Pclose = 1,000

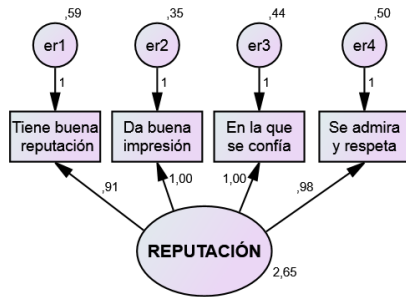
**Invarianza del índice PULSE  
por edad  
Adulto  
Área: EEUU, 2012**



Chi-cuadrado=18,740  
g.l. = 7  
Chi square/df = 2,677  
p-valor = ,009  
CFI = 1,000  
NFI = ,999  
RMSEA=,018  
Pclose = 1,000

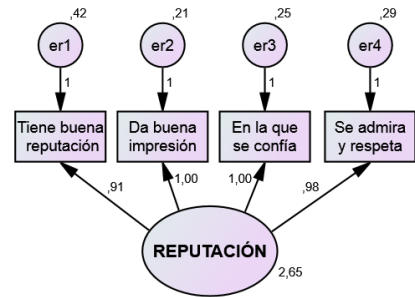
## Invarianza de varianza Pulse

**Invarianza del índice PULSE  
por edad  
Joven  
Área: EEUU, 2012**



Chi-cuadrado=25,853  
g.l. = 8  
Chi square/df = 3,232  
p-valor = ,001  
CFI =,999  
NFI =,999  
RMSEA=,021  
Pclose = 1,000

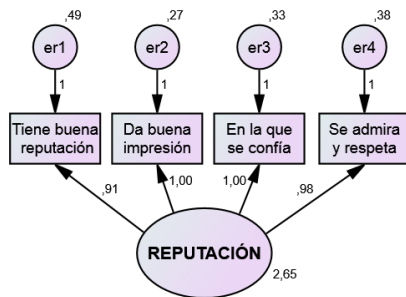
**Invarianza del índice PULSE  
por edad  
Adulto  
Área: EEUU, 2012**



Chi-cuadrado=25,853  
g.l. = 8  
Chi square/df = 3,232  
p-valor = ,001  
CFI =,999  
NFI =,999  
RMSEA=,021  
Pclose = 1,000

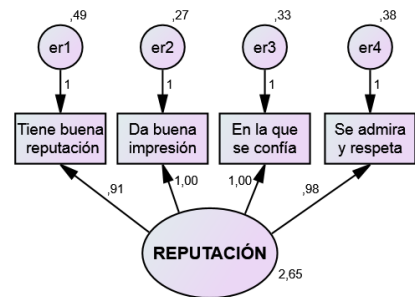
## Invarianza estricta

**Invarianza del índice PULSE  
por edad  
Joven  
Área: EEUU, 2012**



Chi-cuadrado=475,316  
g.l. = 12  
Chi square/df = 39,610  
p-valor = ,000  
CFI =,982  
NFI =,981  
RMSEA=,088  
Pclose = ,000

**Invarianza del índice PULSE  
por edad  
Adulto  
Área: EEUU, 2012**



Chi-cuadrado=475,316  
g.l. = 12  
Chi square/df = 39,610  
p-valor = ,000  
CFI =,982  
NFI =,981  
RMSEA=,088  
Pclose = ,000

Test Bollen Stine asumiendo la invarianza estricta. P-valor = 0,000

Tablas de comparación siendo verdadero el modelo de igualdad de cargas.

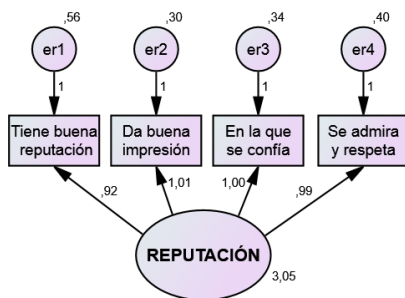
Assuming model Invarianza covarianza errores to be correct:

Model	DF	CMIN	P	NFI Delta-1	IFI Delta-2	RFI rho-1	TLI rho2
cargas, varianza Pulse y errores	4	449,463	,000	,018	,018	,017	,017

### 2.3. Invarianza por nivel de estudios. EEUU.

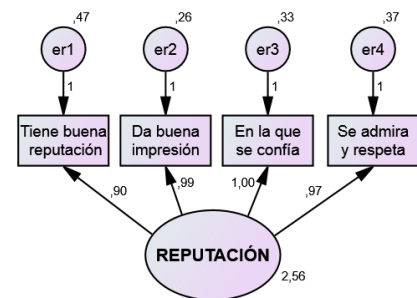
#### Modelo configural

Invarianza del índice PULSE  
por nivel de estudios  
Bajos  
Área: EEUU, 2012



Chi-cuadrado=12,830  
g.l. = 4  
Chi square/df = 3,208  
p-valor = ,012  
CFI = 1,000  
NFI = ,999  
RMSEA = ,021  
Pclose = 1,000

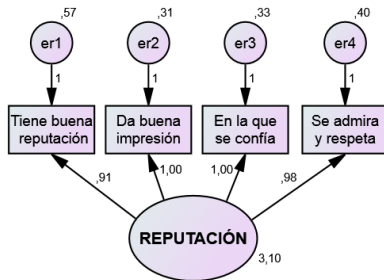
Invarianza del índice PULSE  
por nivel de estudios  
Alto  
Área: EEUU, 2012



Chi-cuadrado=12,830  
g.l. = 4  
Chi square/df = 3,208  
p-valor = ,012  
CFI = 1,000  
NFI = ,999  
RMSEA = ,021  
Pclose = 1,000

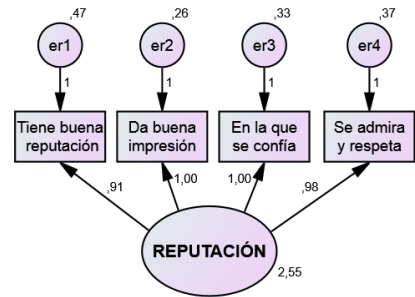
## Modelo invarianza de cargas

**Invarianza del índice PULSE  
por nivel de estudios  
Bajos  
Área: EEUU, 2012**



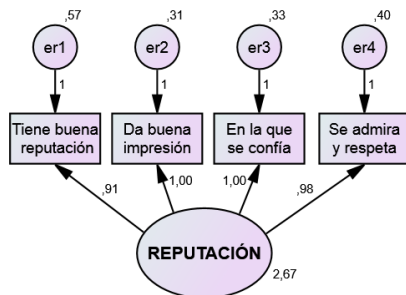
Chi-cuadrado=14,325  
g.l. = 7  
Chi square/df = 2,046  
p-valor = ,046  
CFI =1,000  
NFI =,999  
RMSEA=.014  
Pclose = 1,000

**Invarianza del índice PULSE  
por nivel de estudios  
Alto  
Área: EEUU, 2012**



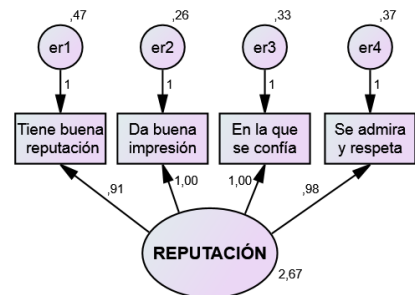
Chi-cuadrado=14,325  
g.l. = 7  
Chi square/df = 2,046  
p-valor = ,046  
CFI =1,000  
NFI =,999  
RMSEA=.014  
Pclose = 1,000

**Invarianza del índice PULSE  
por nivel de estudios  
Bajos  
Área: EEUU, 2012**



Chi-cuadrado=30,896  
g.l. = 8  
Chi square/df = 3,862  
p-valor = ,000  
CFI =,999  
NFI =,999  
RMSEA=.024  
Pclose = 1,000

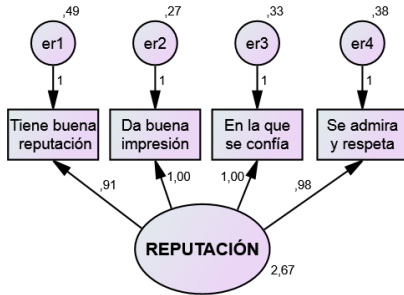
**Invarianza del índice PULSE  
por nivel de estudios  
Alto  
Área: EEUU, 2012**



Chi-cuadrado=30,896  
g.l. = 8  
Chi square/df = 3,862  
p-valor = ,000  
CFI =,999  
NFI =,999  
RMSEA=.024  
Pclose = 1,000

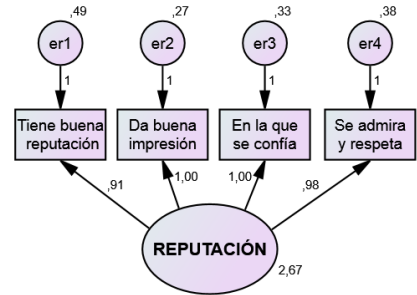
**Invarianza estricta**

**Invarianza del índice PULSE  
por nivel de estudios  
Bajos  
Área: EEUU, 2012**



Chi-cuadrado=52,725  
g.l. = 12  
Chi square/df = 4,394  
p-valor = ,000  
CFI =,998  
NFI =,998  
RMSEA=,026  
Pclose = 1,000

**Invarianza del índice PULSE  
por nivel de estudios  
Alto  
Área: EEUU, 2012**



Chi-cuadrado=52,725  
g.l. = 12  
Chi square/df = 4,394  
p-valor = ,000  
CFI =,998  
NFI =,998  
RMSEA=,026  
Pclose = 1,000

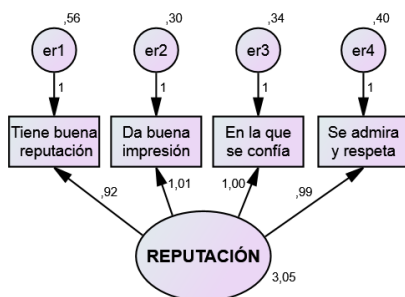
Test Bollen-Stine asumiendo el modelo de invarianza sea verdadero P-valor = 0,07. El modelo verdadero es el de invarianza estricta. No hay tabla de comparación de pérdidas de ajuste.

