

Revista de Psicología del Deporte
2010. Vol. 19, núm. 1, pp. 41-58
ISSN: 1132-239X

Universitat de les Illes Balears
Universitat Autònoma de Barcelona

LA MEDICIÓN DE LA SATISFACCIÓN DEL CONSUMIDOR DE SERVICIOS DEPORTIVOS A TRAVÉS DE LA LÓGICA BORROSA

José Antonio Martínez García y Laura Martínez Caro*

CUSTOMER SATISFACTION WITH SPORTS SERVICES: A FUZZY LOGIC APPROACH

KEYWORDS: Satisfaction, Sports services, Fuzzy logic.

ABSTRACT: This study focuses on measuring customer satisfaction within the context of sports services by using a novel approach in sport psychology: modelling uncertainty through the application of fuzzy logic. Several analytical procedures for dealing with fuzziness have been described in order to verify empirically the extent to which this methodology complements the traditional statistical approach. This study examines the relationships between numerical evaluations of customer satisfaction and the degree of uncertainty linked to these verbal judgements. The proposed approach enriches the information on the attitudes of sport service consumers and overcomes the shortcomings of the “third-person” perspective. In addition, this proposal is robust against two sources of bias: scale coarseness, and interaction between numerical and verbal judgements.

Correspondencia: José Antonio Martínez García. Departamento de Economía de la Empresa. Universidad Politécnica de Cartagena. Facultad de Ciencias de la Empresa. Paseo Alfonso XIII, 50. 30203, Cartagena. España.

E-mail: josean.martinez@upct.es

* Departamento de Economía de la Empresa. Universidad Politécnica de Cartagena

— Fecha de recepción: 10 de Septiembre de 2008. Fecha de aceptación: 8 de Agosto de 2009.

Martínez García, J. A. y Martínez Caro, L.

La medición de la satisfacción del consumidor de servicios...

El estudio de la satisfacción del consumidor es uno de los temas más tratados en la literatura científica en las disciplinas hermanas de marketing y psicología de consumo. La importancia de su estudio se fundamenta en dos razones: (1) su medición sirve como aproximación para calibrar si las organizaciones están actuando acorde a sus objetivos; (2) su conocimiento permite elaborar predicciones sobre el comportamiento futuro del consumidor.

El ámbito del deporte no ha sido ajeno a la proliferación de estudios sobre este tema, con numerosas aportaciones en revistas de psicología general (ej. Mañas, Giménez, Muyor, Martínez-Tur y Moliner, 2008), psicología del deporte (ej. Rodríguez y Barriopedro, 2003), gestión deportiva (ej. Madrigal y Chen, 2008), o marketing deportivo (ej. Greenwell, Brownlee, Jordan, y Popp, 2008).

Aunque ésta es un área de investigación madura, en los últimos años se han realizado interesantes contribuciones a través de la incorporación de metodologías menos convencionales para su análisis, que dan un nuevo enfoque a los estudios sobre satisfacción. De esta forma, métodos asociados a las ciencias de la complejidad y a la inteligencia artificial, se están comenzando a utilizar. Como ejemplos, podemos citar los estudios de Martínez y Martínez (2009), Gondholt y Martensen (2005) o Lozano y Fuentes (2003). En el primer de ellos, se aplica la dinámica de sistemas para investigar las relaciones entre las actitudes del consumidor y su comportamiento efectivo. En el segundo, los autores apuestan por el uso de las redes neuronales artificiales en lugar de los tradicionales análisis estadísticos para estudiar las relaciones entre la satisfacción del consumidor, sus antecedentes y consecuencias. Finalmente, el último de ellos aplica la lógica borrosa para el estudio de las actitudes de los consumidores

como aproximación a la valoración de intangibles.

En este estudio, utilizaremos una de las técnicas comentadas anteriormente, la lógica borrosa, para analizar el problema de la medición de la satisfacción de consumidores de servicios deportivos, con el fin de comprobar empíricamente en qué medida esta metodología puede complementar los resultados derivados de los métodos estadísticos tradicionales. De este modo, se profundiza sobre la relación entre las evaluaciones numéricas de la satisfacción de los consumidores y el grado de incertidumbre asociado a las evaluaciones lingüísticas. El carácter único de este estudio reside en la derivación empírica de las funciones de pertenencia borrosas, diferenciándose de esta forma de la perspectiva de Lozano y Fuentes (2003) o Tsai, Wu y Liang (2006). La replicación de la investigación en una segunda muestra de consumidores, da robustez a los resultados empíricos.

Marco conceptual

La satisfacción del consumidor es considerada como una actitud; una evaluación subjetiva sobre las experiencias de consumo. Así, la satisfacción es entendida como un juicio evaluativo que el consumidor realiza de forma global sobre su experiencia de consumo en un momento posterior al consumo (Fornell, 1992), basado parte en aspectos cognitivos, y parte en la respuesta afectiva a estímulos del producto o servicio (Oliver, 1997). Existen excelentes investigaciones que profundizan ampliamente en la naturaleza de este concepto, sus antecedentes y consecuencias, y los paradigmas asociados a su definición (ej. Anderson y Sullivan, 1993; Fournier y Mick, 1999; Seiders, Voss, Grewal y Godfrey, 2005).

La medición de la satisfacción del consumidor, por su parte, no ha tenido un papel

tan protagonista en la literatura, principalmente porque el debate científico se ha centrado en la medición de actitudes en general. Desde que Likert (1932) presentó su propuesta para medir actitudes, este método ha sido ampliamente utilizado en la investigación social. Sin embargo, se ha discutido apasionadamente acerca de cuantiosos temas relacionados con la implementación de esa medición, por ejemplo, acerca del número adecuado de alternativas de respuesta, la elección de etiquetas verbales para esas alternativas, la fiabilidad y validez de escalas de diferente rango, el efecto de presentar una respuesta neutra, o la preferencia de los consumidores entre diferentes formatos de respuesta. Esa discusión, se ha generalizado a escalas tipo diferencial semántico y, en definitiva, a cualquier escala de medición que categorice o constriña las respuestas de los individuos a un determinado rango de valores. En este campo de investigación también existen magníficas referencias que repasan las aportaciones realizadas y que muestran las numerosas contradicciones que de los estudios empíricos se han derivado (ej. Cox, 1980; Hofmans, Theuns y Mairesse, 2007; Weng, 2004).

En este escenario de discusiones científicas, la mayoría de los autores que investigan sobre satisfacción del consumidor, ya sea por prudencia algunos y por rutina otros, utilizan generalizadamente las escalas Likert de 1 a 5 o de 1 a 7, apoyándose sobre todo en trabajos como los de Lissitz y Green (1975) o Cox (1980). Del mismo modo, las escalas diferencial semántico también son muy populares, gracias principalmente a la distinción entre emociones positivas y negativas de Kahneman y Tversky (1979) en su laureada teoría de las perspectivas.

