

La percepción de la calidad en el sector de la mensajería urgente

Perceived quality in urgent transport sector

JOSE ANTONIO MARTÍNEZ GARCÍA

Laura Martínez Caro

Universidad Politécnica de Cartagena

Recibido el 17 de marzo de 2008, aceptado el 6 de febrero de 2009

Número de clasificación JEL: M31

Resumen:

Esta investigación muestra como, en el servicio de transporte urgente y mensajería, el operador público «Correos» es percibido con una calidad inferior a varios de sus competidores más cercanos: Seur, MRW y Nacex. A pesar de los esfuerzos realizados por esta entidad pública en la mejora de la calidad del servicio, éstos no se ven reflejados en las evaluaciones de los consumidores, al menos al mismo nivel que sus competidores. Estos resultados han sido obtenidos tras la utilización de diferentes metodologías, como la regresión ordinal, la comparación entre grupos o los modelos multinivel. Asimismo, diversas propuestas para futuras investigaciones son comentadas al final del estudio

Abstract:

This research has focused on the customer evaluation of perceived quality in the urgent transport service in Spain. Using different statistical procedures, such as ordinal regression, between-groups comparison or multi-level modeling, this study shows how the perceived service quality of the public company «Correos» is lower than several of the main competitors: Seur, MRW and Nacex. Despite the investment achieved by this public institution in order to improve service quality, these efforts have not been reflected in the customer evaluation, at least to the same extent as competitors. Several recommendations for further research are discussed in the final part of the study.

Palabras clave:

Calidad percibida, transporte urgente y mensajería, Correos

Key Words:

Perceived quality, urgent transport service, Correos.

La dirección de contacto es: Jose Antonio Martínez García, Universidad Politécnica de Cartagena. Facultad de Ciencias de la Empresa. Departamento de Economía de la Empresa, Paseo Alfonso XIII, 50, 30203 CARTAGENA, e-mail: josean.martinez@upct.es

1. INTRODUCCIÓN

El sector del transporte y mensajería es uno de los más dinámicos de la economía española. Su crecimiento en los últimos años ha superado el 8% y a diferencia de otros países, el mercado está dominado por empresas nacionales. El proceso de liberalización y globalización ha incrementado su competitividad, siendo la calidad un elemento esencial de diferenciación (Ali y Seshadri, 1993; Naz, 1996). Prueba de la preocupación por la calidad en este sector es el porcentaje cada vez mayor de empresas que adoptan sistemas de aseguramiento de la calidad como la certificación ISO 9000, el modelo de excelencia EFQM o la aplicación de los principios de la Calidad Total.

Asimismo, el sector público en Europa está haciendo grandes esfuerzos por mejorar su eficiencia de la administración con el fin de alcanzar una orientación al ciudadano, siendo la evaluación del desempeño del servicio una de las necesidades básicas a las que se debe enfrentar la administración pública (Scharitzer y Korunka, 2000).

En España, el sector de transporte urgente y mensajería ha estado dominado hasta su liberación por la empresa pública «Correos y Telégrafos», la cual está inmersa en un proceso estratégico de cambio para adaptar sus servicios a la nueva situación competitiva. Sin embargo, su situación anterior de monopolio y su baja reputación hacen que el esfuerzo por mejorar la percepción de la calidad de los clientes no se vea rápidamente recompensado.

El objetivo de este estudio es comparar la percepción de calidad de los usuarios de la empresa pública «Correos» con respecto a sus principales competidores en el mercado de transporte urgente y mensajería, con el fin de evaluar si los esfuerzos realizados por la compañía pública en los últimos años están viéndose reflejados en las evaluaciones de los individuos.

2. LA CALIDAD DEL SERVICIO

Desde la década de los ochenta la calidad ha sido considerada como una de las prioridades para la dirección de cara a aumentar la competitividad y el crecimiento (Camisón, 1996; Devlin et al., 1994; Sureshghandar et al., 2002). La calidad es considerada como un juicio personal y subjetivo que el consumidor emite sobre la excelencia o superioridad de un servicio o compañía (Bitner y Hubert, 1994), y supone un determinante clave del buen desempeño empresarial.

La calidad es un concepto que, debido a su compleja naturaleza, ha sido siempre difícil de definir, no existiendo consenso sobre su conceptualización y medición (Cronin y Taylor, 1992; Rust y Oliver, 1994). Aún son varios los debates abiertos referentes a determinados aspectos de la calidad del servicio. El primero de ellos hace referencia al uso de expectativas o no para medir la percepción de la calidad. A lo largo de la literatura se han propuesto diversos modelos para medir la percepción de la calidad, confluendo dos perspectivas claramente diferenciadas. La primera perspectiva sugiere que la percepción de la calidad del servicio está basada en el paradigma de la disconfirmación, en el que la calidad percibida es el resultado de la discrepancia entre las percepciones y las expectativas sobre el desempeño del servicio (Parasuraman et al. 1985, 1988; Boulding et al., 1993; Grönroos, 1994). Siguiendo esta corriente, se encuentra el «Modelo Nórdico» desarrollado por Grönroos (1984), el «Modelo

SERVQUAL» propuesto por Parasuraman et al. (1985, 1988) y posteriormente revisado (Parasuraman et al., 1994), el «Modelo de Tres Componentes de la Calidad del Servicio¹» de Rust y Oliver (1994), o el «Modelo de Desempeño Evaluado²» planteado por Teas (1993, 1994). La segunda perspectiva defiende la superioridad de los modelos basados sólo en percepciones frente a los que emplean las diferencias entre percepciones y expectativas (Carman, 1990; Cronin y Taylor, 1992; McDougall y Levesque, 1994; Brady y Cronin, 2001). Esta corriente considera que incluir las expectativas puede ser ineficiente e innecesario, debido a que los individuos tienden a indicar consistentemente altos niveles de expectativas, por lo que sus niveles de percepción raramente son superiores a éstas (Babakus y Boller, 1992). Bajo esta planteamiento se han desarrollado escalas como el «SERVPERF» (Cronin y Taylor, 1992), la «Escala de Calidad del Servicio en un Entorno Minorista³» (Dabholkar et al., 1996) o el «Modelo Jerárquico Multidimensional» (Brady y Cronin, 2001).