**2.4. INVARIANZA POR SECTOR DE ACTIVIDAD. EEUU.**

Invarianza por sector de actividad. EEUU

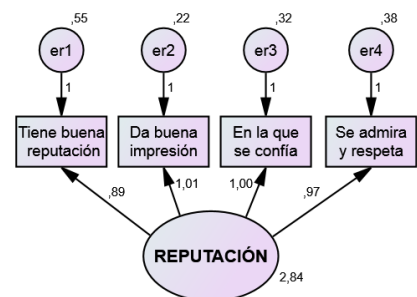
**Modelo configural**

**Invarianza del índice PULSE  
por sector de actividad  
Bancario  
Área: EEUU, 2012**



Chi-cuadrado=12,538  
g.l. = 4  
Chi square/df = 3,134  
p-valor = ,014  
CFI =,999  
NFI =,999  
RMSEA=,030  
Pclose = ,957

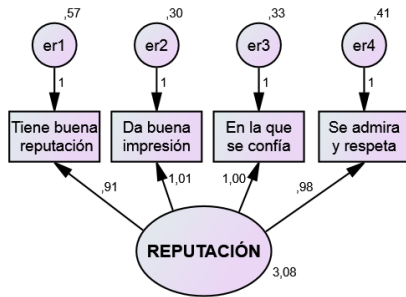
**Invarianza del índice PULSE  
por sector de actividad  
Telecomunicaciones  
Área: EEUU, 2012**



Chi-cuadrado=12,538  
g.l. = 4  
Chi square/df = 3,134  
p-valor = ,014  
CFI =,999  
NFI =,999  
RMSEA=,030  
Pclose = ,957

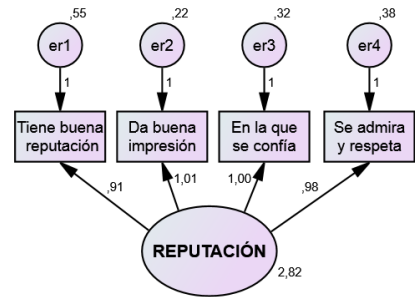
## Modelo invarianza de cargas

**Invarianza del índice PULSE  
por sector de actividad  
Bancario  
Área: EEUU, 2012**



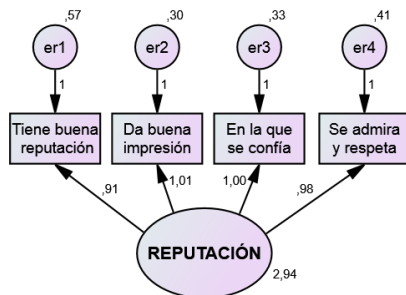
Chi-cuadrado=15,368  
g.l. = 7  
Chi square/df = 2,195  
p-valor = ,032  
CFI = ,999  
NFI = ,999  
RMSEA=,022  
Pclose = ,999

**Invarianza del índice PULSE  
por sector de actividad  
Telecomunicaciones  
Área: EEUU, 2012**



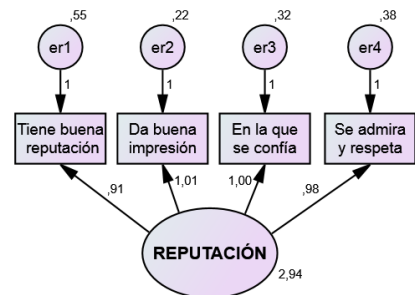
Chi-cuadrado=15,368  
g.l. = 7  
Chi square/df = 2,195  
p-valor = ,032  
CFI = ,999  
NFI = ,999  
RMSEA=,022  
Pclose = ,999

**Invarianza del índice PULSE  
por sector de actividad  
Bancario  
Área: EEUU, 2012**



Chi-cuadrado=17,600  
g.l. = 8  
Chi square/df = 2,200  
p-valor = ,024  
CFI = ,999  
NFI = ,999  
RMSEA=,022  
Pclose = 1,000

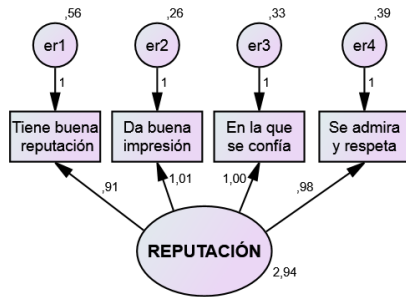
**Invarianza del índice PULSE  
por sector de actividad  
Telecomunicaciones  
Área: EEUU, 2012**



Chi-cuadrado=17,600  
g.l. = 8  
Chi square/df = 2,200  
p-valor = ,024  
CFI = ,999  
NFI = ,999  
RMSEA=,022  
Pclose = 1,000

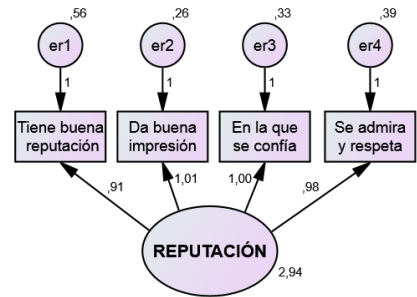
**Invarianza estricta**

**Invarianza del índice PULSE  
por sector de actividad  
Bancario  
Área: EEUU, 2012**



Chi-cuadrado=36,354  
g.l. = 12  
Chi square/df = 3,029  
p-valor = ,000  
CFI =,998  
NFI =,997  
RMSEA=,029  
Pclose = ,999

**Invarianza del índice PULSE  
por sector de actividad  
Telecomunicaciones  
Área: EEUU, 2012**



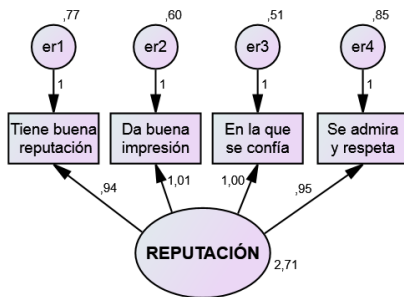
Chi-cuadrado=36,354  
g.l. = 12  
Chi square/df = 3,029  
p-valor = ,000  
CFI =,998  
NFI =,997  
RMSEA=,029  
Pclose = ,999

Test Bollen-Stine asumiendo el modelo de invarianza sea verdadero P-valor = 0,226. El modelo verdadero es el de invarianza estricta. No hay tabla de comparación de pérdidas de ajuste.