En aras de mejorar la utilidad de las mediciones, teóricamente sería más adecuado

que los consumidores respondieran de la forma que les resulte preferida, es decir, no restringiendo las respuestas a una escala dada por el investigador, sino siendo el propio individuo quien responda en la “escala” que más se ajusta a sus preferencias, al contexto determinado de la pregunta y respuesta, y que minimiza el esfuerzo o coste psicológico (Ferrando, 2003; Weng y Cheng, 2000). Por tanto, el hecho de que el individuo definiera su respuesta sin restricciones de categorización sería, en terminología relativista, una aproximación en “primera persona” a la medida de variables psicológicas (Kilpatrick y Cantril, 1960), en contraposición con la llamada perspectiva en “tercera persona”, asociada al positivismo, que es la forma habitual de proceder en la tradicionalmente etiquetada como investigación cuantitativa. Sin entrar, no obstante, en disquisiciones filosóficas, parece evidente que esa aproximación en primera persona incrementa la utilidad de las respuestas, minimizando un posible sesgo sistemático producido por la categorización de la escala. Por ello, es una opción muy deseable para los investigadores, y más teniendo en cuenta que la evidencia empírica ha demostrado que los individuos utilizan un rango determinado de respuestas (Ferrando 2003), o que directamente prefieren dar sus respuestas en escalas de 1 a 10 (Preston y Colman, 2000). Esto indica que puede existir homogeneidad por parte de los individuos en la forma de contestar. No obstante, y tal y como recomiendan Ferrando (2003) o Hofmans, Theuns y Van Acker (en prensa), es conveniente analizar cada contexto de investigación por separado, abogando por la especificidad de los estudios.

Además de la minimización del sesgo de categorización, un formato libre de respuesta permite que las etiquetas verbales comúnmente utilizadas en las escalas ordinales no



Martínez García, J. A. y Martínez Caro, L.

La medición de la satisfacción del consumidor de servicios...

produzcan interacción con la respuesta, es decir, no sean otra fuente de sesgo. Las diferencias culturales pueden producir divergencias en los patrones de respuesta (Hofmans et al., en prensa). Además algunos individuos pueden considerar, por ejemplo, “muy satisfecho” como el punto final de una escala, mientras que otros no (Saris y Gallhofer, 2007), por lo que el estilo de respuesta varía incluso en función de la personalidad del encuestado (Javaras y Ripley, 2007). Dadas estas circunstancias, se podría obtener una medida numérica de la satisfacción del consumidor optimizando la validez de la respuesta al utilizar una escala de formato libre. Además, si el encuestado asocia libremente la valoración numérica con una etiqueta verbal que le asigne un significado lingüístico, se podría modelar el grado de incertidumbre de esa respuesta verbal, con el fin de obtener una descripción de cómo los individuos utilizan esas etiquetas y cómo se distribuyen las respuestas dentro de cada término lingüístico utilizado. En definitiva, obtener una información mucho más rica de las actitudes del consumidor de servicios deportivos, que las derivadas de los habituales formatos “muy en desacuerdo-muy de acuerdo”, “muy insatisfecho-muy satisfecho” escalados en 5 o 7 categorías de respuesta. Para conseguir tal fin, utilizaremos la lógica borrosa.

Lógica borrosa

La lógica borrosa o difusa es una forma de representación basada en la *Teoría de conjuntos borrosos*, propuesta por Zadeh (1965), que trata de capturar la manera en la que los individuos razonan teniendo en cuenta la incertidumbre de sus pensamientos. Como indican Rajasekaran y Vijayalakshmi (2003), esa incertidumbre puede proceder del conocimiento incompleto, la generalidad, la ambigüedad o la imprecisión en la información.

La lógica borrosa es un tipo de lógica que reconoce más que simples valores verdaderos o falsos. Con lógica borrosa, las proposiciones pueden ser representadas con grados de veracidad o falsedad. Por ejemplo, la frase “hoy es un día soleado” puede ser 100% verdad si no hay nubes, 80% de verdad si hay pocas nubes, 50% de verdad si existe neblina y 0% si el cielo está totalmente cubierto (Lozano y Fuentes, 2003). Del mismo modo, el grado de satisfacción de un consumidor de un servicio deportivo puede ser representado a través de una función de pertenencia que relacione cada respuesta lingüística con una serie de valores probables en un intervalo acotado $[0,1]$. Por ejemplo, cuando un individuo dice que está “satisfecho” con el servicio recibido, ese término tiene un significado incierto, ya que la palabra “satisfecho” puede tener un significado heterogéneo entre los consumidores en función de su personalidad, cultura o contexto de investigación. Si a ese significado se le asigna un valor numérico en una escala universal $[0,1]$, se podría representar un conjunto de valores probables, que forman un número borroso (Zwick y Wallsten, 1989). En este caso particular, sería plausible suponer un número borroso cuyo rango de valores estuviera, por ejemplo, entre 0,6 y 0,9, teniendo ese rango de valores una probabilidad asociada. Esa función de pertenencia se suele establecer a priori (ver Lozano y Fuentes, 2003; Tsai, et al., 2006), asignando un número borroso a un término lingüístico propuesto por el investigador en una escala pre-diseñada (perspectiva en tercera persona). Un ejemplo sería el establecimiento de números borrosos para una escala Likert de 1 a 5 etiquetada desde “muy en desacuerdo” a “muy de acuerdo”.

La perspectiva adoptada en esta investigación es diferente. La función de pertenencia es calculada a través de la relación entre los



valores numéricos que representan su grado de satisfacción sobre una escala libre y los términos lingüísticos utilizados para definir esa actitud (perspectiva en primera persona). De este modo, no hay ninguna asunción a priori sobre la distribución de los valores borrosos, sino que es derivada de forma empírica.

La aplicación de esta metodología se realiza en este estudio, formalmente, y siguiendo la notación de Tsai, et al. (2006), de la manera siguiente:

En un conjunto universal de X , un subconjunto borroso de \tilde{A} de X se caracteriza por una función de pertenencia que asocia un número real en el intervalo $[0,1]$ con cada elemento x de X , representando el grado de pertenencia $\mu_{\tilde{A}}(x)$ de x en \tilde{A} . Un número borroso triangular \tilde{A} en la recta real es un número cuya función de pertenencia $\mu_{\tilde{A}}(x) \rightarrow [0,1]$ es:

$$\mu_{\tilde{A}_i}(x) = \begin{cases} \frac{(x-a)}{(b-a)} & , a \leq x \leq b \\ \frac{(x-c)}{(b-c)} & , b \leq x \leq c \\ 0 & , \text{de otra forma} \end{cases}$$

y puede ser representado por la tripleta $\tilde{A} = (a,b,c)$ con $a \leq b \leq c$. El valor medio del número borroso es expresado con b , que corresponde con el máximo de la función de pertenencia $\mu_{\tilde{A}}(x) = 1$, mientras que a y c son los límites mínimo y máximo respectivamente. La función de pertenencia aumenta linealmente de a a b y decrece de b a c . La Figura 1 muestra un ejemplo de número borroso.

