

## **Dimensiones de la calidad percibida en servicios educativos: determinación empírica mediante una escala de valoración del servicio**

**José Luis Palacios Gómez**  
Universidad Autónoma de Madrid

RECIBIDO: 21 de noviembre de 2008

ACEPTADO: 15 de junio de 2009

---

**Resumen:** El objetivo de este trabajo es mostrar la estructura factorial de la calidad percibida en algunos servicios educativos y su estabilidad temporal. Se ha utilizado para ello una escala aplicada en dos ocasiones sucesivas en dos servicios educativos distintos. Los resultados muestran que la fiabilidad de la escala es alta y que su estructura factorial (dimensiones de la calidad percibida) es similar en distintos servicios y se mantiene a lo largo del tiempo, de tal modo que puede concluirse que la escala probablemente constituye un instrumento apropiado y útil, más ventajoso que otros usados, para medir la calidad percibida de servicios de naturaleza semejante a los aquí considerados.

**Palabras clave:** Calidad percibida / Satisfacción del cliente / Escala de medida.

### **Dimensions of Perceived Quality in Educational Services. An Empirical Definition by Means of a Scale to Assess the Service**

**Abstract:** The aim of this work is to show the factorial structure and time stability of perceived quality in some educational services by means of a scale applied two successive times on two educational services. The results show a high reliability of the scale, remaining its factor structure (perceived quality dimensions) all the time in every service, in such a way that we can conclude this is likely a suitable and useful instrument, more advantageous than others, to measure perceived quality in similar services.

**Key Words:** Perceived quality / Customer satisfaction / Measure scale.

---

## **INTRODUCCIÓN**

En la gestión de servicios se tiende a basar la calidad del servicio en la satisfacción que este produce en quien lo recibe. Identificar calidad del servicio con satisfacción del cliente constituye una ventaja considerable, ya que la satisfacción del cliente puede medirse con relativa facilidad, cobrando la forma de un indicador cuantitativo, objetivable y comparable, de la calidad del servicio. En la fundamentación teórica de las principales escalas utilizadas para medir la calidad del servicio, como SERVQUAL (Parasuraman, Berry y Zeithalm, 1988, 1991), SERVPERF (Cronin y Taylor, 1992, 1994) o de *Desempeño Evaluado* (Teas, 1993), sus autores distinguen nominalmente calidad percibida del servicio y satisfacción del cliente, pero señalan que las diferencias entre ambos conceptos no son claras y que no se ha podido establecer si efectivamente son dos cosas distintas, e incluso algunos autores han considerado que la calidad de un servicio puede ser tanto un antecedente como una consecuencia de la satisfacción (Martínez Tur, Peiró y Ramos, 2001). Por otra parte, es práctica común en la gestión de la calidad total de las organizaciones productoras de servicios,

tanto públicas como privadas, la identificación operativa de la calidad percibida con la satisfacción del cliente (Alet i Vilaginé, 2000; Denton, 1991; Heredia, 2001; López y Gadea, 1998; Mas, Azorín y Casado, 2004; Senlle y Vilar, 1996). Igualmente, parece claro que no resulta imprescindible medir las expectativas del cliente, la importancia que confiere a los distintos aspectos del servicio o el nivel ideal de servicio para obtener una medida suficiente de la calidad percibida del servicio que resulte útil y operativa para gestores y técnicos de calidad (Bielen y D'Hoore, 2002; Carman, 1990; Cronin y Taylor, 1994; Hunt, 1977).

Por otro lado, la revisión de la literatura ofrece numerosos ejemplos de medición de la calidad percibida de servicios educativos mediante escalas de tipo *servqual*, *servperf* o inspiradas en estas, especialmente cuando se estudia la calidad de estos servicios sin circunscribirse al aspecto docente (Capelleras y Veciana, 2001; Le Blanc y Nguyen, 1997; Manzebón, Martínez y Pérez, 2007; Olfeld y Baron, 2000; Owlia y Aspinwall, 1996; Rigotti y Pitt, 1992; de Salvador, 2003). En estos casos, además, la calidad percibida del servicio educativo se relaciona casi invariable-

mente con tres grupos de elementos o factores: relativos a las instalaciones, a la interacción con el personal que presta el servicio y a distintos aspectos organizativos, aunque la calidad percibida no siempre se sustancia teórica u operativamente en sólo estos tres factores analíticos o dimensiones (mientras que algunos autores, como Oldfield y Baron, presentan tres dimensiones –requisitos materiales, docentes y funcionales– otros, como Owlia y Aspinwall, proponen cuatro –actitud, contenido, recursos académicos y competencia– y otros, como Leblanc y Nguyen, ofrecen siete –interacción con el profesorado, reputación, aspectos físicos, interacción con el personal administrativo, currículo, respuesta y accesibilidad–).

Por otra parte, la satisfacción con un servicio puede entenderse, en definitiva, como una actitud (Churchill y Surprenant, 1982; Swan, Trawick y Carroll, 1982; Tse, Nicosia y Wilton, 1990), resultado de un proceso mental que sintetiza elementos cognitivos y afectivos (Mano y Oliver, 1993), y que como tal puede ser medida con los instrumentos apropiados a tal efecto.

En consecuencia, pensamos que para medir la calidad del servicio de una forma apropiada y suficiente puede utilizarse una escala tipo *servperf* (sin ponderar por importancia o punto ideal ni considerar expectativas), cuyas dimensiones de calidad se inspiren en las propuestas por *servqual*, que tome la satisfacción del cliente como indicador de la calidad percibida del servicio y que la mida mediante gradientes tipo Likert ordinales o numéricos con ítems relativos a la satisfacción del cliente con distintos aspectos del servicio. La escala que se ha utilizado en la medida de la calidad percibida de dos servicios educativos municipales que estudiamos aquí es una escala de este tipo y sus características concretas se describen en el apartado siguiente. La adaptación conceptual de *servqual* se ha realizado en función de la constatación cualitativa empírica de que los aspectos que interesa evaluar a los gestores de los servicios municipales aquí considerados se circunscriben a los que comprenden las tres dimensiones propuestas.

Sostenemos aquí, como primera hipótesis, que la calidad percibida de los servicios educati-

vos descansa sobre tres dimensiones principales (que abarcan, correlativamente, los conjuntos de aspectos fundamentales involucrados en esta clase de servicios): infraestructura, profesorado y trámites para recibir el servicio (en correspondencia conceptual con las apuntadas más arriba: instalaciones, interacción con el personal y aspectos organizativos), y que una escala que tenga capacidad de medir la calidad percibida en este ámbito debe, pues, construirse con ítems relativos a esas tres dimensiones articuladas en la implementación del servicio. Sostenemos también, como segunda hipótesis, que esta estructura tridimensional debe mantenerse a lo largo del tiempo, al menos en el corto-medio plazo, si efectivamente el constructo de calidad tridimensional propuesto se corresponde con la naturaleza del servicio y con la percepción de sus usuarios y no es efímero. Una escala de medida de la calidad percibida del servicio, construida y evaluada con técnicas factoriales, comparadas sus aplicaciones trasversal y longitudinalmente, permite comprobar ambas hipótesis: si la calidad percibida del servicio educativo consta de las dimensiones propuestas en distintos servicios y si esa estructura dimensional resulta estable durante cierto tiempo. La confirmación de estos extremos comportaría: a) una evidencia a favor de una determinada concepción de la dimensionalidad de la calidad percibida en estos servicios; y b) tomar la escala como un instrumento potencialmente válido para medirla en servicios semejantes a los aquí considerados. En las páginas que siguen se sustancian las operaciones necesarias para llevar a cabo la comprobación de las hipótesis antes enunciadas.

Tal como señalan algunos autores (Marzo y Muntané, 2002), en servicios educativos como los aquí considerados (una Universidad Popular y una Escuela Municipal de Música), a pesar de su abundante implantación entre los servicios públicos municipales, apenas pueden encontrarse mediciones de la calidad percibida del servicio con instrumentos como el que se describe y valida aquí, lo que entendemos justifica mostrarlo y proponerlo como herramienta estándar especialmente para esta clase de servicios.