**2.5. INVARIANZA CONJUNTA POR SEXO, EDAD Y NIVEL DE ESTUDIOS.  
EEUU.**

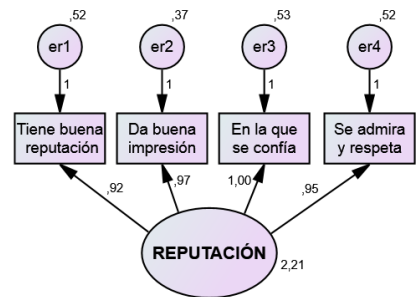
**Modelo configural**

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón joven, nivel bajo  
EEUU, 2012**



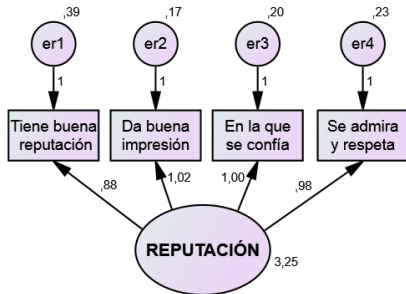
Chi-cuadrado=64,458  
g.l. = 16  
Chi square/df = 4,029  
p-valor = ,000  
CFI =,998  
NFI =,997  
RMSEA=,025  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón joven, nivel alto  
EEUU, 2012**



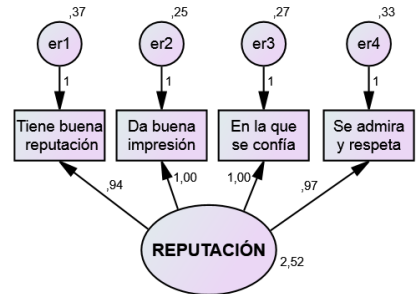
Chi-cuadrado=64,458  
g.l. = 16  
Chi square/df = 4,029  
p-valor = ,000  
CFI =,998  
NFI =,997  
RMSEA=,025  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón adulto, nivel bajo  
EEUU, 2012**



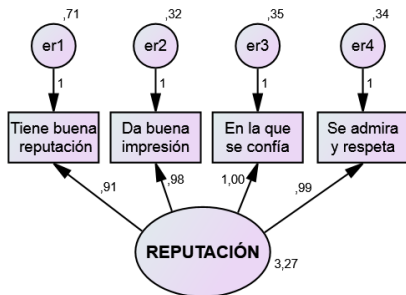
Chi-cuadrado=64,458  
g.l. = 16  
Chi square/df = 4,029  
p-valor = ,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón adulto, nivel alto  
EEUU, 2012**



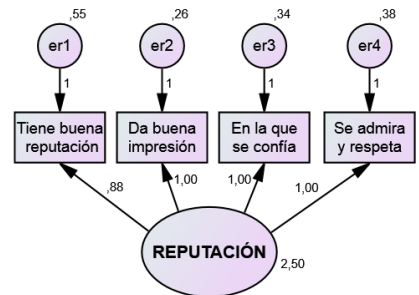
Chi-cuadrado=64,458  
g.l. = 16  
Chi square/df = 4,029  
p-valor = ,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel bajo  
EEUU, 2012**



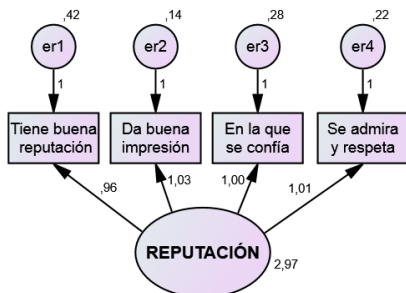
Chi-cuadrado=64,458  
g.l. = 16  
Chi square/df = 4,029  
p-valor = ,000  
CFI = ,998  
NFI = ,997  
RMSEA = ,025  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel alto  
EEUU, 2012**



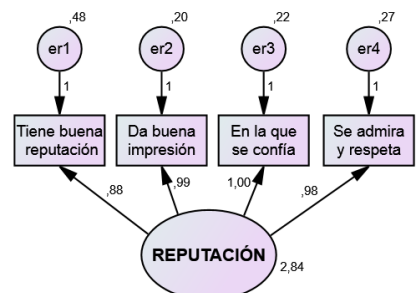
Chi-cuadrado=64,458  
g.l. = 16  
Chi square/df = 4,029  
p-valor = ,000  
CFI = ,998  
NFI = ,997  
RMSEA = ,025

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel bajo  
EEUU, 2012**



Chi-cuadrado=64,458  
g.l. = 16  
Chi square/df = 4,029  
p-valor = ,000  
CFI = ,998  
NFI = ,997  
RMSEA = ,025  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel alto  
EEUU, 2012**

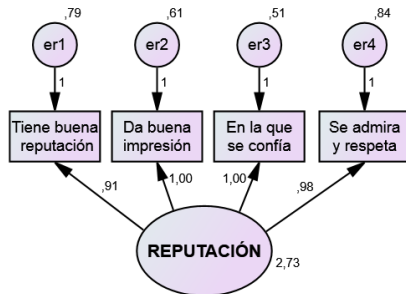


Chi-cuadrado=64,458  
g.l. = 16  
Chi square/df = 4,029  
p-valor = ,000  
CFI = ,998  
NFI = ,997  
RMSEA = ,025  
Pclose = 1,000



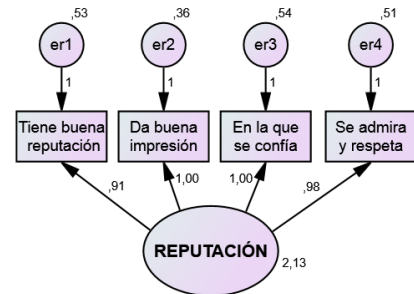
## Modelo invarianza de cargas

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón joven, nivel bajo  
EEUU, 2012**



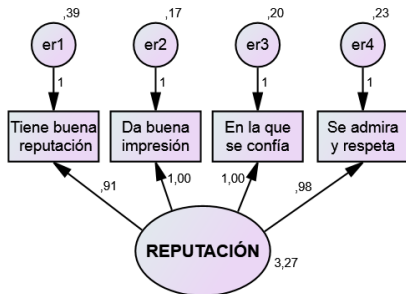
Chi-cuadrado=95,694  
g.l. = 37  
Chi square/df = 2,586  
p-valor = ,000  
CFI = ,998  
NFI = ,996  
RMSEA=,018  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón joven, nivel alto  
EEUU, 2012**



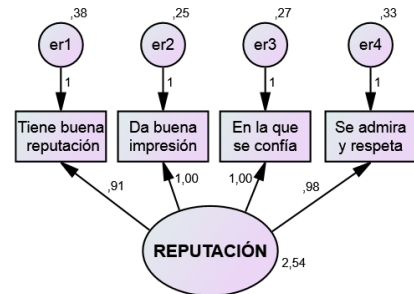
Chi-cuadrado=95,694  
g.l. = 37  
Chi square/df = 2,586  
p-valor = ,000  
CFI = ,998  
NFI = ,996  
RMSEA=,018  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón adulto, nivel bajo  
EEUU, 2012**



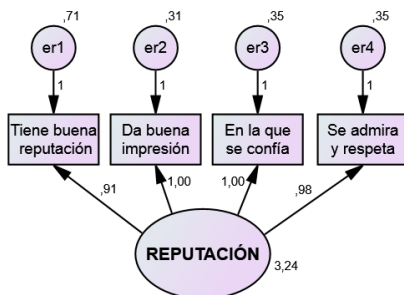
Chi-cuadrado=95,694  
g.l. = 37  
Chi square/df = 2,586  
p-valor = ,000  
CFI =,998  
NFI =,996  
RMSEA=,018  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón adulto, nivel alto  
EEUU, 2012**



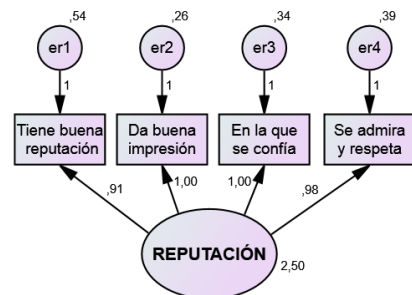
Chi-cuadrado=95,694  
g.l. = 37  
Chi square/df = 2,586  
p-valor = ,000  
CFI =,998  
NFI =,996  
RMSEA=,018  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel bajo  
EEUU, 2012**



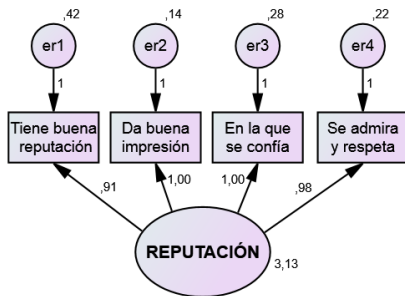
Chi-cuadrado=95,694  
g.l. = 37  
Chi square/df = 2,586  
p-valor = ,000  
CFI =,998  
NFI =,996  
RMSEA=,018  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel alto  
EEUU, 2012**



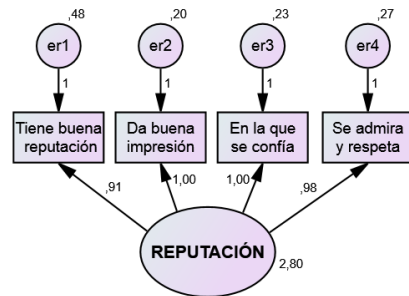
Chi-cuadrado=95,694  
g.l. = 37  
Chi square/df = 2,586  
p-valor = ,000  
CFI =,998  
NFI =,996  
RMSEA=,018

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel bajo  
EEUU, 2012**



Chi-cuadrado=95,694  
g.l. = 37  
Chi square/df = 2,586  
p-valor = ,000  
CFI = ,998  
NFI = ,996  
RMSEA=,018  
Pclose = 1,000

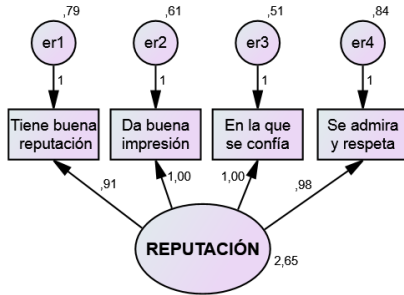
**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel alto  
EEUU, 2012**