Otro de los debates existentes se refiere a la dimensionalidad del constructo calidad del servicio. A pesar de que existe un gran consenso en la literatura sobre su naturaleza multidimensional, algunos autores han apuntado que la calidad percibida es un constructo unidimensional que refleja únicamente las valoraciones sobre el desempeño percibido por el consumidor (Babakus y Mangold, 1992; Brown et al, 1993; Cronin y Taylor, 1992; Johns et al., 2004). Por lo tanto, si la calidad percibida fuera una variable multidimensional, esta se manifestaría a través de una serie de factores con un alto grado de correlación. Estos factores no son universales, tal y como algunos autores han sostenido (ej. Parasuraman et al., 1988) sino que son específicos del tipo de servicio evaluado (ej. Carman, 1990; Buttle, 1996). Por el contrario, si la calidad percibida es medida de forma global, sin hacer referencia a los atributos específicos del servicio, sería considerada como una evaluación de una actitud general sobre la excelencia del desempeño de la empresa. Esta opción de medición se suele realizar en investigaciones donde se analizan diferentes conceptos en redes de relaciones causa-efecto, a través de los modelos de ecuaciones estructurales (ej. Cronin et al., 2000; Brady et al., 2002). La medición global facilita la implementación de modelos causales, y permite una caracterización latente de la calidad percibida, pudiéndose estudiar su fiabilidad. Sin embargo, se pierde información acerca de los atributos específicos del servicio, por lo que desde un punto de vista gerencial puede resultar menos atractivo.

Además, varios autores han medido la calidad de forma global pero utilizando un único ítem (ej. Babakus y Boller, 1992; Bolton y Drew, 1991). Esta opción es criticada por ciertos autores, quienes indican que carece de cierta consistencia psicométrica, ya que la calidad, al igual que otras variables que miden desempeño, como la satisfacción o la imagen corporativa, es considerada como un concepto abstracto, no directamente observable, y que conviene ser aproximado a través varios indicadores. Sin embargo, existen procedimientos para realizar los análisis fijando el error de medida de la variable, como mostraremos en los resultados de esta investigación.

A pesar de las diferentes formas propuestas para medir la calidad y de los debates que continúan todavía abiertos sobre los modelos de calidad percibida, existe consenso sobre

¹ En su nombre original: Tri- Component Model of Service Quality.

² Modelo «EP» (Evaluated Performance).

³ Escala «RSQS» (Retail Service Quality Scale).

la importancia que la percepción de calidad tiene para los resultados de la empresa: cuota de mercado, beneficios, lealtad del consumidor o probabilidad de compra (Brady y Cronin, 2001)

3. ESCENARIO DE LA INVESTIGACIÓN

Desde los años setenta, se ha detectado una cierta incapacidad por parte de los monopolios postales públicos para hacer frente a las nuevas demandas de los usuarios. Estas tendencias comenzaban a apuntar hacia una necesidad de mayor valor añadido en los servicios postales tradicionales, especialmente en aspectos relativos a plazos de entrega, seguridad y seguimiento en la distribución de los envíos.

Tras su liberación, y debido a estas diferencias entre oferta y demanda postal comienza la aparición y paulatina consolidación de un conjunto de nuevos operadores que pasaron a ocupar el incipiente nicho de mercado. Estas empresas, se especializaron en estas nuevas actividades de valor añadido a unos precios muy superiores a los servicios postales básicos que, sin embargo, fueron asumidos por el mercado, dada la necesidad que de ellos se había generado⁴.

Actualmente, este sector está caracterizado por presentar una estructura empresarial muy fragmentada, con un número muy importante de operadores, algunos de ellos pertenecientes a grandes grupos multinacionales, aunque en su mayoría lo constituyen empresas de mediano y pequeño tamaño. Sin embargo, en los últimos años, se ha experimentado una tendencia creciente a la concentración de las empresas del sector, con el fin de aumentar su tamaño, penetrar en nuevas áreas geográficas, ganar grupos de clientes y complementar su cartera de servicios. En virtud de este proceso, los cinco primeros operadores del sector (Seur, MRW, Correos⁵, Guipuzcoana y Azkar) absorben conjuntamente el 48% del volumen de negocio del sector⁶.

La industria del transporte urgente es uno de los sectores de mayor expansión de Europa. Desde 1998, su crecimiento ha sido aproximadamente seis veces superior al de la economía europea en su conjunto. Su aportación directa al PIB de la Unión Europea superó los 10.500 millones de euros en 2003, la mitad de los cuales correspondió a las cuatro empresas: DHL, FedEx, TNT y UPS⁷.

En España, el mercado de transporte urgente y mensajería vio prolongada, a lo largo de 2005, la tendencia de ralentización iniciada en el año 2001. Así, el volumen total de negocio generado por las empresas del sector se situó en 2005 en 6.300 millones de euros, experimentando un crecimiento del 9,9 % respecto al año anterior. Las perspectivas de futuro son favorables. Ello es debido a la fortaleza de la demanda industrial y de servicios, así como

⁴ Extraído del Informe: El sector postal en España (2006). Subdirección General de Regulación de los Servicios Postales. Ministerio de Fomento.

⁵ Correos posee el control exclusivo de las compañías Chronoexpres y Servipack.

⁶ Datos obtenidos del Informe del Servicio de Defensa de la Competencia. N-03064. Investigación Deutsche Post-DHL-Danzas. 2003.

⁷ Información contemplada en el Informe «El impacto económico del servicio de transporte urgente en Europa». Oxford Economic Forecasting. Octubre 2004.

las nuevas oportunidades de negocio surgidas a raíz del cercano proceso de liberalización en la Unión Europea, cuya finalización está prevista, en principio para el año 2009.

Para seguir siendo un referente en este sector, las empresas están atravesando momentos de cambio motivado por una serie de factores estructurales, y por la necesidad de modernizar su gestión para mejorar la calidad y competitividad de su servicio. Su estrategia competitiva se fundamenta en ofrecer un servicio de calidad basado en el transporte garantizado, rápido, fiable, personalizado, integrado, puerta a puerta e internacional, de cuyos envíos se realiza un seguimiento y un control durante el trayecto. Muchas de estas mejoras se han producido gracias a los principios que inspiran los diferentes de sistemas de calidad (ISO 9000, EFQM, 6Sigma o Calidad Total) que han implantado las empresas de este sector. De este modo la calidad es considerada, en esta industria, como un elemento estratégico de competitividad⁸.

La empresa pública, al igual que el resto de privadas, tratan de adelantarse a las nuevas exigencias de los clientes, adaptándose constantemente para prestar un servicio por encima de sus expectativas, desarrollando proyectos a medida, y realizando todo tipo de inversiones necesarias, a nivel tecnológico, de infraestructuras o de capital humano. Sin embargo, su posición anterior de monopolio puede afectar a la percepción de la calidad de los usuarios de este tipo de servicio. Por ello, esta investigación pretende contrastar si existen diferencias en el desempeño de ambos tipos de empresas (pública vs privadas) según la percepción de la calidad de sus clientes, tomando como muestra las empresas que compiten directamente con Correos, es decir, aquellas que pueden considerarse dentro de su mismo grupo estratégico.