## METODOLOGÍA

En la construcción de escalas de medida de actitudes, la utilización de técnicas factoriales es una práctica que cuenta con una ya larga tradición (Brown, 2006; De Vellis, 1991; Hurley *et al.*, 1996; Kline, 1986; Long, 1986; Morales, 1988; Nunnally, 1978; Tukey, 1977) y que, con la ayuda de los modernos paquetes informático-estadísticos, ha llegado a estandarizarse metodológicamente, pues ofrece ventajas operativas considerables: fundamentalmente, sencillez conceptual y potencia analítica. El análisis factorial exploratorio (AFE), entre sus múltiples aplicaciones, resulta útil para determinar la dimensionalidad del rasgo actitudinal que pretende medir la escala y los subconjuntos de variables que aquel encierra (factores), así como para establecer la cantidad de varianza total que pueden explicar. El análisis factorial confirmatorio (AFC) permite, también entre otras aplicaciones, comprobar si una matriz de datos se ajusta a un modelo conceptual multivariable concreto y si efectivamente conforma subconjuntos de variables preestablecidos por ese modelo, y puede constituir una eficaz prueba de validez de constructo de un instrumento de medida como es una escala. El uso combinado de AFE y AFC se ha revelado como una metodología especialmente adecuada para diseñar y validar escalas de actitudes, permitiendo comprobar si los datos se ajustan al modelo y si el modelo se ajusta a los datos. Además, y en lo que aquí nos interesa, la combinación de análisis factorial exploratorio y análisis factorial confirmatorio para diseñar y testar la validez de las escalas de medida de la calidad percibida es una estrategia que cuenta con abundantes antecedentes documentados en la literatura (Brady y Cronin, 2001; Bielen y D'Hoore, 2002; Céspedes y de Burgos, 2004; Chou, Boldy y Lee, 2002; Fernández, López y Mariel, 2005; San Martín, 2003).

## DISEÑO DE LA ESCALA

Tal y como hemos señalado, para medir la calidad percibida del servicio, sobre la base de la satisfacción del usuario con el mismo, puede optarse por utilizar una escala de tipo *servperf*, que sólo recoge la calificación directa del usuario a

los distintos aspectos del servicio, sin ponderar por importancia ni recoger expectativas. Los ítems de la escala aquí utilizada, sin embargo, están inspirados en la escala *servqual*, que distingue entre aspectos tangibles e intangibles del servicio, pero adaptados a las características de los servicios cuya calidad percibida se pretende medir, según la propuesta que defienden diversos autores (Mels, Boshoff y Nel, 1997; Bielen y D'Hoore, 2002; Johnson, Hermann y Gustafsson, 2002; Barrera y Reyes, 2003) y que los propios creadores de *servqual* contemplan como posible (Parasuraman, Zeithalm y Berry, 1994, p. 123). Nuestra escala no diferencia cinco dimensiones de la calidad, sino tres: *instalaciones* (tangibles), *actuación del personal* y *organización de los trámites para usar el servicio* (intangibles). La dimensión *instalaciones* está constituida por cinco ítems: limpieza, accesibilidad, conservación, seguridad y confort; la dimensión *personal* comprende cuatro ítems: simpatía, profesionalidad, motivación y comunicación; la dimensión *organización de los trámites* comprende tres ítems: horario de atención, rapidez de los trámites y comodidad de estos. Los ítems de la escala han sido determinados siguiendo los intereses informativos de los gestores de los servicios considerados. Cada ítem es puntuado por el usuario entre 1 (mínima satisfacción) y 10 (máxima satisfacción) y su descripción semántica es la que sigue:

- Limpieza: pulcritud de las dependencias y dotaciones del edificio donde se presta el servicio.
- Accesibilidad: facilidad de llegar y localizar las distintas dependencias.
- Conservación: cuidado y mantenimiento de las instalaciones.
- Seguridad: ausencia de riesgos físicos y posibilidad de accidentes.
- Confort: climatización, ruidos, comodidad del mobiliario.
- Simpatía: amabilidad, trato agradable.
- Profesionalidad: ejercicio de la función con competencia.
- Motivación: predisposición para realizar su trabajo.
- Comunicación: capacidad de conectar con el usuario.

- Horario de atención: períodos e intervalos de tiempo disponibles para hacer los trámites necesarios para recibir el servicio.
- Rapidez: velocidad de realización de los trámites para recibir el servicio.
- Comodidad: facilidad y confort para la realización de los trámites.

Los ítems referidos a la subdimensión *organización de los trámites* responden en buena medida a la relativa singularidad de los servicios educativos municipales, para cuyo disfrute es necesario seguir un proceso burocrático de cierta entidad, cuya consideración por el usuario constituye un elemento importante de la calidad percibida del servicio.

## MUESTRA

La escala propuesta para medir la satisfacción de los usuarios, como indicador de la calidad percibida, se ha aplicado en los usuarios de una Universidad Popular Municipal y de una Escuela Municipal de Música de un gran municipio de Madrid, con dos encuestas consecutivas realizadas en cada uno de los servicios en la modalidad de cuestionarios autoadministrados. Los tamaños muestrales de las encuestas y sus errores muestrales teóricos (para el supuesto de máxima indeterminación,  $p=q$ , y un nivel de confianza de  $2\sigma$ ) son los mostrados en la tabla 1.

**Tabla 1.-** Tamaños y errores muestrales de las encuestas

	Escuela Música 2004	Escuela Música 2006	Univ. Popular 2003	Univ. Popular 2005
<i>n</i>	256	210	313	429
<i>e</i> ( $\pm\%$ )	5,18	5,81	5,05	4,13

## ANÁLISIS ESTADÍSTICOS

Los análisis estadísticos llevados a cabo para confirmar empíricamente la estructura teórica de la escala propuesta y comprobar su consistencia y estabilidad son los siguientes:

- Análisis descriptivo de las muestras para determinar la linealidad y la normalidad de las variables empleadas en la escala mediante gráficos bivariantes de dispersión e indicadores de

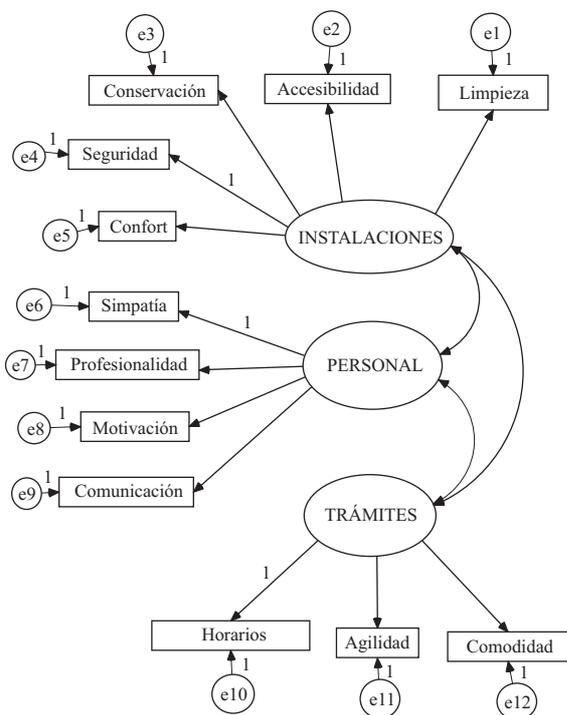
curtosis y asimetría, respectivamente. Toda vez que los análisis factoriales empleados descansan sobre medidas de asociación como son las correlaciones y las covarianzas, la no linealidad impide una correcta interpretación de estos estadísticos. Igualmente, la violación notoria de la exigencia de normalidad en las distribuciones de los valores de las variables que componen la escala comporta que las propiedades de ciertos estimadores o métodos de estimación (como el de máxima verosimilitud, aquí empleado en el AFC) queden afectadas y que los contrastes de hipótesis resulten distorsionados. Comprobar estos extremos, por tanto, resulta una condición previa de los análisis factoriales practicados.