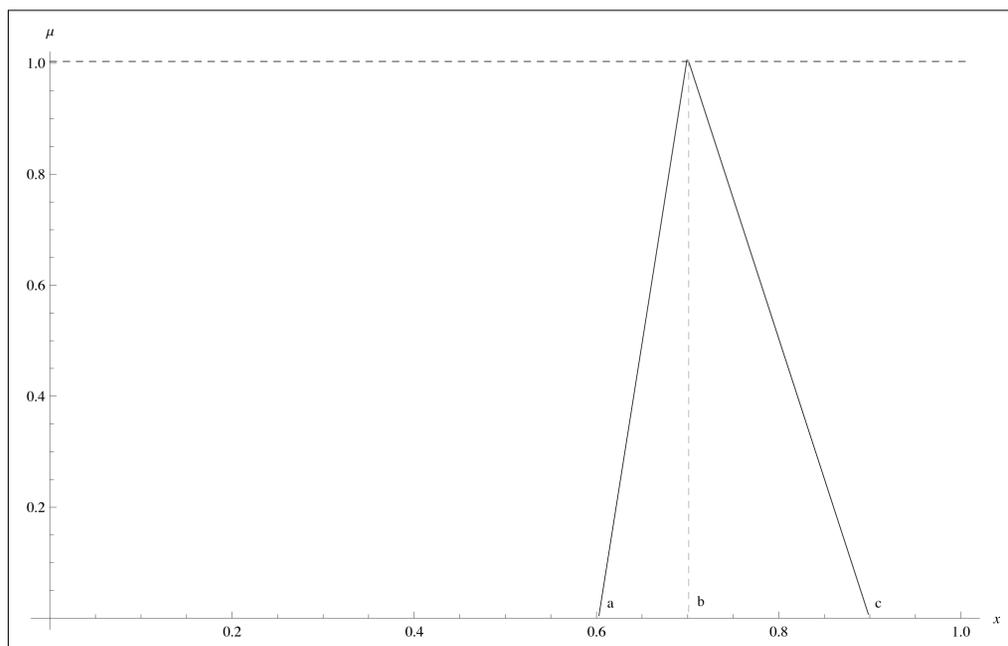


Figura 1. Ejemplo de número borroso triangular.

Martínez García, J. A. y Martínez Caro, L.

La medición de la satisfacción del consumidor de servicios...

Hemos optado por una función discreta triangular, dadas las facilidades de que este tipo de función presenta para realizar operaciones. En Rajasekaran y Vijayalakshmi (2003), pueden consultarse otras posibles formas de la función de pertenencia.

La valoración otorgada por cada individuo puede convertirse en un número borroso derivado empíricamente si se le pide al encuestado que dé esa valoración en un intervalo, si es que así lo estima necesario. De esta forma, es a través de las respuestas de los propios individuos desde donde se derivan los números borrosos. Suponemos entonces, que una persona que da un valor puntual, tiene un grado de incertidumbre mínimo sobre su satisfacción, por lo que el número borroso triangular asociado es simplemente un número *crisp*, o bien definido, es decir, no difuso (a partir de ahora utilizando el término *crisp*, ya que su traducción no es demasiado amigable).

Si los individuos, además, expresan con términos lingüísticos su grado de satisfacción, se puede establecer una relación borrosa entre las etiquetas verbales y los números asignados a ellas. Por ejemplo, las valoraciones numéricas asociadas al término “satisfecho” se distribuyen de forma multinomial en la muestra, ya que los diferentes valores numéricos que toma esa palabra, tienen una probabilidad asociada. El valor esperado de la distribución es el valor máximo de la función de pertenencia al que llamamos b' , mientras que a' y c' coinciden con el mínimo y el máximo valor de la distribución. Cuando no existe variabilidad en la distribución de probabilidades, el número borroso es un número *crisp*.

Formalmente, dado un conjunto de m términos lingüísticos $S_{\mathcal{D}} = \{S_i\} i = 1...m$, existe un conjunto de m números triangulares borrosos $\tilde{A}_{\mathcal{D}} = \{\tilde{A}_i\} i = 1...m$ donde $\tilde{A}_i = \{a_i, b_i, c_i\}$ cuya $\mu_{\tilde{A}_i}(x)$ es:

$$\mu_{\tilde{A}_i}(x) = \begin{cases} \frac{(x - a'_i)}{(b'_i - a'_i)} & , a'_i \leq x \leq b'_i \\ \frac{(x - c'_i)}{(b'_i - c'_i)} & , b'_i \leq x \leq c'_i \\ 0 & , \text{de otra forma} \end{cases}$$

siendo $b'_i = \sum p_{ij} x_{ij}$ donde p_{ij} es la probabilidad que un individuo j responda el valor x_j para cada término lingüístico $i \in m$. Finalmente $a'_i = \min(x_{ij})$ y $c'_i = \max(x_{ij})$.

Los números borrosos pueden ser combinados matemáticamente utilizando simples reglas (ver Rajasekaran y Vijayalakshmi, 2003), y posteriormente se suele realizar una conversión de nuevo a valores *crisp*, o *defuzzificación*, es decir, a valores con interpretación real. En nuestro estudio utilizaremos el ampliamente difundido método del centro de gravedad en el que el valor real x^* es el centro del área ocupada por el conjunto borroso, por lo que de forma general:

$$x^* = \frac{\sum_i x_i \mu(x_i)}{\sum_i \mu(x_i)}$$

Asimismo, los números borrosos pueden compararse en función de su grado de similitud. Aunque se han propuesto multitud de medidas de similitud (ej. Zwick, Carlstein y Budescu, 1987), y en línea con la interpretación geométrica adoptada en la *defuzzificación*, calcularemos la intersección entre dos números borrosos, $a_{\cap} = \tilde{A} \cap \tilde{B}$ y la proporción del área que ocupa esa intersección sobre la suma de las áreas de ambos números borrosos $a_{\cup} = \tilde{A} \cup \tilde{B}$. De esta forma, dos números borrosos serán totalmente distintos si $a_{\cap}/a_{\cup} = 0$, e iguales si a_{\cup}/a_{\cap} .

Estudio empírico 1

Participantes

Se recogió una muestra aleatoria de 116 consumidores del centro deportivo municipal más importante de la ciudad de Cartagena: el Pabellón Municipal de Deportes. Esa muestra constituye alrededor del 6% de usuarios de esos servicios. El factor de imprecisión sobre la escala de medida *-FIEM-* (Martínez y Martínez, 2008) fue del 4,5%, factor que se tomó como criterio para la determinación de un tamaño de muestra mínimo ($<5\%$). La muestra estaba compuesta por un 69% de hombres y un 31% de mujeres, siendo la media de experiencia en el servicio de 31 meses.

