

Validación de un instrumento para medir satisfacción de usuarios en instituciones educativas del sector privado: ciudad de Xalapa-Enríquez (México)

Validation of an instrument to measure user satisfaction in private sector education institutions: city of Xalapa-Enríquez (Mexico)

Judith Guadalupe Montero-Mora
Profesora, Proyecto Aula
Universidad Veracruzana (UV)
Veracruz, México
jmontero@uv.mx

Rosa María Cantón-Croda
Decana, Posgrados de Ingeniería y Negocios
Universidad Popular Autónoma del Estado de Puebla (UPAEP)
Puebla, México
rosamaria.canton@upaep.mx

Recibido - Received: 05 / 12 / 2019

Corregido - Revised: 20 / 03 / 2020

Aceptado - Accepted: 15 / 04 / 2020

DOI: <https://doi.org/10.22458/ie.v22i32.2727>

URL: <https://revistas.uned.ac.cr/index.php/innovaciones>

Resumen: La satisfacción del cliente es un tema de gran importancia para toda organización que desee lograr una ventaja competitiva, ya que los clientes satisfechos serán fieles a la empresa. Las instituciones educativas no son ajenas a esta realidad. En el contexto empresarial, existen diferentes instrumentos para medir satisfacción de clientes, algunos de uso mundial, que se han adaptado al contexto educativo, como las escalas SERVQUAL y SERVPERF. Existen instrumentos diseñados específicamente para medir la satisfacción de usuarios de instituciones de educación enfocados sólo en universidades públicas y privadas, donde el entrevistado es el usuario directo (estudiante). El objetivo del trabajo fue desarrollar y validar un instrumento para medir satisfacción de usuarios de instituciones educativas del sector privado, aplicable a usuarios indirectos: padres de familia de estudiantes en niveles básicos, a usuarios directos: estudiantes en bachillerato, nivel superior (licenciatura y posgrado) y a capacitación para el trabajo. El diseño y la validación del instrumento se realizaron aplicando el método Delphi en dos rondas, con la participación de 12 expertos. El instrumento final se aplicó a una muestra piloto de 50 usuarios; con esta muestra se validó utilizando el Alfa de Cronbach y se obtuvo un valor de 0,98. El resultado fue un instrumento con 24 ítems, 22 en una escala diferencial semántico con valores de 1 a 7 y dos preguntas abiertas. Se considera que este instrumento es un recurso apropiado que aporta elementos para la planeación de la estrategia de negocio a quienes toman decisiones en el sector educativo privado.

Palabras clave: satisfacción del cliente, calidad de la educación, instituciones educativas privadas, instrumentos de medición, método Delphis

Summary: Customer satisfaction is a matter of great importance for any organization that wants to achieve a competitive advantage, since satisfied customers will be loyal to the company. Education institutions are no stranger to this reality. In the business context, there are different instruments to measure customer satisfaction, some of them worldwide, which have been adapted to the education context, such as the SERVQUAL and SERVPERF scales. There are instruments specifically designed to measure the satisfaction of users of education institutions focused only on public and private universities, where the interviewee is the direct user (student). The objective of the work was to develop and validate an instrument to measure satisfaction of users of education institutions in the private sector, applicable to indirect users: parents of students at basic levels, to direct users: students in high school, higher level (undergraduate and graduate)) and to job training. The design and validation of the instrument were carried out by applying the Delphi method in two rounds, with

the participation of 12 experts. The final instrument was applied to a pilot sample of 50 users. This sample was validated using Cronbach's Alpha and a value of 0.98 was obtained. The result was an instrument with 24 items, 22 on a semantic differential scale with values from 1 to 7 and two open questions. This instrument is considered an appropriate resource that provides elements for business strategy planning to decision makers in the private education sector.

Key Words: customer satisfaction, quality of education, private educational institutions, measurement instruments, Delphi method

INTRODUCCIÓN

Vivimos en un mundo cada vez más competitivo, las empresas desarrollan sus actividades en un entorno que cambia continuamente y donde, en muchas ocasiones, la oferta de un determinado producto o servicio es superior a su demanda; así ante la gama de ofertas el cliente puede elegir entre las múltiples opciones que el mercado ofrece. Por tanto, en el ambiente de competencia en el que se encuentran inmersas las empresas, la satisfacción del cliente se convierte en un factor clave de su éxito (Racet, Espinosa, Suárez, Alfonso y Martínez, 2017).

Medina-Merodio et al. (2014) señalan algunas estadísticas importantes sobre los clientes, comentan que 96% de los clientes que no quedaron satisfechos, no se quejan; de los que no quedaron satisfechos 91% no regresa; finalmente, cada cliente insatisfecho se lo dice a 10 personas. Así la satisfacción del cliente deriva en una posición competitiva para las empresas, y como resultado una alta cuota de mercado y ganancias (Turkyilmaz, Oztekin, Zaim, & Fahrettin Demirel, 2013).

Como campo de investigación, la satisfacción del cliente surgió a finales de los años 70, centrándose particularmente en el estudio del comportamiento del consumidor. La base fundamental de la investigación en la calidad del servicio se encuentra principalmente en dos escuelas: la nórdica, reunida alrededor del Service Research Center de la Universidad de Karlstad, encabezada por Grönroos, Gummesson y Lehtinen, que enfocan el tema de la calidad de servicio desde el punto de vista del producto. La otra escuela, la norteamericana, se encuentra guiada desde el Interstate Center for Services Marketing de la Universidad del Estado de Arizona por Parasuraman, Zeithaml y Berry, centrados en el estudio desde la óptica de la percepción de los clientes (Pérez, Martínez, Noda y Miguel, 2015).

Según Mora (2011), desde el enfoque de pensamiento de la escuela nórdica, sintetizada en el denominado Modelo de la Imagen de Grönroos, la calidad percibida por el cliente es el resultado de tres factores: la calidad técnica, la calidad funcional y la imagen corporativa. En la escuela norteamericana Parasuraman, Zeithaml y Berry diseñaron un modelo de la calidad de servicio, en el que ésta se define por la medida en que el servicio satisface las expectativas de los clientes.