- Análisis factorial exploratorio con el fin de comprobar la unidimensionalidad de la escala y la existencia de las tres subdimensiones teóricas previstas (validez de constructo). Hemos optado por el procedimiento de extracción de factores por el método de factorización de ejes principales para explicar la varianza común de las variables de la escala (Tabachnik y Fidell, 1989), y por el procedimiento de rotación *oblimin directo*, presuponiendo que los factores se encuentran relacionados y no imponiendo restricciones de ortogonalidad (Hair, Anderson, Tatham y Black, 1999).
- Análisis de la fiabilidad de la escala mediante el coeficiente  $\alpha$  de Cronbach, observando el aumento o pérdida de fiabilidad de la escala si se suprime cada uno de sus ítems en ocasiones sucesivas
- Análisis factorial confirmatorio mediante modelos de ecuaciones estructurales (SEM) (con el programa AMOS 5.0) para comprobar si se reproduce la estructura de la escala propuesta y se mantiene en sucesivas aplicaciones en el mismo servicio. En el uso que aquí hacemos de esta metodología, básicamente se trata de comprobar si los ítems de cada subdimensión correlacionan con la suficiente fuerza con cada una de estas y si las covarianzas entre las subdimensiones (constructos *instalaciones*, *personal* y *trámites*) son lo suficientemente altas como para permitirnos pensar que efectivamente son subdimensiones de una dimensión que es la que mide la escala (constructo *calidad percibida*), pero no tanto como para concluir que no

constituyen subdimensiones distintas entre sí. En la terminología de SEM, los ítems de la escala reciben el nombre de *variables observables*, mientras que las subdimensiones o factores se denominan *variables latentes*. El modelo estructural especificado puede verse en la figura 1, donde se observan las restricciones impuestas en el mismo para que resulte “identificado” (“1” en las flechas del diagrama). Hay que señalar que el análisis factorial con SEM se ha llevado a cabo suprimiendo todos los casos con valores perdidos al efecto de calcular algunos estadísticos de bondad de ajuste del modelo (RMR, GFI) que precisan una matriz de datos sin valores perdidos.

- Análisis de la fiabilidad compuesta en el AFC, así como pruebas de validez convergente, mediante el cálculo de la varianza extraída media, y de validez discriminante, comparando varianza extraída y correlaciones entre constructos (Fornell y Larcker, 1981).

Figura 1.- Modelo factorial especificado



## RESULTADOS

El análisis de la estadística descriptiva de las puntuaciones dadas a las variables que compo-

nen la escala, mediante el análisis de gráficos de dispersión bivariable y de indicadores de asimetría y curtosis, pone de relieve que los presupuestos de linealidad y normalidad se cumplen dentro de unos márgenes tolerables. Aunque algunos valores de los indicadores de normalidad son altos, ninguna distribución de frecuencias presenta estructuras plurimodales ni diferencias notables entre los estadísticos de posición que permitan concluir alejamiento significativo de la gaussiana (obviamos reproducir aquí, por razones de espacio, los gráficos de dispersión bivariable y las tablas de coeficientes de asimetría y curtosis).

El análisis factorial exploratorio (AFE) resulta pertinente en todos los casos, puesto que el estadístico KMO alcanza un valor cercano o superior a 0,85 (Hair *et al.*, 1999, pp. 88-89) y el test de esfericidad de Barlett, que contrasta la hipótesis nula de que la matriz de correlaciones es una matriz de identidad, se muestra estadísticamente significativo (tabla 2). Podemos observar (tablas 3 y 4) que en los análisis factoriales con los datos de las dos aplicaciones de la escala en ambos servicios educativos aparecen siempre tres factores, con autovalores superiores a la unidad, que explican entre el 71% y el 75% de la varianza total. Como puede verse (tabla 6) en la matriz de estructura (que es la relevante para nuestro interés) de los análisis de los datos de las dos encuestas en la Universidad Popular, el primer factor aglutina las cuatro variables relativas al Personal del servicio; el segundo factor comprende las cinco variables correspondientes a las Instalaciones; y, finalmente, el tercer factor engloba las tres variables referidas a los Trámites para usar el servicio. Sólo encontramos algún problema para incluir a la variable *Limpieza* dentro del factor correspondiente en el caso de la matriz obtenida con los datos de la encuesta de 2005, pues su carga factorial (0,460) es notoriamente menor que las de las otras cuatro variables que se agrupan como segundo factor, y menor incluso que las de otras variables de la escala en la columna de este factor, pero el hecho de que esta variable presente cargas mucho menores en los otros dos factores nos lleva a concluir que la *Limpieza* forma parte, efectivamente, del factor relativo a las Instalaciones. En los análisis realizados con los datos de las encuestas en la Escue-

la de Música encontramos que, en el caso de los correspondientes a la encuesta de 2004, el primer factor aglutinaría a las variables relativas al Personal, el segundo a las variables relativas a las Instalaciones y el tercer factor a las relativas a los Trámites, mientras que en el caso del análisis practicado con los datos de la encuesta de 2006 se reproduciría la relación entre factores y variables que hemos observado en el caso del análisis con los datos de la Universidad Popular. Ahora también encontramos algún problema para incluir la variable *Accesibilidad* dentro del factor correspondiente en el caso de la matriz obtenida con los datos de la encuesta de 2004, pues su carga factorial (0,570) es menor que las de las otras cuatro variables que se agrupan en el primer factor, y menor incluso que las de otras variables de la escala en la columna de este factor, pero el hecho de que esta variable presente cargas mucho menores en los otros dos factores nos lleva a concluir, como lo hicimos con el ítem *Limpieza* en el caso del análisis con los datos de

la Universidad Popular, que la *Accesibilidad* forma parte, asimismo, del factor relativo a las Instalaciones.

Por último, si observamos la matriz de correlaciones entre los factores de los AFE practicados con los conjuntos de datos de todas las encuestas (tablas 8 y 9), vemos que los coeficientes toman valores que no son muy altos ni muy bajos, indicativos de que los tres factores son diferentes pero relacionados.

En definitiva, el análisis factorial exploratorio ofrece unos resultados en plena coherencia con la teoría: a) la escala parece ser unidimensional, ya que las saturaciones o cargas de todas las variables sobre el primer factor superan el valor de 0,30 establecido para llegar a esa conclusión (Comrey, 1985; García, Gil y Rodríguez, 2000; Yela, 1997); b) el constructo *calidad percibida* posee tres subdimensiones (factores) que comprenden cada una el correspondiente conjunto de variables hipotetizado en la teoría.

**Tabla 2.-** Estadístico de Kaiser-Mayer-Olkin y test de Bartlett

ENCUESTA	UNIV. POP. 2003	UNIV. POP. 2005	ESC. MÚSICA 2004	ESC. MÚSICA 2006
KMO	0,847	0,881	0,871	0,869
$\chi^2$ Test de Bartlett	2038,98	2564,18	1431,35	1735,14
Signif. Test Bartlett	0,000	0,000	0,000	0,000

**Tabla 3.-** Varianza total explicada (caso Universidad Popular)

Factor	ENCUESTA 2003			ENCUESTA 2005		
	Total	% varianza	% acum.	Total	% varianza	% acum.
1	5,069	42,244	42,244	5,435	45,292	45,292
2	2,005	16,707	58,952	1,750	14,584	59,876
3	1,549	12,911	71,863	1,305	10,871	70,747
4	,671	5,593	77,456	,780	6,503	77,250
5	,633	5,278	82,735	,549	4,573	81,823
6	,495	4,124	86,859	,485	4,045	85,868
7	,436	3,634	90,493	,416	3,470	89,339
8	,391	3,256	93,750	,364	3,037	92,376
9	,253	2,109	95,858	,280	2,336	94,711
10	,206	1,717	97,575	,257	2,140	96,852
11	,153	1,271	98,846	,204	1,698	98,549
12	,138	1,154	100,000	,174	1,451	100,000

**Tabla 4.-** Varianza total explicada (caso Escuela de Música)

Factor	ENCUESTA 2003			ENCUESTA 2005		
	Total	% varianza	% acum.	Total	% varianza	% acum.
1	6,194	51,615	51,615	5,684	47,369	47,369
2	1,703	14,189	65,804	1,740	14,501	61,870
3	1,098	9,149	74,953	1,344	11,199	73,069
4	,698	5,815	80,768	,731	6,088	79,157
5	,503	4,194	84,962	,587	4,889	84,046
6	,421	3,504	88,466	,473	3,946	87,991
7	,365	3,044	91,510	,397	3,305	91,297
8	,287	2,393	93,903	,332	2,765	94,062
9	,238	1,980	95,883	,316	2,635	96,697
10	,214	1,786	97,669	,189	1,576	98,273
11	,166	1,385	99,054	,137	1,145	99,418
12	,114	,946	100,000	,070	,582	100,000

