

MEDICIÓN DE LA CALIDAD PERCIBIDA DE LOS ESPECTADORES EN LOS EVENTOS DEPORTIVOS: PROPUESTA METODOLÓGICA

Ana Milena Álvarez Cano, Universidad Autónoma de Occidente

RESUMEN

El interés por diseñar un instrumento de medición adecuado, válido, confiable, de utilidad práctica, replicable en otros eventos deportivos y que la información obtenida facilite a los directivos tomar decisiones más adecuadas; surge cuando al revisar los modelos de medición de la calidad percibida del servicio no se encuentra un modelo que se ajuste a las necesidades del sector deportivo. Por tanto, el objetivo de este trabajo es presentar una propuesta metodológica para la selección adecuada de las variables a incluir en un instrumento de medición de la calidad percibida de los servicios deportivos, resaltando el uso del Método Delphi para comprobar la validez del contenido mediante la participación de expertos. La propuesta metodológica se desarrolló en tres fases. En primera fase, se elabora una lista preliminar de variables mediante la revisión de literatura de investigaciones previas, modelos teóricos validados y entrevistas con expertos del área de estudio; en la segunda fase, se valida el contenido del instrumento mediante el uso del método Delphi y estadísticas descriptivas; y en la tercera fase, se calcula la confiabilidad de consistencia interna, para ello se utilizan, con el propósito de contrastar los resultados y asegurar que el instrumento mida lo que realmente se pretende medir, los métodos estadísticos de Alpha de Cronbach (1951); Dos mitades, corregido por la fórmula de Spearman-Brown (1919); Rulón (1939) y Hoyt (1941) (Anova). Como resultado final, se logró un instrumento conformado por 18 ítems, el cual podrá ser empleado por quienes se interesen por la evaluación de la calidad de sus eventos deportivos. Asimismo, se resalta como hallazgo principal, la inclusión del ítem “la seguridad ofrecida durante el desarrollo del evento” porque no se encontró en los modelos teóricos consultados sobre de medición de la calidad percibida de los servicios.

PALABRAS CLAVES: Calidad Percibida, Validación, Método Delphi, Confiabilidad

MEASUREMENT OF PERCEIVED QUALITY SPECTATOR SPORTS EVENTS: PROPOSED METHODOLOGY

ABSTRACT

This study develops a model to measure perceived service quality in sporting events. Even though there are models to measure perceived quality, none are designed specifically for sporting events. We use the Delphi method, Alpha Cronbach, Spearman–Brown prediction formula, and ANOVA to develop and test the model. The methodology was developed in three phases. In the first phase, a preliminary list of variables is elaborated by literature review, validated theoretical models and interviews with experts. In the second phase, the content of the instrument is validated using the Delphi method and descriptive statistics; and the third phase, the internal consistency reliability is calculated. As a result, an 18-item instrument is proposed to evaluate perceived quality of sporting events.

JEL: C35, C39, M31, L83

KEYWORDS: Delphi Method, Perceived Quality, Reliability, Validation

INTRODUCCIÓN

La nominación a la ciudad de Cali como “Ciudad deportiva de América” ha suscitado el interés de los directivos del comité deportivo en diseñar modelos de gestión eficiente que les permita ser competitivos y poder sostener en el tiempo su liderazgo como la casa deportiva del país. Dado que las evaluaciones de la calidad de los servicios deportivos han sido diagnosticadas desde las opiniones de los espectadores que expresan sus percepciones de manera informal o desde los comentaristas y dirigentes deportivos, y que en la revisión de la literatura, se pudo evidenciar que en la ciudad de Cali no existen estudios formales dirigidos a la medición de la calidad de los servicios deportivos, se detectó la necesidad de realizar un instrumento de medición para conocer la percepción de los espectadores de la calidad de los servicios ofrecidos en los eventos deportivos.

Dado lo anterior, este trabajo se sistematiza en cuatro apartes: la revisión de la literatura, la metodología, los resultados y las conclusiones. En la revisión de la literatura, se presentan las investigaciones previas que se consideraron para la elaboración de la primera lista de variables. Entre los aportes más significativos se encuentran: las dimensiones de tangibilidad o evidencia física (Hightower *et al.*, 2002; Yusof & Lee, 2008), la accesibilidad en toda la instalación, estética del estadio, marcadores, comodidad de los asientos, suficiencia de espacio, señalización y la calidad del equipo (Kim & Kim, 1995) y la actitud de los empleados (Calabuig, Quintanilla & Mundina, 2008; Calabuig & Crespo, 2009 y Kelley & Turley, 2001). Las otras variables incluidas en la lista son recurrentes en los estudios revisados, tales como: fiabilidad, información, programación, precio, exclusividad, privacidad, conveniencia, estimulación y oportunidad social.

Seguidamente, se muestra en la metodología empleada para la validez de contenido y confiabilidad de consistencia interna del instrumento, empleando técnicas tanto cualitativas como cuantitativas de investigación. Para iniciar y con los resultados de la literatura revisada, se elabora una escala preliminar tipo likert con 71 ítems. Para Validar el contenido de la escala, se emplearon el método cualitativo Delphi, ampliamente utilizado para la validación del contenido del instrumento de medición (Hung, Altschuld y Lee, 2007). Para calcular la confiabilidad de la consistencia interna del instrumento, se aplicó el cuestionario preliminar a una muestra exploratoria de 89 personas que asistieron a eventos deportivos en los últimos seis meses, para ello se utilizaron los métodos estadísticos de Alpha de Cronbach (1951); Dos mitades, corregido por la fórmula de Spearman-Brown (1919); Rulón (1939) y Hoyt (1941) (Anova) con el propósito de contrastar los resultados y asegurar que el instrumento realmente mida lo que se pretende medir. Como resultado final, se logró un instrumento conformado por 18 ítems, el cual podrá ser empleado por quienes se interesen por la evaluación de la calidad de sus eventos deportivos. Asimismo, se resalta como hallazgo principal, la inclusión del ítem “la seguridad ofrecida durante el desarrollo del evento” porque no se considera en los modelos teóricos consultados sobre de medición de la calidad percibida de los servicios. Este trabajo permitió concluir que los modelos de medición no son aplicables directamente en cualquier tipo de servicio, siendo necesario la construcción de escalas de medición de la calidad percibida del servicio acorde con las necesidades particulares del sector estudiado.