4. METODOLOGÍA

Este estudio se realizó durante el año 2005 a una muestra de usuarios de servicios de transporte y mensajería de la Región de Murcia que habían utilizado el servicio en los doce meses previos⁹. Las limitaciones presupuestarias de la investigación obligaron a establecer el límite en 375 cuestionarios. De ellos, un total de 71 no fueron considerados siguiendo el siguiente criterio: (1) cuestionarios que evaluaban empresas con baja cuota de mercado, o (2) cuestionarios referentes a empresas importantes (ej. UPS, Halcourier o DHL) que estaban marginalmente representadas, es decir, con un nivel muy bajo de respuestas.

Finalmente, además de Correos con 142 cuestionarios, tres empresas quedaron representadas: Seur, MRW y Nacex, con 116, 31 y 15 observaciones respectivamente. Estas empresas, junto con las descartadas Halcourier, DHL y UPS poseen la mayor cuota de mercado de servicios de transporte urgente y mensajería¹⁰.

⁸ Como muestra de ello se recogen las declaraciones realizadas por D. Víctor Calvo Sotelo (Presidente de la Sociedad Anonima Estatal de Correos y Telégrafos, 2003). «La calidad no solo es una obligación legal o un compromiso con los ciudadanos, sino que también es una clave de competitividad presente y futura en un mercado en progresiva liberalización».

⁹ Se tomó como referencia este periodo de tiempo, teniendo en cuenta el frecuente uso de los servicios seleccionados (Ryan y Cliff, 1997); la frecuencia media de uso de este servicio es de tres veces al año.

¹⁰ Estudio de Sectores de DBK. Mensajería y Paquetería, 2003. Madrid: DBK-Sectores. Disponible en: <http://www.dbk.es>.

4.1. Medición de la calidad percibida

La medición de la percepción de la calidad fue realizada por un sólo ítem (Babakus y Boller, 1992; Bolton y Drew, 1991: «*La compañía XXX ofrece un excelente servicio*»). Para ello se utilizó una escala ordinal con cinco opciones de respuesta (1 a 5; muy en desacuerdo – muy de acuerdo).

Aunque las mediciones de un constructo latente son tradicionalmente implementadas con escalas multi-ítem (ej. Mulaik y Millsap, 2000), si existe un profundo compromiso entre la definición del concepto abstracto y la manifestación observable de ese concepto, es posible proponer medidas de constructos latentes con dos e incluso un ítem (Hayduk, 1996; Hayduk y Glaser, 2000), asumiendo el posible error cometido en la medición. Dadas las ventajas a nivel de coste que supone diseñar cuestionarios cortos, teniendo en cuenta otros estudios que abogan por la simplicidad en las mediciones de calidad (Martínez y Martínez, 2008), y considerando que se pretende obtener una evaluación global del consumidor acerca de la excelencia del servicio (y no juicios sobre atributos particulares), la medición con un solo ítem se ha considerado como adecuada. La reciente investigación de Bergkvist y Rossiter (2007) muestra resultados muy interesantes en favor de las mediciones con un solo ítem frente a las realizadas con múltiples indicadores.

5. RESULTADOS

La simplicidad de los objetivos planteados permite una más cuidada utilización de los análisis estadísticos para mostrar la evidencia empírica. Esa defensa de la simplicidad es una práctica altamente recomendable cuando los objetivos de la investigación lo permiten (Cohen, 1990), y lo es aún más teniendo en cuenta las limitaciones y peligros de la estadística aplicada comúnmente a la investigación social (ej. Harlow et al., 1997; Barrett, 2005). Es por ello que han utilizado diferentes opciones metodológicas para analizar el objetivo de esta investigación.

5.1. Consideración de la escala ordinal como categorías discretas

En la literatura de marketing, y en muchas áreas de las ciencias sociales, las escalas ordinales (ej. Likert o diferencial semántico) son tratadas como de naturaleza continua y no discreta. Sin embargo, los problemas asociados a su distinción se siguen discutiendo (Ato y López, 1996) y existen diferencias importantes en cuanto a su aplicación en análisis estadísticos (ej., Coenders y Saris, 1995; Jöreskog y Sörbom, 2001; Vermunt y Magidson, 2005).

El tratamiento de la medida de calidad percibida como variable ordinal discreta permite la utilización de modelos loglineales para analizar las diferencias en las evaluaciones de los consumidores de diferentes empresas de transporte urgente y mensajería. La distribución de frecuencias de respuesta se muestra en la siguiente tabla de contingencia (Tabla 1).

Tabla 1
Distribución de frecuencias de respuesta

Percepción de calidad	Empresa				Total
	MRW	Seur	Correos	Nacex	
Muy en desacuerdo	0	0	10	0	10
En desacuerdo	1	14	22	3	40
Indiferente	10	36	72	1	119
De acuerdo	16	53	33	9	111
Muy de acuerdo	4	13	5	2	24
Total	31	116	142	15	304

Para analizar si el servicio proporcionado por cada empresa afectaba a la percepción del consumidor se realizó una regresión ordinal de clases latentes con el programa Latent Gold 4.0 (Vermunt y Magidson, 2005). Esta herramienta informática permite, además, la identificación de heterogeneidad no observable en los datos, es decir, la posible detección de diferentes segmentos de consumidores en función de su patrón de respuestas.

La Tabla 2 muestra los estadísticos de los modelos estimados. El modelo sin clases latentes (M_1) presenta un mejor ajuste, ya que los índices BIC(LL) y AWE tienen un valor más bajo (Vermunt y Magidson, 2005). No obstante, la comparación entre modelos puede testarse también a través de una diferencia entre los valores del logaritmo de la verosimilitud (-2LL). En este caso, el modelo de dos clases latentes (M_2) tiene una ligera ganancia significativa de ajuste (Bootstrap p -valor para la diferencia menor que 0,05), pero, dado que está prácticamente en el límite (0,038) y teniendo en cuenta que la interpretación posterior carece de capacidad discriminante, el modelo óptimo es el que no presenta heterogeneidad, es decir, M_1 .