**Tabla 5.-** Matriz factorial: cargas sobre el primer factor de las variables de la escala

ITEMS	ENCUESTA U.P. 2003	ENCUESTA U.P. 2005	ENCUESTA E.M. 2004	ENCUESTA E.M. 2006
Limpieza	,542	,387	,726	,584
Accesibilidad	,498	,572	,520	,530
Conservación	,530	,582	,781	,621
Seguridad	,514	,594	,665	,587
Confort	,389	,549	,682	,461
Simpatía	,786	,644	,772	,834
Profesionalidad	,807	,796	,705	,824
Motivación	,792	,716	,720	,868
Comunicación	,770	,740	,805	,826
Horarios	,559	,720	,615	,536
Rapidez	,648	,728	,658	,577
Comodidad	,550	,690	,713	,634

**Tabla 6.-** Matriz de estructura factorial en caso Univ. Popular

ITEMS	ENCUESTA 2003			ENCUESTA 2005		
	Factor			Factor		
	1	2	3	1	2	3
Limpieza	,315	,684	,353	,243	,460	-,277
Accesibilidad	,273	,702	,273	,371	,651	-,421
Conservación	,257	,817	,275	,318	,814	-,353
Seguridad	,329	,608	,323	,369	,736	-,399
Confort	,200	,596	,184	,330	,729	-,334
Simpatía	,880	,380	,396	,738	,367	-,415
Profesionalidad	,912	,363	,422	,924	,439	-,509
Motivación	,900	,369	,388	,857	,382	-,437
Comunicación	,902	,294	,399	,852	,402	-,490
Horarios	,390	,304	,699	,515	,477	-,816
Rapidez	,393	,353	,928	,488	,462	-,891
Comodidad	,315	,313	,805	,448	,424	-,878

**Tabla 7.-** Matriz de estructura factorial en caso Escuela Música

ITEMS	ENCUESTA 2004			ENCUESTA 2006		
	Factor			Factor		
	1	2	3	1	2	3
Limpieza	,790	-,451	,499	,342	,768	,360
Accesibilidad	,570	-,319	,355	,338	,613	,392
Conservación	,908	-,428	,524	,363	,821	,376
Seguridad	,709	-,489	,358	,399	,737	,286
Confort	,731	-,448	,445	,341	,545	,213
Simpatía	,616	-,823	,368	,909	,479	,471
Profesionalidad	,457	-,887	,303	,935	,422	,454
Motivación	,449	-,932	,297	,955	,501	,466
Comunicación	,602	-,861	,449	,915	,477	,426
Horarios	,539	-,293	,799	,383	,337	,721
Rapidez	,478	-,355	,961	,384	,339	,877
Comodidad	,650	-,438	,723	,461	,493	,695

**Tabla 8.-** Correlaciones entre los factores extraídos en caso Univ. Popular

Factores	DATOS ENCUESTA 2003			DATOS ENCUESTA 2005		
	1°	2°	3°	1°	2°	3°
1°	1,000	0,386	0,392	1,000	0,470	0,513
2°	0,386	1,000	0,446	0,470	1,000	0,548
3°	0,392	0,446	1,000	0,513	0,548	1,000

**Tabla 9.-** Correlaciones entre los factores extraídos en caso E. Música

Factores	DATOS ENCUESTA 2004			DATOS ENCUESTA 2006		
	1°	2°	3°	1°	2°	3°
1°	1,000	-0,550	0,570	1,000	,497	,472
2°	-0,550	1,000	-0,357	,497	1,000	,443
3°	0,570	-0,357	1,000	,472	,443	1,000

Por su parte, el análisis de la fiabilidad de la escala nos muestra valores muy altos del coeficiente alfa de Cronbach: 0,874 y 0,883 para las escalas aplicadas en 2003 y 2005, respectivamente, en la Universidad Popular y 0,912 y 0,893, para las escalas aplicadas en 2004 y 2006, respectivamente, en la Escuela de Música, muy superior al establecido comúnmente como mínimo admisible (Knapp, 1976; Miller, 1977; Nunnally, 1978). La práctica de retirar ítems de la escala uno a uno no nos ofrece valores de alfa que indiquen mejora sustancial de la fiabilidad de la escala en ningún caso. El uso del coeficiente de fiabilidad como indicador de la unidimensionalidad de una escala presenta riesgos notables bien descritos en la literatura (Kline, 1994; Nunnally, 1978; Nunnally y Berstein, 1994), pero puede afirmarse que un coeficiente alto de fiabilidad es una condición necesaria, aunque no suficiente, para afirmar que una escala mide una única dimensión o rasgo (Morales, Urosa y Blanco, 2003, p. 100). Así pues, podemos concluir que los coeficientes de fiabilidad hallados nos indican, por un lado, una alta fiabilidad de la escala, y, por otro, que la escala cumple con una de las exigencias comunes para establecer que efectivamente es unidimensional.

En el análisis de los resultados del análisis factorial confirmatorio (AFC) nos hemos centrado en la bondad del ajuste del modelo especifi-

cado a los datos empíricos, ya que es la cuestión más importante para determinar si el modelo es correcto y si es útil para nuestros objetivos confirmatorios, entendiéndose por “correcto” el modelo que incorpora las restricciones y supuestos implícitos que se cumplen en la población y, por lo tanto, especifica correctamente las relaciones entre las variables sin omitir parámetros (Batista Foguet y Coenders, 2000, p. 84). También hemos considerado especialmente los valores que toman los coeficientes entre las variables del modelo (observables y latentes) para comprobar nuestros supuestos sobre la estructura de la escala (tablas 10 y 11). En este sentido, hay que señalar que en ninguna de las aplicaciones con el modelo utilizado existen “casos Heywood”, es decir, parámetros que toman valores impropios (varianzas negativas o correlaciones  $>1$  o  $<-1$ ) y que todos los coeficientes resultan estadísticamente significativos ( $t \geq 1,96$ ). En los resultados mostrados en la tabla 12 se han omitido, por razones de espacio, los estadísticos de los modelos saturado y de independencia. Las figuras 2 y 3 ilustran las representaciones gráficas del modelo estructural con las soluciones no estandarizada y estandarizada en el caso del AFC con los datos de la Encuesta de Música en 2004 (el más problemático a efectos confirmatorios). Obviamos reproducir las representaciones gráficas de los otros casos por razones de espacio.

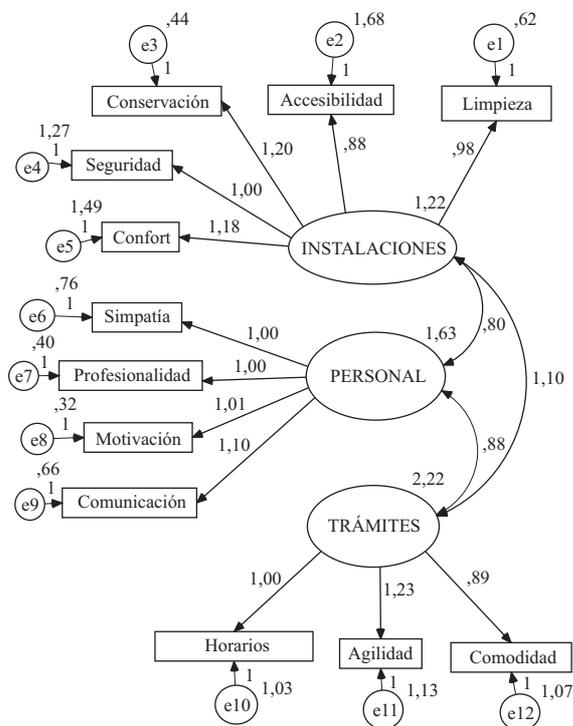
**Tabla 10.-** Coeficientes de regresión entre variables latentes y variables observables del AFC (solución estandarizada)

ITEMS	INSTALACIONES				PERSONAL				TRÁMITES			
	UP '03	UP '05	EM '04	EM '06	UP '03	UP '05	EM '04	EM '06	UP '03	UP '05	EM '04	EM '06
Limp.	,70	,46	,81	,76								
Acces.	,72	,68	,59	,64								
Cons.	,83	,80	,89	,72								
Segur.	,73	,74	,70	,54								
Conf.	,66	,74	,73	,90								
Simp.					,91	,80	,83	,93				
Profes.					,92	,93	,89	,96				
Motiv.					,92	,85	,91	,92				
Comun.					,92	,84	,86	,72				
Horar.									,69	,83	,83	,81
Rapid.									,97	,89	,87	,76
Comod.									,79	,87	,79	,82