Aunque existen diferentes métodos estadísticos para determinar la validez de un instrumento, en este trabajo se resaltan las bondades del método Delphi, que siendo una propuesta cualitativa, permite determinar la validez de contenido de un instrumento de medición. Asimismo, Este método también es un buen indicador para depurar la lista preliminar de las características listadas. Adicionalmente, y con el ánimo de asegurar la confiabilidad y consistencia interna, se calculó el coeficiente de confiabilidad en cuatro métodos diferentes, obteniendo un coeficiente promedio de 0.859, lo que indica que el instrumento es apropiado y satisfactorio.

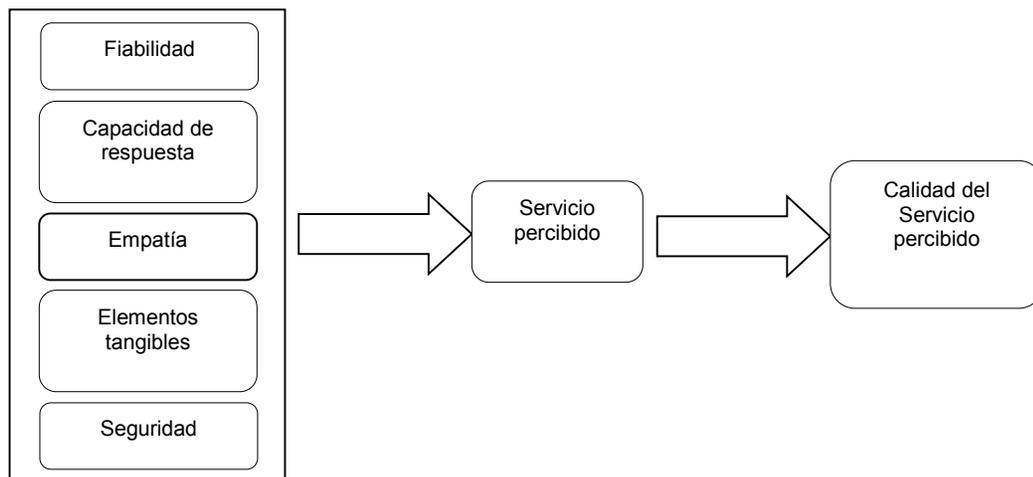
REVISIÓN LITERARIA

Con la medición de calidad percibida, entendida como “el proceso de evaluación donde el cliente compara la experiencia del servicio con determinadas expectativas previas” Pérez (2006, p.29), los dirigentes deportivos podrán conocer las percepciones, necesidades y expectativas de los espectadores (Delgado (2004, p.115), aspectos que permiten la identificación del nivel de satisfacción de sus usuarios y el diseño de mejores estrategias de marketing de servicios deportivos. Aunque existen gran variedad de propuestas para medir la calidad percibida de los servicios en diferentes sectores económicos, para el desarrollo de este trabajo se tomaron como referentes los modelos SERVQUAL, SERVPERF y el Modelo conceptual del servicio con base en las percepciones del cliente. Así mismo, se revisaron estudios donde han surgido nuevas dimensiones para evaluar la calidad percibida del servicio en los eventos deportivos. En el modelo SERVQUAL desarrollado por Parasuraman, Zeithaml & Berry (1988) consistente en una escala conformada por 22 ítems agrupados en cinco dimensiones del servicio: Fiabilidad, garantía, tangibilidad, empatía y sensibilidad, se mide la satisfacción de los clientes a partir de las brechas identificadas entre las expectativas y las percepciones del servicio. “La escala propuesta por los autores se ha convertido en una de las medidas más utilizadas para evaluar calidad del servicio” (Curubeto, 2001, p.138). Christian Grönroos (1984, 1988, 2001, 2007) desarrolla y prueba el modelo conceptual del servicio con base en las percepciones del cliente denominado Modelo Nórdico, puesto que para él, la calidad del servicio es influenciada por las expectativas de la calidad (conformada por las comunicaciones de la empresa, su imagen, las experiencias y las necesidades y deseos del cliente) y por la calidad experimentada, que es la imagen formada después del contacto con el servicio. En este enfoque, la calidad total percibida es la etapa final de la transacción entre la organización y un cliente. La percepción es una función de la calidad técnica (ítems del servicio) y de la calidad funcional (la calidad de la interacción con la empresa).

En el año 1992, Joseph Cronin y Steven Taylor critican fuertemente al modelo SERVQUAL porque consideran que la escala no se ajusta a todas las clases de servicios, y sugieren que la calidad de servicio debería medirse a partir de las actitudes. Es así, que como resultado de sus investigaciones surge una escala alternativa denominada SERVPERF (*Service Performance*). Como se observa en la Figura 1, este modelo se basa en las percepciones de la calidad del servicio a partir de las cinco dimensiones de la calidad del servicio: fiabilidad, capacidad de respuesta, empatía, elementos tangibles y seguridad. Aunque los autores se basan en el modelo SERVQUAL, solo se concentran en medir las percepciones, excluyendo las expectativas porque consideran que son factores externos que pueden sesgar la percepción de los beneficiarios del servicio.

Kim & Kim (1995) siguiendo el modelo de Parasuraman, Zeithamal y Berry realizan una investigación para evaluar la calidad de los servicios prestados en los centros deportivos de Corea. Como resultado de la investigación surge el instrumento QUESC conformado por 33 ítems agrupados en las siguientes 12 dimensiones: el ambiente, actitud de los empleados, fiabilidad, información, programación, consideración personal, precio, exclusividad, privacidad, conveniencia, estimulación y oportunidad social. Calabuig, Quintanilla & Mundina (2008) se basaron en los modelos SERVQUAL, SERVPERF y en el instrumento Neptuno-1 utilizado en investigaciones realizadas por Luna, Mundina & Gómez (1998), Mundina *et al.* (1998) y Mundina & Calabuig (1999). La aplicación del instrumento les permitió determinar el grado de calidad percibida en los servicios deportivos de carácter público como son las escuelas de actividades náuticas e identificar diez dimensiones relevantes: las clases, el tiempo libre, el entorno náutico, los horarios, la comida, el material náutico, la comida complementaria, la conserjería, la limpieza y la teoría.

