

Propiedades métricas de las escalas de competencia profesional y trato del personal sanitario de servicios de consultas externas hospitalarias

Rosa M^a Bermejo Alegría¹, M^a Dolores Hidalgo Montesinos¹, Pedro Parra Hidalgo², Adelia Más Castillo²
y Rafael Gomis Cebrián²

¹ Universidad de Murcia, ² Consejería de Sanidad y Consumo de la Región de Murcia

El objetivo de este trabajo es el estudio de las propiedades psicométricas de dos escalas de valoración de la calidad percibida y satisfacción de usuarios de consultas externas de Servicios de Salud. En concreto estas escalas evalúan la Competencia Profesional (CP) y Trato del Personal (TP). Se trabajó con una muestra de 2.017 sujetos usuarios de consultas externas del Servicio Murciano de Salud durante los años 2008 y 2009. Los resultados muestran una consistencia interna de 0,68 para la escala CP y de 0,75 para la escala TP. Los resultados del análisis factorial exploratorio indican que ambas medidas son unidimensionales. Las puntuaciones obtenidas en estas escalas correlacionaron positivamente con la puntuación en satisfacción global de los pacientes. Las escalas han mostrado su utilidad para detectar áreas de mejora y planificar estrategias de intervención.

Psychometric properties of scales of professional competence and treatment by health personnel in outpatient hospital surgeries. The aim of this study was to analyze the psychometric properties of two scales that assess the perceived quality and patient satisfaction with outpatient surgery in the Health Service of Murcia. These scales assess the degree of Professional Competence (PC) and Personnel Treatment (PT). The scales were administered to a sample of 2017 users of outpatient surgery in the Health Service of Murcia during the years 2008 and 2009. Exploratory factor analysis indicates a unidimensional structure for each scale. Internal consistency was adequate: .68 for PC and .75 for PT. The correlation between the PC scale and patients' global satisfaction was positive and statistically significant. The correlation between the PT scale and patients' global satisfaction was also statistically significant. The scales have shown their utility to detect areas of improvement and to plan intervention strategies.

Los estudios para conocer el punto de vista de los pacientes sobre la calidad de un servicio sanitario son cada vez más numerosos, lo que no es de extrañar puesto que este aspecto influye en la fidelidad de los usuarios (Olsen, 2002). Ésta es la razón de que los responsables de dichas organizaciones se preocupen por saber qué aspectos de la prestación del servicio son determinantes para obtener buenas valoraciones, y qué instrumentos de medida aseguran que evalúan estos aspectos con garantía.

Para determinar los elementos que componen la calidad es necesario estudiar la dimensionalidad de este constructo. Así, Grönroos (1984) nos habla de dos dimensiones básicas: *calidad técnica* (referida a los aspectos tangibles del servicio) y *calidad funcional* (relacionada con la forma en que se recibe el servicio). Parasuraman, Zeithaml y Berry (1988) referencian cinco dimensiones: fiabilidad, capacidad de respuesta, seguridad, empatía y tangibles. No obstante, cuatro de estas dimensiones reflejan calidad funcional. Price, Arnould y Tierney (1995) destacan los aspectos de rela-

ción que devienen de la expresión de emociones que los empleados desarrollan (Dube y Menon, 1998). Aunando estas líneas de investigación, Peiró, Martínez-Tur y Ramos (2005) confirmaron la existencia de dos grandes factores en el estudio de la calidad del servicio. El primero continúa con la denominación de *calidad funcional*, pero concreta su definición haciendo referencia a la eficiencia con que se atiende a los usuarios en los aspectos centrales del servicio; y el segundo, llamado *calidad relacional*, da cuenta de la obtención, por parte del usuario, de beneficios emocionales (Mañas, Giménez, Muyor, Martínez-Tur y Moliner, 2008).