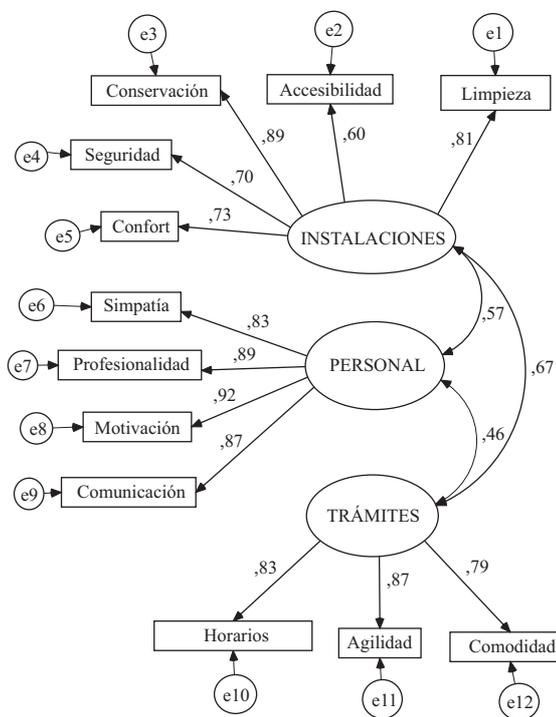
**Tabla 11.-** Correlaciones entre variables latentes del AFC (solución estandarizada)

	INSTALACIONES				PERSONAL				TRÁMITES			
	UP '03	UP '05	EM '04	EM '06	UP '03	UP '05	EM '04	EM '06	UP '03	UP '05	EM '04	EM '06
Instal.	1,00	1,00	1,00	1,00	,40	,46	,57	,50	,41	,53	,67	,52
Pers.	,40	,46	,57	,50	1,00	1,00	1,00	1,00	,39	,57	,46	,55
Trám.	,41	,53	,67	,52	,39	,57	,46	,55	1,00	1,00	1,00	1,00

**Figura 2.-** Modelo factorial E. Música 04 (solución no estandarizada)



**Figura 3.-** Modelo factorial E. Música 04 (solución estandarizada)



**Tabla 12.-** Estadísticos de ajuste del AFC

CMIN					
Modelo especificado	NPAR	CMIN	DF	P	CMIN/DF
Datos Enc. U.P. 2003	27	90,630	51	0,053	1,777
Datos Enc. U.P. 2005	27	94,949	51	0,001	1,862
Datos Enc. E.M. 2004	27	170,173	51	0,000	3,337
Datos Enc. E.M.2006	27	101,331	51	0,000	1,987
RMR, GFI					
Modelo especificado	RMR	GFI	AGFI	PGFI	
Datos Enc. U.P. 2003	,212	,943	,913	,616	
Datos Enc. U.P. 2005	,108	,960	,939	,628	
Datos Enc. E.M. 2004	,181	,865	,794	,566	
Datos Enc. E.M.2006	,099	,928	,889	,607	
Baseline Comparisons					
Modelo especificado	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Datos Enc. U.P. 2003	,764	,694	,881	,838	,875
Datos Enc. U.P. 2005	,963	,953	,983	,978	,983
Datos Enc. E.M. 2004	,894	,862	,923	,899	,922
Datos Enc. E.M.2006	,943	,926	,971	,962	,971
RMSEA					
Modelo especificado	RMSEA	LO 90	HI 90	PCLOSE	
Datos Enc. U.P. 2003	,054	,035	,071	,034	
Datos Enc. U.P. 2005	,049	,033	,064	,031	
Datos Enc. E.M. 2004	,113	,094	,132	,000	
Datos Enc. E.M.2006	,069	,049	,088	,059	

En los resultados del AFC el estadístico el estadístico CMIN (*minimum discrepancy*) fundamentado en  $\chi^2$  indica que, salvo en el caso de los datos UP '03 existe riesgo de discrepancia entre

los datos y la especificación del modelo, puesto que la hipótesis nula que se contrasta es que las matrices de datos efectiva y prevista son iguales ( $H_0: \Sigma(\theta) = \Sigma$ ). Sin embargo, como el estadístico

es sensible a los grados de libertad, la ratio CMIN/DF (Chi-cuadrado normada) resulta más valiosa para apreciar la existencia de discrepancia con el modelo: aunque no hay plena concordancia entre los especialistas, se suele aceptar que valores de la ratio inferiores a 2 o incluso a 3 indican un ajuste aceptable (Hair *et al.*, 1999, pp. 684-685). Consiguientemente, sólo en el caso EM '04 encontraríamos un desajuste más notable. Los principales estadísticos de bondad de ajuste complementarios como GFI (*goodness of fit index*) y AGFI (*adjusted goodness of fit index*), superiores al valor mínimo aconsejado de 0,90 (Bentler, 1992, p. 402), nos indican, asimismo, un grado de ajuste aceptable con residuos perfectamente tolerables, salvo de nuevo en el caso de EM '04, donde desciende por debajo de ese valor mínimo. El estadístico RMSEA (*root mean square error of approximation*), por su parte, con un valor inferior al máximo aconsejado de 0,08 (Hair *et al.*, 1999, p. 686), nos indica también un grado de ajuste aceptable en todos los casos salvo otra vez en EM '04. Otros índices de ajuste, como el NFI (*normed fit index*) o los derivados del mismo, sólo muestran valores indicativos de buen ajuste en los casos de UP '05 y EM '06, pero hay que tener en cuenta que algunos autores desaconsejan estos índices porque al no tener en cuenta los grados de libertad favorecen la adopción de modelos sobreparametrizados (Batista Foguet y Coenders, 2000, p. 91). En los resultados referidos a los datos de la encuesta de la Universidad Popular en 2005 es donde observamos un mejor ajuste al modelo: los estadísticos de bondad de ajuste GFI y AGFI arrojan valores en torno a 0,95 y el NFI y los índices derivados también muestran en esta ocasión un buen ajuste. El estadístico RMSEA indica en este caso el mejor ajuste de los conjuntos de datos sometidos a AFC, con un valor de 0,049 y una *pclose* con un valor superior al mínimo aconsejado de 0,5 (Jöreskog y Sörborm, 1996).

Por otro lado, los coeficientes correspondientes nos muestran pesos notables entre las variables observables y las latentes y, lo que es más importante, los coeficientes entre las variables latentes (factores), ni muy altos ni muy bajos, nos indican que, efectivamente, *Instalaciones*, *Personal* y *Trámites*, son dimensiones diferen-

tes, pero relacionadas entre sí, del constructo que denominamos *calidad percibida*. El análisis factorial confirmatorio practicado pone de relieve, con razonable consistencia, que efectivamente la escala para medir la calidad percibida del servicio comprende tres dimensiones distintas pero vinculadas entre sí, tal como parecía indicar el análisis factorial exploratorio.

Finalmente, el cálculo de la fiabilidad compuesta de la escala en el AFC (Werts, Linn y Joreskog, 1974; Bagozzi y Yi, 1988) arroja coeficientes  $\rho$  considerablemente altos, que no descienden de 0,82 en el caso de *Instalaciones*, no descienden de 0,85 en el caso de *Personal*, y no descienden de 0,84 en el caso de *Trámites*, dándonos noticia de la notable consistencia interna de cada constructo (subdimensión) en todas las aplicaciones. Las pruebas de validez convergente, realizadas siguiendo el procedimiento de calcular la varianza extraída media (VEM, AVE en inglés) solamente ofrece un resultado inferior al 0,50 recomendado (Fornell y Larcker, 1981) en el caso del constructo *Instalaciones* en la aplicación de la escala en la Universidad Popular en 2005, aunque llega a 0,49, de tal modo que en casi la totalidad de los casos al menos el 50% de la varianza de cada constructo es debida a sus indicadores (oscila entre 52% y 85% en el resto de ocasiones). Por su parte, la validez discriminante ha sido probada comparando la varianza extraída media de cada constructo (subdimensión) con la varianza compartida entre cada constructo y el resto de constructos del modelo (Fornell y Larcker, 1981), arrojando unos resultados que permiten concluir que efectivamente el modelo la posee pues la raíz cuadrada de todas las VEM's calculadas, oscilando entre 0,72 y 0,92, siempre supera las correlaciones entre constructos halladas (tabla 11).