Figura 1: Representación Gráfica del Modelo SERVPERF



Dimensiones de la calidad. Ésta figura hace alusión al desempeño que tiene el proveedor en el encuentro de servicios, lo que permite medir de forma objetiva la calidad del servicio. En la representación, se pueden observar las cinco dimensiones propuestas por Cronin & Taylor que agrupan los 22 ítems definidos para evaluar la percepción. Fuente: Adaptado de Losada & Rodríguez (2007)

Calabuig & Crespo (2009) utilizan el método Delphi para la elaboración de una escala de medida de la calidad percibida de los espectadores de eventos deportivos denominada EVENTQUAL, obteniendo como resultado 22 ítems agrupados en cinco factores: los tangibles, la accesibilidad, el personal, los servicios complementarios y la relación entre espectadores. Kelley & Turley (2001) utilizaron el modelo SERVQUAL para medir la calidad del servicio en un evento deportivo. En este estudio se identifica la experiencia del juego y el papel de los empleados como ítems significativos del servicio. Hightower, Brady & Baker (2002) centraron su investigación en la influencia de los distintos aspectos del entorno físico durante un evento deportivo de béisbol encontrando que el ambiente del servicio (la evidencia física) es un atributo determinante para que las personas asistan en un futuro a eventos deportivos.

Asimismo, Greenwell, Fink & Pastore (2002) observan que el ambiente del servicio influye en la satisfacción de los espectadores de hockey hielo de un modo global más que evaluándolos como elementos individuales (como son la comodidad o los accesos). Finalmente, Yusof y Lee (2008) realizaron un estudio en los estadios de fútbol conociendo que los elementos tangibles del evento (accesibilidad en toda la instalación, estética del estadio, marcadores, comodidad de los asientos, suficiencia de espacio, señalización, etc.) y la calidad del equipo (buen juego del equipo, jugadores estrella, registro victorias/derrotas, historia y palmarés del club, calidad del oponente y jugadores estrella del oponente, etc.) son predictores de la satisfacción de los espectadores. Del análisis realizado, se encontró que los estudios tomados como referencia se basaron en el Modelo SERVQUAL elaborado por Parasuraman, Zeithaml & Berry. Otros de los autores más citados en los estudios son Kim & Kim (1995) quienes elaboraron una escala más ajustada a la medición de la calidad de servicios. Para la elaboración de la escala preliminar, se consideraron las dimensiones de tangibilidad o evidencia física (Hightower *et al.*, 2002; Yusof & Lee, 2008), la accesibilidad en toda la instalación, estética del estadio, marcadores, comodidad de los asientos, suficiencia de espacio, señalización y la calidad del equipo (Kim & Kim, 1995) y la actitud de los empleados (Calabuig, Quintanilla & Mundina, 2008; Calabuig & Crespo, 2009 y Kelley & Turley, 2001). Las otras dimensiones consideradas son recurrentes en los estudios revisados, tales como: fiabilidad, información, programación, precio, exclusividad, privacidad, conveniencia, estimulación y oportunidad social.

METODOLOGÍA

Para lograr un instrumento idóneo y que se pueda aplicar con toda confianza, se requiere que cumpla con dos requisitos: Validez y confiabilidad (González, 2008). Dadas las características cualitativas de las variables a evaluar en el estudio, la validez del contenido se realizó mediante el uso del método Delphi y algunos estadísticos descriptivos y para calcular la confiabilidad interna se consideraron cuatro métodos estadísticos que al comparar los resultados se logra comprobar que existe consistencia entre los ítems evaluados. Para ello se utilizaron los métodos estadísticos de Alpha de Cronbach (1951); Dos mitades, corregido por la fórmula de Spearman-Brown (1919); Rulón (1939) y Hoyt (1941) (Anova). Para calcular el coeficiente de confiabilidad, se realizó un sondeo mediante la aplicación del instrumento a 89 personas que han asistido a eventos deportivos en los últimos seis meses. Para el cálculo de estos índices se empleó el paquete estadístico SPSS y el Excell.

Primera Fase: Construcción de la Escala Preliminar de Variables

Identificar los ítems principales que serán incluidos en la lista preliminar, es la primera función que debe realizar el comité investigador, puesto que de ésta depende el éxito de la investigación. Para la construcción de la primera lista de ítems se desarrollaron las siguientes actividades: Interpretación y análisis de los modelos SERVQUAL, SERVPERF y Nórdico, revisión de algunos trabajos sobre medición de la calidad percibida del servicio y cuatro entrevistas a dirigentes deportivos. Como resultado, en esta primera fase, el comité investigador elaboró la primera lista preliminar conformada por 71 ítems, la cual se organizó en un cuestionario con escala tipo Likert de cinco puntos, siendo el 1 equivalente a *totalmente en desacuerdo* y el 5 *totalmente de acuerdo*. Al finalizar el cuestionario, se incluyó la opción a los entrevistados de modificar o eliminar ítems.

Segunda Fase: Validez de Contenido del Instrumento de Medición

Para la validez del contenido del instrumento se empleó el método Delphi y para la selección se las variables, se emplearon la media y desviación estándar. Aunque el método Delphi surge para realizar análisis prospectivo, es una técnica cualitativa fácilmente aplicable a otro tipo de investigaciones, tales como unificar criterios sobre problemas organizacionales y definir factores o indicadores de evaluación, entre otros. Es así que, dadas sus características de anonimato, respuesta numérica y retroalimentación de criterios hasta llegar a un consenso, se consideró como la metodología que más se ajusta al objetivo del proyecto, consistente en la definición de los ítems para medir de manera objetiva la calidad percibida del servicio de eventos deportivos. Según Landeta (citado por Ortega, 2008) la técnica Delphi consiste en “la compilación de las opiniones y comentarios de uno o varios grupos de personas que tienen una estrecha relación sobre la cuestión, sector, tecnología, [...] objeto de la investigación” (p.34). Los miembros del comité de expertos, al conocer las opiniones de los otros miembros, pueden mantener, modificar o ampliar sus opiniones hasta que se logre un consenso en las respuestas. En este proyecto, el método Delphi se desarrolló en tres etapas así: a) conformación del comité de expertos, b) recolección de la información y c) análisis y resultados.

Conformación del Comité de Expertos

Para la selección de las personas que conformaron el comité de expertos se consideraron los siguientes aspectos: experiencia en el área los eventos deportivos, la disponibilidad de tiempo e interés en participar en la investigación. Según Landeta (citado por Ortega, 2008), el experto es “aquel individuo cuya situación y recursos personales le posibiliten contribuir positivamente a la consecución del fin que ha motivado la iniciación del trabajo Delphi” (p.36), y que, de acuerdo con lo indicado por Pill (citado por Ortega, 2008), puede a su vez aportar una perspectiva relevante a la investigación. Respecto al número de expertos, Malla & Zabala (citados por Ortega, 2008) “consideran que un grupo de 10 a 20 expertos puede ser

metodológicamente fiable” (pp. 36-37). Para el desarrollo de este trabajo, el comité de expertos se conformó un equipo de diez personas: cuatro empleados con cargos directivos, cuatro espectadores y dos estudiantes de marketing de servicios.