El modelo propuesto por los autores citados anteriormente se denomina SERVQUAL y considera cinco dimensiones de calidad del servicio: aspectos tangibles (físicos), confiabilidad (cumplimiento y consistencia), velocidad de respuesta ante las demandas del cliente, aseguramiento de lo ofrecido y empatía con el cliente. El modelo así como el instrumento asociado al mismo se ha difundido y utilizado ampliamente en todo el mundo.

A partir de estos modelos, se han realizado diferentes investigaciones tomando como punto de partida el modelo SERVQUAL, adaptándolo al contexto educativo con la finalidad de evaluar las percepciones de los estudiantes sobre la calidad del servicio que reciben.

En este contexto, autores como Abdullah (2006), Capelleras y Venciana (2001) y Álvarez, Chaparro y Reyes (2015), por mencionar algunos, realizaron investigaciones en el ámbito de la educación superior, que como resultado han derivado en propuestas de nuevos modelos e instrumentos de medición aplicados sólo a estudiantes (usuarios directos del servicio). Sin embargo, después de una revisión de

literatura, se encontró que investigaciones en niveles básicos son escasas. Aunado a lo anterior, como señala Abdullah (2006) los instrumentos que se han diseñado, como el SERVQUAL o SERVPERF fueron diseñados como medidas genéricas de la calidad del servicio aplicables en toda la industria, por lo que los instrumentos pueden considerarse como "esqueletos" básicos que requieren modificaciones para ser adaptados al contexto de aplicación específica.

En definitiva, se evidencia la necesidad de mejorar los modelos de evaluación existente, así como la importancia de medir la satisfacción de los usuarios de instituciones de educación, con instrumentos acordes con el contexto de aplicación. Por tanto, el objetivo de la presente investigación es desarrollar y validar un instrumento de medición relacionado con el contexto de instituciones de educación del sector privado ubicadas en la ciudad de Xalapa-Enríquez (México); que abarque los niveles básico (inicial, preescolar, primaria y secundaria), medio superior (bachillerato), superior (licenciatura y posgrado) y la capacitación para el trabajo (cursos de inglés, belleza, computación, etc.); aplicable no sólo a estudiantes, sino también a usuarios indirectos del servicio, en este caso padres de estudiantes del nivel básico.

MODELOS DESARROLLADOS EN INVESTIGACIONES SOBRE SATISFACCIÓN DE USUARIOS

Las investigaciones de Parasuraman, Zeithaml y Berry con su modelo SERVQUAL que define la calidad del servicio como la diferencia entre las percepciones de los clientes y sus expectativas (22 ítems) y de Cronin y Taylor, con el modelo SERVPERF que se compone de los 22 ítems de la escala SERVQUAL, pero mide exclusivamente las percepciones de la calidad del servicio, son considerados como los trabajos pioneros en la medición de la calidad del servicio y satisfacción del cliente y han sido aplicados en un gran número de empresas de servicios (Mora, 2011). Sin embargo, se encontraron investigaciones que realizaron adaptaciones a los modelos SERVQUAL y SERVPERF para medir la percepción de la calidad de los servicios de estudiantes en universidades privadas; e investigaciones en las que se diseñaron los instrumentos también en instituciones de educación superior. En la tabla 1, se presenta un resumen de las investigaciones analizadas, se destacan las dimensiones que utilizaron en su medición ya que a partir de ellas se definió lo que en la presente investigación se considera son las principales dimensiones que intervienen en la medición de la calidad y satisfacción de usuarios de instituciones educativas privadas.

El Instituto Nacional para la Evaluación de la Educación (INEE, 2006) asume un modelo denominado CIPP para la evaluación de la calidad de la educación, que considera cuatro componentes: contexto, insumos, procesos y productos. El contexto tiene que ver con aspectos sociales que condicionan la operación y desempeño del sistema educativo. Los insumos y recursos están relacionados con la operación de las escuelas y las aulas (alumnos, docentes, directores y escuelas), se incluyen los recursos financieros destinados a la operación del sistema. En relación con los procesos escolares, se considera el acceso y trayectoria de los estudiantes en los distintos niveles educativos, la gestión, funcionamiento y organización de los planteles.