Tabla 2
Estadísticos de los modelos estimados

	LL	BIC(LL)	Npar	L ²	Bootstrap p -valor (L ²)	gl	AWE	-2LL diferencia	Bootstrap p -valor (-2LL)
M_1	-382,01	804,05	7	18,58	0,022	9	865,07		
M_2	-373,89	833,54	15	2,32	0,148	1	1217,40	16,25	0,038

La Tabla 3 muestra los parámetros estimados del modelo óptimo. La significación del estadístico de Wald indica que el patrón de respuestas entre los diferentes niveles de la variable dependiente es significativamente diferente, así como que existe al menos una categoría de la variable dependiente que afecta a la calidad percibida. El estadístico Z proporciona resultados más concretos sobre la significación individual; como se puede apreciar los niveles de percepción de calidad se ven afectados negativamente por los consumidores de la empresa Correos. Es decir, Correos discrimina negativamente entre los clientes del

servicio de transporte urgente y mensajería para los niveles «Muy en desacuerdo», «Indiferente» y «De acuerdo». Por tanto, se puede afirmar que el patrón de percepciones de los clientes de Correos es diferente al del resto de empresas, obteniendo valoraciones con un sentido más negativo.

Tabla 3

Parámetros estimados

Variable dependiente	Parámetro	Z	p-valor	Wald	p-valor
Muy en desacuerdo	-2,12	-6,387	< 0,001	127,93	< 0,001
En desacuerdo	-0,251	-1,482	n.s		
Indiferente	1,195	9,814	< 0,001		
De acuerdo	1,327	8,493	< 0,001		
Muy de acuerdo	-0,151	-0,648	n.s		
<hr/>					
Variable independiente					
MRW	0,346	1,816	n.s	29,76	< 0,001
Seur	0,077	0,593	n.s		
Correos	-0,651	-4,907	< 0,001		
Nacex	0,227	0,924	n.s		

5.2. Consideración de la escala ordinal como variable continua

Bajo este supuesto, el análisis más básico de comparación es el de las estimaciones medias entre empresas. Es decir, comparar los intervalos de confianza de los valores medios de respuesta entre consumidores de las cuatro compañías. Para ello, en primer lugar se ha realizado un test de normalidad utilizando el análisis conjunto de asimetría y curtosis proporcionado por el programa STATA 8. Los resultados indican que la distribución de respuestas puede considerarse como normal para los cuatro grupos (χ^2 no significativa al 95%) (Tabla 4).

Tabla 4

Análisis de normalidad

Variable dependiente	Prob. Asimetría	Prob. Curtosis	χ^2	p-valor
Nacex	0,156	0,949	2,33	0,321
MRW	0,793	0,927	0,08	0,962
Seur	0,285	0,166	3,13	0,208
Correos	0,028	0,825	4,89	0,087

Una vez aceptado el supuesto de normalidad, es sencilla la comparación de los intervalos de confianza para las medias entre los grupos. En la Figura 1, se muestra esa comparativa transformando la métrica de la variable dependiente al intervalo (0,1) (Cohen et al., 1999). Como puede observarse, Correos muestra valores diferentes con respecto a MRW y Seur, solapándose ligeramente con Nacex.

Los análisis realizados hasta ahora asumen que la percepción de calidad es medida sin error. Una simple simulación podría dar una idea de cómo variarían los resultados considerando ciertos niveles de error en la medida de la variable dependiente. A medida que aumenta el error de medida (o lo que es lo mismo, disminuye la fiabilidad del ítem), se incrementan las varianzas de las estimaciones entre grupos, por lo que es necesario corregir el tamaño de efecto (o diferencia sustantiva) entre grupos por un índice de fiabilidad (Ree y Carreta, 2006). Hay que recordar que el error de medida no afecta a los valores medios bajo la hipótesis de aleatoriedad.

Figura 1

Estimación de las medias para cada empresa (Intervalos de confianza al 95%)

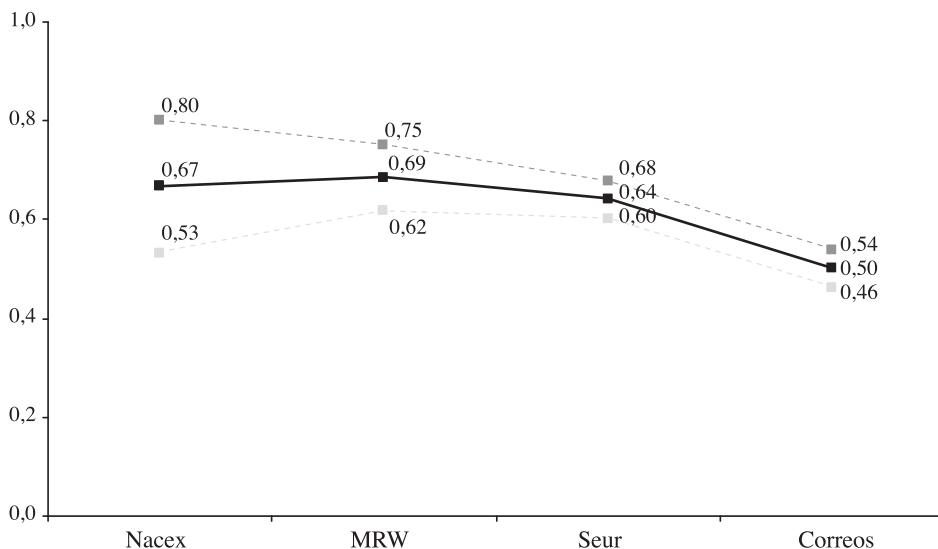


Tabla 5

Tamaños de efecto *d* de Cohen^a y error de medida

	Sin error			90% fiabilidad			80% fiabilidad			70% fiabilidad		
	MRW	Seur	Correos	MRW	Seur	Correos	MRW	Seur	Correos	MRW	Seur	Correos
Nacex	0,09	0,13	0,73	0,08	0,12	0,69	0,08	0,11	0,65	0,07	0,11	0,61
MRW		0,22	0,84		0,20	0,79		0,19	0,75		0,18	0,70
Seur			0,63			0,59			0,56			0,52

^aConvenciones de Cohen (1988) sobre el tamaño de efecto *d*: pequeño (0.20), mediano (0.50), grande (0.80).

Como puede contemplarse en la Tabla 5, una disminución en la fiabilidad de la medida de calidad disminuye el tamaño del efecto entre grupos. Aún así, en el caso más desfavorable de fiabilidad, los tamaños de efecto entre Correos y el resto de empresas no son marginales¹¹.

Dos aspectos adicionales podrían ser considerados para completar los análisis estadísticos. En primer lugar, la muestra es de conveniencia y se ha circunscrito a una zona geográfica concreta, y en segundo lugar, las observaciones dentro de cada empresa podrían no considerarse independientes, y estar agrupadas en niveles de información más amplios: tipo de empresa y tipo de propiedad (pública o privada). Los modelos multinivel ofrecen una interesante opción metodológica para manejar estas situaciones, ya que permiten considerar efectos aleatorios para los grupos y tener en cuenta las covarianzas entre los diferentes niveles. Siguiendo la propuesta de Pardo et al. (2007), se ha desarrollado la siguiente secuencia de análisis.