Recolección de la Información

En este trabajo, la información se recolecta de manera individual mediante el uso del correo electrónico entre el 1 de junio y 30 de agosto de 2013. En la primera ronda, el comité de investigación se encarga de enviar el cuestionario, recoger las respuestas y compararlas con los otros miembros del comité de expertos; las respuestas afines no se alteran pero las preguntas en donde aún existe divergencia se envían de nuevo para solicitar a los expertos que revisen y reconsideren la respuesta. El comité de investigación depura la lista, selecciona las respuestas en consenso y las aprueban. Aquellas donde hay divergencia se organizan y se regresan para que el comité vuelva a evaluarlas. A la relación inicial de las preguntas se pueden adicionar las propuestas hechas por algunos participantes. En la próxima ronda se debe informar a todos los miembros de comité que estas nuevas propuestas provienen de algunos participantes del grupo, por tanto, se les recomienda evaluarla y expresar si se deben o no incluir en la escala de medición. En la Tabla 1 se detallan las actividades y el periodo de tiempo de recolección de la información.

Tabla 1: Cronograma de Recolección de Información y Análisis de Resultados

Proceso	Actividades Relacionadas	Fecha de Aplicación
Primera ronda	Se excluyen 20 ítems (de 71 a 51)	Del 1 al 8 de junio de 2013
Segunda ronda	Se excluyen 8 ítems (de 51 a 43)	Del 17 al 21 de junio de 2013
Tercera ronda	Se excluyen 16 ítems (de 43 a 27)	Del 1 al 30 de julio de 2013
Análisis de confiabilidad	Se realiza el análisis de confiabilidad de consistencia interna	Del 1 al 30 de agosto de 2013
Aprobación de instrumento definitivo	Revisión y aprobación entre los miembros del comité de expertos	15 de septiembre de 2013

En la Tabla se detallan las etapas y plazos (cronograma) que transcurren para la recolección de información y su respectivo análisis.

Tercera Fase: La Confiabilidad de Consistencia Interna

Sánchez & Sarabia (citados por Calabuig, Molina & Núñez, 2012) exponen que “una escala es fiable cuando permite obtener medidas similares en distintos momentos o, posee una baja variabilidad interna”. Para ello se utilizaron los métodos estadísticos de Alpha de Cronbach (1951); Dos mitades, corregido por la fórmula de Spearman-Brown (1919); Rulón (1939) y Hoyt (1941) (Anova). Para calcular el coeficiente de confiabilidad, se realizó un sondeo mediante la aplicación del instrumento a 89 personas que han asistido a eventos deportivos en los últimos seis meses y para su análisis se empleó el programa estadístico SPSS. Dentro de la categoría de coeficientes, Alfa de Cronbach es el más empleado por los investigadores donde se estima el límite inferior del coeficiente de fiabilidad y se expresa como:

$$\alpha = \left[\frac{k}{k-1} \right] \left[1 - \frac{\sum S_i^2}{S_{sum}^2} \right] \quad (1)$$

Donde k es el número de ítems de la prueba, S_i^2 es la varianza de los ítems (desde 1...i), S_{sum}^2 es la varianza de la prueba total. El coeficiente mide la fiabilidad del test en función de dos términos: el número de ítems (o longitud de la prueba) y la proporción de varianza total de la prueba debida a la covarianza entre sus partes (ítems). “Ello significa que la fiabilidad depende de la longitud de la prueba y de la covarianza entre sus ítems” (Ledezma, Molina y Valero, 2002). El valor mínimo aceptable para el coeficiente de Alfa de Cronbach es de 0,7; por debajo de ese valor la consistencia interna de la escala utilizada es baja (Celina y campo 2005). Este valor muestra la correlación entre cada uno de los ítems; un valor superior a 0,7 revela

una fuerte relación entre los ítems, un valor inferior revela una débil relación entre ellas. Lucero y Meza (2002) mencionan que el valor mínimo aceptable del coeficiente de fiabilidad depende de la utilización del instrumentos, es decir dependiendo de la exactitud requerida por la disciplina que lo requiere. El Método de Dos Mitades (corregido por la fórmula de Spearman–Brown) fue publicado independientemente por Spearman (1910) y Brown (1910) expresado en la siguiente formula:

$$r_{tt} = \frac{2r_{hh}}{1+r_{hh}} \quad (2)$$

En donde, r_{tt} es el coeficiente de confiabilidad y r_{hh} es la correlación entre las dos mitades. Primero, se dividen los ítems de la prueba en dos partes iguales; seguidamente, se correlacionan las puntuaciones totales de las dos mitades y se multiplica el coeficiente obtenido por 2; finalmente, se dividen por el término 1 (uno) más la correlación de las dos mitades. Rulón (1939) desarrollo un modelo alternativo para calcular la confiabilidad por el método de las dos mitades. En este caso se calcula la diferencia de varianza de las dos mitades para cada sujeto, se divide por la varianza de los puntajes y se resta esta proporción de la unidad (1,0). La fórmula es:

$$r_{tt} = 1 - \frac{S_d^2}{S_t^2} \quad (3)$$

En donde: r_{tt} es el coeficiente de confiabilidad, S_d^2 es la varianza diferencial entre las dos mitades del test y S_t^2 es la varianza total de la prueba. El método de Hoyt (1941) para calcular la confiabilidad de consistencia interna, consiste en aplicar el método de análisis de varianza (Anova), para lo cual, se elabora una matriz de ítem por sujeto; se calcula la varianza total de la distribución de puntajes totales; se computan los diferentes componentes de la varianza total: varianza debida a los sujetos (V_s), varianza debida a los ítems (V_i) y varianza residual o de error (V_e); se divide V_s/V_e para obtener la F_s ; y V_i/V_e para obtener la F_i ; y se determina la significación estadística de cada F. La fórmula de Hoyt es la siguiente:

$$r_{tt} = 1 - \frac{S_e^2}{S_s^2} \quad (4)$$

En donde: r_{tt} es el coeficiente de confiabilidad, S_e^2 es la varianza de error y S_s^2 es la varianza debida a os sujetos.