Instrumento

El instrumento utilizado fue un breve cuestionario diseñado en función de las necesidades de esta investigación particular y de otro proyecto de investigación paralelo. Así, se les pedía a los encuestados que expresaran su grado de satisfacción con el servicio deportivo, a través de un término lingüístico y a través de un valor numérico, utilizando una única pregunta para ambos requerimientos. Además de las preguntas sociodemográficas, se recogieron otros datos de interés relativos a los motivos de elección del centro. Para el análisis de datos y la construcción de gráficos se utilizaron los programas Microsoft Excel y Mathematica 6.0.

Procedimiento

Estudiantes de marketing previamente entrenados y formados fueron los encargados de recoger la información, entrevistando personalmente a los sujetos participantes en la puerta del centro deportivo. Al preguntarles por su grado de satisfacción, cuando el encuestado expresaba ese valor numérico se

le pedía que indicara la escala en la que ese valor cobraba sentido, y si prefería dar ese valor puntual en un intervalo para expresar así su incertidumbre sobre ese juicio. La recogida de datos se realizó durante la primavera de 2008.

Resultados

En primer lugar, se realizó un análisis descriptivo de las respuestas obtenidas (Tabla 1). De esta primera descripción de los datos puede deducirse: (1) que el rango de términos lingüísticos utilizado es muy amplio; (2) que se prefiere por igual las escalas de 1 a 10 y de 0 a 10, pudiendo utilizar en esta última valores decimales; y (3) que los encuestados casan con certidumbre su valoración lingüística y numérica.

Dada la sorprendente exactitud con la que los consumidores expresan su grado de satisfacción, la conversión de esos valores a números borrosos carece prácticamente de sentido. No obstante, podemos obtener una medida puntual (con sus intervalos de confianza asociados, por supuesto) con alto grado de validez, utilizando los valores *crisp* con las técnicas estadísticas habituales. Para ello se necesita, en primer lugar, convertir esos valores en una escala universal $[0,1]$ (ver Cohen, Cohen, Aiken y West, 1999). Una vez realizada esa homogeneización, el valor medio calculado de satisfacción de los consumidores es 0.74, con un intervalo de confianza al 95%: (0.70, 0.78).

El siguiente paso de análisis es la obtención de los números borrosos asociados a los términos lingüísticos utilizados por los encuestados. Siguiendo el procedimiento descrito en páginas anteriores, se obtienen los números borrosos que aparecen en la Tabla 1. Asimismo, en la Figura 2 pueden verse ilustrados los 5 términos más utilizados.



Descripción		Números borrosos			
Términos lingüísticos	%	a'	b'	c'	Evaluación lingüística
Bueno	52.59	0.50	0.79	1.00	+
Muy bueno	9.48	0.75	0.91	1.00	+
Malo	6.90	0.00	0.25	0.50	-
Satisfecho	6.03	0.44	0.74	1.00	+
Excelente	5.17	0.89	0.95	1.00	+
Mejorable	2.59	0.40	0.45	0.50	Neutra
Regular	2.59	0.44	0.50	0.56	Neutra
Contento	2.59	0.67	0.80	0.89	+
Bien	1.72	0.60	0.73	0.85	+
Insatisfecho	1.72	0.33	0.50	0.67	-
Medio	1.72	0.50	0.52	0.55	Neutra
Muy contento*	1.72		1.00		
Bastante buena*	0.86		0.89		
Cumple lo exigido*	0.86		0.70		
Insuficiente*	0.86		0.22		
Mucho*	0.86		0.89		
Muy satisfecho*	0.86		0.90		
Suficiente*	0.86		0.60		
Escala preferida					
Escala de 1 a 10	51.72				
Escala de 0 a 10	47.41				
Escala de 1 a 5	0.86				
Incertidumbre en la respuesta					
No	99.14				
Si	0.86				

* La varianza en la distribución de las respuestas es 0, por lo que son números *crisp*.

Tabla 1. Descripción de las respuestas y relación entre términos lingüísticos y números borrosos.

Una vez calculados los números borrosos, se puede operar con ellos, utilizando simples reglas aritméticas. Es por ello que creemos de interés, por ejemplo, crear dos nuevos números borrosos \tilde{P} y \tilde{N} como combinación de los ya existentes, que representen respectivamente evaluaciones positivas y negativas del servicio. Para ello, hemos clasificado los términos lingüísticos como positi-

vos, negativos y neutros, quedando en esta última categoría los términos que su interpretación ciertamente ofrecía dudas (Tabla 1). El cálculo de \tilde{P} y \tilde{N} y se realiza como mostramos.

$$\tilde{P} = \cup \tilde{A}_i (+)$$

$$\tilde{N} = \cup \tilde{A}_i (-)$$



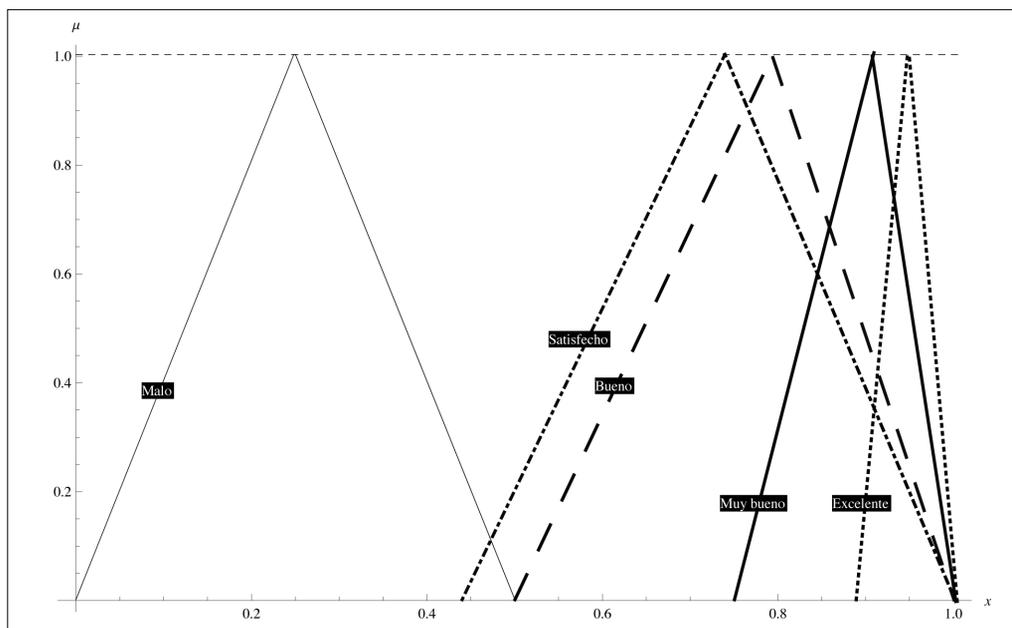


Figura 2. Borrosidad de los 5 términos lingüísticos más utilizados.