1. Estimación de un modelo con efectos aleatorios, o modelo incondicional nulo (M_1). En este nivel, los consumidores están agrupados en función de la empresa evaluada, considerando a las cuatro organizaciones como una muestra aleatoria de empresas¹². Se utilizó el procedimiento «Mixed» de SPSS 13.0 con estimación por máxima verosimilitud restringida. La Tabla 6 muestra los estadísticos de ajuste (mejores cuanto más pequeños). Analizando los parámetros de covarianza (Tabla 8) se obtiene que la varianza poblacional del factor «empresa» no puede considerarse distinta de cero. Este resultado supondría afirmar que la percepción de la calidad no se ve afectada por el tipo de empresa, de la que las cuatro analizadas son una muestra aleatoria de toda la población. Sin embargo, el estadístico de Wald es menos fiable que la comparación entre modelos utilizando la diferencia en el logaritmo de la verosimilitud (-2LL) que se distribuye según una chi-cuadrado con grados de libertad igual a la diferencia entre parámetros (Pardo et al., 2007). Dado que el coeficiente de correlación intra-clase (CCI), o tamaño de efecto del factor es distinto de cero (0,12), se analizó el cambio observado en -2LL al estimar un modelo en el que no se incluye el factor «empresa» (M_2). Este modelo presenta peores índices de ajuste (Tabla 6) y una diferencia significativa en el estadístico -2LL (52,38; $p < 0,001$). Por tanto, y con ciertas reservas, se puede afirmar que el factor «empresa» explica mejor la variabilidad de los datos, por lo que existe un efecto no nulo de este factor sobre la percepción de calidad.
2. Estimación de un modelo multinivel donde el tipo de propiedad (pública o privada) actúe como factor fijo para testar la diferencia entre empresas (M_3). En este caso, los sujetos están agrupados en empresas (nivel 1), y el factor de interés es el tipo de propiedad. Como se muestra en la Tabla 7, la propiedad afecta significativamente a la percepción de la calidad de los sujetos agrupados por empresas, es decir, existen diferencias entre las empresas privadas y Correos.

¹¹ Este tipo de comparaciones dos a dos entre grupos presenta ciertas carencias estadísticas puesto que no se ha hecho ninguna corrección para ganar potencia ni ningún contraste enfocado (Steiger, 2004). Sin embargo, sirve de base para, a nivel exploratorio, ayudar a detectar el posible efecto del error de medida sobre los resultados.

¹² Al recogerse los datos en una zona geográfica determinada, los resultados son dependientes de ese diseño de investigación. El considerar a esas empresas como una muestra aleatoria en función de la localización puede ayudar a modelar esa limitación. Es decir, por ejemplo, Correos de la Región de Murcia es una muestra de todos los servicios de Correos a nivel nacional.

Tabla 6
Estadísticos de los modelos estimados

	-2LL	BIC(LL)	Npar
M ₁	786,86	798,297	3
M ₂	813,05	818,764	2
M ₃	781,81	787,52	3

Tabla 7
Estadísticos de los parámetros de efectos fijos

	Parámetro	Estimación	Error típico	t	p-valor
M ₁	Intersección	3,47	0,17	19,58	< 0.001
M ₂	Intersección	3,32	0,05	63,23	< 0.001
M ₃	Intersección	2,40	0,16	14,97	< 0.001
	Propiedad	0,59	0,09	5,99	< 0.001

Tabla 8
Estadísticos de los parámetros de covarianza

	Parámetro	Estimación	Error típico	Wald	p-valor
M ₁	Residuos	0,75	0,06	12,25	< 0.001
	Empresa ^a	0,10	0,09	1,09	0,272
M ₂	Residuos	0,84	0,07	12,30	< 0.001
M ₃	Residuos	0,75	0,06	12,28	< 0.001

^aCCI: 0,12

6. CONCLUSIONES

Esta investigación ha mostrado cómo la empresa pública Correos es percibida como de menor calidad con respecto a tres de sus principales competidores: Seur, Nacex y MRW. A pesar de la simplicidad del estudio en cuanto a su planteamiento, se ha tratado de ser muy cauto en la aplicación de los análisis estadísticos, con el fin de intentar minimizar los posibles sesgos que éstos pudieran producir debido a su mal uso; es decir, lo que algunos autores llaman ya Error Tipo IV, o deficiente aplicación de los métodos para analizar los datos empíricos. Los resultados de esta investigación muestran diferencias en la calidad percibida tras la utilización de diferentes métodos y asunciones sobre los datos. Es por ello que se puede afirmar que existe evidencia empírica de diferencias entre Correos y el resto de empresas.

Ciertamente, Correos está cambiando su estrategia de gestión. Ha adoptado sistemas de gestión de la calidad, como la certificación ISO 9000, y ha desarrollado nuevos servicios en una evidente orientación al marketing (atención al cliente, quejas y sugerencias). Además ha mejorado sus equipamientos y sistemas de información (sitio web muy completo, servicio de localizador de envíos) y demuestra su compromiso con la sociedad a través de su programa de Responsabilidad Social Corporativa. Sin embargo estos esfuerzos parecen no verse reflejados de momento en la percepción de calidad de los consumidores. La baja reputación de este servicio público podría ser una explicación de estos resultados constituyendo un atractivo contexto de análisis para futuras investigaciones. En una fase previa de esta investigación, tras unas pocas entrevistas a consumidores, éstos mostraban cómo tenían una pobre percepción sobre el historial de esta compañía pública, sobre todo en relación a los retrasos y pérdidas de envíos. Como bien indican Markwick y Fill (1997), esta reputación es más difícil de cambiar que la imagen corporativa actual, y crea un efecto halo sobre los juicios de los consumidores sobre el desempeño del servicio.

Sin embargo, conviene recordar que este estudio es de naturaleza transversal, es decir, se ha realizado en un momento puntual del tiempo sobre una muestra de consumidores. Gollob y Reichardt (1991) advierten del peligro de sacar conclusiones precipitadas acerca de ciertos efectos «causales» sobre las variables dependientes, ya que esos efectos pueden tardar diferentes periodos de tiempo en materializarse. Esta explicación puede ser alentadora para la empresa pública, ya que indicaría que los esfuerzos por mejorar aún no han tenido tiempo de hacerse efectivos. Indudablemente habría que realizar un seguimiento, en la línea que ha realizado este estudio, sobre la evolución de la calidad percibida en el futuro más cercano.