RESULTADOS

Primera Fase: Construcción Preliminar de la Escala de Medición

Al revisar los modelos SERVQUAL, SERVPERF y Nórdico se decidió tomar como referente inicial el modelo SERVPERF dado que se concentra, a diferencia de los otros modelos, en evaluar exclusivamente las percepciones; este modelo propone 22 ítems agrupados en cinco dimensiones: tangibilidad, confiabilidad, capacidad de respuesta, seguridad y empatía. Asimismo, se consideraron los opiniones de los expertos deportivos y los aspectos teóricos. Para la construcción de la primera lista de ítems, además de los propuestos en el modelo SERVPERF, el comité investigador estuvo a cargo de adicionar aquellos criterios que consideraba importantes y que podían ayudar a complementar la evaluación de la calidad percibida de los eventos deportivos. Por otro lado, el mismo equipo realizó una revisión detallada de cada ítem, teniendo en cuenta aspectos como la ortografía, redacción, palabras o frases repetidas. En esta etapa se obtuvo como resultado una lista conformada por 71 ítems, los cuales fueron enviados al comité de expertos (vía correo electrónico) para su respectiva evaluación, ajuste y aprobación.

Primera Fase: Validez de Contenido del Instrumento de Medición.

Para someter la escala de medición a la primera evaluación se construyó una lista de chequeo que utilizaba una escala de respuesta tipo Likert de cinco puntos, donde una evaluación entre uno (1) y tres (3) significa que el atributo se debe excluir de la lista y una evaluación entre cuatro (4) y cinco (5) significa que el atributo es altamente relevante y debe continuar en la lista de ítems. En esta misma lista de chequeo los expertos también tenían la posibilidad de incluir, modificar o eliminar ítems. Los criterios utilizados para realizar la depuración de la lista de ítems fueron (ver Tabla 2) la exclusión de los ítems con una calificación promedio inferior a 3.0 y la unificación de los ítems que presentaban similitudes en su redacción y evaluaban el mismo aspecto de la calidad percibida de eventos deportivos.

Tabla 2: Criterios Utilizados Para la Depuración de Ítems de la Escala de Medición V1

No. de Ítems Iniciales	Criterio de Depuración	No. de Criterios Excluidos
71	Calificación promedio inferior a 3.0	10
61	Unificación de acuerdo a los comentarios de los expertos	10
51	Lista de chequeo resultante de la primera ronda	

En la Tabla se detallan los criterios utilizados por el comité investigador para depurar la escala de medición V1, que para este caso se excluyeron los ítems con una calificación promedio inferior a 3,0 y se unificaron los que presentaban similitudes teniendo en cuenta los comentarios del comité de expertos.

Resultados de la Segunda Ronda

La nueva lista de chequeo conformada por 51 ítems (escala de medición V2), fue revisada por los integrantes del comité investigador, quienes se encargaron de organizarla y enviarla nuevamente al comité de expertos para su respectiva revisión, ajuste y aprobación. En esta versión se incluyeron las calificaciones asignadas en la ronda anterior con el fin de suministrar la información suficiente a los expertos, la cual les permitiera llegar a un consenso. Después de recibir las respuestas por parte de los expertos, el comité investigador procedió a realizar un análisis a través del promedio de las calificaciones de cada ítem. Considerando lo anterior, el comité investigador decidió excluir aquellos ítems que presentaban calificaciones promedio inferiores a 3.5. Posteriormente, se unificaron los ítems que presentaban semejanzas en su redacción teniendo en cuenta los comentarios realizados por el panel de expertos. Dichos criterios de depuración dieron como resultado un listado de 43 ítems (ver Tabla 3).

Tabla 3: Criterios Utilizados Para la Depuración de Ítems de la Escala de Medición V2

No. de Ítems Iniciales	Criterio de Depuración	No. de Criterios Excluidos
51	Calificación promedio inferior a 3.5	2
49	Unificación de acuerdo a los comentarios de los expertos	6
43	Lista de chequeo resultante de la segunda ronda	

En la anterior Tabla se realiza un resumen de los criterios utilizados por el comité investigador para depurar la escala de medición V1, excluyendo aquellos ítems con calificación inferior a 3,5. Igualmente, se unificaron los ítems similares.

Resultados de la Tercera Ronda

En esta última ronda, el comité investigador unificó los criterios que presentaban semejanzas en su redacción a partir de las recomendaciones dadas por el panel de expertos (Tabla 4). Igualmente, se adicionaron dos criterios que los expertos consideraron relevantes, dadas las experiencias obtenidas en eventos realizados con anterioridad. Finalmente se obtuvo una escala de 27 ítems (ver Tabla 5)

Tabla 4: Criterios Utilizados Para la Depuración de Ítems de la Escala de Medición V3

No. de Ítems Iniciales	Criterio de Depuración	No. de Criterios
43	Unificación de acuerdo a los comentarios de los expertos	18
25	Se adicionaron ítems	2
27	Lista de chequeo resultante de la tercera ronda	

En la anterior Tabla se realiza un resumen de los criterios utilizados por el comité investigador para depurar la escala de medición V2, donde se unificaron ítems similares y se adicionaron dos nuevos ítems.

Tabla 5: Escala de Medición de la Calidad Percibida de Eventos Deportivos

No	Ítems	No	Ítems
1	La instalación es ideal para la práctica de la actividad deportiva	14	La señalización de las instalaciones facilitó la orientación y ubicación de los espectadores
2	El arbitraje de la actividad deportiva ha sido justo	15	La instalación es ideal para la práctica de la actividad deportiva
3	El nivel ofrecido por los deportistas fue de alto rendimiento	16	El resultado de la competición superó sus expectativas
4	La calidad del evento deportivo superó sus expectativas	17	Recibió más de lo que esperaba por el precio cancelado
5	Los empleados le atendieron con amabilidad, respeto y paciencia	18	La instalaciones permanecieron limpias durante el evento
6	Le permitieron expresar sus emociones sin problemas	19	Las instalaciones cuentan con servicio de cafetería y variedad de alimentos
7	La información durante el transcurso del evento fue permanente	20	Se sintió identificado con los demás seguidores de su equipo/deportista
8	La ubicación de la gradería le facilitó la visibilidad del evento	21	No se generan colas para aparcar
9	Le impresionaron las instalación deportivas	22	La instalación física es cómoda y acogedora
10	La instalación le permitió desplazarse con facilidad por el interior	23	El número de empleados fue suficientes para atender a los espectadores
11	Tuvo la oportunidad de comprar la entrada con anticipación	24	Los puertas de accesos permitieron que la cola circulara con rápidamente
12	La información sobre disponibilidad y precios de las entradas fue clara	25	Le ofrecieron diferentes opciones para comprar las boletas de entrada
13	El parqueadero está ubicado cerca a las instalaciones del evento	26	El parqueadero fue suficiente para los espectadores que asistieron al evento
		27	La seguridad ofrecida durante el desarrollo del evento

Como resultado de la fase cualitativa, se conformó una escala para medir la calidad percibida de los eventos deportivos que incluye los ítems propuestos por el grupo de individuos consultado. Esta escala se le realizó la prueba piloto y el análisis de confiabilidad.