TABLA 1
Investigaciones analizadas en el sector educativo

Autores	Descripción	Dimensiones	Ítems	Campo de aplicación
(Abdullah, 2006)	Realiza instrumento propio HEdPERF (Higher Education PERFORMANCE only)	<ul style="list-style-type: none"> • Aspectos no académicos • Aspectos académicos • Reputación • Acceso • Problemas de los programas 	45	Instituciones de educación superior, Malasia
(Kwek, Lau, & Tan, 2010)	Adaptaron el Modelo SERVQUAL	<ul style="list-style-type: none"> • Calidad de bibliotecarios • Receptividad del personal de la División de Exámenes y Premios • Plan de estudios • Cantidad de actividades recreativas 	16	Universidad privada en Malasia
(Tumino & Poitevin, 2013)	Adaptaron la escala de (Capelleras & Venciana, 2001) basados en el modelo SERVQUAL	<ul style="list-style-type: none"> • Competencia del cuerpo docente • Organización de la enseñanza o de la carrera (Contenido del plan de estudios) • Comunicación y clima organizacional • Actitudes y comportamiento del cuerpo docente • Instalaciones y equipamientos • Evaluaciones • Asignaturas optativas • Efectividad de los procesos administrativo 	42	Universidad privada, Argentina
(Álvarez Botello, Chaparro Salinas, y Reyes Pérez, 2015)	Realizan su propio instrumento	<ul style="list-style-type: none"> • Plan de estudios • Capacitación y habilidad para la enseñanza • Métodos de enseñanza y evaluación • Nivel de autorrealización del estudiante • Servicios de apoyo • Servicios administrativos • Ambiente propicio • Infraestructura 	41	Universidades públicas y privadas, Valle de Toluca, México
(Alvarado-Lagunas et al., 2015)	Adaptaron el Modelo SERVPERF	<ul style="list-style-type: none"> • Componente físico • Planta docente • Habilidad docente • Desarrollo integral 	22	Universidades privadas en Monterrey, México
(Clifford, Reszka, & Rossbach, 2010)	Escala ECERS-R, mide la calidad global de programas basados en centros para la primera infancia.	<ul style="list-style-type: none"> • Espacio y mobiliario • Rutinas de cuidado personal • Lenguaje-razonamiento • Actividades • Interacción • Estructura del programa • Padres y personal 	43	Preescolar y jardín de niños (2,5-5 años)
(Stipek & Byler, 2004)	Escala ECCOM, evalúa la naturaleza y la calidad de la instrucción, así como el clima social y la gestión del aula.	<ul style="list-style-type: none"> • Instrucción • Gestión • Clima social 	32	Preescolar, jardín de niños y primer grado (Edades 3-7)
(Pianta, Paro, & Hamre, 2008)	Escala CLASS, examina la relación entre profesores y sus alumnos, y evalúa la calidad de la instrucción y la gestión del comportamiento en el aula.	<ul style="list-style-type: none"> • Apoyo emocional: a) clima positivo, b) clima negativo, c) respeto por la perspectiva del estudiante, d) sensibilidad del maestro • Organización del aula: a) gestión del comportamiento, b) productividad, c) formatos de aprendizaje instructivo • Apoyo educativo: a) desarrollo de conceptos, b) calidad de la retroalimentación, c) modelado del lenguaje 	41	Versiones para jardín de niños primaria y secundaria

Fuente: Elaboración propia con información de (Abdullah, 2006), (Kwek et al., 2010), (Tumino & Poitevin, 2013), (Álvarez Botello et al., 2015), (Alvarado-Lagunas et al., 2015), (Clifford et al., 2010), (Pianta et al., 2008), (Stipek & Byler, 2004).

Se agregó además una dimensión que se llamó *evaluación general*, con el objetivo de tener información de la satisfacción de los usuarios con el servicio. Se propusieron estas dimensiones debido a que la medición de la satisfacción no sólo sería realizada en estudiantes, sino también, en el caso de los niveles iniciales, a los padres de los niños.

Es importante resaltar que las investigaciones de la tabla 1 evalúan la satisfacción sólo del estudiante y su enfoque es evaluar la calidad del servicio educativo, considerando también variables relacionadas, por ejemplo, con el plan de estudios, con los métodos de enseñanza, con la autorrealización como estudiante, etcétera; cuestiones que los padres entrevistados no podrían evaluar.

Como resultado de esta primera revisión, en la tabla 2 se muestran las dimensiones y los ítems propuestos, en una primera versión, para medir la percepción de la satisfacción de usuarios de instituciones educativas privadas, la escala de medición utilizada fue el diferencial semántico con una codificación de 1 a 7, donde 7 corresponde a una calificación totalmente positiva y 1 a una calificación totalmente negativa (Hernández et al., 2014). La primera versión fue enviada a los expertos para su valoración.

TABLA 2
Primera versión de variables e ítems

Dimensiones	Ítems								
Infraestructura (instalaciones y equipamiento)	1. Las instalaciones de la institución (aulas, biblioteca, baños, deportivos, cafetería) se encuentran en buen estado.	7	6	5	4	3	2	1	Instalaciones en mal estado
	2. Los espacios destinados a la enseñanza (aulas, talleres, laboratorios, área de cómputo) son apropiados (buena iluminación, bien equipados, espacio apropiado).	7	6	5	4	3	2	1	Espacios inapropiados
	3. Las instalaciones siempre se encuentran limpias.	7	6	5	4	3	2	1	Instalaciones sucias
	4. El mobiliario se encuentra en buen estado.	7	6	5	4	3	2	1	Mobiliario en mal estado
Personal docente (competencia)	5. El personal docente tiene conocimiento y dominio de los temas que atiende.	7	6	5	4	3	2	1	El personal no tiene conocimiento.
	6. El personal docente está capacitado y actualizado.	7	6	5	4	3	2	1	El personal no está capacitado.
	7. El personal docente es responsable (puntual, buenas clases, comunicativo, creativo, paciente).	7	6	5	4	3	2	1	Personal irresponsable
	8. El personal docente tiene la capacidad de buena comunicación.	7	6	5	4	3	2	1	Mala comunicación
Servicio administrativo	9. Los procesos administrativos (inscripción, altas, bajas, trámites varios, etc.) son ágiles.	7	6	5	4	3	2	1	Procesos administrativos son lentos.
	10. El personal administrativo siempre es amable y dispuesto a ayudar.	7	6	5	4	3	2	1	El personal no es amable.
	11. El personal administrativo es eficaz para resolver problemas.	7	6	5	4	3	2	1	Personal ineficaz
Ambiente (ambiente dentro de la institución)	12. Todo el personal que labora en la institución (apoyo, mantenimiento, docente, administrativo) brinda un trato amable.	7	6	5	4	3	2	1	Trato grosero
	13. En la institución existe buen ambiente, el cual favorece la buena convivencia.	7	6	5	4	3	2	1	Mal ambiente
	14. Existen procesos y canales de comunicación apropiados para el intercambio de información.	7	6	5	4	3	2	1	Procesos y canales inapropiados
	15. Cuando surgen dificultades las soluciones y alternativas son acertadas.	7	6	5	4	3	2	1	Soluciones no acertadas
Evaluación general	16. En general estoy muy satisfecho con la institución.	7	6	5	4	3	2	1	Muy insatisfecho
	17. Esta institución supera mis expectativas.	7	6	5	4	3	2	1	No supera expectativas.

Fuente: Elaboración propia (2019).