Chi-cuadrado=95,694  
g.l. = 37  
Chi square/df = 2,586  
p-valor = ,000  
CFI = ,998  
NFI = ,996  
RMSEA=,018  
Pclose = 1,000

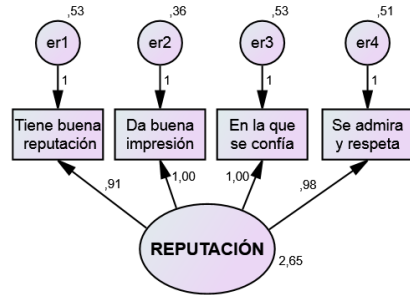
## Invarianza de varianza Pulse

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón joven, nivel bajo  
EEUU, 2012**



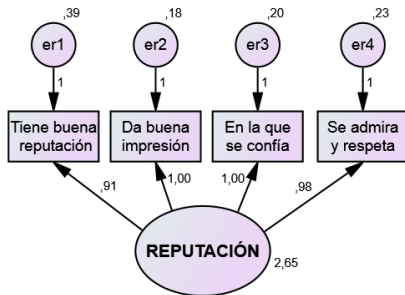
Chi-cuadrado=132,681  
g.l. = 44  
Chi square/df = 3,015  
p-valor = ,000  
CFI =,996  
NFI =,995  
RMSEA=,020  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón joven, nivel alto  
EEUU, 2012**



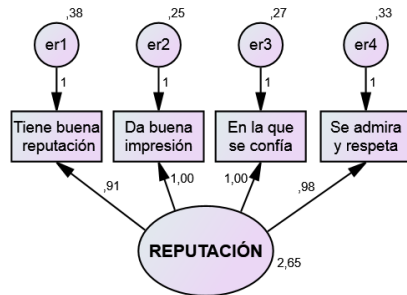
Chi-cuadrado=132,681  
g.l. = 44  
Chi square/df = 3,015  
p-valor = ,000  
CFI =,996  
NFI =,995  
RMSEA=,020  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón adulto, nivel bajo  
EEUU, 2012**



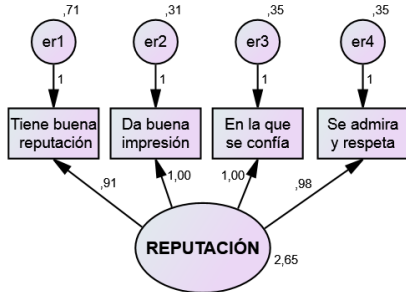
Chi-cuadrado=132,681  
g.l. = 44  
Chi square/df = 3,015  
p-valor = ,000  
CFI =,996  
NFI =,995  
RMSEA=,020  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón adulto, nivel alto  
EEUU, 2012**



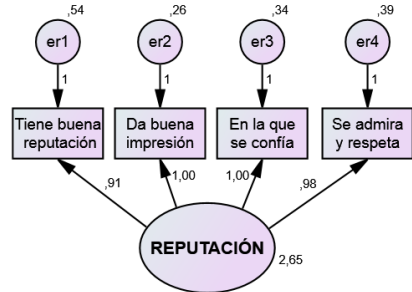
Chi-cuadrado=132,681  
g.l. = 44  
Chi square/df = 3,015  
p-valor = ,000  
CFI =,996  
NFI =,995  
RMSEA=,020  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel bajo  
EEUU, 2012**



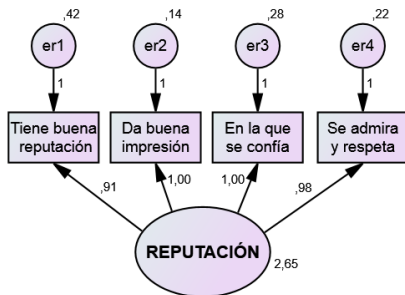
Chi-cuadrado=132,681  
g.l. = 44  
Chi square/df = 3,015  
p-valor = ,000  
CFI =,996  
NFI =,995  
RMSEA=,020  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel alto  
EEUU, 2012**



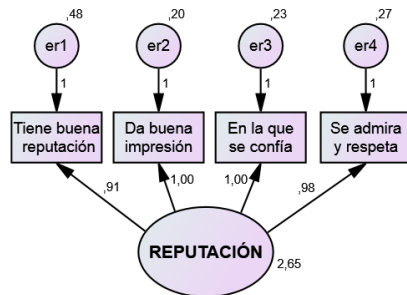
Chi-cuadrado=132,681  
g.l. = 44  
Chi square/df = 3,015  
p-valor = ,000  
CFI =,996  
NFI =,995  
RMSEA=,020

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel bajo  
EEUU, 2012**



Chi-cuadrado=132,681  
g.l. = 44  
Chi square/df = 3,015  
p-valor = ,000  
CFI =,996  
NFI =,995  
RMSEA=,020  
Pclose = 1,000

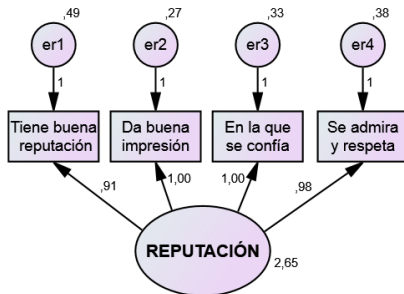
**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel alto  
EEUU, 2012**



Chi-cuadrado=132,681  
g.l. = 44  
Chi square/df = 3,015  
p-valor = ,000  
CFI =,996  
NFI =,995  
RMSEA=,020  
Pclose = 1,000

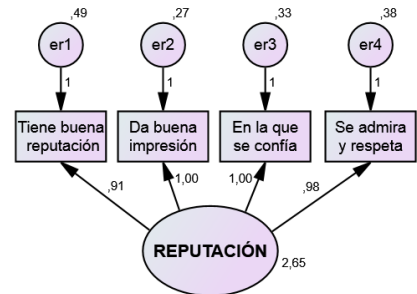
**Invarianza estricta**

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón joven, nivel bajo  
EEUU, 2012**



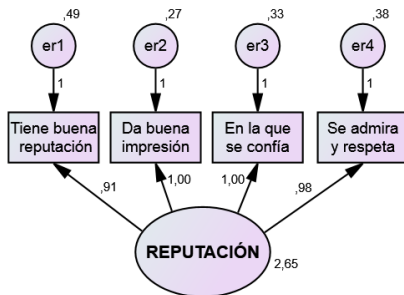
Chi-cuadrado=803,289  
g.l. = 72  
Chi square/df = 11,157  
p-valor = ,000  
CFI =,971  
NFI =,968  
RMSEA=,045  
Pclose = ,998

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón joven, nivel alto  
EEUU, 2012**



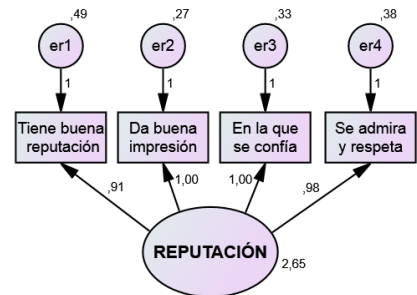
Chi-cuadrado=803,289  
g.l. = 72  
Chi square/df = 11,157  
p-valor = ,000  
CFI =,971  
NFI =,968  
RMSEA=,045  
Pclose = ,998

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón adulto, nivel bajo  
EEUU, 2012**



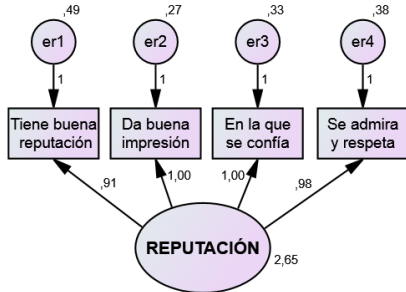
Chi-cuadrado=803,289  
g.l. = 72  
Chi square/df = 11,157  
p-valor = ,000  
CFI =,971  
NFI =,968  
RMSEA=,045  
Pclose = ,998

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Varón adulto, nivel alto  
EEUU, 2012**



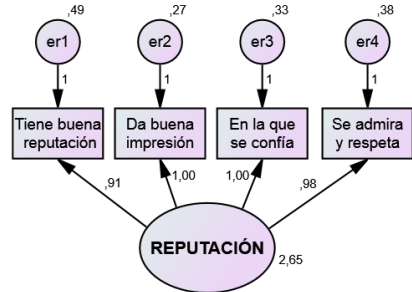
Chi-cuadrado=803,289  
g.l. = 72  
Chi square/df = 11,157  
p-valor = ,000  
CFI =,971  
NFI =,968  
RMSEA=,045  
Pclose = ,998