La Confiabilidad de Consistencia Interna del Instrumento

Con el propósito de realizar la revisión final de la lista de ítems y teniendo en cuenta los resultados de la ronda anterior, se construyó un cuestionario piloto que fue aplicado a 89 individuos, los cuales se caracterizaban porque asistieron, como mínimo a un evento deportivo en los últimos seis meses. Asimismo, En el instrumento, se asignó un espacio para que los espectadores describieran si se deben adicionar, modificar o eliminar algún atributo. En este caso, como se muestra en la Tabla 5 se obtuvo un Alfa de 0.816, lo que indica que los ítems del instrumento son consistentes y homogéneos.

Tabla 5: Estadísticos de Fiabilidad

Alfa de Cronbach	Alfa de Cronbach Basada en los Elementos Tipificados	No. de Elementos
0.816	0.819	27

En Alfa De Cronbach Obtenido Es De 0.86 Para Los 27 Ítems Incluidos En La Escala.

En esta etapa, se eliminaron aquellos ítems que generan un incremento significativo en el Alpha de Cronbach (1951) del total de ítems (superior a 0.816 para obtener mayor fiabilidad del instrumento) y se agruparon, de acuerdo a los comentarios de los individuos, dos ítems que presentaban similitudes en su redacción y evaluaban el mismo aspecto de la calidad percibida de eventos deportivos (ver Tabla 6) La Tabla 5 muestra que el Alfa de Cronbach total del instrumento fue de 0.816 superior al mínimo aceptado para un total de 27 ítems (elementos) evaluados, lo anterior refleja un alto grado de congruencia interna de la escala. La media obtenida en cada uno de los elementos indica el índice de importancia de cada uno de ellos para la escala con respecto a la opciones de respuesta, las medias oscilan entre 2.85 (ítem 23) y 4.85(ítem 15). La desviación típica o estándar mide el grado de dispersión de las observaciones individuales alrededor de la media. En este caso, seis ítems (2, 3, 6, 16, 19, 20, 21) presentan una media inferior a 4.0 y la mayor desviación estándar. Desde esta perspectiva, estos ítems fueron modificados o eliminados, como se muestra en la Tabla 6.

Tabla 6: Criterios Utilizados Para la Depuración de Ítems de la Escala de Medición V4

No. de Ítems Iniciales	Criterio de Depuración	No. de Ítems Excluidos
27	Calificación promedio inferior a 4.0	6
21	Según alpha de Cronbach	1
20	Criterios agrupados	2
18	Escala resultante	

En la anterior Tabla se realiza un resumen de los criterios utilizados por el comité investigador para depurar la escala de medición V3, excluyendo aquellos ítems con calificación inferior a 4.0, con un coeficiente de Cronbach igual a 0.816; finalmente, se unificaron los ítems similares.

Para comprobar estos resultados se calculó el coeficiente de confiabilidad de consistencia interna mediante el uso de los métodos de Dos Mitades (Spearman-Brown, 1910), método de Hoyt (1941) y el método de Rulón (1939) obteniendo coeficientes de 0.897; 0.816 y 0.9 respectivamente. En la Tabla 7 y Tabla 8 se presentan los estadísticos de fiabilidad de los métodos Dos Mitades (Spearman-Brown, 1910) y Hoyt (1941)

Tabla 7: Estadísticos de Fiabilidad Método de Dos Mitades (Spearman-Brown, 1910)

Alfa De Cronbach	Parte 1	Valor	0.715
		N de elementos	14
	Parte 2	Valor	0.587
		N de elementos	13
	N total de elementos		27
Correlación entre formas			0.814
Coeficiente de Spearman-Brown	Longitud igual		0.897
	Longitud desigual		0.897
Dos mitades de Guttman			0.890

En Alfa de Cronbach se calcula para dos partes, las cuales corresponden en dividir los ítems en dos mitades, la parte 1 contiene 14 ítems y la parte 2 contiene 13 ítems para un total de 27 ítems. Las Dos mitades de Spearman-Brown corresponden al coeficiente confiabilidad de 0.897.

Tabla 8: Resumen de ANOVA, Hoyt (1941)

Fuente de Variación	Suma de Cuadrados	Gl	Media Cuadrática	F	Sig.
Sujetos	443.178	86	5.153		
Ítem	510.407	26	19.631	20.656	0.000
Residual	2125.074	2236	0.950		
Total	3078.660	2348	1.311		

En la fuentes de variación, los sujetos se refieren al número de personas encuestadas, los ítems al número de variables que conforma la escala, en la columna media cuadrática, se observa la varianza del sujeto (5.153) y la varianza del error (0.95). Estos valores son los utilizados para calcular el coeficiente de confiabilidad.

En conclusión, se resalta que los diferentes métodos empleados para estimar la confiabilidad de consistencia interna dan resultados similares que oscilan entre 0.816 y 0.908, lo que indican que los ítems incluidos en el instrumento tienen un coeficiente de confiabilidad alto, que en promedio da 0.859

Tabla 9: Coeficientes Promedio de Confiabilidad de Consistencia Interna Estimada Por Cuatro Métodos

Método	Coeficiente de Confiabilidad
Alpha de Cronbach	0.816
Spearman-Brown (Dos mitades)	0.897
Rulón	0.908
Hoyt	0.816
Promedio	0.859

En la tabla se presentan los coeficientes obtenidos a partir del cálculo de cada uno de los métodos analizados, los cuales se promediaron obteniendo un valor de alfa de 0.859. Lo que indica que la consistencia interna del instrumento es altamente confiable.