La evidencia empírica muestra como Correos es la única empresa que obtiene datos extremos negativos. Esto puede ser debido a que ciertos consumidores podrían haber sufrido fallos importantes en el servicio recibido. Como indican Mittal et al. (1998), estos errores pueden llevar a una evaluación global muy negativa, considerando la asimetría y no linealidad entre la polaridad de sentimientos extremos (negativos y positivos) y las evaluaciones del servicio. Los sentimientos negativos cuentan más para el consumidor que los positivos en sus actitudes y comportamientos. Futuras investigaciones deberían ahondar más en este aspecto.

El bajo número de cuestionarios analizados para Nacex y MRW es otra limitación de esta investigación. Aunque este aspecto se ha tenido en cuenta para las comparaciones entre empresas (al contar con los intervalos de confianza), y se ha comprobado la normalidad de sus distribuciones, un tamaño tan pequeño de cada grupo es peligroso porque cuestiona la asunción de compensación de los términos de error, es decir, que los errores negativos y positivos se equilibran en un valor medio de cero. Además, la representatividad de los datos podría también ser discutida.

La replicación de esta investigación se considera, por tanto, fundamental. Podrían analizarse más empresas y más localizaciones de aquellas susceptibles de análisis. La metodología de modelos multinivel y el análisis de la varianza de efectos aleatorios ayudarían a describir las posibles diferencias entre distintas empresas o localizaciones.

Por otro lado, futuros estudios podrían considerar dos aspectos importantes en el análisis de la calidad percibida, el papel de las expectativas del consumidor y el efecto placebo del precio sobre la evaluación de la calidad. Rust et al. (1999) muestra como la experiencia

es un factor crucial para los juicios sobre calidad; los consumidores con menor nivel de experiencia son más sensibles a la disconfirmación de expectativas que los de mayor grado de experiencia. Por otro lado, estudios como el de Shiv et al. (2005) indican cómo el precio de los productos ejerce una influencia no consciente sobre las evaluaciones de la calidad, de forma similar a un efecto placebo. La banda de precios de Correos está un poco por debajo de sus principales competidores en algunos servicios, por tanto, queda margen para profundizar en este aspecto.

Debido a la solidez de la posición adquirida por su condición inicial de operador monopolístico, Correos seguirá siendo en el futuro el gran referente de los servicios postales en nuestro país, sobre todo en el negocio de la correspondencia tradicional. Motivos para ello pueden ser su elevado grado de implantación y volumen de envíos, el alto nivel de inversiones o la densidad de su red de centros operativos, transporte y distribución, en comparación con el resto de operadores nacionales, sobre todo, dado el carácter más especializado que han adquirido sus competidores hacia determinados servicios o clientes. Sin embargo, la preocupación por la eficiencia y la excelencia en el servicio seguirán siendo un elemento importante en este mercado, y una necesidad para la modernización de los servicios públicos. Por tanto, aunque Correos disponga de ventajas competitivas heredadas de su situación de monopolio, deberá continuar esforzándose en mejorar la calidad de su servicio de transporte urgente y mensajería.

7. IMPLICACIONES PARA LA GESTIÓN

¿Por qué Correos tiene una calidad percibida peor que el resto de sus principales competidores? Hay que considerar que esta empresa pública está haciendo un gran esfuerzo en los últimos años por realizar prácticas de gestión encaminadas a la mejora continua y a la optimización de los procesos, así como por diseñar un programa de responsabilidad social corporativa y patrocinios, en aras de mejorar su imagen. Sin embargo, estas prácticas son también realizadas por las empresas competidores analizadas.

Así, prácticamente la totalidad de oficinas de Correos cuenta con la certificación ISO 9001 desde el año 2005. Unos años antes, en 1999, inició su programa de Responsabilidad Social Corporativa, donde actualmente acomete diversas acciones relacionadas con el medio ambiente, la infancia, los inmigrantes o las personas con discapacidad. Además, dentro de sus acciones de relaciones públicas destacan patrocinios importantes como los de la Vuelta Ciclista a España y la Expo de Zaragoza 2008.

Pero tanto Seur, como MRW y Nacex realizan unas acciones similares. Estas tres empresas están certificadas con la Norma ISO 9000 desde 1995, 1995 y 1998 respectivamente. Además, Seur y MRW poseen la norma ISO 140001 de medio ambiente, teniendo MRW la primera certificación EFR en España como empresa familiarmente responsable. En cuanto a las acciones de responsabilidad social. Seur creó la Fundación Seur en 2004 para canalizar las actuaciones de la compañía en el ámbito de la acción social, MRW tiene un ambicioso plan de acción social, y Nacex colabora con varias organizaciones no gubernamentales en diversas labores sociales y medioambientales. Por lo que respecta a los patrocinios, también existen estrategias similares. Por ejemplo, Seur patrocina a la Federación Española de Baloncesto, Nacex ha patrocinado la Liga ASOBAL de balonmano y la

Liga Nacional de Fútbol Sala. MRW, sin embargo, parece que enfoca más su línea de patrocinio sobre acciones culturales. Sin embargo, la estrategia global de relaciones públicas (acción social y patrocinios) de MRW tiene un gran reconocimiento. Así, en 2001 ha sido la primera empresa europea del sector en obtener la Norma SA 8000 de responsabilidad social, y en 2006 ocupaba la posición 25 del Ranking del Monitor Español de Reputación Corporativa (muy por encima de Seur, que es la única empresa competidora que aparece en ese ranking).

Dado que las cuatro empresas sobre las que hemos enfocado este estudio tienen estrategias muy similares en cuanto a gestión de procesos y relaciones públicas, la divergencia que existe entre la percepción de calidad de Correos y el resto puede ser debida a la diferencia en el tiempo de materialización de los efectos de esas acciones. Como indican Mitra y Golder (2006), los efectos de los cambios sobre la calidad objetiva tardan un tiempo en ser percibidos por los consumidores. Y la amplitud de ese tiempo es también dependiente de la reputación de las empresas, es decir, si una empresa tiene mejor reputación que otra, los cambios en la calidad objetiva del producto o servicio que ofrece serán más rápidamente percibidos. Como nuestro estudio se ha desarrollado en 2005, no existe ese periodo necesario para que esos efectos sean perceptibles en Correos. Recordemos que las otras empresas han acometido acciones de mejora de los procesos, como las certificaciones de calidad, mucho antes que la entidad pública. Es más, tal y como Martínez y Martínez (2009) recogen en su estudio sobre entidades aseguradoras, las empresas que poseen la certificación ISO 9000 son mejor percibidas que las que no la poseen, en cuanto a calidad, satisfacción e imagen.