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel bajo  
EEUU, 2012**



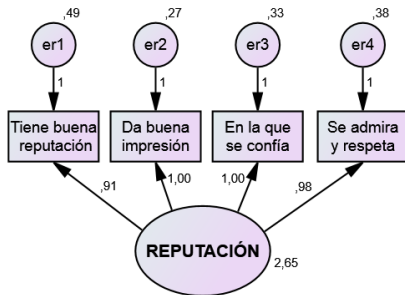
Chi-cuadrado=803,289  
g.l. = 72  
Chi square/df = 11,157  
p-valor = ,000  
CFI = ,971  
NFI = ,968  
RMSEA=,045  
Pclose = ,998

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer joven, nivel alto  
EEUU, 2012**



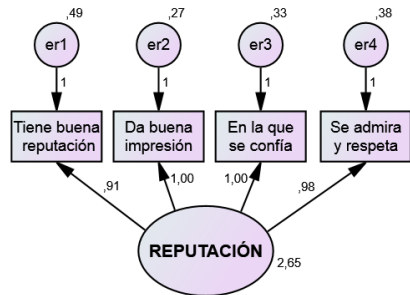
Chi-cuadrado=803,289  
g.l. = 72  
Chi square/df = 11,157  
p-valor = ,000  
CFI = ,971  
NFI = ,968  
RMSEA=,045

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel bajo  
EEUU, 2012**



Chi-cuadrado=803,289  
g.l. = 72  
Chi square/df = 11,157  
p-valor = ,000  
CFI = ,971  
NFI = ,968  
RMSEA=,045  
Pclose = ,998

**Invarianza del índice PULSE  
por sexo, edad y nivel de estudios  
Mujer adulta, nivel alto  
EEUU, 2012**



Chi-cuadrado=803,289  
g.l. = 72  
Chi square/df = 11,157  
p-valor = ,000  
CFI = ,971  
NFI = ,968  
RMSEA=,045  
Pclose = ,998

## 2.6. MATRICES RESIDUALES ESTANDARIZADAS POR SEXO. EEUU.

**Standardized Residual Covariances (Mujeres - configural)**

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
Se admira y respeta	0			
En la que se confía	0,225	0		
Da buena impresión	-0,074	-0,096	0	
Tiene buena reputación	-0,291	-0,185	0,36	0

**Standardized Residual Covariances (varones - configural)**

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
Se admira y respeta	0			
En la que se confía	-0,063	0		
Da buena impresión	-0,046	0,081	0	
Tiene buena reputación	0,168	-0,062	-0,063	0

**Standardized Residual Covariances (Mujeres - cargas factoriales)**

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
Se admira y respeta	0,567			
En la que se confía	0,509	-0,071		
Da buena impresión	0,257	-0,12	0,032	
Tiene buena reputación	-0,529	-0,762	-0,182	-0,929

**Standardized Residual Covariances (varones - cargas factoriales)**

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
Se admira y respeta	-0,787			
En la que se confía	-0,463	0,094		
Da buena impresión	-0,529	0,109	-0,046	
Tiene buena reputación	0,246	0,526	0,447	0,9

**Standardized Residual Covariances (Mujeres - cargas y varianza Pulse)**

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
Se admira y respeta	2,201			
En la que se confía	2,242	1,558		
Da buena impresión	1,986	1,595	1,695	
Tiene buena reputación	1,111	0,872	1,488	0,513

**Standardized Residual Covariances (varones - cargas y varianza Pulse)**

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
Se admira y respeta	-2,266			
En la que se confía	-2,079	-1,457		
Da buena impresión	-2,151	-1,555	-1,622	
Tiene buena reputación	-1,363	-1,107	-1,193	-0,573

**Standardized Residual Covariances (Mujeres - invarianza estricta)**

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
Se admira y respeta	2,164			
En la que se confía	2,217	1,059		
Da buena impresión	2	1,54	1,143	
Tiene buena reputación	1,183	0,877	1,529	0,852

**Standardized Residual Covariances (varones - invarianza estricta)**

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
Se admira y respeta	-2,391			
En la que se confía	-2,101	-0,92		
Da buena impresión	-2,136	-1,611	-1,097	
Tiene buena reputación	-1,29	-1,104	-1,153	-0,818



## 2.7. MATRICES RESIDUALES ESTANDARIZADAS POR GRUPOS HOMOGÉNEOS. EEUU.

**MODELO CONFIGURAL**  
**(varones jóvenes bajos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	0			
Ítem 3	0,113	0		
ítem 2	0,059	-0,184	0	
ítem 1	-0,171	0,084	0,102	0

**MODELO CONFIGURAL**  
**(varones jóvenes altos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	0			
Ítem 3	0,436	0		
ítem 2	-0,317	-0,076	0	
ítem 1	-0,156	-0,334	0,434	0

**MODELO CONFIGURAL**  
**(varones adultos bajos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	0			
Ítem 3	0,023	0		
ítem 2	-0,02	0,001	0	
ítem 1	0,003	-0,039	0,029	0

**MODELO CONFIGURAL**  
**(varones adultos altos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	0			
Ítem 3	0,159	0		
ítem 2	-0,062	-0,094	0	
ítem 1	-0,247	-0,076	0,309	0

**MODELO CONFIGURAL**  
**(mujeres jóvenes bajos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	0			
Ítem 3	0,237	0		
ítem 2	-0,1	0,011	0	
ítem 1	0,03	-0,381	0,132	0

**MODELO CONFIGURAL**  
**(mujeres jóvenes altos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	0			
Ítem 3	0,114	0		
ítem 2	-0,053	-0,041	0	
ítem 1	-0,128	-0,121	0,204	0

**MODELO CONFIGURAL**  
**(mujeres adultas bajos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	0			
Ítem 3	0,091	0		
ítem 2	-0,114	0,023	0	
ítem 1	0,104	-0,331	0,232	0

**MODELO CONFIGURAL**  
**(mujeres adultas altos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	0			
Ítem 3	0,195	0		
ítem 2	-0,043	-0,066	0	
ítem 1	-0,23	-0,1	0,183	0

**MODELO INVARIANZA DE CARGAS**  
**(varones jóvenes bajos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	0,01			
Ítem 3	-0,286	-0,665		
ítem 2	-0,023	-0,646	-0,14	
ítem 1	0,352	0,211	0,558	0,875

**MODELO INVARIANZA DE CARGAS**  
**(varones jóvenes altos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	-0,226			
Ítem 3	0,671	0,583		
ítem 2	-0,851	-0,161	-0,731	
ítem 1	0,084	0,376	0,365	0,621

**MODELO INVARIANZA DE CARGAS**  
**(varones adultos bajos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	-0,239			
Ítem 3	-0,057	0,099		
ítem 2	-0,136	0,076	0,034	
ítem 1	-0,051	0,093	0,126	0,134

**MODELO INVARIANZA DE CARGAS****(varones adultos altos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	0,261			
Ítem 3	0,288	-0,04		
ítem 2	-0,116	-0,314	-0,363	
ítem 1	0,09	0,103	0,308	0,346

**MODELO INVARIANZA DE CARGAS****(mujeres jóvenes bajos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	-0,151			
Ítem 3	-0,018	-0,264		
ítem 2	-0,131	-0,089	0,102	
ítem 1	0,096	-0,381	0,361	0,26

**MODELO INVARIANZA DE CARGAS****(mujeres jóvenes altos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	0,498			
Ítem 3	0,489	0,13		
ítem 2	0,199	-0,002	-0,054	
ítem 1	-0,427	-0,62	-0,419	-0,942

**MODELO INVARIANZA DE CARGAS****(mujeres adultas bajos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	0,099			
Ítem 3	0,278	0,236		
ítem 2	0,047	0,268	0,208	
ítem 1	-0,724	-1,072	-0,554	-1,45

**MODELO INVARIANZA DE CARGAS****(mujeres adultas altos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	-0,293			
Ítem 3	-0,149	-0,296		
ítem 2	-0,013	-0,037	0,356	
ítem 1	-0,351	-0,223	0,431	0,069

**MODELO INVARIANZA DE VARIANZA****(varones jóvenes bajos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	0,278			
Ítem 3	-0,003	-0,413		
ítem 2	0,269	-0,367	0,131	
ítem 1	0,647	0,502	0,861	1,15

**MODELO INVARIANZA DE VARIANZA****(varones jóvenes altos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	-1,477			
Ítem 3	-0,805	-0,784		
ítem 2	-2,202	-1,59	-2,035	
ítem 1	-1,312	-1,063	-1,083	-0,664

**MODELO INVARIANZA DE VARIANZA****(varones adultos bajos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	1,608			
Ítem 3	1,937	1,999		
ítem 2	1,866	2,108	1,98	
ítem 1	1,895	2,062	2,115	1,879

**MODELO INVARIANZA DE VARIANZA****(varones adultos altos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	0,427			
Ítem 3	0,467	0,133		
ítem 2	0,06	-0,136	-0,192	
ítem 1	0,265	0,281	0,49	0,51

**MODELO INVARIANZA DE VARIANZA****(mujeres jóvenes bajos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	-0,344			
Ítem 3	-0,226	-0,449		
ítem 2	-0,343	-0,298	-0,106	
ítem 1	-0,108	-0,576	0,15	0,082