Aunque se han creado algunos instrumentos para valorar la calidad asistencial, con propiedades métricas buenas o aceptables (Aletras, Papadopoulos y Niakas, 2006; Feldman, Vivas, Lugli, Alviarez, Pérez y Bustamante, 2007; González, Quintana, Bilbao, Escobar, Aizpuru, Thompson et al., 2005; Granada, Rodríguez, Olmedo, Chacón, Vigil y Rodríguez, 2007; Mallou, Rial y García, 2003; Mira, Rodríguez-Marín, Buil, Castell y Vitaller, 1998; Rodríguez-Fernández, Ramon, Montserrat-Capella, Sain, Cebrián, García-Artiaga et al., 2008), existen lagunas en el área de la evaluación, puesto que no se han generado pruebas en base a la evolución conceptual del constructo calidad asistencial.

Este trabajo pretende dar un paso en este aspecto, presentando dos escalas de *screening* de calidad funcional y relacional en el ámbito de las consultas externas hospitalarias. Se espera que resulten

unidimensionales, con niveles adecuados de consistencia interna, con evidencias de validez convergente respecto a la satisfacción del usuario y discriminante en referencia a la fidelidad del mismo.

Método

Participantes

Los sujetos de este estudio eran pacientes atendidos en consultas externas de los hospitales generales del Servicio Murciano de Salud. Se seleccionaron a partir del listado de los pacientes que asistieron a consulta durante el período del 17 al 31 de enero de 2008 y del 19 de mayo al 1 de junio de 2009. Se realizó un muestreo sistemático con arranque aleatorio hasta completar 40 encuestas diarias por hospital. La selección fue estratificada, asegurando que todos los hospitales tuvieran la misma representatividad. Se excluyó a los pacientes seleccionados inicialmente cuando: tras dos llamadas a distintas horas del día no descuelgan o la persona indicada no está, el paciente rechaza ser entrevistado, el número de teléfono no corresponde a la persona especificada, corresponde a fax o contestador, el paciente no asistió a consulta o existía registro repetido.

El número de encuestas cumplimentadas fue de 1.800 en cada uno de los dos años del estudio. Con el fin de homogeneizar la muestra para el análisis de las propiedades métricas se eliminaron los casos con datos ausentes y los encuestados con edades inferiores a los 18 años.

La muestra final utilizada fue de 2.017 sujetos. De ellos, un 13,3% no fue el propio paciente sino su acompañante el respondiente. Un 39% (787) fueron hombres y un 61% (1.230) mujeres, la edad media fue de 55,08 [desviación típica (DT)= 18,06]. Un 27,9% informaron que no tenían estudios; un 29,9% completaron sus estudios primarios; un 17,6% bachiller elemental; un 8,9% realizaron estudios de formación profesional (primer o segundo grado); un 7,9% cursaron estudios de bachiller superior o BUP; un 4,1% eran diplomados, arquitectos o ingenieros técnicos; y un 3,7% licenciados universitarios, arquitectos o ingenieros superiores. La tabla 1 presenta la distribución de frecuencias conjuntas en función del

	18-44 años	45-64 años	>= 65 años	Total
Hombres	207 (10,3%)	257 (12,7%)	323 (16,0%)	787 (39,0%)
Mujeres	415 (20,6%)	419 (20,8%)	396 (19,6%)	1.230 (60,0%)
Total	622 (30,9%)	676 (33,5%)	719 (35,6%)	2.017 (100%)

	Nivel de estudios 1	Nivel de estudios 2	Nivel de estudios 3	Total
Hombres	451 (22,7%)	257 (12,9%)	66 (3,3%)	787 (39,0%)
Mujeres	701 (35,2%)	427 (21,4%)	89 (4,5%)	1.230 (60,0%)
Total	1.152 (57,9%)	684 (34,4%)	155 (7,8%)	2.017 (100%)

Nota: Nivel de estudios 1= Sin estudios/estudios primarios; Nivel de estudios 2= Bachiller/FP; Nivel de estudios 3= Diplomado/licenciado

sexo y la edad agrupada en tres categorías (18 a 44, 45 a 64 y 65 o más años). La tabla 2 indica la distribución conjunta de frecuencias de las variables sexo y nivel de estudios. La información de esta última variable se presenta agrupada en tres niveles (Nivel 1: sin estudios y estudios primarios; Nivel 2: bachiller elemental, formación profesional y bachiller superior; y Nivel 3: diplomado y licenciado). Los resultados de la tabla 2 muestran que un 57% de los encuestados no tienen estudios o solamente han realizado estudios primarios.