En un sector donde los principales competidores ofrecen un abanico de servicios muy similares, es difícil la diferenciación. De hecho sólo unos pocos servicios pueden considerarse como diferentes en alguna empresa, como el de entrega de envíos los sábados que ofrece Nacex o Seur, o la especialidad en transporte de mascotas de MRW. Es por ello que creemos que Correos debería acometer acciones de comunicación específicas para intentar mejorar su imagen corporativa, yendo más allá de las acciones de patrocinio o responsabilidad social. Se trataría de comunicar expresamente cómo ha conseguido en los últimos años modernizar su estructura tecnológica, y que es una empresa en la que se puede confiar, con el fin de ir eliminando esa mala reputación histórica que es un hándicap para los servicios públicos en España. No obstante, Correos está consiguiendo parcialmente ya ese objetivo, ya que es el servicio público mejor valorado, según el Centro de Investigaciones Sociológicas – CIS - (octubre, 2008). En esa encuesta, son precisamente los más jóvenes (18 a 24 años) los que dan una mejor valoración, lo que refuerza la idea de que las personas con más edad son las que aún tienen en mente esas reminiscencias históricas sobre el funcionamiento de los servicios públicos.

En definitiva, aunque utilizan estrategias muy similares, nuestra investigación ha mostrado que Correos es percibida como de peor calidad que sus principales competidores. La causa más probable de este resultado es la diferencia que existe en cuanto a la adopción de procesos de mejora continua y comunicación corporativa, donde Correos ha empezado su camino más tarde que el resto. Dada la reputación histórica de los servicios públicos en España, los esfuerzos de Correos deben de continuar en la misma línea de actuación, incidiendo además en realizar acciones específicas de comunicación para potenciar su imagen. Es de esperar, que en los próximos años la percepción de calidad de Correos se equipare

con la de sus competidores. Futuros estudios deberán comprobar este hecho, que de no ser corroborado, podría ser indicativo de problemas estructurales importantes en el funcionamiento del servicio.

8. REFERENCIAS BIBLIOGRÁFICAS

- ALI, A., y SESHADRI, S. (1993). «Customer perception and competitive quality strategy». *Managerial and Decision Economics*, 12, pp. 235-246.
- ATO, M., y LÓPEZ, J. J. (1996). *Análisis estadístico para datos categóricos*. Madrid, Síntesis.
- BABAKUS, E., y BOLLER, G. W. (1992). «An empirical assessment of the SERVQUAL scale». *Journal of Business Research*, 24, pp. 253-268.
- BABAKUS, E., y MANGOLD, D. W. G. (1992). «Adapting the SERVQUAL Scale to Hospital Services: and Empirical Investigating». *Health Services Research*, 26 (6), pp. 767-786.
- BARRETT, P. T. (2005). «What if there were no psychometrics? Commentary on Robert McGrath's «Constructs, complexity, and measurement» target article». *Journal of Personality Assessment*, 85 (2), pp. 134-140.
- BERGKVIST, L., y ROSSITER, J. R. (2007). «The predictive validity of multiple-item vs single item measures of the same construct». *Journal of Marketing Research*, 44 (2), Mayo, pp. 175-184.
- BITNER, M. J., y HUBBERT, A. R. (1994). «Encounter Satisfaction versus Overall Satisfaction versus Quality». En R. T. Rust, & R. L. Oliver, (Eds.), *Service Quality: New Directions in Theory and Practice* (pp. 72-94). Thousand Oaks, CA: Sage Publications.
- BOULDING, W., KALRA, A., STAELIN, A., y ZEITHAML, V. A. (1993). «A Dynamic Process Model of Service Quality: From Expectations to Behavioural Intentions». *Journal of Marketing Research*, 30 (February), pp. 7-27.
- BOLTON, R., y DREW, J. (1991). «A Longitudinal Analysis of the Impact of Service Changes on Customer Attitudes». *Journal of Marketing*, 55 (January), pp. 1-9.
- BRADY, M. K., y CRONIN, J. J. (2001). «Some new thoughts on conceptualizing perceived service quality: a hierarchical approach». *Journal of Marketing*, 5, pp. 34-49.
- BRADY, M. K., CRONIN, J. J., y Brand, R. R. (2002). «Performance-Only Measurement of Service Quality: A Replication and Extension». *Journal of Business Research*, 55, pp. 17-31.
- BROWN, S. W., CHURCHILL, G., y PETER, J. P. (1993). «Improving the Measurement of Service Quality». *Journal of Retailing*, 69 (1), pp. 127-139.
- BUTTE, F. (1996). «SERVQUAL: review, critique, research agenda». *European Journal of Marketing*, 30 (1), pp. 8-25.
- CAMISÓN, C. (1996). «La Calidad como Factor de Competitividad en Turismo: Análisis del Caso Español». *Información Comercial Española*, 749 (Enero), pp. 99-122.
- CARMAN J. (1990). «Consumer perceptions of service quality: an assessment of the SERVQUAL dimensions». *Journal of Retailing*, 66, pp. 33-35.
- COENDERS, G., y SARIS, W. E. (1995). «Categorization and Measurement Quality. The Choice Between Pearson and Polychoric Correlations». En Saris, W. E. y Münnich,