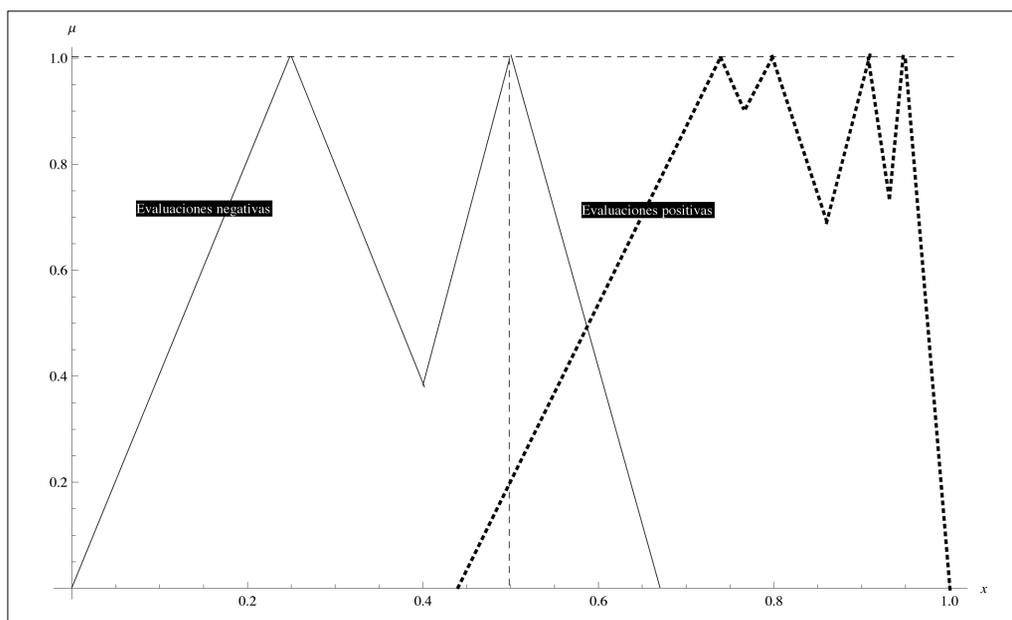


Figura 3. Borrosidad de las evaluaciones negativas y positivas.



La Figura 3 muestra los valores de la función de pertenencia para los dos números borrosos creados. Cabe destacar que el grado de pertenencia de un valor numérico entre 0.5 y 0.6, es mayor para el número borroso que representa términos negativos que para el que representa términos positivos, por lo que valoraciones numéricas obtenidas en ese rango no tienen porqué significar una evaluación positiva del grado de satisfacción del consumidor.

Podemos realizar, asimismo, un análisis heurístico del sesgo que se habría producido en los resultados si se hubiera utilizado un método habitual de medición, por ejemplo, una escala Likert de 1 a 5 mostrando las etiquetas de cada opción de la escala. Para utilizar el mismo lenguaje que han utilizado los consumidores en este estudio, vamos a suponer que la pregunta hubiera sido: *Califique su grado de satisfacción con el servicio del que es usted usuario en una escala de 1 a 5, siendo: 1- Muy malo; 2-Malo; 3-Neutro; 4-Bueno; 5-Muy bueno.*

Para ello comparamos los valores borrosos con los números *crisp* que se hubieran obtenido si cada individuo hubiera sido coherente en sus asociaciones lingüísticas con los números de la escala Likert. La Tabla 2 muestra como, para los 3 términos lingüísticos sobre los que es posible hacer la comparativa, el sesgo estimado ofrece valores ciertamente dispares, siendo inexistente para el

caso de “malo” y considerable para el caso de “muy bueno”. Aunque la interpretación de la importancia del sesgo puede depender de la subjetividad del investigador, podríamos convenir que alrededor del 5% sería admisible, y que el valor del 10% sería ya importante. Hay que resaltar que el sesgo medio total producido, considerando la ponderación por el número de respuestas, es de sólo 2%, lo que puede resultar engañoso, ya que indica que los errores cometidos se van compensando por exceso y por defecto. Aunque el efecto sobre la varianza es mayor (0.041 frente a 0.034), no existen diferencias significativas entre ambas, por lo que de nuevo el comportamiento es muy homogéneo.

Finalmente, el bajo porcentaje de respuesta para los términos “muy bueno” y “malo”, nos hace ser cautos sobre la utilidad del proceso de *defuzzificación*, o conversión de los números borrosos a valores *crisp*, ya que no consideramos que exista suficiente consistencia con submuestras tan pequeñas. No obstante, y a modo indicativo, si hubiéramos asumido que el porcentaje de respuestas es suficiente para operar fiablemente con las tres etiquetas lingüísticas, tras aplicar el método del centro de gravedad, el valor *crisp* resultante hubiera sido de 0.52. Este valor, distaría mucho del valor de 0.67 que se hubiera obtenido promediando estimaciones consistentes y fiables en una escala de 1 a 5.

Descripción		Números borrosos			Estimación escala 1 a 5	
		a'	b'	c'	Crisp	Sesgo*
Términos lingüísticos	%				Crisp	Sesgo*
Bueno	52.59	0.50	0.79	1.00	0.75	4%
Muy bueno	9.48	0.75	0.91	1.00	1	-9.9%

* Es el porcentaje de error cometido sobre la longitud de la escala [0,1], tomando como referencia el valor borroso b' más probable.

Tabla 2. Descripción de las respuestas y relación entre términos lingüísticos y números borrosos.



Estudio empírico 2

Participantes, instrumento y procedimiento

El primer estudio fue replicado en otra muestra aleatoria del segundo centro deportivo municipal más importante de la ciudad: la Piscina Municipal, con el fin de contrastar la consistencia de los resultados en un contexto diferente, es decir, ante lo que Hitchcock (2002) llama otras situaciones test. Asumimos que puede existir heterogeneidad a priori, y por tanto, diferenciación en el test, porque este centro es de gestión compartida con una empresa privada. Por tanto, partimos de la hipótesis de que el grado de satisfacción puede ser diferente en este segundo centro. Dada esta situación, creemos conveniente estudiar si los resultados del primer estudio se confirman con esta segunda muestra.

De nuevo se estableció como criterio para obtener el tamaño de muestra el obtener un *FIEM* menor del 5%. En este caso se obtuvieron 98 cuestionarios válidos; un 60% de los encuestados eran hombres y un 40% mujeres, siendo la media de experiencia en el servicio de 40 meses. Se aplicó el mismo instrumento y procedimiento que en el Estudio 1.

Resultados

La descripción de las respuestas se contempla en la Tabla 3. Podemos observar que: (1) 14 de los 18 términos lingüísticos utilizados coinciden con los del primer estudio, constituyendo el 95,7% de las respuestas; (2) Las escalas de 1 a 10 y de 0 a 10, son las preferidas por los encuestados, aunque en esta ocasión, la primera de ellas lo hace en un porcentaje sensiblemente superior a la segunda; y (3) los encuestados vuelven a casar con certidumbre su valoración lingüística y numérica (96% de las respuestas)

El valor esperado de la satisfacción del consumidor considerando sus juicios numéricos es de 0.70 con un intervalo de confianza al 95%: (0.66, 0.73). Por tanto, el grado de satisfacción del consumidor es homogéneo en ambas muestras, ya que se solapan los intervalos de confianza, es decir, el centro deportivo con gestión compartida no produce niveles de satisfacción diferentes del centro con gestión exclusivamente pública.