**MODELO INVARIANZA DE VARIANZA****(mujeres jóvenes altos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	-1,155			
Ítem 3	-1,283	-1,522		
ítem 2	-1,548	-1,736	-1,713	
ítem 1	-2,043	-2,225	-2,051	-2,315

**MODELO INVARIANZA DE VARIANZA****(mujeres adultas bajos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	1,132			
Ítem 3	1,384	1,299		
ítem 2	1,145	1,386	1,28	
ítem 1	0,278	-0,082	0,469	-0,623

**MODELO INVARIANZA DE VARIANZA****(mujeres adultas altos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	-1,071			
Ítem 3	-0,981	-1,089		
ítem 2	-0,857	-0,883	-0,481	
ítem 1	-1,153	-1,034	-0,413	-0,668

**MODELO INVARIANZA ESTRICTA****(varones jóvenes bajos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	0,72			
Ítem 3	0,018	0,697		
ítem 2	0,319	-0,328	1,211	
ítem 1	0,682	0,554	0,94	1,055

**MODELO INVARIANZA ESTRICTA****(varones jóvenes altos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	-0,386			
Ítem 3	-0,811	-0,105		
ítem 2	-2,235	-1,577	-1,439	
ítem 1	-1,318	-1,041	-1,04	-0,411

**MODELO INVARIANZA ESTRICTA****(varones adultos bajos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	1,493			
Ítem 3	1,962	1,989		
ítem 2	1,92	2,18	1,986	
ítem 1	1,931	2,116	2,197	1,852

**MODELO INVARIANZA ESTRICTA****(varones adultos altos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	0,118			
Ítem 3	0,484	-0,599		
ítem 2	0,108	-0,073	-0,147	
ítem 1	0,296	0,323	0,558	-0,021

**MODELO INVARIANZA ESTRICTA****(mujeres jóvenes bajos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	-0,043			
Ítem 3	-0,214	0,447		
ítem 2	-0,308	-0,256	-0,348	
ítem 1	-0,084	-0,558	0,204	0,506

**MODELO INVARIANZA ESTRICTA****(mujeres jóvenes altos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	-1,276			
Ítem 3	-1,254	-1,736		
ítem 2	-1,501	-1,674	-1,789	
ítem 1	-2,017	-2,184	-1,996	-2,145

**MODELO INVARIANZA ESTRICTA****(mujeres adultas bajos)**

Standardized Residual Covariances

	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	0,543			
Ítem 3	1,387	0,636		
ítem 2	1,184	1,436	0,993	
ítem 1	0,316	-0,032	0,55	0,229

**MODELO INVARIANZA ESTRICTA****(mujeres adultas altos)**

Standardized Residual Covariances

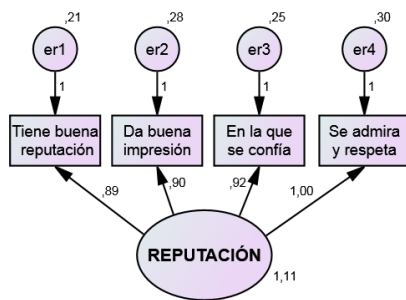
	Ítem 4	Ítem 3	Ítem 2	Ítem 1
ítem 4	-1,421			
Ítem 3	-0,944	-1,543		
ítem 2	-0,796	-0,806	-1,012	
ítem 1	-1,1	-0,968	-0,332	-1,266

### 3. ANÁLISIS DE INVARIANZA POBLACIONAL EN CHINA.

#### 3.1. INVARIANZA POR SEXO. CHINA.

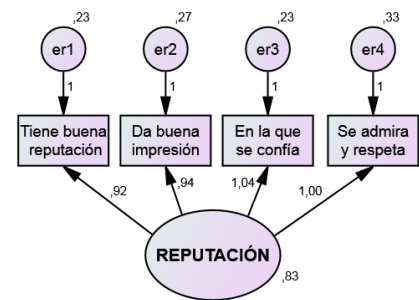
##### Modelo configural

**Invarianza del índice PULSE  
por sexo  
Varón  
Área: CHINA, 2012**



Chi-cuadrado=9,338  
g.l. = 4  
Chi square/df = 2,335  
p-valor = ,053  
CFI =,998  
NFI =,997  
RMSEA=,038  
Pclose = ,689

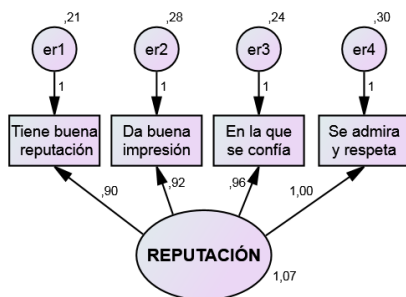
**Invarianza del índice PULSE  
por sexo  
Mujer  
Área: CHINA, 2012**



Chi-cuadrado=9,338  
g.l. = 4  
Chi square/df = 2,335  
p-valor = ,053  
CFI =,998  
NFI =,997  
RMSEA=,038  
Pclose = ,689

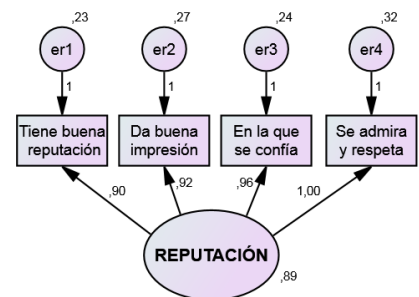
##### Modelo invarianza de cargas

**Invarianza del índice PULSE  
por sexo  
Varón  
Área: CHINA, 2012**



Chi-cuadrado=15,805  
g.l. = 7  
Chi square/df = 2,258  
p-valor = ,027  
CFI =,997  
NFI =,995  
RMSEA=,037  
Pclose = ,792

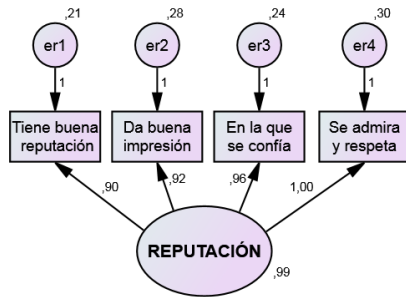
**Invarianza del índice PULSE  
por sexo  
Mujer  
Área: CHINA, 2012**



Chi-cuadrado=15,805  
g.l. = 7  
Chi square/df = 2,258  
p-valor = ,027  
CFI =,997  
NFI =,995  
RMSEA=,037  
Pclose = ,792

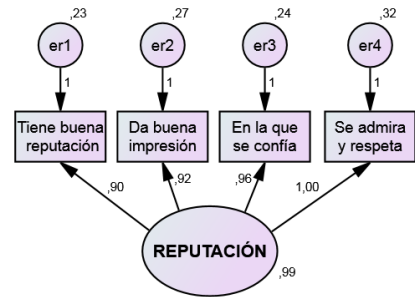
## Invarianza de varianza Pulse

**Invarianza del índice PULSE  
por sexo  
Varón  
Área: CHINA, 2012**



Chi-cuadrado=19,273  
g.l. = 8  
Chi square/df = 2,409  
p-valor = ,013  
CFI =,996  
NFI =,994  
RMSEA=,039  
Pclose = ,768

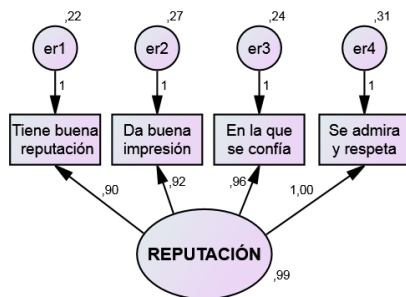
**Invarianza del índice PULSE  
por sexo  
Mujer  
Área: CHINA, 2012**



Chi-cuadrado=19,273  
g.l. = 8  
Chi square/df = 2,409  
p-valor = ,013  
CFI =,996  
NFI =,994  
RMSEA=,039  
Pclose = ,768

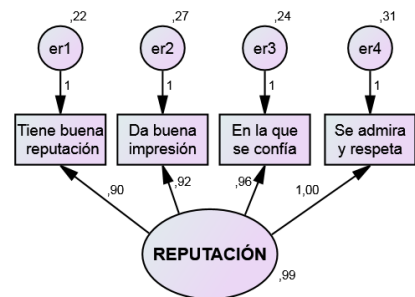
## Invarianza estricta

**Invarianza del índice PULSE  
por sexo  
Varón  
Área: CHINA, 2012**



Chi-cuadrado=19,978  
g.l. = 12  
Chi square/df = 1,665  
p-valor = ,067  
CFI =,997  
NFI =,993  
RMSEA=,027  
Pclose = ,974

**Invarianza del índice PULSE  
por sexo  
Mujer  
Área: CHINA, 2012**



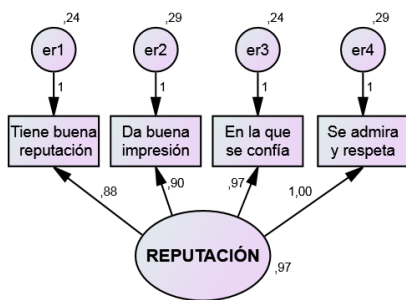
Chi-cuadrado=19,978  
g.l. = 12  
Chi square/df = 1,665  
p-valor = ,067  
CFI =,997  
NFI =,993  
RMSEA=,027  
Pclose = ,974

El modelo verdadero es el de invarianza estricta. No hay tabla de comparación de pérdidas de ajuste.