Segunda fase: Selección del panel de expertos. Con respecto al número de expertos que se deben incluir, Skulmoski & Hartman (2007) sugieren que cuando el grupo de expertos es homogéneo la muestra puede ser de entre 10 y 15; si el grupo es heterogéneo, la muestra requerida es mayor, pero debe tenerse cuidado porque si es demasiado grande se podría complicar la recolección de la información y llegar a un consenso. La misma idea es compartida por Hsu & Sandford (2005) pues sugieren que una muestra muy pequeña (menor de siete) no ofrecería información representativa y si es demasiado grande podrían aparecer inconvenientes asociados con los tiempos requeridos para recabarla. Por su parte, López-Gómez (2018) concluye en su investigación que el grupo no debe ser menor de diez personas y León y Montero (2004) sugieren que deberían considerarse entre 10 y 30 expertos. De esta revisión se consideró entonces trabajar con un panel de 20 expertos, sin embargo, sólo 12 expertos apoyaron en la investigación, pero se consideró este número apropiado, ya que el grupo de expertos era homogéneo y como lo sugiere Skulmoski & Hartman (2007), cuando el grupo es homogéneo, una muestra de entre 10 o 15 expertos es suficiente.

Para la definición de las características de los expertos se consideró lo propuesto por Skulmoski & Hartman (2007) quienes plantean que el panel de expertos debe cumplir cuatro requisitos: a) conocimiento y experiencia del tema que se investigan, b) capacidad y voluntad de participar, c) tiempo suficiente para participar y d) habilidades de comunicación. Así el panel de expertos quedó conformado por cuatro investigadores con experiencia en la medición de la calidad en los servicios (INVESTIGADOR), cuatro expertos en la implementación de sistemas de gestión de calidad (SGC) y por consiguiente familiarizados con la medición de la satisfacción del cliente, dos funcionarios encargados del área de calidad en sus empresas (FUNCIONARIO) y dos profesores de cursos en una maestría en Gestión de la Calidad (ACADÉMICO).

Para evaluar el panel se consideró el coeficiente de competencia experta (Cabero y Barroso, 2013), que evalúa la opinión del experto sobre su nivel de conocimiento acerca del problema de investigación. El coeficiente se obtuvo para cada experto con la siguiente ecuación:

$$K = \frac{1}{2}(kc + ka) \quad (1)$$

Donde Kc es el coeficiente de conocimiento que posee el experto acerca del tema. Se calculó a partir de la valoración que realizó el experto en una escala del 0 al 10, multiplicado por 0,1. Ka es el coeficiente de argumentación de los criterios de los expertos (Cabero y Barroso, 2013). Este coeficiente se obtuvo a partir de la asignación de una serie de puntuaciones a diversos factores (fuentes de argumentación de los expertos) como son experiencia en su actividad y práctica, conocimiento del tema, conocimientos del estado del arte a nivel nacional e internacional, conocimiento sobre publicaciones del tema, entre otros (García y Lena, 2018). La tabla 3 muestra la asignación de puntuaciones a cada uno de los factores, con los cuales se determinó el coeficiente de argumentación; la asignación de las puntuaciones se realizó considerando la importancia de cada una de las fuentes.

TABLA 3
Puntuaciones asignadas a cada una de las fuentes de argumentación

Fuentes de argumentación	Grado de influencia de cada una de las fuentes		
	Bajo	Medio	Alto
Análisis teóricos realizados por usted	0,10	0,20	0,30
Experiencia obtenida a través de su actividad y práctica	0,20	0,40	0,60
Estudio de trabajos y publicaciones sobre el tema de autores nacionales	0,04	0,06	0,10
Estudio de trabajos y publicaciones sobre el tema de autores extranjeros	0,04	0,06	0,10
Conocimiento sobre el estado del arte a nivel nacional e internacional	0,10	0,20	0,30
Intuición sobre el tema abordado y conocimiento sobre tecnología	0,04	0,06	0,10

Fuente: Elaboración propia (2019).

La tabla 4 presenta el resultado de la determinación del coeficiente de competencia experta (Ecuación 1) para cada uno de los expertos. Es importante mencionar que el coeficiente de competencia experta ideal tendría un valor de 1,25, asumiendo que la autovaloración de cada experto es 10 y el coeficiente de conocimiento es alto en todas las fuentes de argumentación.

TABLA 4
Resultados de la evaluación de la competencia de los expertos

Perfil	Kc	Autoevaluación	Ka	K
Investigador	0,86	8	0,8	0,83
Investigador	1,5	9	0,9	1,2
Investigador	1,36	10	1	1,18
Investigador	1,34	9	0,9	1,12
SGC	1,32	9	0,9	1,11
SGC	0,96	9	0,9	0,93
SGC	1,06	8	0,8	0,93
SGC	1,3	8	0,8	1,05
Académico	0,98	8	0,8	0,89
Académico	0,96	8	0,8	0,88
Funcionario	1,28	8	0,8	1,04
Funcionario	0,96	8	0,8	0,88

Fuente: Elaboración propia (2019).

Tercera fase: Desarrollo práctico- obtención de las observaciones del panel de expertos e interpretación. Esta fase se llevó a cabo en dos rondas. En la primera ronda, se hizo llegar a cada experto la primera propuesta de cuestionario (tabla 2); se anexó un espacio para que realizaran el registro de su valoración. Se solicitó que valoraran cada ítem en escala Likert con un rango de valores de 1 a 5, donde el valor "1" significó "no adecuado" y "5" "muy adecuado". Se les pidió también que en caso de haber propuestas, sugerencias o modificaciones, las indicaran para cada ítem.