Aunque no existe esa heterogeneidad que a priori podía esperarse en la satisfacción media entre las dos muestras, el proceso de replicación sigue siendo igualmente de interés, ya que, dadas estas circunstancias, cabría esperar entonces una alta similitud entre los números borrosos de ambas situaciones test. Para realizar ese análisis, seguiremos con la propuesta comparativa frente al hipotético escenario de medidas sobre escala de Likert de 1 a 5.

Como puede verse en la Tabla 4, los resultados son bastante similares a los del primer estudio. Es más, el nivel de sesgo medio es también muy bajo (3%) y las varianzas no son significativamente diferentes (0.028 frente a 0.026).

Al igual que en el primer estudio, el bajo porcentaje de respuesta para los términos “muy bueno” y “malo”, nos hace ser cautos sobre la utilidad del proceso de *defuzzificación*. No obstante, y de nuevo a modo indicativo, si hubiéramos asumido que el porcentaje de respuestas es suficiente para operar fiablemente con las tres etiquetas lingüísticas, el valor *crisp* resultante hubiera sido de 0.70. Esta vez sí, resulta un valor muy similar (0.67) al que se hubiera obtenido promediando estimaciones consistentes y fiables en una escala de 1 a 5.

Finalmente, el grado de solapamiento entre estos tres números borrosos se mues-





Descripción		Números borrosos			
Términos lingüísticos	%	a'	b'	c'	Evaluación lingüística
Bueno	43.88	0.44	0.72	1.00	+
Regular	12.24	0.33	0.56	1.00	Neutro
Bien	10.20	0.66	0.75	0.89	+
Muy bueno	6.12	0.78	0.90	1.00	+
Satisfecho	4.08	0.66	0.72	0.78	+
Malo	4.08	0.22	0.31	0.35	-
Contento	3.06	0.78	0.89	1.00	+
Excelente	3.06	0.78	0.92	1.00	+
Aceptable	2.04	0.67	0.68	0.70	+
Bastante satisfecho	2.04	0.67	0.68	0.70	+
Media*	1.02				
Mejorable*	1.02				
Buenísima*	1.02				
Mucho*	1.02				
Muy satisfecho*	1.02				
Insatisfecho*	1.02				
Normal*	1.02				
Poco satisfecho*	1.02				
Suficiente*	1.02				
Escala preferida					
Escala de 1 a 10	73.46				
Escala de 0 a 10	24.48				
Escala de 1 a 5	2.04				
Incertidumbre en la respuesta					
No	95.92				
Si	4.08				

* La varianza en la distribución de las respuestas es 0, por lo que son números *crisp*.

Tabla 3. Descripción de las respuestas y relación entre términos lingüísticos y números borrosos (estudio 2).

Descripción		Números borrosos			Estimación escala 1 a 5	
Términos lingüísticos	%	a'	b'	c'	<i>Crisp</i>	Sesgo*
Bueno	43.88	0.44	0.72	1.00	0.75	3%
Muy bueno	6.12	0.78	0.90	1.00	1	-10%
Malo	4.08	0.22	0.31	0.35	0.25	6%

* Es el porcentaje de error cometido sobre la longitud de la escala [0,1], tomando como referencia el valor borroso b' más probable.

Tabla 4. Descripción de las respuestas y relación entre términos lingüísticos y números borrosos (estudio 2).



tra en la Tabla 5. Insistimos en que el bajo número de respuestas en “muy bueno” y “malo” hace aventurado realizar una interpretación concluyente, aunque podemos vislumbrar que cuando los términos lingüísticos hacen referencia a valores más extremos (caso de “muy bueno”) la concordancia debe ser mayor que cuando se refieren a términos susceptibles de más ambigüedad (caso de “bueno” o “mala”). Sí que, no obstante, el número de respuestas etiquetadas con “bueno” son suficientes en ambas muestras para realizar una comparación fiable. En este

caso, el grado de similitud alcanza el 70%, lo que interpretamos como un índice alto, ya que a medida que aumenta la incertidumbre asociada con cada concepto será más difícil obtener concordancias cercanas al valor máximo (100%). Ese grado de incertidumbre es mayor a medida que se incrementa, cuyo valor máximo sería el área que delimita el intervalo de recta real $[0,1]$ y la función de pertenencia $[0,1]$, es decir: 1. Que el término “bueno” ocupe prácticamente un tercio de toda el área de incertidumbre da una idea de su “gran borrosidad”.

Descripción	Estudio 1	Estudio 2	Similitud		
			Respuestas (%)	Respuestas (%)	a_{\cap}
Términos lingüísticos					
Bueno	52.59	43.88	0.22	0.31	70
Muy bueno	9.48	6.12	0.11	0.12	88
Malo	6.90	4.08	0.06	0.25	25

Tabla 5. Similitud entre términos lingüísticos.

Discusión

Esta investigación ha planteado el problema de medir la satisfacción de los consumidores de servicios deportivos desde un enfoque novedoso en psicología del deporte: la modelización de la incertidumbre a través del uso de la lógica borrosa. Se han descrito sencillos procedimientos de análisis para operar con técnicas borrosas, que pueden ser aplicados por una amplia audiencia de investigadores y gestores deportivos. La principal limitación de esta investigación estriba en la carencia de un tamaño muestral mayor en ambos estudios que hubiera garantizado la consis-

tencia y estabilidad en la creación de los números borrosos dada la cantidad de términos lingüísticos utilizados. Pese a ello, se pueden obtener implicaciones importantes:

En primer lugar, los consumidores de servicios deportivos prefieren expresar su grado de satisfacción en escalas de 1 a 10 o de 0 a 10, esta última con la posibilidad de dar valores decimales. La escala de 1 a 10 se muestra como claramente preferida en el segundo estudio, coincidiendo con los resultados de Preston y Colman (2000). Esto indica que, aunque las actitudes tengan teóricamente un carácter latente y continuo, existe una gran unanimidad en la expresión o mani-

festación de las mismas en este tipo de escalas, lo que muestra que existe un marco de referencia común sobre el cual los consumidores hacen sus evaluaciones. Futuras investigaciones deberán plantearse con más profundidad si la asunción de ese *continuum* sigue siendo realista en la medición de actitudes en general, y más teniendo en cuenta trabajos como los de Ferrando (2003), donde se evidencia que los consumidores tienden a utilizar un limitado número de opciones de respuesta para las escalas continuas. Como bien indica ese autor, es preceptivo estudiar cada contexto y variable medida; en el caso de la satisfacción del consumidor de servicios deportivos, las preferencias de los individuos son diáfanos.