Instrumentos

Para la construcción de estas escalas se utilizó información obtenida a través de la realización de grupos focales sobre satisfacción con la hospitalización en pacientes del Servicio Murciano de Salud, así como la revisión de cuestionarios utilizados en otros servicios regionales de salud.

Se conceptualizó la competencia profesional como: percepción que refiere el entrevistado referida a la capacidad de los profesionales para atender su problema de salud, y trato como: los aspectos de la relación entre pacientes y profesionales referidos a la amabilidad, apoyo emocional, intimidad y respeto.

Las tablas 3 y 4 detallan los ítems y sus formatos de respuesta.

La satisfacción fue valorada globalmente con una escala de 0 a 10 y con otro ítem sobre si el paciente, en caso de poder elegir, volvería a la misma consulta.

Procedimiento

La aplicación de las escalas, que formaban parte de una encuesta más amplia, se realizó por una empresa especializada en entrevistar.

Los pacientes seleccionados eran encuestados antes de que transcurrieran 48 horas desde su visita a la consulta correspondiente,

Ítem	Enunciado	Respuesta
1	¿Cree usted que su médico se preocupó por aclararle todas las dudas sobre su enfermedad?	Sí No
2	En el momento de reconocerle (explorarle) ¿sintió que se respetaba su intimidad?	Sí No
3	¿Cree que el médico le ha dedicado el tiempo que usted necesitaba?	Sí No
4	¿Cómo valora usted la profesionalidad del personal que le atendió? Médicos	Muy buena Buena Regular Mala Muy mala
5	¿Cómo valora usted la profesionalidad del personal que le atendió? Personal de enfermería	Muy buena Buena Regular Mala Muy mala
6	¿Cómo valora usted la profesionalidad del personal que le atendió? Personal administrativo	Muy buena Buena Regular Mala Muy mala

con el fin de que lo ocurrido fuera recordado con la mayor exactitud, minimizando así, en lo posible, el sesgo del recuerdo de una vivencia, generalmente, poco impactante (Fowler, 1993).

La administración del instrumento fue telefónica y se llevó a cabo durante 5 días de dos semanas de 2008 y dos de 2009. Cuando los pacientes seleccionados tenían problemas de salud, que les impedían contestar la encuesta, ésta se realizaba al acompañante.

Análisis de datos

El tratamiento estadístico de los datos se ha realizado mediante el programa informático SPSS 15.0 (SPSS, 2006). Se obtuvieron las propiedades métricas de la Escala de Competencia Profesional (ECP) y de la Escala de Trato del Personal Sanitario (ETP). Se realizó un análisis clásico de los ítems y un análisis de la fiabilidad mediante el coeficiente de consistencia interna de Cronbach (1951). Para el estudio de la dimensionalidad se ha utilizado como técnica el Análisis Factorial Exploratorio (AFE) de mínimos cuadrados no ponderados. La solución factorial fue determinada a partir del procedimiento «scree-test» y considerando el porcentaje de varianza explicado por los factores. Además se realizó un Análisis Factorial Confirmatorio (AFC) utilizando el programa LISREL, versión 8.20 (Jöreskog y Sörbom, 1999). Se empleó el método de estimación de mínimos cuadrados. Siguiendo los criterios de Hu y Bentler (1999) y Dimitrov (2006), se seleccionaron los siguientes índices de ajuste: GFI (Goodness of Fit Index; Índice de Bondad de Ajuste) y CFI (Comparative Fit Index; Índice de Ajuste Comparativo) >0,95 y RMSA (Root Mean Square Approximation; Error de Aproximación Cuadrático Medio) <0,06 (buen ajuste). Como estrategia para obtener evidencias de validez externa se calculó la correlación lineal entre las escalas (ítems y puntuación total) y el grado de satisfacción general informado por el usuario. Cuando utilizamos una medida de correlación empleando un tamaño muestral elevado se puede encontrar un resultado estadísticamente significativo aun cuando el tamaño de la relación sea pequeño, por lo que los resultados obtenidos se interpretaron siguiendo los criterios de Cohen (1988) para el tamaño de la relación entre variables. Así, se considera un tamaño de la relación pequeño cuando ésta es menor o igual