## DISCUSIÓN Y CONCLUSIONES

Aunque no tiene lugar una plena concordancia entre los especialistas sobre la relación entre la calidad del servicio y la satisfacción del cliente con el mismo, en la revisión de la literatura la tendencia dominante presenta la última como el indicador por excelencia de la primera (calidad percibida). Tampoco existe acuerdo definitivo

sobre cuál es la mejor manera de medir la calidad percibida del servicio, y las investigaciones, sobre todo en el campo de la mercadotecnia, ofrecen resultados en ocasiones contradictorios, pero parece claro, en todo caso, que no resulta estrictamente necesario medir las expectativas del cliente, la importancia concedida a los distintos aspectos del servicio o el nivel ideal de servicio para obtener datos que representen la calidad percibida o la satisfacción del cliente. En este sentido, las escalas tipo *servperf* sin ponderar aportan una información suficiente para predecir el comportamiento del cliente con razonable robustez. La elaboración de dimensiones tangibles e intangibles del servicio, tal como proponen las escalas tipo *servqual*, constituyen sin embargo una aportación conceptual que puede tenerse en cuenta a la hora de diferenciar los distintos componentes de la calidad de un servicio. Tomando ambas argumentaciones en consideración, hemos entendido que una escala que diferencie aspectos tangibles e intangibles del servicio y que mida la satisfacción del usuario como indicador de la calidad percibida sobre la base del desempeño, sin ponderar por expectativas u otras consideraciones, constituye un instrumento de medida suficientemente útil y adecuado.

Esta clase de escala para medir la satisfacción del usuario con el servicio es la que hemos empleado para aplicarla en dos servicios educativos, sometiéndola a diversas clases de análisis estadísticos con objeto de verificar la consistencia estructural y su estabilidad temporal del constructo calidad percibida hipotetizado. La escala comprende doce ítems que se refieren a tres dimensiones teóricas del servicio: instalaciones, personal (profesorado) y trámites, las cuales son, por así decir, subdimensiones de la calidad percibida del servicio (una única dimensión que es la que mide el conjunto de la escala). El propósito de los distintos análisis practicados es observar si efectivamente la escala contiene esas tres dimensiones (subdimensiones) y si su aplicación en distintos servicios repetidamente presenta características métricas semejantes y estables, de tal manera que si así fuera podría constituir un instrumento estándar para medir la calidad (percibida) del servicio en otros servicios de naturaleza similar.

Usar el análisis factorial exploratorio en combinación con el análisis factorial confirmatorio constituye una estrategia frecuente para probar la validez del constructo medido y las propiedades métricas de la escala: los datos empíricos sugieren una estructura factorial que a su vez puede refrendarse con datos empíricos. Los análisis factoriales exploratorios, han ofrecido en todos los casos porcentajes de explicación de la varianza total empírica superiores al 71% y autovalores para los tres primeros factores extraídos siempre superiores a la unidad; además la matriz factorial en todos los análisis nos indica unidimensionalidad de la escala (saturaciones  $>0,30$  en todos los ítems) y la matriz de correlaciones entre factores nos señala relación y diferencia entre los mismos conjuntamente. Todo lo cual puede considerarse congruente con la teoría arriba expresada.

El análisis de la fiabilidad de la escala ha ofrecido resultados claros y rotundos, pues en todos los casos (todas las aplicaciones) el coeficiente  $\alpha$  de Cronbach alcanza valores altos y nunca desciende de 0,85 cuando se suprimen unos u otros ítems, lo cual nos da noticia de la alta fiabilidad de la escala. La fiabilidad compuesta en el AFC también ha sido probada con éxito, pues el coeficiente  $\rho$  alcanza siempre valores por encima de 0,80 (nivel adecuado) y muestra que la consistencia interna del bloque de indicadores de cada constructo es alta.

El análisis factorial confirmatorio resulta en términos generales positivo, en el sentido de que los datos empíricos se ajustan razonablemente bien al modelo de medida y que, por consiguiente, la escala ofrece propiedades estructurales que confirman la teoría: la dimensión *calidad percibida* está compuesta de tres subdimensiones, distintas pero relacionadas entre sí. Los usuarios encuestados distinguen la subdimensión a la que se refiere cada ítem de la escala (correlaciones altas o muy altas entre variables correspondientes a ítems y variables latentes o factores). Los coeficientes entre las variables latentes presentan valores medios, que resultan idóneos para concluir que efectivamente la escala mide un “constructo” (calidad percibida) que está compuesto de tres factores claramente identificables (Lévy Mangin, 2003). Los estadísticos de bondad de

ajuste calculados con el programa AMOS (GFI, AGFI, RMSEA) nos indican, en general, un ajuste satisfactorio del modelo especificado a los datos. Otros estadísticos no indican ajuste tan bueno, pero hemos señalado su valor relativo al depender del número de parámetros introducidos en el modelo y del tamaño muestral. Concretamente, el mejor ajuste entre matriz de datos empírica y matriz de datos teórica lo encontramos en el caso de la encuesta en la Universidad Popular en 2005. El peor ajuste lo encontramos en el caso de la encuesta en la Escuela de Música en 2004, que arroja bastantes valores de los estadísticos de ajuste en el AFC que rebasan los límites señalados comúnmente para aceptar un modelo. Cabe sospechar que en este caso el problema se deba a que hemos incluido en el modelo estructural la variable *Accesibilidad*, que mostraba cierta dificultad para ser incluida dentro de la subdimensión correspondiente, y, ciertamente, suprimirla del AFE hace que los tres primeros factores con autovalores superiores a la unidad alcancen a explicar tres puntos porcentuales más de la varianza total que el AFE con esa variable y si la retiramos en el test de fiabilidad de Cronbach la fiabilidad de la escala llega a ser de una décima más en promedio que con ella. Pero si reespecificamos el modelo factorial confirmatorio, suprimiendo el ítem *Accesibilidad*, los estadísticos de ajuste apenas mejoran sustancialmente y siguen ofreciéndonos dudas sobre la pertinencia del modelo.

En consecuencia, cabría pensar que en el caso de la aplicación de la escala entre los usuarios de la Escuela de Música en 2004 el modelo escalar utilizado es inapropiado, y eso permitiría dudar de la efectividad del instrumento para medir la calidad percibida en estos servicios, pero pensamos que ésta sería una conclusión precipitada y probablemente desafortunada. Si atendemos al análisis de fiabilidad de la escala, es precisamente en la aplicación realizada entre los usuarios de la Escuela de Música en 2004 donde encontramos el alpha de Cronbach más elevado, es decir, donde la escala se muestra más precisa para medir la actitud investigada. Además, el AFC mediante ecuaciones estructurales no es una técnica estadística exenta de controversia y no puede decirse que exista unanimidad entre los especialistas a la hora de establecer cuáles son los valores

de los test de bondad de ajuste fuera de discusión (Bollen y Long, 1993; Byrne, 2001; Schumacker y Lomas, 1996), y son numerosos los investigadores que señalan que incluso índices de bondad de ajuste comunes, como el NFI, que carece de distribución conocida, implican una interpretación totalmente subjetiva (García Cueto, Gallo y Miranda, 1998, p. 718). Algunos autores (Suhr, 2006) estiman que si el AFC no confirma la estructura factorial de la hipótesis es entonces cuando resulta pertinente llevar a cabo un AFE para determinar qué estructura factorial pueden sugerir los datos, es decir, invertir el proceso aquí empleado. La misma inversión procedimental sugieren otros investigadores (Blanco y Díaz, 2005) para depurar los ítems escalares, al intentar garantizar la validez escalar desde el criterio de convergencia, descartando los que no saturan suficientemente en los factores correspondientes hallados en un AFE practicado *después* del AFC. Otros autores (Barbero, 1999; Morales et al., 2003) defienden el uso del AFE también *a posteriori*, como forma de asegurar la validez de constructo de una escala, después de que ésta haya sido elaborada con recursos metodológicos en los que no interviene ni el AFE ni el AFC. Finalmente, otros (Hair *et al.*, 1999, pp. 618-617; Pérez Gil, Chacón y Moreno, 2000, pp. 445-446), apuntan que el AFC resulta en realidad un ejercicio tautológico, porque en general las soluciones factoriales obtenidas usando AFE resultan validadas usando AFC y que el caso contrario puede explicarse normalmente acudiendo al error de tipo I, inherente al procedimiento analítico en sí mismo, o a problemas de representatividad de las muestras empleadas. Por otra parte, si bien es cierto que el AFC con los datos de la encuesta en la Escuela de Música de 2004 nos hace albergar alguna duda respecto del modelo factorial hallado con el AFE (que por sí también presenta algún problema, como hemos visto), el resto de los análisis confirmatorios más bien la disipan, incluido el practicado con los datos de 2006 obtenidos en el mismo servicio. Si partiendo de un modelo meramente teórico hubiésemos practicado en primer lugar un AFC en el que hallamos incongruencia en el caso EM 2004, el AFE con esos mismos datos, usado como prueba de la adecuación del modelo a los datos, nos sugiere con bastante consistencia una estructura factorial como la de la hipótesis pro-

puesta. Por otra parte, restar parámetros del modelo, sustrayéndole capacidad informativa práctica, o sobreparametrizarlo, llevando a cabo un artificio meramente estadístico, no parece conveniente cuando las variables que incorpora logran explicar el 75% de su varianza total. Entender prevalente, pues, el resultado del AFE sobre el del AFC, en caso de encontrar una no muy importante incongruencia entre ambos, no parece que sea una opción injustificada.