### 3.2. INVARIANZA POR EDAD. CHINA.

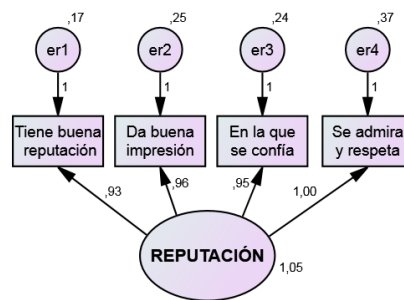
#### Modelo configural

**Invarianza del índice PULSE por edad Jóvenes CHINA, 2012**



Chi-cuadrado=10,645  
g.l. = 4  
Chi square/df = 2,661  
p-valor = ,031  
CFI =,998  
NFI =,997  
RMSEA=,042  
Pclose = ,606

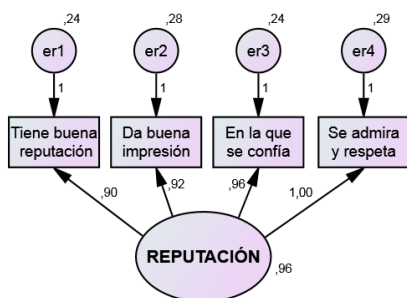
**Invarianza del índice PULSE por edad Adultos CHINA, 2012**



Chi-cuadrado=10,645  
g.l. = 4  
Chi square/df = 2,661  
p-valor = ,031  
CFI =,998  
NFI =,997  
RMSEA=,042  
Pclose = ,606

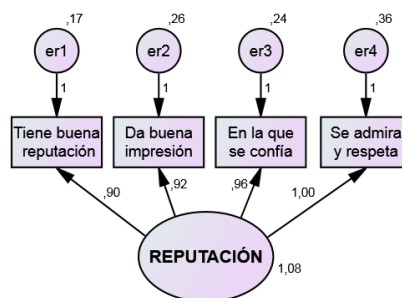
#### Modelo invarianza de cargas

**Invarianza del índice PULSE por edad Jóvenes CHINA, 2012**



Chi-cuadrado=14,457  
g.l. = 7  
Chi square/df = 2,065  
p-valor = ,044  
CFI =,998  
NFI =,995  
RMSEA=,034  
Pclose = ,843

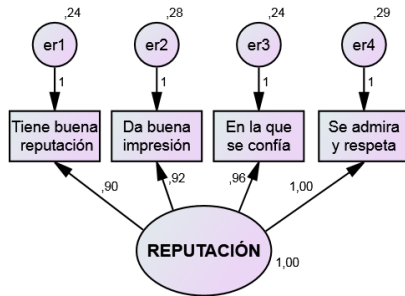
**Invarianza del índice PULSE por edad Adultos CHINA, 2012**



Chi-cuadrado=14,457  
g.l. = 7  
Chi square/df = 2,065  
p-valor = ,044  
CFI =,998  
NFI =,995  
RMSEA=,034  
Pclose = ,843

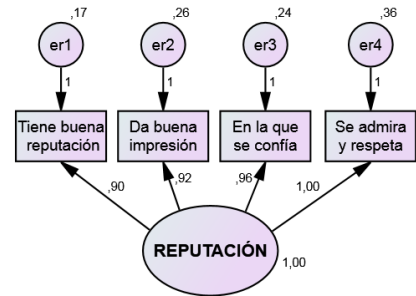
## Invarianza de varianza Pulse

**Invarianza del índice PULSE  
por edad  
Jóvenes  
CHINA, 2012**



Chi-cuadrado=15,673  
g.l. = 8  
Chi square/df = 1,959  
p-valor = ,047  
CFI = ,997  
NFI = ,995  
RMSEA=,032  
Pclose = ,887

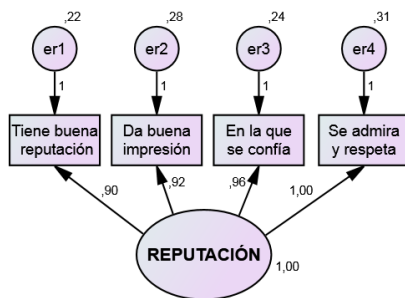
**Invarianza del índice PULSE  
por edad  
Adultos  
CHINA, 2012**



Chi-cuadrado=15,673  
g.l. = 8  
Chi square/df = 1,959  
p-valor = ,047  
CFI = ,997  
NFI = ,995  
RMSEA=,032  
Pclose = ,887

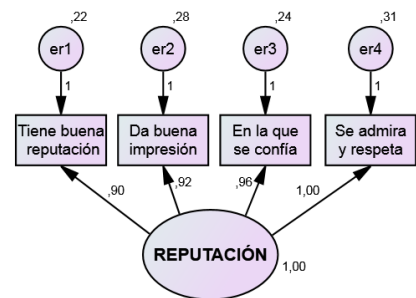
## Invarianza estricta

**Invarianza del índice PULSE  
por edad  
Jóvenes  
CHINA, 2012**



Chi-cuadrado=24,174  
g.l. = 12  
Chi square/df = 2,015  
p-valor = ,019  
CFI = ,996  
NFI = ,992  
RMSEA=,033  
Pclose = ,926

**Invarianza del índice PULSE  
por edad  
Adultos  
CHINA, 2012**



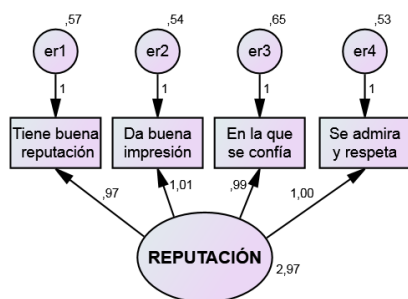
Chi-cuadrado=24,174  
g.l. = 12  
Chi square/df = 2,015  
p-valor = ,019  
CFI = ,996  
NFI = ,992  
RMSEA=,033  
Pclose = ,926

El modelo verdadero es el de invarianza estricta. No hay tabla de comparación de pérdidas de ajuste.

### 3.3. INVARIANZA CONJUNTA POR SEXO Y EDAD. CHINA.

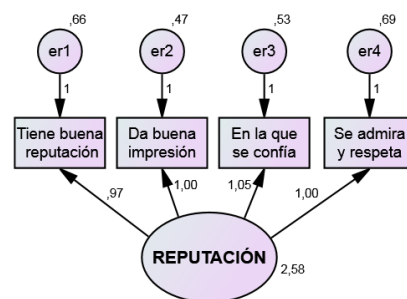
#### Modelo configural

**Invarianza del índice PULSE  
por sexo y edad  
Varón joven  
CHINA, 2012**



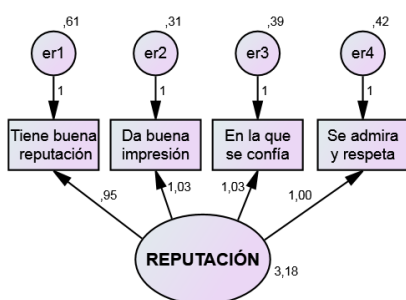
Chi-cuadrado=60,993  
g.l. = 8  
Chi square/df = 7,624  
p-valor = ,000  
CFI =,995  
NFI =,995  
RMSEA=,052  
Pclose = ,370

**Invarianza del índice PULSE  
por sexo y edad  
Varón adulto  
CHINA, 2012**



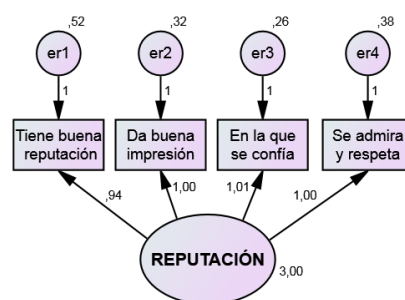
Chi-cuadrado=60,993  
g.l. = 8  
Chi square/df = 7,624  
p-valor = ,000  
CFI =,995  
NFI =,995  
RMSEA=,052  
Pclose = ,370

**Invarianza del índice PULSE  
por sexo y edad  
Mujer joven  
CHINA, 2012**



Chi-cuadrado=60,993  
g.l. = 8  
Chi square/df = 7,624  
p-valor = ,000  
CFI =,995  
NFI =,995  
RMSEA=,052