También Posteriores Estudios, deberán poner atención sobre si realmente no existe distinción entre ambos formatos de respuesta. Y es que la distinción entre 1 y 0 en el límite inferior podría ser debida a la inexistencia de grados de satisfacción extremadamente negativos. De hecho, sólo un individuo respondió con el peor valor posible (y lo hizo utilizando la escala de 0 a 10 y no la de 1 a 10).

En segundo lugar, existe una gran disparidad de términos lingüísticos que los consumidores utilizan libremente para expresar su grado de satisfacción, aunque la palabra “bueno” es, sin duda, la más destacada. Centrándonos en este término, que es el único donde el porcentaje de respuestas permite una interpretación más fiable, los resultados muestran como posee gran borrosidad, lo que indica que tiene significados diferentes para individuos dispares. Ello refrenda otras investigaciones que muestran la heterogeneidad de la significación de los términos lingüísticos (Wallsten, Budescu, Zwick, 1993), por lo que cabría esperar que las escalas usadas habitualmente donde se identifican las etiquetas lingüísticas con números produje-

ran una distorsión importante. Sin embargo, los resultados de nuestra investigación no concuerdan exactamente con esa línea de razonamiento, ya que como se muestra en las Tablas 2 y 4, el sesgo producido parece no ser importante. Aunque sí existen diferencias entre el valor *crisp* original y el valor *crisp defuzzificado* en el primer estudio, la replicación ofrece resultados contradictorios. Es por ello, que recomendamos que se teste empíricamente ese sesgo en otras investigaciones, preguntando específicamente al individuo por su nivel de satisfacción usando diferentes tipos de escala. Este hecho ayudaría a clarificar por fin si el error sistemático que se comete al constreñir una variable latente en una escala ordinal puede considerarse verdaderamente sistemático o si se comporta de forma aleatoria. Los resultados de nuestro estudio van en esa segunda línea, por lo que podría decirse que, en el caso de una escala de 1 a 5, el valor medio y la varianza no se verían estadísticamente distorsionados. Sin embargo, una escala de 1 a 5 con etiquetas diferentes, por ejemplo, “muy en desacuerdo-muy de acuerdo” o “muy insatisfecho-muy satisfecho”, podría producir resultados diferentes porque los términos lingüísticos no tendrían un significado idéntico.

La formación de nuevos números borrosos que recogen las evaluaciones positivas y negativas indica que valores numéricos entre 0.5 y 0.6, es decir, asociados a una interpretación positiva, tienen más probabilidad de ser ligados a un término lingüístico con connotaciones negativas. Este resultado es importante, porque demuestra que no basta con conocer el grado de satisfacción del consumidor con un índice numérico si no sabemos qué significa ese número.

En tercer lugar, e inesperadamente, los consumidores expresan su nivel de satisfacción numéricamente sin incertidumbre, lo



Martínez García, J. A. y Martínez Caro, L.

La medición de la satisfacción del consumidor de servicios...

que significa que los individuos son capaces de emitir un juicio preciso que refleja sus pensamientos. Los encuestados saben perfectamente (al menos en cuanto a la asignación de un valor numérico se refiere), qué significa que, por ejemplo, su grado de satisfacción sea bueno. Individualmente no parece que exista incertidumbre en ese juicio, lo que contradice los habituales enfoques de estudios de este tipo (ej. Lozano y Fuentes, Tsai et al., 2006). Aquí reside la ventaja de nuestro enfoque en “primera persona”, donde permitimos al individuo que sea él quien otorgue significado a sus juicios, para luego agregar esos significados formando una termino lingüístico borroso. Por tanto, los estudios que utilizan la lógica borrosa con una perspectiva en “tercera persona” pueden llegar a conclusiones ciertamente sesgadas, ya que la formación de los números borrosos puede distar mucho del valor borroso derivado empíricamente.

Finalmente, y admitiendo de nuevo las limitaciones de este estudio, podemos concluir que el enfoque propuesto ofrece la posibilidad de obtener información más rica sobre las actitudes del consumidor de servi-

cios deportivos, superando las desventajas de los enfoques en “tercera persona” y evitando la posibilidad de la aparición de sesgo por categorización y por interacción entre juicios verbales y numéricos. Este análisis podría ser obviamente extendido a la medición de otras actitudes, como la calidad percibida. De hecho, el interés por la medición de la calidad percibida es también creciente en psicología del deporte (ej. Hernández-Mendo, 2001; Morales, Hernández-Mendo y Blanco, 2009). Medir y analizar la satisfacción de la forma que proponemos es relativamente sencillo a nivel operativo y con bajo coste. La inclusión de más ítems para medir este concepto (escalas multi-ítem) complicaría un poco esta forma de proceder, aunque nuestra postura al respecto es apostar por medir conceptos utilizando el menor número de ítems, incluso sólo con un ítem como en el caso de nuestra investigación. Los trabajos de Rossiter (2002), Hayduk (1996), o Hayduk et al. (2007) ofrecen una brillante defensa de esta perspectiva. Pero este, claro está, es otro de esos apasionados debates cuya discusión supera los objetivos de esta investigación.

LA MEDICION DE LA SATISFACCION DEL CONSUMIDOR DE SERVICIOS DEPORTIVOS A TRAVES DE LA LOGICA BORROSA

PALABRAS CLAVE: Satisfacción, Servicios deportivos, Lógica borrosa.

RESUMEN: Esta investigación plantea el problema de medir la satisfacción de los consumidores de servicios deportivos desde un enfoque novedoso en psicología del deporte: la modelización de la incertidumbre a través del uso de la lógica borrosa. Se han descrito sencillos procedimientos de análisis para operar con técnicas borrosas, con el fin de comprobar empíricamente en qué medida esta metodología puede complementar los resultados derivados de los métodos estadísticos tradicionales. De este modo, se profundiza sobre la relación entre las evaluaciones numéricas de la satisfacción de los consumidores y el grado de incertidumbre asociado a las evaluaciones lingüísticas. El enfoque propuesto ofrece la posibilidad de obtener información más rica sobre las actitudes del consumidor de servicios deportivos, superando las desventajas de los enfoques en “tercera persona” y evitando la posibilidad de la aparición de sesgo por categorización y por interacción entre juicios verbales y numéricos.





A MEDIAÇÃO DA SATISFAÇÃO DO CONSUMIDOR DE SERVIÇOS DESPORTIVOS ATRAVÉS DA LÓGICA "FUZZY"

PALAVRAS-CHAVE: Satisfação, Serviços desportivos, Lógica "fuzzy".

RESUMO: Esta investigação debruça-se sobre a avaliação da satisfação dos consumidores de serviços desportivos através de uma abordagem inovadora em psicologia do desporto: a modelagem da incerteza através do uso da lógica "fuzzy". Para tal, foram descritos procedimentos simples de análise para lidar com este tipo de lógica, com o objectivo de comprovar empiricamente em que medida esta metodologia pode complementar os métodos estatísticos tradicionais. Deste modo, aborda-se de forma mais aprofundada a relação entre as avaliações numéricas da satisfação dos consumidores e o grau de incerteza associado a avaliações linguísticas. A abordagem proposta sugere a possibilidade de obter informação mais rica sobre as atitudes do consumidor de serviços desportivos, superando as desvantagens dos enfoques na "terceira pessoa" e evitando a possibilidade da ocorrência de enviesamento por categorização e por interacção entre avaliações verbais e numéricas.