Finalmente tanto las pruebas de validez convergente como de validez discriminante muestran resultados en general satisfactorios, pues pude observarse que en casi la totalidad de los casos el 50% de la varianza de cada constructo es debida a sus indicadores y que las VEM's calculadas, oscilando entre 0,72 y 0,92, siempre superan las correlaciones cuadráticas entre constructos encontradas en el AFC.

En definitiva, el análisis estadístico practicado sobre la escala para medir la calidad (percibida) del servicio, sobre el concepto de satisfacción del cliente con el mismo, pone de manifiesto que es un instrumento que presenta notables fortalezas relativas a la solidez de su estructura y a su estabilidad. El hecho de que, en general, las dimensiones (subdimensiones) escalares sean las previstas teóricamente y que su fiabilidad sea considerablemente alta, en todas las aplicaciones realizadas, nos lleva a pensar que constituye un instrumento pertinente y útil, capaz de dar cuenta suficiente de la calidad percibida del servicio en este tipo de servicios considerados. El hecho de que la misma escala haya sido aplicada en ocasiones sucesivas en distintos servicios, presentando propiedades psicométricas semejantes y estables, constituye una prueba de la robustez del instrumento que pocas veces se lleva a cabo. Pensamos además que, en relación con otras escalas utilizadas, presenta ventajas derivadas de su sencillez conceptual y operativa, estableciendo las dimensiones estrictamente fundamentales de la calidad percibida y sustanciándolas con pocos ítems, lo cual supone mejor comprensión del objeto estudiado y simplifica y agiliza la medida de éste, tanto para gestores como para clientes: la evitación de la complejidad innecesaria en los cuestionarios de satisfacción de usuarios de servicios redundando positivamente en la abundancia y calidad de los datos obtenidos, siendo éste un ex-

tremo al que se le presta especial énfasis en los textos de metodología de la investigación social con encuesta (Alvira, 2004; Díaz de Rada, 2002; Palacios, 1999; Rojas, Fernández y Pérez, 1998) y al que no siempre se le concede la merecida atención.

Es evidente, sin embargo, que la escala diseñada, aun presentando las virtudes señaladas, debería aplicarse en más servicios, en distintos municipios o ámbitos geográficos y en repetidas ocasiones para contar con mayores garantías de constituir un instrumento estándar que, sin perjuicio de evaluarse permanentemente, pudiese ser empleada comúnmente en servicios educativos públicos (o privados) del tipo de los aquí considerados. Las dificultades que entraña conseguir muestras apropiadas al efecto de validación es una limitación práctica muy considerable que no podemos obviar. En todo caso, abundar en la aplicación comparada en distintas instituciones prestadoras de servicios de naturaleza semejante perfilará mejor su adecuación para la medida de la calidad percibida del servicio como instrumento estandarizado.

## BIBLIOGRAFÍA

- ALET I VILAGINÉS, J. (1994): *Cómo obtener clientes leales y rentables*. Barcelona: Gestión 2000
- ALVIRA, F. (2004): *La encuesta: una perspectiva general metodológica*. Madrid: Centro de Investigaciones Sociológicas
- BAGOZZI, R.P.; YI, T. (1988): "On the Evaluation of Structural Equation Models", *Journal of the Academy of Marketing*, vol. 16, núm. 1, pp. 74-94
- BARBERO GARCÍA, M.I. (1999): *Psicometría II. Métodos de elaboración de escalas*. Madrid: Universidad Nacional de Educación a Distancia
- BARRERA, R.; REYES, M.C. (2003): "Análisis comparado de las escalas de medición de la calidad del servicio", *Actas de las XIII Jornadas Hispano-Lusas de Gestión Científica*. Lugo.
- BATISTA FOGUET, J.M.; COENDERS, G. (2000): *Modelos de ecuaciones estructurales*. Madrid: La Muralla/Hespérides.
- BENTLER, P.M. (1992): "On the Fit of Models to Covariances and Methodology to the Bulletin", *Psychological Bulletin*, 112, pp. 400-404.
- BIELÉN, F.; D'HOORE, W. (2002): "Qualité perçue du service et satisfaction du patient à l'égard de l'hôpital", *Actes du XIII Congrès Annuel de*

- L'Association Latine pour l'Analyse des Systèmes de Santé*. Louvain.
- BRADY, M.K.; CRONIN, J.J. (2001): "Some New Thoughts on Conceptualizing Perceived Service Quality: A Hierarchical Approach", *Journal of Marketing*, vol. 65, núm. 3, pp. 34-49.
- BROWN, T.A. (2006): *Confirmatory Factor Analysis for Applied Research*. New York: Guilford Press.
- BYRNE, B.M. (2001): *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming*. Mahwah, NJ: Erlbaum.
- CAPELLERAS, J.L.; VECIANA, J.M. (2001): *Calidad de servicio en la enseñanza universitaria: desarrollo y validación de una escala de medida*. (Documents de Treball, 2001/4, PP. 3-34). Barcelona: Universitat Autònoma de Barcelona, Dept. d'Economía de l'Empresa.
- CARMAN, J.M. (1990): "Consumer Perceptions of Service Quality: An Assessment of the Servqual Dimensions", *Journal of Retailing*, 66, pp. 33-55.
- CÉSPEDES, J.; BURGOS, J. DE (2004): "Un análisis de las dimensiones de la gestión ambiental en los servicios hoteleros", *Dirección y Organización*, 30, pp. 5-15.
- CHOU, S.; BOLDY, D.P.; LEE, A. (2002): "Measuring Job Satisfaction in Residential Aged Care", *International Journal for Quality in Health Care*, vol. 14, núm. 1, pp. 49-54.
- CHURCHILL, G.A.; SURPRENANT, C. (1982): "An Investigation on the Determinants of Satisfaction", *Journal of Marketing Research*, vol. 19 (November), pp. 491-504.
- TSE, D.; NICOSIA, F.; WILTON, P. (1990): "Consumers Satisfaction as a Process", *Psychology and Marketing*, núm. 7, pp. 177-193.
- COMREY, A.L. (1985): *Manual de análisis factorial*. Madrid: Cátedra.
- CRONIN, J.J.; TAYLOR, S.A. (1992): "Measuring Service Quality: A Reexamination and Extension", *Journal of Marketing*, vol. 56, núm. 3, pp. 55-69.
- CRONIN, J.J.; TAYLOR, S.A. (1994): "Reassessment of Expectations as a Comparison Standard in Measuring Service Quality: Implications for Further Research", *Journal of Marketing*, vol. 58, núm. 1, pp. 111-124.
- DENTON, D.K. (1991): *Calidad en el servicio a los clientes*. Madrid: Díaz de Santos.
- DEVELLIS, R.F. (1991): *Scale Development: Theory and Applications*. Newbury Park, CA: Sage.
- DÍAZ DE RADA, V. (2002): *Tipos de encuesta y diseños de investigación*. Pamplona: Universidad Pública de Navarra.
- FERNÁNDEZ, K.; LÓPEZ, C.; MARIEL, P. (2005): "La satisfacción del consumidor: una aplicación del análisis factorial confirmatorio a la industria automovilística española", *Estadística Española*, 158, pp. 117-141.
- FORNELL, C.; LARCKNER, D.F. (1981): "Evaluating Structural Models with Unobservable Variables and Measurement Error", *Journal of Marketing Research*, núm. 18 (February), pp. 39-50.
- GARCÍA, E.; GIL, J.; RODRÍGUEZ, G. (2000): *Análisis factorial*. Madrid: La Muralla/Hespérides.
- GRÖNROOS, C. (1984): "A Service Quality Model and its Marketing Implications", *European Journal of Marketing*, vol. 18, núm. 4, pp. 36-44.
- HAIR, J.F.; ANDERSON, R.E.; TATHAM, R.L.; BLACK, W.C. (1999): *Análisis multivariante*. Madrid: Prentice Hall.
- HEREDIA, J.A. (2001): *Sistema de indicadores para la mejora y el control integrado de la calidad de los procesos*. Castellón de la Plana: Universitat Jaume I.
- HUNT, H.K. (1977). CS/D Overview and Future Research Direction", en H.K. Hunt [ed.]: *Conceptualization and Measurement of Customer Satisfaction and Disatisfaction*. Cambridge, MA: Marketing Science Institute.
- HURLEY, A.E.; SCANDURA, T.A.; SCHRIESHEIM, CH.A.; BRANNICK, M.T.; SEERS, A.; VANDENBERG, R.J.; WILLIAMS, L.J. (1997): "Exploratory and Confirmatory Factor Analysis: Guidelines, Issues, and Alternatives", *Journal of Organizational Behaviour*, vol. 18, núm. 6, pp. 667-683.
- JOHNSON, M.D.; HERMANN, A.; GUSTAFSSON, A. (2002): "Comparing Customers Satisfaction Across Industries and Countries", *Journal of Economic Psychology*, 23, pp. 749-769.
- JÖRESKOG, K.G.; SÖRBOM, D. (1996): "Structural equation modelling", *Workshop NORC Social Science Research Professional Development Training Sessions*. Chicago.
- KLINE, P. (1994): *An Easy Guide to Factor Analysis*. Newbury Park, CA: Sage.
- KNAPP, R.P. (1976): *Handbook for the Personal Orientation Inventory*. San Diego, CA: S.D. Edits.
- LEBLANC, G.; NGUYEN, N. (1997): "Searching for Excellence in Business Education: An Exploratory Study of Customers Impressions of Service Quality", *The International Journal of Educational Management*, vol. 11, núm. 2, pp. 72-79.
- LÉVY MANGIN, J.P. (2003): "Modelización y análisis con ecuaciones estructurales", en J.P. Lévy Mangin y J. Varela [comp.]: *Análisis multivariable para las Ciencias Sociales*, pp. 767-814. Madrid: Prentice.
- LONG, J.S. (1986): *Confirmatory Factor Analysis*. (Versión original de 1983). Beverly Hills, CA: Sage.