que 0,10, medio cuando está en torno a 0,30 y grande cuando es igual o mayor que 0,50 (Cohen, 1988; Lipsey y Wilson, 2001). Como otra evidencia de validez externa se procedió a clasificar a los encuestados en dos grupos en función de las respuestas a la pregunta «En caso de poder elegir, ¿volvería usted a esta consulta?». El primer grupo estuvo definido por los sujetos que habían contestado «Sí, volvería» y el segundo grupo por aquellos que eligieron una de las restantes opciones de respuesta. Se calcularon las diferencias entre medias, usando una prueba t de Student, en función de los grupos para cada una de las puntuaciones totales en las escalas. Por último, se realizaron ANOVAS para observar las posibles diferencias en valoración de la competencia y el trato del personal sanitario en función del sexo y el nivel de estudios. Se calculó el índice d como medida del tamaño del efecto para valorar la magnitud de las diferencias entre medias, se interpretó un tamaño del efecto bajo cuando $d \leq 0,50$, moderado si $0,51 \leq d \leq 0,79$ y alto si $d \geq 0,80$.

Resultados

Análisis de ítems y fiabilidad

La tabla 5 presenta el análisis clásico para cada uno de los ítems de la ECP y de la ETP. Se muestran los estadísticos descriptivos (media y desviación típica) y el índice de homogeneidad (correlación ítem-total y correlación ítem-total corregida). En los ítems de respuesta dicotómica (1, 2 y 3) más de un 90% de los sujetos dieron una respuesta positiva al ítem. Por otro lado, la media para los ítems de respuesta graduada se situó por encima del valor central de la escala, obteniéndose una valoración más alta en el ítem 4 y la más baja en el ítem 6. El rango de correlación ítem-total osciló entre 0,16 para el ítem 2 y 0,84 para el ítem 4.

Los ítems de la ETP presentaron promedios elevados, obteniendo la valoración más alta el ítem 1 (referido al trato del personal médico) y la más baja el ítem 3 (referido al trato del personal administrativo).

La fiabilidad del test, obtenida con el coeficiente alfa de Cronbach, obtuvo un valor de 0,68 para la ECP y de 0,75 para la ETP.

Tabla 4
Escala de trato del personal sanitario

Ítem	Enunciado	Respuesta
1	¿Cómo calificaría el trato humano que recibió de parte del personal que le atendió? Médicos	Muy bueno Bueno Regular Malo Muy malo
2	¿Cómo calificaría el trato humano que recibió de parte del personal que le atendió? Personal de enfermería	Muy bueno Bueno Regular Malo Muy malo
3	¿Cómo calificaría el trato humano que recibió de parte del personal que le atendió? Personal administrativo	Muy bueno Bueno Regular Malo Muy malo

Tabla 5
Análisis clásico de los ítems ECP y ETP

	Míni-mo	Máxi-mo	Media	Desviación típica	Correlación ítem-total	Correlación ítem-total corregida
ECP						
Ítem 1	0	1	0,92	0,27	0,49	0,36
Ítem 2	0	1	0,98	0,15	0,16	0,07
Ítem 3	0	1	0,94	0,24	0,47	0,36
Ítem 4	0	4	3,35	0,65	0,84	0,65
Ítem 5	0	4	3,29	0,61	0,81	0,63
Ítem 6	0	4	3,14	0,63	0,72	0,46
ETP						
Ítem 1	0	4	3,35	0,62	0,81	0,55
Ítem 2	0	4	3,29	0,59	0,87	0,69
Ítem 3	0	4	3,15	0,61	0,76	0,50

Análisis de la dimensionalidad

Se analizó la estructura factorial de cada una de las escalas empleando un procedimiento de validación cruzada. La idea esencial que subyace a este procedimiento es la de replicar los resultados obtenidos en un primer análisis para someterlos a una contrastación posterior. Así se dividió la muestra total en dos submuestras aleatorias ($N_1 = 992$ y $N_2 = 1.025$). Con la primera submuestra se efectuó el AFE, y con la segunda se realizó el AFC para los ítems de la ECP y un nuevo AFE para los de la ETP, teniendo en cuenta que para esta escala no es posible obtener índices de ajuste del modelo confirmatorio por el pequeño número de variables.