Escala Definitiva

La escala de 18 ítems obtenida hasta el momento, se envió nuevamente al comité de expertos, logrando un consenso, lo que permite llegar a una escala definitiva conformada por 18 ítems (ver Tabla 10).

Tabla 10: Escala de Medición de la Calidad Percibida de Eventos Deportivos

No	Ítems	No	Ítems
1	La instalación es ideal para la práctica de la actividad deportiva	10	El parqueadero está ubicado cerca a las instalaciones del evento
2	La calidad del evento deportivo superó sus expectativas	11	La señalización de las instalaciones facilitó la orientación y ubicación de los espectadores
3	Los empleados le atendieron con amabilidad, respeto y paciencia	12	La instalación es ideal para la práctica de la actividad deportiva
4	La información durante el transcurso del evento fue permanente	13	Recibió más de lo que esperaba por el precio cancelado
5	La ubicación de la gradería le facilitó la visibilidad del evento	14	La instalaciones permanecieron limpias durante el evento
6	Le impresionaron las instalación deportivas	15	La instalación física es cómoda y acogedora
7	La instalación le permitió desplazarse con facilidad por el interior	16	El número de empleados fue suficientes para atender a los espectadores
8	Tuvo la oportunidad de comprar la entrada con anticipación	17	Los puertas de accesos permitieron que la cola circulara con rápidamente
9	La información sobre disponibilidad y precios de las entradas fue clara	18	La seguridad ofrecida durante el desarrollo del evento

Como resultado final se conformó una escala definitiva para medir la calidad percibida de los eventos deportivos. Con esta escala se construyó el instrumento que será aplicado en la fase cuantitativa de esta investigación.

Adicional a esta información, también se incluyeron en el cuestionario definitivo preguntas demográficas y se asignó un espacio para que los usuarios presenten sus sugerencias o recomendaciones.

CONCLUSIONES

La necesidad de los directivos deportivos de diseñar e implementar proyectos que se ajusten a las expectativas de los usuarios, propicia la elaboración de este trabajo investigativo consistente en la creación de una escala de medición de la calidad percibida de los eventos deportivos internacionales mediante el uso del método Delphi. El método Delphi es una técnica que se puede aplicar en la fase previa de cualquier investigación que requiera la construcción de un instrumento de medición con escalas tipo Likert. Este método ofrece la oportunidad de llegar a consenso con la participación de personas conocedoras de la actividad de la empresa. Para el análisis definitivo, el método Delphi recurre a la media y desviación estándar. Para el desarrollo de este trabajo se revisaron los modelos SERVQUAL, SERVPERF y el Modelo conceptual del servicio con base en las percepciones del cliente. El modelo SERVQUAL desarrollado por Parasuraman, Zeithaml & Berry (1988) se mide la satisfacción de los clientes a partir de las brechas identificadas entre las expectativas y las percepciones del servicio mediante una escala conformada por 22 ítems agrupados en cinco dimensiones del servicio: Fiabilidad, garantía, tangibilidad, empatía y sensibilidad. Dado que el interés del trabajo es medir las percepciones de los espectadores, se seleccionó la escala SERVPERF (Service Performance) propuesta por Joseph Cronin y Steven Taylor (1992) porque ellos excluyen las expectativas pues consideran que son factores externos que pueden sesgar la percepción de los beneficiarios del servicio. Esta escala se basa en cinco dimensiones: fiabilidad, capacidad de respuesta, empatía, elementos tangibles y seguridad.

En cuanto a los eventos deportivos, los estudios realizados con anterioridad presentan una gran variedad de escalas que oscilan entre cuatro y doce dimensiones. Estos estudios se han basado en los modelos de SERVQUAL Y SERPERF para construir escalas propias de medición. Entre los más consultados se encuentran: la escala QUESC de Kim & Kim (1995) conformada por 33 ítems agrupados en las siguientes 12 dimensiones: el ambiente, actitud de los empleados, fiabilidad, información, programación, consideración personal, precio, exclusividad, privacidad, conveniencia, estimulación y oportunidad social y el instrumento Neptuno-1 elaborado por Calabuig, Quintanilla & Mundina (2008) quienes se concentran en medir las clases, el tiempo libre, el entorno náutico, los horarios, la comida, el material náutico, la comida complementaria, la conserjería, la limpieza y la teoría. Adicionalmente, se resalta el estudio de Calabuig & Crespo (2009) por el uso del método Delphi para la elaboración de una escala de medida de la calidad percibida de los espectadores de eventos deportivos denominada EVENTQUAL, obteniendo como resultado 22 ítems agrupados en cinco factores: los tangibles, la accesibilidad, el personal, los servicios complementarios y la relación entre espectadores.

De los ítems que conformaron la escala definitiva, 17 de ellos se identificaron en los estudios previos y en el modelo SERVPERF, el cual fue tomado como referencia porque se concentra en la evaluación de las percepciones de los usuarios con relación a la calidad del servicio. El atributo No. 18 denominado “La seguridad ofrecida durante el desarrollo del evento” es consistente con la necesidad de incrementar la vigilancia interna en los escenarios deportivos. Finalmente, como resultado en esta fase del estudio, se elabora una versión del cuestionario con 18 ítems, con validez de contenido y confiabilidad de consistencia interna para medir el constructo de la calidad percibida de los servicios deportivos en escenarios al cubierto. Para realizar la validez del contenido, se realizaron tres rondas con el comité de expertos. Cabe señalar que las principales limitaciones para desarrollar el proyecto fueron la poca disponibilidad de tiempo de algunos expertos convocados para la evaluación de la escala, debido a las múltiples ocupaciones que manejan y los cortos plazos para la remisión de los resultados de la evaluación; y para determinar la consistencia interna de la escala, se realizó una prueba piloto una muestra de 89 usuarios del servicio, la cual se analizó mediante el uso de cuatro métodos de confiabilidad de consistencia interna. La confiabilidad promedio es de 0.859, valor que reflejan una buena confiabilidad del instrumento.

Igualmente, se espera que esta metodología sirva de referente para futuras investigaciones sobre medición de la calidad percibida del servicio en diferentes sectores de servicios, puesto que la escala diseñada surge ante la ausencia de un instrumento adecuado que permita evaluar dicho aspecto en los eventos deportivos que se celebran en Colombia.

ANEXOS

El cuestionario y escalas preliminares están disponibles bajo petición. Se recomienda contactarse con los autores.