- LÓPEZ, J.; GADEA, A. (1995): *Servir al ciudadano. Gestión de la calidad en la Administración Pública*. Barcelona: Gestión 2000.
- MANCEBÓN, M.J.; MARTÍNEZ CARBALLO, N.; PÉREZ XIMÉNEZ, J. (2007): "Un análisis de la calidad percibida por los estudiantes en los centros públicos y privados de enseñanza secundaria", *XVI Jornadas de Economía de la Educación*. Gran Canaria.
- MANO, H.; OLIVER, R.L. (1993). "Assessing the Dimensionality and Structure of the Consumption Experience: Evaluation, Feeling and Satisfaction". *Journal of Consumer Research*, vol. 20 (December), pp. 451-466.
- MARTÍNEZ TUR, V.; PEIRÓ, J.M.; RAMOS, J. (2001): *Calidad del servicio y satisfacción del cliente: una perspectiva psicosocial*. Madrid: Síntesis.
- MARZO, M.; MUNTANÉ, S. (2002): "Herramientas para la evaluación de la calidad", *Educación Social*, núm. 21 (mayo), pp. 115-130.
- MELS, G.; BOSHOFF, C.; NEL, D. (1997): "The Dimensions of Service Quality: The Original European Perspective Revisited", *Services Industries Journal*, vol. 17, núm. 1, pp. 173-189.
- MILLER, D.C. (1977): *Handbook of Research Design*. New York: David Mc Kay.
- MORALES, P.; UROSA, B.; BLANCO, A. (2003): *Construcción de escalas de tipo Likert*. Madrid: La Muralla/Hespérides.
- MUÑOZ, A. (1999): *La gestión de la calidad total en la administración pública*. Madrid: Díaz de Santos.
- NUNNALLY, J.; BERSTEIN, I.H. (1994): *Psychometric Theory*. New York: McGraw Hill.
- NUNNALLY, J.C. (1978): *Psychometric Theory*. New York: McGraw-Hill.
- OLFIELD, B.; BARON, S. (2000): "Student Perceptions of Service Quality in a UK University Business and Management Faculty", *Quality Assurance in Education*, vol. 8, núm. 2, pp. 85-95
- OWLIA, M.S.; ASPINWALL, E.M. (1996): "A Framework for the Dimensions of Quality in Higher Education", *Quality Assurance in Education*, vol. 4, núm. 2, pp. 12-20
- PALACIOS GÓMEZ, J.L. (1999): *Técnicas de investigación social para servicios socioculturales*. Granada: CEMCI.
- PARASURAMAN, A.; ZEITHALM, V.A.; BERRY, L.L. (1988): "SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality", *Journal of Retailing*, vol. 64, núm. 1, pp. 12-40.
- PARASURAMAN, A.; ZEITHALM, V.A.; BERRY, L.L. (1991): "Refinement and Reassessment of the SERVQUAL Scale", *Journal of Retailing*, vol. 67, núm. 4, pp. 420-450.
- PARASURAMAN, A.; ZEITHALM, V.A.; BERRY, L.L. (1994): "Servperf versus Servqual: Reconciling Performance-Based and perceptions-Minus-Expectations Measurement or Service Quality", *Journal of Marketing*, vol. 58, núm.1, pp. 125-131.
- PÉREZ GIL, J.A.; CHACÓN, S.; MORENO, R. (2000): "Validez de constructo: el uso del análisis factorial exploratorio-confirmatorio para obtener evidencias de validez", *Psicothema*, vol. 12, núm. 2, supl., pp. 442-446.
- RIGOTTI, S.; PITT, L. (1992): "Servqual as a Measuring Instrument for Service Provider Gaps in Business Schools", *Management Research News*, vol. 15, núm. 3, pp. 9-17.
- ROJAS, A.J.; FERNÁNDEZ, J.S.; PÉREZ, C. (1998): *Investigar mediante encuestas. Fundamentos teóricos y aspectos prácticos*. Madrid: Síntesis.
- SALVADOR GONZÁLEZ, X. DE (2003): "Calidad en servicios educativos: análisis del instrumento SERVQUAL", *Revista Española de Pedagogía*, vol. 61, núm. 224, pp. 5-21.
- SENLE, A.; VILAR, J. (1996): *ISO 9000 en empresas de servicios*. Barcelona: Gestión 2000.
- SWAN, J.E.; TRAWICK, I.F.; CARROLL, S. (1982). "Satisfaction Related to Predictive, Desired Expectations: A Field Study", en H.K. Hunt y R.L. Day [ed.]: *New Findings on Consumer Satisfaction and Complaining*, pp. 15-22. Bloomington, IN: Indiana University Press.
- TABACHNICK, B.G.; FIDELL, L.S. (1989): *Using Multivariate Statistics*. New York: Harper.
- TEAS, R.K. (1993): "Expectations, Performance, Evaluation, and Consumers' Perceptions of Quality", *Journal of Marketing*, vol. 57, núm. 4, pp. 18-34.
- TUKEY, J.W. (1977): *Exploratory Data Analysis*. Reading: Addison-Wesley.
- VILLEGAS, M.M.; DÍAZ, I.M.R. (2003): "La calidad asistencial: concepto y medida", *Dirección y Organización*, 29, pp. 50-58.
- WALBRIGDE, S.; DELENE, L. (1993): "Measuring Physician Attitudes on Service Quality", *Journal of Health Care Marketing*, (Winter), pp. 6-15.
- WERTS, C.E.; LINN, R.L.; JORESKÖG, K.G. (1974): "Interclass Reliability Estimates: Testing Structural Assumptions", *Educational and Psychological Measurement*, 34, pp. 25-33.
- YELA, M. (1997): *La técnica del análisis factorial*. Madrid: Biblioteca Nueva.