En la primera submuestra el índice de adecuación muestral de Kaiser-Meyer-Olkin (KMO) obtuvo un valor de 0,69 para los ítems de ECP y de 0,66 para los de ETP, indicando una correcta adecuación de los datos para efectuar el AFE. La prueba de esfericidad de Barlett resultó significativa tanto para la ECP ($\chi^2 = 1221,92$, $gl = 15$ y $p < 0,001$) como para la ETP ($\chi^2 = 895,56$, $gl = 3$ y $p < 0,001$).

Para los ítems de ECP se encontró que el primer factor explicó un 39,64% de la variación total (autovalor = 2,38), y para los tres siguientes factores los autovalores fueron de 1,24, 0,94 y 0,56. A la vista de los resultados se optó por una solución unifactorial. La tabla 6 presenta las saturaciones factoriales obtenidas. Con el fin de corroborar la estructura de un solo factor se realizó un AFC a partir de la segunda submuestra. Los resultados obtenidos no indican un buen ajuste del modelo ($\chi^2 = 1210,97$, $gl = 9$, $p < 0,01$; RMSEA > 0,10; GFI = 0,96; CFI = 0,95).

Por otro lado, el AFE de los ítems de la ETP, realizado en la primera submuestra, mostró que el primer factor explicaba un 69,46% de la varianza total. Al analizar la segunda submuestra, el porcentaje de varianza explicado por el primer factor fue similar (65,14%). En la primera submuestra las cargas factoriales fueron de 0,714, 0,905 y 0,601 para los ítems 1, 2 y 3. Estos resultados se mantuvieron en la segunda submuestra.

Por último, la correlación entre las puntuaciones observadas de ambas escalas resultó ser de 0,77.

Validez convergente

Para obtener evidencias de validez convergente se pidió a los encuestados que valoraran de forma global el grado de satisfacción con la consulta. Se calcularon las correlaciones entre cada uno de los ítems de la escala ECP y la puntuación total con las respuestas que los sujetos dieron a esta pregunta. Los resultados obtenidos se presentan en la tabla 7. Las correlaciones fueron de tamaño medio y medio alto, excepto para el ítem 3 y para el ítem 2, que el tamaño

Ítem	Cargas factoriales
1	0,415
2	0,112
3	0,426
4	0,817
5	0,703
6	0,535

de la relación, aunque estadísticamente significativo, fue pequeño, y para el total de la escala que fue de un tamaño del efecto alto.

En cuanto a los ítems de la ETP se encontró que todas las correlaciones fueron mayores de 0,30. Además, se obtuvo una relación moderada-alta entre la puntuación total en esta escala y la valoración de la satisfacción general con la consulta.

Estos resultados fueron similares en cada uno de los años evaluados.

Validez discriminante

Se obtuvieron evidencias de validez discriminante analizando las puntuaciones obtenidas en la ECP y la ETP con los resultados de la pregunta *En caso de poder elegir, ¿volvería usted a esta consulta?*

Se compararon las medias en la ECP y la ETP de los encuestados que volverían a la consulta con la de aquellos que probablemente volverían o no volverían. Se encontraron diferencias estadísticamente significativas en la ECP ($t = 15,40$; $gl = 240,64$; $p < 0,01$), donde los sujetos que sí volverían a la consulta puntuaron en promedio 2,32 puntos por encima de los que no volverían (tabla 8), siendo el tamaño de las diferencias alto ($d = 1,46$).

En cuanto a las puntuaciones en la ETP también se encontraron diferencias estadísticamente significativas ($t = 14,78$; $gl = 2015$; $p < 0,01$). De nuevo los sujetos que sí volverían a la consulta puntuaron en promedio más alto, aunque el tamaño de estas diferencias fue moderado ($d = 0,79$).