Referencias

- Anderson, E. W. y Sullivan, M. W. (1993). The antecedents and consequences of customer satisfaction. *Marketing Science*, 12 (2), 125-143.
- Cohen, P., Cohen, J., Aiken, L., y West, S. (1999). The problem of units and the circumstance for POMP. *Multivariate Behavioral Research*, 34 (3), 315-346.
- Cox, E. P. (1980). The optimal number of response alternatives for a scale: a review. *Journal of Marketing Research*, 17(Noviembre), 407-422.
- Ferrando, P. J. (2003). A Kernel density analysis of continuous typical-response scales. *Educational and Psychological Measurement*, 63 (5), 809-824.
- Fornell, C. (1992). A national customer satisfaction barometer: The Swedish experience. *Journal of Marketing*, 56 (1), 6-21..
- Fournier, S. y Mick, D. G. (1999). Rediscovering satisfaction. *Journal of Marketing*, 62 (Octubre), 5-23.
- Greenwell, C. T., Brownlee, E., Jordan, J. S. y Popp, N. (2008). Service fairness in spectator sport: the importance of voice and choice on customer satisfaction. *Sport Marketing Quarterly*, 17 (2), 71-78
- Hayduk, L. A. (1996). *LISREL Issues, Debates and Strategies*. Baltimore, MA: Johns Hopkins University Press.
- Hayduk, L. A., Pazderka-Robinson, H., Cummings, G. G., Boadu, K., Verbeek, E. L. y Perks, T. A. (2007). The weird world, and equally weird measurement models: Reactive indicators and the validity revolution. *Structural Equation Modeling*, 14 (2), 280-310.
- Hernández-Mendo, A. (2001). Un cuestionario para evaluar la calidad en programas de actividad física. *Revista de Psicología del Deporte*, 10 (2), 179-196.
- Hitchcock, C. (2002). *Probabilistic causation*. Stanford Encyclopedia of Philosophy.
- Hofmans, J., Theuns, P. y Van Acker, F. (En prensa). Combining quality and quantity. A psychometric evaluation of the self-anchoring scale. *Quality & Quantity*.
- Hofmans, J., Theuns, P., y Mairesse, O. (2007). On the impact of the number of response categories on linearity and sensitivity of 'Self Anchoring Scales'. A Functional Measurement approach. *Methodology*, 3, (4), 160-169.





- Javaras, K. N. y Ripley, B. D. (2007). An “unfolding” latent variable model for likert attitude data: drawing inferences adjusted for response style. *Journal of the American Statistical Association*, 102 (478), 454-463
- Kahneman, D. y Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, 47 (2), 263- 291.
- Kilpatrick, F. P. y Cantril, H. (1960). Self-anchoring scaling: A measure of individuals’ unique reality worlds. *Journal of Individual Psychology*, 16, 158-173.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 140, 1-55.
- Lissitz, R. W. y Green, S. B. (1975). Effect of the number of scale points on reliability: A Monte Carlo approach. *Journal of Applied Psychology*, 60 (1), 10-13.
- Lozano, M. C. y Fuentes, F. (2003). *Tratamiento borroso del intangible en la valoración de empresas de Internet*. Extraído el 1 de Junio de 2008 desde <http://www.eumed.net/cursecon/libreria/>
- Madrigal, R. y Chen, J. (2008). Moderating and mediating effects of team identification in regard to causal attributions and summary judgments following a game outcome. *Journal of Sport Management*, 22 (6), 717-733.
- Mañas, M. A., Giménez, G., Muyor, 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
- Martense, A. y Gronholdt, L. (2005). Analysing customer satisfaction data: a comparison of regression and artificial neural networks. *International Journal of Market Research*, 47 (2), 121-130.
- Martínez, J. A. y Martínez, L. (2008). Determinación de la máxima varianza para el cálculo del Factor de Imprecisión sobre la Escala de Medida, y extensión a diferentes tipos de muestreo. *Psicothema*, 20 (2), 305-310.
- Martínez, J. A. y Martínez, L. (2009). Understanding customer loyalty through system dynamics; the case of a public sports service in Spain. *Management Decision*, 47 (1), 151-172.
- Morales, V., Hernández-Mendo, A. y Blanco, A. (2009). Evaluación de la calidad en organizaciones deportivas: adaptación del modelo Servqual. *Revista de Psicología del Deporte*, 18 (2).
- Oliver, R. L. (1997). *Satisfaction: A Behavioural Perspective on the Customer*. New York: McGraw-Hill Companies.
- Preston, C. C. y Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104 (1), 1-15.
- Rajasekharan S. y Vijayalakshmi G. A. (2003). *Neural networks, Fuzzy logic and genetic algorithms*. New Delhi: Prentice Hall.
- Rodríguez, G. y Barriopedro, M. I. (2003). Niveles de satisfacción en usuarios de piscinas cubiertas con la profundidad del vaso utilizado para la práctica. *Revista de Psicología del Deporte*, 12 (2), 147-164.
- Rossiter, J. R. (2002). The coarse procedure for scale development in marketing. *International Journal of Research in Marketing*, 19 (4), 305-335.
- Saris, W. E. y Gallhofer, I. (2007). Estimation of the effects of measurement characteristics on the quality of survey questions. *Survey Research Methods*, 1 (1), 29-43.
- Seiders, K., Voss, G. B., Grewal, D. y Godfrey, A. L. (2005). Do satisfied customers buy more? Examining moderating influences in a retailing context. *Journal of Marketing*, 69 (4), 26-43.





- Tsai, M. T., Wu, H. L. y Liang, W. K. (2008). Fuzzy decision making for market positioning and developing strategy for improving service quality in department stores. *Quality & Quantity*, 42 (3), 303-319.
- Wallsten, T. S., Budescu, D.V. y Zwick, R. (1993). Comparing the calibration and coherence of numerical and verbal probability judgments. *Management Science*, 39 (2), 176-190.
- Weng, L.-J. y Cheng, C.-P. (2000). Effects of response order on Likert-type scales. *Educational and Psychological Measurement*, 60 (6), 908-924.
- Zadeh, L. (1965). Fuzzy sets. *Information and Control*, 8 (3), 338-353.
- Zwick, R. y Wallsten, T. S. (1989). Combining stochastic uncertainty and linguistic inexactness: Theory and experimental evaluation of four fuzzy probability models, *International Journal of Man Machine Studies*, 30, 69-111.
- Zwick, R., Carlstein, E. y Budescu, D. V. (1987). Measures of similarity among fuzzy concepts: A comparative analysis. *International Journal of Approximate Reasoning*, 1 (2), 221-242.