Una vez recuperados los cuestionarios con las observaciones se realizó el análisis. Con los puntajes asignados se calculó la suma para cada ítem. El puntaje ideal sería 60, considerando que todos los expertos hubieran asignado un valor de 5 (muy adecuado) a cada ítem y un puntaje de 48 indicaría que todos los ítems fueron evaluados con 4 (bastante adecuado). Con estos valores de referencia se observaron los puntajes para cada ítem, seis tuvieron puntajes menores a 48, primer referente para hacer modificaciones (tabla 5). Otro aspecto por considerar fue la validez de contenido de los ítems. Aiken (1980) propone un coeficiente para la evaluación de la validez de contenido a través de la valoración hecha por expertos, el Coeficiente V de Aiken, el cual toma valores en un rango de 0 y 1, a medida que el valor del índice se acerca a uno se considera que el ítem posee una mayor validez de contenido, lo cual significa que existe un acuerdo perfecto en la valoración de los expertos. El valor de referencia para este índice varía en función del número de expertos, del número de ítems y del número de categorías por evaluar, en la investigación se consideró un valor aceptable para $V = 0,70$, la ecuación para obtener este coeficiente es (Aiken, 1980):

$$V = \frac{x - l}{K} \quad (2)$$

Donde x es la media de las calificaciones emitida por los expertos para cada ítem; l es la calificación más baja que se puede seleccionar en un ítem, en este caso es 1; finalmente k , el rango de valores en la escala que se utilizó, es decir, $k = 5 - 1 = 4$. García y Lena (2018) y López (2018) sugieren utilizar otra medida para evaluar el consenso en las observaciones de los expertos: el coeficiente de variación. La ecuación para obtener el coeficiente de variación es (Spiegel & Stephens, 2009):

$$C_v = \frac{s}{x} \quad (3)$$

Donde s es la media de las calificaciones emitida por los expertos para cada ítem, y su desviación estándar. Este valor suele multiplicarse por 100 para tener un referente en porcentaje (Spiegel & Stephens, 2009). Al ser una medida de variación, los valores grandes indican que no existe consenso en las evaluaciones de los expertos. Se consideró como valor de referencia de este coeficiente 40%, por lo que valores menores indicaron un nivel de consenso aceptable por parte de los expertos. El coeficiente de validez de contenido V de Aiken y el coeficiente de variación fueron calculados para cada uno de los ítems, tal como se muestra en la tabla 5.

De la tabla 5 puede comentarse que 16 de los ítems propuestos cumplen con los valores de referencia $V = 0,70$, sólo el ítem 13 no lo cubre. Para el caso del coeficiente de variación, todos los ítems son menores al valor de referencia de 40%. De este análisis se concluyó que los 17 ítems son pertinentes por lo que se decidió no eliminar ninguno. El primer análisis se complementó con las observaciones o mejoras que los expertos realizaron para los ítems y que fueron considerados para la modificación del cuestionario tales como agregar ítems relacionados con la seguridad en la institución, evaluar sólo una característica en cada ítem, agregar ítems para evaluar los resultados del servicio (incluido el fomento de valores), agregar al menos una pregunta abierta para que el encuestado pueda expresar su opinión del servicio. Con estas observaciones se realizó el ajuste del instrumento, y resultó una segunda versión con 24 ítems, se agregaron ítems para valorar resultados y satisfacción y dos preguntas abiertas.

TABLA 5
Evaluación de los ítems

Ítem	Puntaje	Media	Desviación estándar	Coefficiente V de Aiken	Coefficiente de variación (%)
P1	48	4,00	0,85	0,75	21,3
P2	50	4,17	1,03	0,79	24,7
P3	47	3,92	1,08	0,73	27,7
P4	47	3,92	0,79	0,73	20,2
P5	52	4,73	0,47	0,93	9,9
P6	53	4,42	0,90	0,85	20,4
P7	54	4,50	0,80	0,88	17,7
P8	52	4,33	0,89	0,83	20,5
P9	50	4,17	1,03	0,79	24,7
P10	47	3,92	1,00	0,73	25,4
P11	46	3,83	1,40	0,71	36,6
P12	48	4,00	1,04	0,75	26,1
P13	45	3,75	0,97	0,69	25,7
P14	46	3,83	1,19	0,71	31,1
P15	49	4,08	0,90	0,77	22,0
P16	51	4,25	0,87	0,81	20,4
P7	49	4,08	0,90	0,77	22,0

Fuente: Elaboración propia (2019).

En la segunda ronda, se envió a los expertos la segunda versión del instrumento. Nuevamente, se les pidió la valoración de cada ítem (valorar en la escala Likert con un rango de valores de 1 a 5, donde el valor "1" es "no adecuado" y "5" "muy adecuado") y además se incluyó una pregunta para consultarles si estaban de acuerdo con los cambios realizados en el cuestionario, con dos respuestas "sí" y "no" (García y Lena, 2018).

Cuarta fase. Finalización del método y conclusiones. Debido a que se observó en la segunda ronda que las evaluaciones de los expertos tuvieron muy poca variación, se decidió concluir con la aplicación del método. Para apoyar la decisión se consideró lo propuesto por López (2018) quien señala que para finalizar con la aplicación del método Delphi debe considerarse el consenso, el cual está relacionado con la convergencia de las opiniones de los expertos. Así mismo, López (2018) sugiere el coeficiente de variación (Ecuación 3) como la prueba más adecuada para evaluar el consenso, entre otras pruebas. Por lo que se calculó el estadístico con las nuevas valoraciones para cada uno de los nuevos 22 ítems de la segunda versión del instrumento, el coeficiente de variación fue menor de 17% en todos los ítems. Con respecto a la pregunta final que se incluyó en el cuestionario para los expertos (¿está de acuerdo con los cambios realizados?), 100% de los expertos indicó estar de acuerdo con los cambios realizados en el nuevo cuestionario, por lo que con los dos antecedentes se concluyó con las rondas. Finalmente, como resultado de las observaciones hechas por los expertos en las dos rondas se construyó una nueva versión con 22 ítems en cinco dimensiones, cuatro de ellas ya habían sido propuestas (tabla 2): *infraestructura, personal docente, servicio administrativo, ambiente* y como resultado de las observaciones se cambió la dimensión *evaluación general*, por la dimensión de *resultados* y se agregó la dimensión *satisfacción*. Esta versión del instrumento es la que se utilizó para recabar la muestra piloto.