Diferencias de sexo y nivel de estudios en valoración de la competencia profesional y el trato del personal sanitario

Se estudiaron las posibles diferencias en valoración de la competencia profesional y el trato del personal sanitario en función del sexo y el nivel de estudios. Se realizó un ANOVA 2 (sexo) \times 3 (nivel de estudios) para cada una de las escalas. En la ECP no se encontraron diferencias estadísticamente significativas ni en fun-

ECP	Satisfacción global		
	2008-2009 (N= 2017)	2008 (N= 1030)	2009 (N= 987)
Ítem 1	0,308	0,303	0,314
Ítem 2	0,121	0,130	0,120
Ítem 3	0,284	0,296	0,272
Ítem 4	0,441	0,465	0,420
Ítem 5	0,400	0,411	0,389
Ítem 6	0,324	0,311	0,337
Total escala	0,499	0,530	0,466
ETP			
Ítem 1	0,427	0,414	0,442
Ítem 2	0,405	0,388	0,423
Ítem 3	0,345	0,328	0,373
Total escala	0,475	0,465	0,487

Nota: todas las correlaciones son significativas $p < 0,01$

ción del sexo ($F(1,1991)= 0,234, p= 0,629$) ni el nivel de estudios ($F(2, 1991)= 1,737; p= 0,176$).

En la ETP se encontraron diferencias estadísticamente significativas en función del nivel de estudios (NE) ($F(2,1991)= 3,014; p= 0,049$), puntuando ligeramente más alto los sujetos con niveles de estudios superiores (tabla 9); sin embargo, estas diferencias fueron de un tamaño del efecto bajo ($d(NE1 vs NE2)= 0,05; d(NE1 vs NE3)= -0,17; d(NE2 vs NE3)= -0,21$). No se encontraron diferencias significativas en función de la variable sexo ($F(1,1991)= 0,674, p= 0,412$).

Diferencias en competencia profesional y trato del personal sanitario en función del año evaluado

Para conocer si la valoración en competencia profesional y trato del personal sanitario de los Servicios de Consultas Externas cambió a través de los años evaluados, se realizó una prueba t de Student de comparación entre medias. No se encontraron diferencias estadísticamente significativas ni en la ECP ($t= -0,704, gl= 2015, p= 0,485$) (año 2008: $M= 11,64$ y $DT= 1,71$; año 2009: $M= 11,66$ $DT= 1,77$) ni en la de ETP ($t= -0,295, gl= 2015, p= 0,768$). Así, en el año 2008 la media fue de 9,76 ($DT=1,49$) y en el año 2009 fue de 9,81 ($DT= 1,47$).

Tabla 8
Estadísticos descriptivos para la ECP y la ETP en función de si volverían o no a la consulta

Escala de competencia profesional			
	N	Media	Desviación típica
Volver a la consulta			
Sí	1.801	11,90	1,50
No	216	9,58	2,15
Escala de trato del personal			
	N	Media	Desviación típica
Volver a la consulta			
Sí	1.801	9,94	1,38
No	216	8,84	1,56

Tabla 9
Medias (y desviación típica) para la ECP y la ETP en función del nivel de estudios y sexo

	Nivel de estudios 1	Nivel de estudios 2	Nivel de estudios 3	Total
ECP				
Hombres	11,60 (1,63)	11,63 (1,73)	11,83 (2,09)	11,63 (1,70)
Mujeres	11,67 (1,70)	11,59 (1,80)	11,97 (1,95)	11,67 (1,76)
Total	11,65 (1,68)	11,61 (1,77)	11,91 (2,01)	11,65 (1,74)
ETP				
Hombres	9,76 (1,38)	9,79 (1,49)	10,15 (1,46)	9,80 (1,46)
Mujeres	9,82 (1,46)	9,68 (1,51)	9,97 (1,61)	9,78 (1,49)
Total	9,80 (1,43)	9,72 (1,50)	10,05 (1,71)	9,79 (1,48)

Nota: Nivel de estudios 1= Sin estudios/estudios primarios; Nivel de estudios 2= Bachiller/FP; Nivel de estudios 3= Diplomado/licenciado

Discusión y conclusiones

Este trabajo pone a prueba dos escalas de *screening* sobre la percepción de los usuarios de la competencia y el trato que ofrecen los profesionales en el ámbito de las consultas externas hospitalarias, estudiando su dimensionalidad, consistencia interna, validez convergente y discriminante.