- Á. (Eds.) *The Multitrait-Multimethod Approach to Evaluate Measurement Instruments* (pp. 125-144). Budapest, Eötvös University Press.
- COHEN, J. (1990). «Things I Have Learned (so far) ». *American Psychologist*, 49 (12), pp. 997-1003.
- COHEN, J. (1988). *Statistical power analysis for the behavioural sciences* (2ª ed.). Hillsdale, N. J.: Erlbaum.
- COHEN, P., COHEN, J., AIKEN, L., y WEST, S. (1999). «The problem of units and the circumstance for POMP». *Multivariate Behavioral Research*, 34 (3), pp. 315-346.
- CRONIN, J. J., BRADY, M. K., y HULT, G. T. M. (2000). «Assessing the Effects of Quality, Value, and Customer Satisfaction on Consumer Behavioural Intentions in Service environments». *Journal of Retailing*, 76 (2), pp. 193-218.
- CRONIN, J., y TAYLOR, S. (1992). «Measuring Service Quality: a Reexamination and Extension». *Journal of Marketing*, July, pp. 55-68.
- DABHOLKAR, P. A., THORPE, D. I., y RENTZ, J. O. (1996). «A measure of service quality for retail stores: Scale development and validation». *Journal of the Academy of Marketing Science*, 24(1), pp. 3-16.
- DEVLIN, S. J., y DONG, H. K. (1994). «Service Quality from the Customers' Perspective». *Marketing Research*, 6 (1), pp. 4-10.
- GOLLOB, H. F. y REICHARDT, C. S. (1991). «Interpreting and estimating indirect effects assuming time lags really matter». En L. M. Collins y Horn, J. L. (Eds.), *Best methods for the analysis of change: Recent advances, unanswered questions, future directions* (pp. 243-259). Washington, DC: American Psychological Association.
- GRÖNROOS C. (1994). *Marketing y Gestión de Servicios: la Gestión de los momentos de la verdad y la competencia en los servicios*. Madrid: Díaz de Santos.
- GRÖNROOS, C. (1984). «A service quality model and its marketing implications». *European Journal of Marketing*, 18 (4), pp. 36-44.
- HARLOW, L. L., MULAİK, S. A., y STEIGER, J. H. (1997). *What if there were no significance tests?*. Hillsdale: Lawrence Erlbaum.
- HAYDUK, L. A. (1996). *LISREL Issues, Debates and Strategies*. Baltimore, MD.: Johns Hopkins University Press.
- HAYDUK, L. A., y GLASER, D. N. (2000). «Jiving the four-step, waltzing around factor analysis, and other serious fun». *Structural Equation Modeling*, 7, pp. 1-35.
- JÖRESKOG, K. G. y SÖRBOM, D. (2001). *LISREL 8.50*. Chicago: Scientific Software International.
- JOHNS, N, AVCI, T., y KARATEPE, O. (2004). «Measuring Service Quality of Travel Agents: Evidence from Northern Cyprus». *The Service Industries Journal*, 24 (3), pp. 82-100.
- MARKWICK, N., y FILL, C. (1997). «Towards a framework for managing corporate identity». *European Journal of Marketing*, 31 (5/6), pp. 396-409.
- MARTÍNEZ, J. A. y MARTÍNEZ, L. (2008). «Conceptualizing and measuring service quality; Some insights into an ongoing debate». *Journal of Retailing and Consumer Services*. Bajo revisión.
- MARTÍNEZ, L. y MARTÍNEZ, J. A. (2009). «Does ISO 9000 certification affect consumer perceptions of the service provider? ». *Managing Service Quality*, 19 (2), pp. 140-161.

- MITRA, D., y GOLDBERGER, P. N. (2006). «How does objective quality affect perceived quality? Short-term effects, long-term effects, and asymmetries». *Marketing Science*, 25, pp 230-247.
- MCDUGALL, G. H., y LEVESQUE, T. J. (1994). «A revised view of service quality dimensions: an empirical investigation». *Journal of Professional Service Marketing*, 11 (1), pp. 189-209.
- MITTAL, V., ROSS JR., W. T., y BALDASER, P. M. (1998). «The asymmetric impact of negative and positive attribute-level performance on overall satisfaction and repurchase intentions». *Journal of Marketing*, 62 (January), pp. 33-47.
- MULAIK, S. A., y MILLSAP, R. E. (2000). «Doing the four-step right». *Structural Equation Modeling*, 7, pp. 36-73.
- NAZ, R. (1996). «Experiencia Española en la Aplicación de las Normas ISO de la Serie 9000». *Información Comercial Española (ICE)*, 724 (Diciembre).
- PARASURAMAN, A., ZEITHAML, V., y BERRY, L. (1988). «Servqual: A Multiple-Item Scale for measuring Consumer Perceptions of Service Quality». *Journal of Retailing*, 64 (1), pp. 12-40.
- PARASURAMAN, A., ZEITHAML, V., y BERRY, L. (1985). «A Conceptual Model of Service Quality and its Implications for Future Research». *Journal of Marketing*, 49 (Fall), pp. 41-50.
- PARASURAMAN, A., ZEITHAML, V., y BERRY, L. (1994). «Alternative Scales for Measuring Service Quality: A Comparative Assessment Based on Psychometric and Diagnostic Criteria». *Journal of Retailing*, 70 (3), pp. 193-194.
- PARDO, A., RUIZ, M. A., y SAN MARTÍN, R. (2007). «Cómo ajustar e interpretar modelos multinivel con SPSS». *Psicothema*, 18 (2), pp. 308-321.
- REE, M. C. y CARRETA, T. R. (2006). «The role of measurement error in familiar statistics». *Organizational Research Methods*, 9 (1), pp. 99-112.
- RUST, R. T., y OLIVER, R. L. (1994). «Service quality: insights and managerial implications from the frontier». En Rust, R.T. y Oliver, R.L. (Eds.), *Service Quality: New Directions in Theory and Practice* (pp. 1-19). Thousand Oaks, Ca: Sage Publications.
- RYAN, C., y CLIFF, A. (1997). «Do Travel Agents Measure Up Customer Expectation? An Empirical Investigation of Travel Agencies Service Quality As Measured By Servqual». *Journal of Travel and Tourism Marketing*, 6 (2), pp. 1-31.
- RUST, R. T., INMAN, J. J., JIA, J., y ZAHORIK, A. (1999). «What you don't know about customer-perceived service quality: The role of customer expectations distributions». *Marketing Science*, 18 (1), pp. 77-92.
- SCHARITZER, D., y KORUNKA, C. (2000). «New Public Management: Evaluating the Success of Total Quality Management and Change Management Interventions in Public Services from the Employees' and Customers' Perspectives». *Total Quality Management*, 11 (September), pp. 941-954.
- SHIV, B., CARMON, Z., y ARICLY, D. (2005). «Placebo effects of marketing actions: consumers may get what they pay for». *Journal of Marketing Research*, 42 (November), pp. 383-393.
- STEIGER, J. H. (2004). «Beyond the F-test: Effect Size Confidence Intervals and Tests of Close Fit in the Analysis of Variance and Contrast Analysis». *Psychological Methods*, 9, pp. 164-182.

- SURESHGHANDAR G. S., CHANDRASEKHARAN R., y ANANTHARAMAN R. N. (2002). «The Relationship between Management's Perception of Total Quality Service and Customer Perceptions of Service Quality». *Total Quality Management*, 13 (1), pp. 69-88.
- TEAS, R. (1993). «Expectations, performance evaluation, and consumer's perceptions of quality». *Journal of Marketing*, 57, pp. 18-34.
- TEAS, R. (1994). «Expectations as a Comparison Standard in Measuring Service Quality: An Assesment of a Reassessment». *Journal of Marketing*, 58 (January), pp. 132-139.
- VERMUNT, J. K. y MAGIDSON, J. (2005). *Latent GOLD 4.0 User's Guide*. Belmont, Massachusetts: Statistical Innovations Inc.