MUESTRA PILOTO

La población objetivo fueron 5,145 usuarios de 622 instituciones educativas privadas ubicadas en la ciudad de Xalapa-Enríquez, México (sólo cabecera municipal), divididas en siete niveles: inicial, preescolar, primaria, secundaria, bachillerato, superior y capacitación para el trabajo.

En una primera etapa, mediante un muestreo aleatorio estratificado con asignación proporcional al tamaño del estrato (se consideró como estrato cada uno de los siete niveles), se determinó una muestra de 83 escuelas distribuidas en cada uno de los niveles como sigue: 4 de inicial, 13 de preescolar, 9 de primaria, 6 de secundaria, 13 de bachillerato, 13 de nivel superior y 25 de capacitación para el trabajo. En la segunda etapa, nuevamente mediante un muestreo aleatorio estratificado con asignación proporcional al tamaño del estrato (ahora se consideró como estrato cada una de las 83 escuelas), se determinó un tamaño de muestra de usuarios, la muestra fue de $n = 358$ usuarios. En ambas etapas se consideró un margen de error de 10% y un nivel de confiabilidad de 95%.

Una vez determinado el tamaño de muestra de usuarios y dadas las características de la población, para la muestra piloto se consideró un total de 50 usuarios, ya que como lo señalan Mora, Amelia, Carrasco, Paola, y Muñoz (2015) para una prueba piloto se requieren entre 30 y 50 usuarios que además posean los atributos que se desean de la población objetivo. De esta manera, se realizó el levantamiento de la muestra piloto con 50 usuarios y para asegurar que la muestra incluía los usuarios de los diferentes niveles, estudiantes y padres (ya que el instrumento se aplicaría a ambos tipos de usuarios), la muestra piloto incluyó 15 padres, 10 estudiantes de bachillerato, 15 estudiantes de nivel superior, 10 de capacitación para el trabajo.

Con la muestra piloto se realizó el análisis de confiabilidad del instrumento. En el lenguaje de programación R (R Core Team, 2018) se realizó el cálculo de los estadísticos Alfa de Cronbach y Lambda 6 de Guttman (λ_6). El valor del Alfa de Cronbach obtenido fue de 0,98, el cual es considerado como un valor excelente (Cronbach, 1951). El valor de Lambda 6 de Guttman (Guttman, 1945) para fines prácticos puede interpretarse de la misma manera que el Alfa de Cronbach. El valor de Lambda 6 de Guttman fue de 0,99. Ambos estadísticos indicaron una buena consistencia interna del instrumento. La prueba piloto permitió detectar algunos problemas de comprensión en el instrumento por parte de los encuestados, los cuales fueron considerados para realizar las modificaciones pertinentes. Como resultado se obtuvo la versión final del instrumento.

RESULTADOS Y DISCUSIÓN

La tabla 6 muestra la versión del cuestionario después de haber sido aplicado a la muestra piloto y posteriormente validado. Se incluyó una sección llamada Datos Generales, para recabar información de los entrevistados; 22 ítems en escala diferencial semántico agrupados en cinco dimensiones: *infraestructura*, ítems 1-6; *personal docente*, ítems 7-10; *servicio administrativo*, ítems 11-13; *ambiente*, ítems 14-17; *resultados*, ítems 18-20; *satisfacción*, ítems 21-22 y dos preguntas abiertas.

Con respecto al número de ítems que incluye la versión final del cuestionario, se tomó en cuenta lo siguiente: si bien se reconoce que cuanto más información se obtenga del cliente, más se conocerá acerca de sus expectativas, necesidades y comportamientos, se consideró el que cuestionarios muy extensos suelen ser tediosos para el encuestado, aunado al hecho de que en ocasiones no dispone del tiempo necesario para responder una encuesta, lo cual impacta directamente en la calidad de la información que se obtendría al final; además no se encontró un referente que indicara un número específico de ítems apropiados para elaborar un cuestionario. Por estas razones se optó por lo que se considera un cuestionario claro y sencillo, como lo es el cuestionario final propuesto de 24 ítems.