Respecto a la dimensionalidad, si revisamos la bibliografía sobre las escalas de calidad y satisfacción que presentan propiedades métricas, observamos que el proceso de validación comienza con la realización de un análisis factorial de todos los indicadores. Dicho procedimiento de análisis es plenamente justificable debido a la variabilidad de la definición de calidad, pero no contribuye al esclarecimiento de la misma. En este sentido, este estudio constituye un avance, ya que parte de un marco teórico que diferencia calidad funcional y relacional. No obstante, la asunción de una conceptualización en calidad entraña riesgos, ya que cada tipo de organización o servicio tiene características propias que repercuten en los procesos y resultados que se generan de los clientes (Cunningham, Young y Lee, 1997; Mañas, Jiménez, Mayor, Martínez-Tur y Moliner, 2008). Por ello se relacionaron ambas escalas, obteniendo una alta correlación. Esto puede deberse a que los usuarios de consultas externas perciban el trato como una parte del desarrollo profesional del sanitario, lo cual concuerda con estudios sobre la competencia profesional realizados en otras áreas, como la educación universitaria, donde los indicadores de trato forman parte de escalas unidimensionales sobre la competencia (Muñiz, García y Virgos, 1991). Para estudiar más detalladamente este hecho podría resultar de interés analizar los ítems de las escalas en su conjunto, observando si se distribuyen, como se espera, en las dos variables propuestas o conforman una única dimensión.

En cuanto a la consistencia interna, los valores de nuestras escalas son moderados, frente a los elevados resultados de los estudios de Granado et al. (2007) y Mira et al. (1998). Este hecho se puede explicar atendiendo al número de ítems. Nuestras escalas tienen un escaso número de ítems en base a su propósito de medida, que es la detección y comparación de áreas problema en los diferentes hospitales de la comunidad, objetivo que creemos alcanzado. Evidentemente, con 3 o 6 ítems no pretendemos dar una información precisa sobre la competencia y el trato del personal sanitario de consultas externas. Para ello, se deberían de elaborar escalas con más ítems que muestrearan el universo de indicadores que conforman la descripción de estos procesos.

En cuanto al análisis de los ítems, en la ECP el ítem 2 hace referencia a una actividad médica que puede no realizarse en algunas visitas del paciente. Sin embargo, una redacción alternativa podría mejorar la precisión y validez interna de esta escala, planteando una pregunta genérica sobre si se respeta la intimidad detallando ejemplos.

En referencia a la validez, los instrumentos muestran adecuada validez convergente, corroborada con el ítem criterio sobre satisfacción. Para valorar la aportación de estos factores en la predicción de la satisfacción en los servicios de consultas externas resultaría de interés realizar un *screening* de los tangibles, ya que puede existir un gran número de usuarios que conozcan y utilicen las instalaciones con asiduidad, siendo sensibles no solo a las condiciones funcionales y relacionales, sino también a las condiciones físico-técnicas de la organización (Bitner, 1990; Dubé y Menon, 1998). El análisis de validez discriminante obtiene resultados coin-

cidentes con el estudio de Cabana y Jee (2004). Aun así, consideramos necesario realizar otros estudios de validez que permitan ampliar el conocimiento respecto al funcionamiento de estas escalas, tales como la correlación con otras escalas de calidad asistencial o con indicadores objetivos de competencia profesional.

En definitiva, y a pesar de las limitaciones presentes, la actual presentación de estas escalas es útil para detectar si existen áreas susceptibles de mejora y comparar hospitales. Asimismo, este estudio plantea la realización de futuras investigaciones que arrojen algo más de luz sobre la calidad de los servicios.