BIBLIOGRAFÍA

Alcaldía de Santiago de Cali. (s.f.). Historia del deporte vallecaucano. Recuperado el Marzo de 2014, de http://calimundial.gov.co/Site/Historia_Deportiva.php

Calabuig, F. (2006). La calidad percibida en los servicios náuticos de la Generalitat Valenciana (Tesis doctoral). Universidad de Valencia, España.

Calabuig, F., & Crespo, J. (2009). Uso del método Delphi para la elaboración de una medida de la calidad percibida de los espectadores de eventos deportivos. *Retos. Nuevas tendencias en Educación Física, Deporte y Recreación*(15), 21-25.

Calabuig, F., Burillo, P., Crespo, J., Mundina, J. J., & Gallardo, L. (2010). Satisfacción, calidad y valor percibido en espectadores de atletismo. *Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte*, 10(40), 577-593.

Calabuig, F., Molina, N., & Núñez, J. (2012). Una aplicación del modelo Tridimensional de Calidad de Servicio en centros deportivos privados. *Revista de Ciencias del Deporte*, 8(1), 67-81.

Calabuig, F., Quintanilla, I., & Mundina, J. (2008). La calidad percibida de los servicios deportivos: diferencias según instalación, género, edad y tipo de usuario en servicios náuticos. *International Journal of Sport Science*, 4(10), 25-43.

Cronin, J. J., & Taylor, S. A. (1992). Measuring Service Quality: A Reexamination an Extension. *Journal of Marketing*, 56(3), 55-68.

Curubeto, C. (2001). La marca universitaria. Ítems, beneficios o valores, ¿qué utilizan las universidades de Buenos Aires para construir su identidad de marca y su propuesta de valor?, (Tesis doctoral). Universidad Austral de Argentina, Buenos aires.

Delgado, E. (2004). La marca como activo estratégico. En A. I. Jiménez, H. Calderón, E. Delgado, J. C. Gásquez, M. A. Gómez, C. Lorenzo, & L. M. Zapico, Dirección de productos y marcas (págs. 101-134). Barcelona: Editorial UOC.

Dorado, A., & Gallardo, L. (2007). La satisfacción del Consumidor Deportivo en los Servicios Públicos. *Investigación y Marketing*(83), 26-30.

Greenwell, T. C., Fink, J., & Pastore, D. (2002). Assessing the Influence of the Physical Sports Facility on Customer Satisfaction within the Context of the Service Experience. *Sport Management Review*, 5(2), 129-148.

Grönroos, C. (1984). A Service Quality Model and Its Marketing Implications. *European Journal of Marketing*, 18(4), 36-44.

Grönroos, C. (1988). Service Quality: The Six Criteria of Good Perceived service Quality. *The El Paso business review*, 9, 10-13.

Grönroos, C. (2001). The perceived service quality concept - a mistake? *Managing Service Quality*, 11(3), 150-152.

Grönroos, C. (2007). *Service Management and Marketing. Customer Management in Service Competition*. West Sussex: Wiley.

Hightower, R., Brady, M., & Baker, T. (2002). Investigating the role of the physical environment in hedonic service consumption: an exploratory study of sporting events. *Journal of Business Research*, 55, 697-707.

Kelley, S., & Turley, L. W. (2001). Consumer perceptions of service quality attributes at sporting events. *Journal of Business Research*, 54(2), 161-166.

Kim, D., & Kim, S. Y. (1995). QUESC: An Instrument for Assessing the Service Quality of Sport Centers in Korea. *Journal of Sport Management*, 9(2), 208-220.

Ledesma, R., Molina, G., & Valero, P. (2002). Análisis de consistencia interna mediante Alfa de Cronbach. Un programa basado en gráficos dinámicos. *Revista Psico-USF*, 7(2), 143-152.

Losada, M., & Rodríguez, A. (2007). Calidad del servicio de salud: Una revisión de la literatura desde la perspectiva del marketing. *Cuadernos de Administración*, 20(34), 237-258.

Luna, R., Mundina, J., & Gómez, A. (1998). La creación de una escala para medir la calidad del servicio y la satisfacción: el Neptuno-1. En J. Martínez del Castillo, *Deporte y Calidad de Vida* (págs. 279-290). Madrid: Librerías Deportivas Esteban Sanz.

Medina, R. E., Ceballos, O., Ginér, E., & Márques, L. (2007). La autoevaluación de la calidad total en una organización deportiva municipal. Recuperado el Marzo de 2014, de http://www.aiesepguadalajara2007.com/.../Administracion%20y%20Gestion/La_autoev.%20municipal%20UANL.doc

Mundina, J. J., & Calabuig, F. (1999). El marketing social al servicio de la gestión de la calidad. El caso de los Servicios Públicos Náuticos de la Generalitat Valenciana. *Apunts. Educación Física y Deportes*, 57, 77-83.

Mundina, J., García, J., Calabuig, F., & Balibrea, K. (1998). Calidad de servicio en las escuelas del mar de la Generalitat Valenciana. En J. Martínez del Castillo, *Deporte y Calidad de Vida* (págs. 291-298). Madrid: Librerías Deportivas Esteban Sanz.

Ortega, F. (2008). El Método Delphi, prospectiva en Ciencias Sociales a través del análisis de un caso práctico. *Revista EAN*(64), 31-54.

Parasuraman, A., Zeithaml, V., & Berry, L. (1988). Communication and Control Processes in the Delivery of Service Quality. *Journal of Marketing*, 52, 35-48.

Pérez, V. C. (2006). *Calidad Total en la atención al cliente*. Vigo: Ideaspropias.

Spearman, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, 271-295.

Yusof, A., & Lee, H. S. (2008). Spectator perceptions of physical facility and team quality : a study of a Malaysian Super League soccer match. *Research Journal of International Studies*, 8(2), 132-140.

BIOGRAFÍA

Ana Milena Álvarez Cano es Magister en Administración de Empresas con énfasis en Mercadeo por la Universidad Magister de San José, Costa Rica. Profesora Asistente II en la Universidad Autónoma de Occidente, adscrita al Departamento de Ciencias Administrativas. Se puede contactar en la Facultad de Ciencias Económicas y Administrativas, Universidad Autónoma de Occidente, Calle 25 115 - 85, Km. 2 vía Cali – Jamundí. Correo electrónico: amalvarez@uao.edu.co. PBX: (57-2) 318 80 00 ext. 11545. www.uao.edu.co.