TABLA 6
Cuestionario final

Datos generales:									
Del entrevistado:		a) Femenino () Masculino ()			b) Edad: _____			c) Estudiante () Padre o tutor ()	
d) Nivel educativo del estudiante:		Inicial ()		Preescolar ()		Primaria ()		Secundaria ()	
		Bachillerato ()		Superior ()		Capacitación para el trabajo ()			
Instrucciones: Basado en sus experiencias como usuario de los servicios que ofrece la institución educativa, en una escala de 1 a 7, en la que 1 representa lo menos adecuado o muy inapropiado hasta 7, que representa lo más adecuado o más apropiado, elija el número que mejor refleje su opinión.									
1	Las instalaciones de la institución son inadecuadas para el desarrollo de actividades.	1	2	3	4	5	6	7	Las instalaciones de la institución son adecuadas para el desarrollo de actividades.
2	Las aulas son espacios inapropiados para el desarrollo de las actividades.	1	2	3	4	5	6	7	Las aulas son espacios apropiados para el desarrollo de las actividades.
3	El equipamiento de las instalaciones de la institución es obsoleto.	1	2	3	4	5	6	7	El equipamiento de las instalaciones de la institución es moderno.
4	Las instalaciones de la institución suelen estar sucias.	1	2	3	4	5	6	7	Las instalaciones de la institución se mantienen limpias.
5	El inmueble es inseguro.	1	2	3	4	5	6	7	El inmueble es seguro.
6	Las adecuaciones de la institución para la movilidad de personas con capacidades diferentes son inapropiadas	1	2	3	4	5	6	7	Las adecuaciones de la institución para la movilidad de personas con capacidades diferentes son apropiadas.
7	El personal docente tiene un nivel insuficiente de conocimientos.	1	2	3	4	5	6	7	El personal docente tiene un nivel suficiente de conocimientos.
8	El personal docente es incompetente.	1	2	3	4	5	6	7	El personal docente es competente.
9	El personal docente está desactualizado.	1	2	3	4	5	6	7	El personal docente está actualizado.
10	El personal docente es irresponsable.	1	2	3	4	5	6	7	El personal docente es responsable.
11	Los procesos administrativos (inscripción, altas, trámites varios, etc.) son generalmente lentos.	1	2	3	4	5	6	7	Los procesos administrativos (inscripción, altas, trámites varios, etc.) son generalmente rápidos.
12	El personal administrativo es descortés.	1	2	3	4	5	6	7	El personal administrativo es amable.
13	El personal administrativo es ineficaz.	1	2	3	4	5	6	7	El personal administrativo es eficaz.
14	El ambiente en la institución es desfavorable para la buena convivencia.	1	2	3	4	5	6	7	El ambiente en la institución favorece la buena convivencia.
15	La comunicación en la institución es ineficiente.	1	2	3	4	5	6	7	La comunicación en la institución es efectiva.
16	Los medios de comunicación para el intercambio de información en la institución son ineficientes.	1	2	3	4	5	6	7	Los medios de comunicación para el intercambio de información en la institución son eficientes.
17	En caso de conflicto, las soluciones propuestas por la institución son equivocadas.	1	2	3	4	5	6	7	En caso de conflicto, las soluciones propuestas por la institución son acertadas.
18	En esta institución se frena el desarrollo de las capacidades (poder hacer) del alumno.	1	2	3	4	5	6	7	En esta institución se estimula el desarrollo de las capacidades (poder hacer) del alumno.
19	En esta institución se frena el desarrollo de habilidades (saber hacer) del alumno.	1	2	3	4	5	6	7	En esta institución se estimula el desarrollo de habilidades (saber hacer) del alumno.
20	La institución no fomenta valores.	1	2	3	4	5	6	7	La institución fomenta valores.
21	Esta institución incumple con su promesa de servicio.	1	2	3	4	5	6	7	Esta institución cumple con su promesa de servicio.
22	Estoy insatisfecho con la institución.	1	2	3	4	5	6	7	Estoy satisfecho con la institución.
23	¿Qué haría falta para que estuvieras completamente satisfecho/a con los servicios de la institución?								
24	¿Qué consideras nos distingue de otras instituciones que ofrecen el mismo servicio?								

Fuente: Elaboración propia (2019).

CONCLUSIONES

Los servicios se han convertido en un sector muy fuerte, han adquirido un papel prioritario en la economía de la mayoría de los países. La relación de este concepto con los costos, los beneficios, la satisfacción y la lealtad de los clientes es un tema interesante para investigación y un arma estratégica de gran valor para las empresas (Delgado, 2001). De ahí que, ante la marcada necesidad de conocer tanto las percepciones como las expectativas de los clientes, en los últimos años diversos autores han reconocido la importancia del concepto de calidad en el servicio considerando la percepción del cliente y han desarrollado un concepto operativo de la calidad en el servicio que ha dado la pauta para la elaboración de diferentes escalas de medición aplicables a diferentes tipos de servicios (Capelleras y Venciana, 2001). En este contexto, Álvarez et al. (2015) señalan que los diversos instrumentos provenientes del sector empresarial privado y público suelen ser utilizados con las correspondientes adaptaciones al sector educativo, sin embargo, los resultados obtenidos de estos instrumentos, al ser diseñados en un contexto diferente al educativo, se consideran insuficientes para dar una valoración específica del sector educativo pues no contemplan variables propias del sector. Por tanto, el objetivo del presente estudio fue desarrollar un instrumento de medición para un contexto específico, instituciones de educación privada, y que además su aplicación no sólo se enfoque en los usuarios directos (estudiantes), sino que además sea un instrumento que los usuarios indirectos del servicio en los niveles básicos (padres) puedan responder.

El uso del Método Delphi permitió recibir el aporte de expertos en la medición de la satisfacción del cliente. Sin embargo, como lo señala García y Lena (2018) todo instrumento de medición que haya sido sometido al juicio de expertos debe ser evaluado para verificar que cubre los requisitos mínimos de calidad necesarios en los instrumentos, siendo estos confiabilidad y validez. Se considera la validez como el grado en que un instrumento mide la variable que debe medir y la confiabilidad como el grado en que un instrumento produce resultados consistentes y coherentes y además muestra ausencia de errores de medida (Supo, 2013). Así, dados los resultados obtenidos después de la validación del instrumento para medir la satisfacción de usuarios de instituciones educativas del sector privado, se considera un instrumento confiable y válido.

La intención es aportar una herramienta de medición útil y que permita, a quienes hagan uso de ella, el obtener información confiable, válida, entendible y manejable, que derive en la apropiada toma de decisiones.

REFERENCIAS

- Abdullah, F. (2006). The development of HEdPERF: a new measuring instrument of service quality for the higher education sector. *International Journal of Consumer Studies*, 30(6), 569-581. <https://doi.org/10.1111/j.1470-6431.2005.00480.x>
- Aiken, L. R. (1980). Content validity and reliability of single items or questionnaires. *Educational and Psychological Measurement*, 40(4), 95-959. <https://doi.org/https://doi.org/10.1177/001316448004000419>
- Alvarado-Lagunas, E., Luyando-Cuevas, J.-R., & Picazzo-Palencia, E. (2015). Universidades privadas en Monterrey, vi, 58-76.
- Álvarez Botello, J., Chaparro Salinas, E. M., & Reyes Pérez, D. E. (2015). Estudio de la satisfacción de los estudiantes con los servicios educativos brindados por instituciones de educación superior del valle de Toluca. *REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en la Educación*, 13(2), 5 a 26. Recuperado de <http://www.redalyc.org/pdf/551/55138743001.pdf>