Referencias

- Aletras, V.H., Papadopoulos, E.A., y Niakas, D.A. (2006). Development and preliminary validation of a Greek-language outpatient satisfaction questionnaire with principal components and multi-trait analyses. *BMC Health Services Research*, 6(66), 1-11.
- Bitner, M.J. (1990). Evaluation service encounters: The effects of physical surroundings and employee responses. *Journal of Marketing*, 54, 69-82.
- Cabana, M.D., y Jee, S.H. (2004). Does continuity of care improve patient outcomes? *The Journal of Family Practice*, 53(12), 974-980.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Erlbaum, Hillsdale, NJ.
- Cronbach, L. (1951). Coefficient alpha and the internal consistency of tests. *Psychometrika*, 16, 297-334.
- Cunningham, L.F., Young, C.E., y Lee, M. (1997). A customer-based taxonomy of services: Implications for service marketers. En T.A. Swartz, D.E. Bowen y D. Iacobucci (Eds.), *Advances in services marketing and management* (vol. 6, pp. 189-202). Greenwich: JAI Press.
- Dimitrov, D. (2006). Comparing groups on latent variables: A structural equation modeling approach. *Work*, 26, 429-436.
- Dubé, L., y Menon, K. (1998). Why would certain types of in-process negative emotions increase post-purchase consumer satisfaction with services? En T.A. Swartz, D.E. Bowen y S. Brown (Eds.), *Advances in service marketing and management* (vol. 7, pp. 131-158). Greenwich: JAI Press.
- Feldman, L., Vivas, E., Lugli, Z., Alvarez, V., Pérez, M.G., y Bustamante, S. (2007). La satisfacción del paciente hospitalario: una propuesta de evaluación. *Revista de Calidad Asistencial*, 3, 133-140.
- Fowler, F.J. (1993). *Survey research methods*. California: Sage Publications (pp. 87-88).
- González, N., Quintana, J.M., Bilbao, A., Escobar, A., Aizpuru, F., Thompson, A., Esteban, C., San Sebastián, J.A., y De la Sierra, E. (2005). Development and validation of an in-patient satisfaction questionnaire. *International Journal for Quality in Health Care*, 17(6), 465-472.
- Granado, S., Rodríguez, C., Olmedo, M.C., Chacón, A., Vigil, D., y Rodríguez, P. (2007). Diseño y validación de un cuestionario para evaluar la satisfacción de los pacientes atendidos en las consultas externas de un hospital de Madrid en 2006. *Revista Española de Salud Pública*, 81, 637-645.
- Grönroos, C. (1984). A service quality model and its marketing implications. *European Journal of Marketing*, 18, 36-44.
- Hu, L., y Bentler, P.M. (1999). Cut-off criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- Jöreskog K.G., y Sörbom D. (1999). *LISREL 8.20*. Scientific Software International, Chicago, IL.
- Lipsey, M., y Wilson, D. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Mallou, Rial y García (2003). Presentación de una escala de satisfacción con los servicios sanitarios en atención primaria. *Psicothema*, 15(4), 656-661.
- Mañas, M.A., Jiménez, G., Mayor, J.M., Martínez-Tur, V., y Moliner, C.P. (2008). Los tangibles como predictores de la satisfacción del usuario en servicios deportivos. *Psicothema*, 20(2), 243-248.
- Mira, J.J., Aranaz, J., Rodríguez-Marín, J., Buil, J.A., Castell, M., y Vítaller, J. (1998). SERVQHOS: un cuestionario para evaluar la calidad percibida de la asistencia hospitalaria. *Medicina Preventiva*, IV, 12-18.
- Muñiz, J., García, A., y Virgos, J.M. (1991). Escala de la Universidad de Oviedo para la evaluación del profesorado. *Psicothema*, 3(2), 269-281.
- Olsen, S.O. (2002). Comparative evaluation of the relationship between quality, satisfaction and repurchase loyalty. *Journal of the Academy of Marketing Science*, 30, 240-249.
- Parasuraman, A., Zeithaml, V., y Berry, L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, 64, 12-40.
- Peiró, J.M., Martínez-Tur, V., y Ramos, J. (2005). Employees' overestimation of functional and relational service quality: A gap analysis. *The Service Industries Journal*, 25, 1-17.
- Price, L.L., Arnould, E.J., y Tierney, P. (1995). Going to extremes: Managing service encounters and assessing provider performance. *Journal of Marketing*, 59, 83-97.
- Rodríguez-Fernández, J., Ramón Martínez, J., Montserrat-Capella, D., Sainz, A., Cebrián, M., García-Artiaga, C., et al. (2008). Utilización de la opinión de los pacientes para la mejora: 5 años de seguimiento (2001-2005). *Revista de Calidad Asistencial*, 23(2), 57-64.
- SPSS (2006). *Statistical Package for Social Sciences for Windows*. Version 15. Chicago, IL: SPSS Inc.