**Invarianza del índice PULSE  
por sexo y edad  
Mujer adulta  
CHINA, 2012**

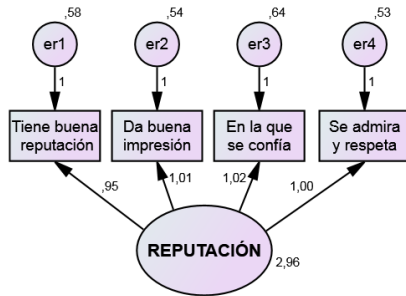


Chi-cuadrado=60,993  
g.l. = 8  
Chi square/df = 7,624  
p-valor = ,000  
CFI =,995  
NFI =,995  
RMSEA=,052  
Pclose = ,370



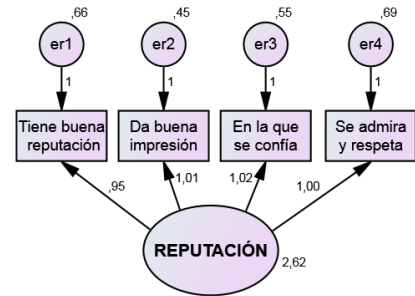
## Modelo invarianza de cargas

**Invarianza del índice PULSE  
por sexo y edad  
Varón joven  
CHINA, 2012**



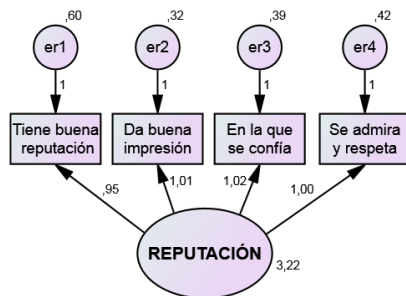
Chi-cuadrado=68,500  
g.l. = 17  
Chi square/df = 4,029  
p-valor = ,000  
CFI = ,996  
NFI = ,994  
RMSEA=,035  
Pclose = ,997

**Invarianza del índice PULSE  
por sexo y edad  
Varón adulto  
CHINA, 2012**



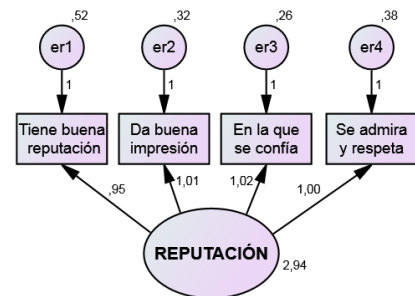
Chi-cuadrado=68,500  
g.l. = 17  
Chi square/df = 4,029  
p-valor = ,000  
CFI = ,996  
NFI = ,994  
RMSEA=,035  
Pclose = ,997

**Invarianza del índice PULSE  
por sexo y edad  
Mujer joven  
CHINA, 2012**



Chi-cuadrado=68,500  
g.l. = 17  
Chi square/df = 4,029  
p-valor = ,000  
CFI = ,996  
NFI = ,994  
RMSEA=,035

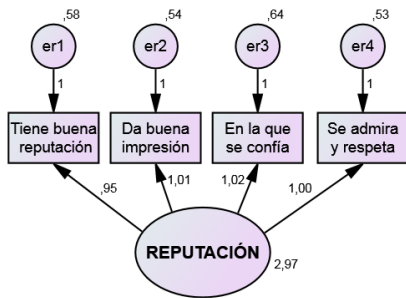
**Invarianza del índice PULSE  
por sexo y edad  
Mujer adulta  
CHINA, 2012**



Chi-cuadrado=68,500  
g.l. = 17  
Chi square/df = 4,029  
p-valor = ,000  
CFI = ,996  
NFI = ,994  
RMSEA=,035  
Pclose = ,997

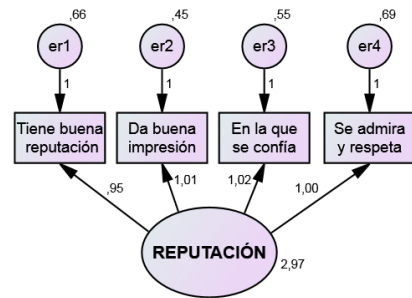
**Invarianza de varianza Pulse**

**Invarianza del índice PULSE  
por sexo y edad  
Varón joven  
CHINA, 2012**



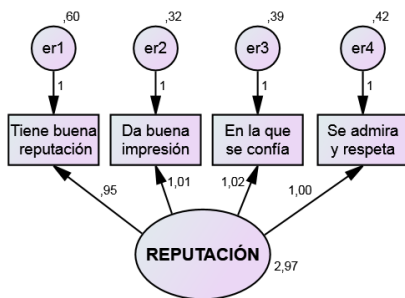
Chi-cuadrado=74,539  
g.l. = 20  
Chi square/df = 3,727  
p-valor = ,000  
CFI =,995  
NFI =,994  
RMSEA=,033  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo y edad  
Varón adulto  
CHINA, 2012**



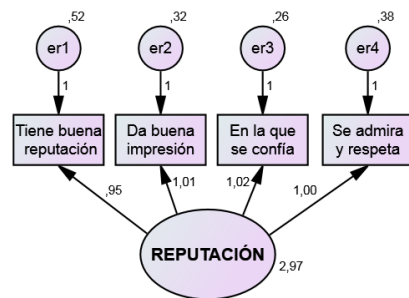
Chi-cuadrado=74,539  
g.l. = 20  
Chi square/df = 3,727  
p-valor = ,000  
CFI =,995  
NFI =,994  
RMSEA=,033  
Pclose = 1,000

**Invarianza del índice PULSE  
por sexo y edad  
Mujer joven  
CHINA, 2012**



Chi-cuadrado=74,539  
g.l. = 20  
Chi square/df = 3,727  
p-valor = ,000  
CFI =,995  
NFI =,994  
RMSEA=,033

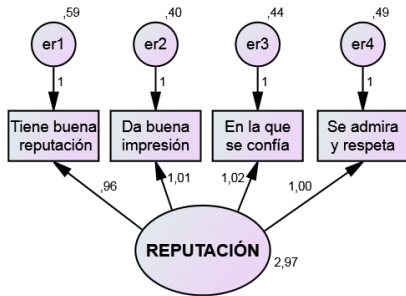
**Invarianza del índice PULSE  
por sexo y edad  
Mujer adulta  
CHINA, 2012**



Chi-cuadrado=74,539  
g.l. = 20  
Chi square/df = 3,727  
p-valor = ,000  
CFI =,995  
NFI =,994  
RMSEA=,033  
Pclose = 1,000

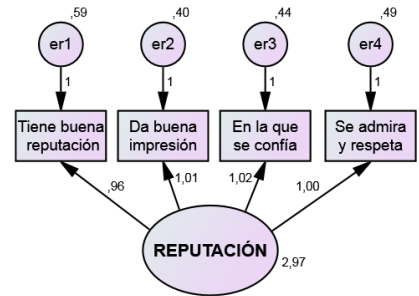
**Invarianza estricta**

**Invarianza del índice PULSE  
por sexo y edad  
Varón joven  
CHINA, 2012**



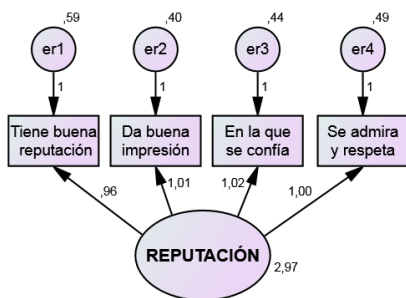
Chi-cuadrado=257,159  
g.l. = 32  
Chi square/df = 8,036  
p-valor = ,000  
CFI = ,981  
NFI = ,978  
RMSEA=,054  
Pclose = ,159

**Invarianza del índice PULSE  
por sexo y edad  
Varón adulto  
CHINA, 2012**



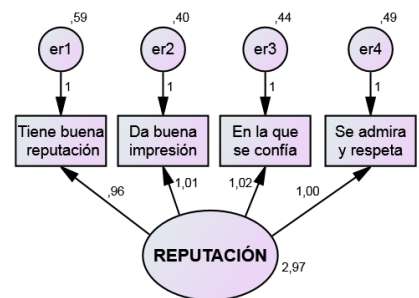
Chi-cuadrado=257,159  
g.l. = 32  
Chi square/df = 8,036  
p-valor = ,000  
CFI = ,981  
NFI = ,978  
RMSEA=,054  
Pclose = ,159

**Invarianza del índice PULSE  
por sexo y edad  
Mujer joven  
CHINA, 2012**



Chi-cuadrado=257,159  
g.l. = 32  
Chi square/df = 8,036  
p-valor = ,000  
CFI = ,981  
NFI = ,978  
RMSEA=,054

**Invarianza del índice PULSE  
por sexo y edad  
Mujer adulta  
CHINA, 2012**



Chi-cuadrado=257,159  
g.l. = 32  
Chi square/df = 8,036  
p-valor = ,000  
CFI = ,981  
NFI = ,978  
RMSEA=,054  
Pclose = ,159

El modelo verdadero es el de invarianza estricta. No hay tabla de comparación de pérdidas.