- Cabero-Almenara, J., & Barroso Osuna, J. (2013). La utilización del juicio de experto para la evaluación de TIC: el coeficiente de competencia experta. *Bordón*, 65(2), 75-96. <https://doi.org/10.13042/brp.2013.65202>
- Cabero Almenara, J., & Infante Moro, A. (2014). Empleo del método Delphi y su empleo en la investigación en comunicación y educación. *Revista Electrónica de Tecnología Educativa*, (48), 1-16.
- Capelleras, J. L., & Venciana, J. M. (2001). Calidad de servicio en la enseñanza universitaria: desarrollo y validación de una escala de medida. *Document de treball*, (4), 1-36. Recuperado de <http://www.uab.cat/servlet/BlobServer?blobtable=Document&blobcol=urldocument&blobheader=application/pdf&blobkey=id&blobwhere=1345650483792&blobnocache=true>
- Clifford, R., Reszka, S., & Rossbach, H.-G. (2010). Reliability and validity of the Early Childhood Environment Rating Scale. *Chapel Hill: FPG Child Development Institute*, (January), 1-27. <https://doi.org/10.1080/03004430.2013.818991>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- Delgado, H. C. (2001). *Desarrollo de una cultura de calidad*. (McGraw-Hill, Ed.) (Primera Ed). México.
- García-Ruiz, M. E., & Lena-Acebo, F. J. (2018). Aplicación del metodo delphi en el diseño de una investigación cuantitativa sobre el fenómeno FABLAB, 129-166.
- Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, 10, 255-282.
- Hernández, R., Fernández, C., & Baptista, P. (2014). *Metodología de la investigación*. *Journal of Chemical Information and Modeling* (Vol. 53). <https://doi.org/10.1017/CBO9781107415324.004>
- Hsu, C.-C., & Sandford, B. A. (2005). The Delphi technique. *Dublin: National Children's Office*, 7(2), 120-125. <https://doi.org/10.1576/toag.7.2.120.27071>
- INEE, (Instituto Nacional para la Evaluación de la Educación). (2006). *Indicadores del Sistema Educativo Nacional*. México. Recuperado de <http://publicaciones.inee.edu.mx/buscadorPub/P1/I/241/P11241.pdf>
- Kwek, L. C., Lau, T. C., & Tan, H. P. (2010). Education Quality Process Model and Its Influence on Students ' Perceived Service Quality. *International Journal of Business and Management*, 5(8), 154-165.
- León, O. G., & Montero, I. (2004). *Métodos de investigación en Psicología y Educación* (3a.). Madrid: McGraw-Hill Interamericana.
- López-Gómez, E. (2018). El método delphi en la investigación actual en educación: una revisión teórica y metodológica. *Educación XX1*, 21(1), 17-40. <https://doi.org/10.5944/educXX1.15536>
- Medina-Merodio, J. A., Pablos-Heredero, C. de, Jiménez-Rodríguez, M. L., Marcos-Ortega, L. de, Barchino-Plata, R., Rodríguez-García, D., & Gómez-Aguado, D. (2014). Análisis de la satisfacción de cliente mediante el uso de cuestionarios con preguntas abiertas. *Dyna*, 81(188), 92-99. <https://doi.org/10.15446/dyna.v81n188.40144>
- Mora Contreras, C. E. (2011). La calidad del servicio y la satisfacción del consumidor. *Revista Brasileira de Marketing*, 10(2), 146-162. <https://doi.org/10.5585/remark.v10i2.2212>
- Mora, E. A., Amelia, A., Carrasco, S., Paola, V., & Muñoz, M. (2015). Características de la prueba piloto : revisión de artículos publicados en enfermería, 14(3), 169-175.
- Pérez Campdesuñer, R., Martínez Vivar, R., Noda Hernández M.y Miguel Guzmán, M. (2015). Redalyc. La satisfacción del cliente, influencia en la lealtad. Análisis del destino Holguín. *Redalyc*, XXI, 1-17. Recuperado de <http://www.redalyc.org/articulo.oa?id=181541051007>
- Pianta, R. C., Paro, K. M. La, & Hamre, B. K. (2008). *The Classroom Assessment Scoring System Manual, K-3*. (Brookes Publishing Co, Ed.).

- R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria. Recuperado de <https://www.r-project.org/>
- Racet-Valdés, A., Espinosa-González, L., Suárez-Quintana, J., Alfonso-Robaina, D., & Martínez-Pérez, E. (2017). Modelo matemático para medir el nivel de servicio al cliente basado en la lógica difusa compensatoria Mathematical model to evaluate the customer service based on compensatory fuzzy logic, *XXXVIII*(2), 193-200.
- Skulmoski, G. J., & Hartman, F. T. (2007). The Delphi Method for Graduate Research. *Journal of Information Technology Education*, *6*(1), 1-21. Recuperado de <https://doi.org/10.1.1.151.8144>
- Spiegel, M. R., & Stephens, L. j. (2009). *Estadística*. (McGRAW-HILL, Ed.) (Cuarta). México.
- Stipek, D., & Byler, P. (2004). The early childhood classroom observation measure. *Early Childhood Research Quarterly*, *19*(3), 375-397. Recuperado de <http://dx.doi.org/10.1016/j.ecresq.2004.07.007>
- Supo, J. (2013). *Cómo validar un instrumento: La guía para validar un instrumento en 10 pasos*. México DF.
- Tumino, M. C., & Poitevin, E. R. (2013). Evaluación de la calidad de servicio universitario desde la percepción de estudiantes y docentes: caso de estudio Quality assessment of university service from students and teacher's perceptions: a case study. *REICE. Revista Iberoamericana sobre Calidad*, *12*(2), 63-84. Recuperado de www.rinace.net/reice/
- Turkyilmaz, A., Oztekin, A., Zaim, S., & Fahrettin Demirel, O. (2013). Universal structure modeling approach to customer satisfaction index. *Industrial Management & Data Systems*, *113*(7), 932-949. Recuperado de <https://doi.org/10.1108/IMDS-12-2012-0444>

Para citar este documento:

Montero-Mora, J. y Cantón-Croda, R. (2020). Validación de un instrumento para medir satisfacción de usuarios en instituciones educativas del sector privado: ciudad de Xalapa-Enríquez (México). *Revista Innovaciones Educativas*. *22*(32), 122-136. DOI: 10.22458/ie.v22i32.2727