

Uso y aceptación de la evaluación automática: estudio de traducción y confiabilidad de un instrumento para medirla

*Use and acceptance of automatic evaluation:
study of translation and reliability of an instrument to measure it*

Julián Gerardo Torres Kauffman • Iván Olmos Pineda • Juan Manuel González Calleros

RESUMEN

El rápido desarrollo de las tecnologías de la información y la comunicación –TIC– ha abierto nuevas posibilidades para las prácticas de enseñanza y evaluación en la educación superior. Este estudio cuantitativo, de diseño cuasiexperimental, examina la confiabilidad de la versión en español de una escala para evaluar la aceptación de la evaluación basada en computadora –*Computer Based Assessment*, CBA– entre estudiantes universitarios. Para llevar a cabo la investigación se utilizaron herramientas de análisis descriptivo, como el coeficiente alfa de Cronbach, para evaluar la confiabilidad del *Modelo de aceptación de evaluación basada en computadora* –*Computer Based Assessment Acceptance Model*, CBAAM– en una universidad privada de México. La investigación se dividió en dos fases: 1) aplicación de una evaluación con preguntas de respuesta corta evaluadas automáticamente, y 2) análisis de la confiabilidad de la versión final del CBAAM con 84 estudiantes de primer semestre de ingeniería. El CBAAM mostró un coeficiente de consistencia interna alto y estadísticamente significativo (α de Cronbach = 0.9190). Los coeficientes de consistencia interna de las subescalas variaron entre 0.6838 y 0.9325, lo que sugiere que la escala es confiable para su uso en este contexto. Se recomienda aplicar el CBAAM en otros contextos para mayor confirmación.

Palabras clave: Influencia de la tecnología, estudiantes universitarios, pruebas asistidas por computadora, actitudes de los estudiantes, automatización.

ABSTRACT

The rapid development of Information and Communication Technologies (ICT) has opened new possibilities for teaching and assessment practices in higher education. This quantitative study, with a quasi-experimental design, examines the reliability of the Spanish version of a scale for evaluating the acceptance of Computer-Based Assessment (CBA) among university students. To conduct the research, descriptive analysis tools, such as Cronbach's alpha coefficient, were used to evaluate the reliability of the Computer-Based Assessment Acceptance Model (CBAAM) in a private University in Mexico. The research was divided into two phases: 1) application of an assessment with automatically graded short-answer questions, and 2) analysis of the reliability of the final version of the CBAAM with 84 first-semester engineering students. The CBAAM showed a high and statistically significant internal consistency coefficient (Cronbach's alpha = 0.9190). The internal consistency coefficients of the subscales ranged from 0.6838 to 0.9325, suggesting that the scale is reliable for use in this context. It is recommended to apply the CBAAM in other contexts for further confirmation.

Keywords: Influence of technology, college students, computer assisted testing, student attitudes, automation.

INTRODUCCIÓN

Durante los años recientes las modalidades de enseñanza y aprendizaje sincrónicas y asincrónicas han atraído considerable atención debido a la contingencia originada por la pandemia de COVID-19 (Chau et al., 2021). En esta situación excepcional la digitalización en las instituciones educativas se ha acelerado, ya que muchas universidades han optado por implementar la enseñanza electrónica –*e-teaching*– y el aprendizaje electrónico –*e-learning*– para sustituir las actividades educativas tradicionales en el aula (Tang et al., 2021).

En realidad, la enseñanza y el aprendizaje electrónicos han sido adoptado durante más de una década en muchas universidades de todo el mundo, impulsados por el rápido desarrollo de las tecnologías de la información y la comunicación –TIC– (Alexiou-Ray y Bentley, 2015). Este fenómeno también ha llamado enormemente la atención en investigaciones anteriores (Mohammadyari y Singh, 2015). Las tecnologías de enseñanza y aprendizaje electrónicos abarcan una variedad de herramientas, incluyendo dispositivos móviles (Tang y Yu, 2018), plataformas de aprendizaje en línea, cursos *online* masivos y abiertos –*Massive Online Open Courses*, MOOC– (Bralić y Divjak, 2018) y tecnologías de realidad virtual y mixta, entre otras (Tang et al., 2020; Tang y Yu, 2018). La aceptación o rechazo de estas aplicaciones tecnológicas por parte de los usuarios, así como sus correspondientes instrumentos de medición en diferentes campos, incluida la educación, han sido temas de gran interés para los investigadores (Granic y Marangunic, 2019; Marangunic y Granic, 2015).

Aunque han surgido numerosos modelos de investigación en este campo, el *Modelo de aceptación de tecnología* –*Technology Acceptance Model*, TAM–, introducido por Davis (1989), se ha consolidado como uno de los modelos más importantes para entender los

Julián Gerardo Torres Kauffman. Benemérita Universidad Autónoma de Puebla, México. Actualmente se desempeña como profesor y coordinador de Innovación Educativa en el Tecnológico de Monterrey, Campus Puebla, y es estudiante del Doctorado en Sistemas y Ambientes Educativos. Participa en proyectos de innovación educativa tanto a nivel regional como nacional y promueve activamente el uso de tecnología en la enseñanza y las plataformas institucionales. Correo electrónico: julian.torresk@alumno.buap.mx. ID: <https://orcid.org/0000-0001-7247-2769>.

Iván Olmos Pineda. Profesor Investigador de la Facultad de Ciencias de la Computación de la Benemérita Universidad Autónoma de Puebla, México. Es Doctor en Ciencias de la Computación por el Instituto Nacional de Astrofísica, Óptica y Electrónica. Miembro del Sistema Nacional de Investigadoras e Investigadores, Nivel I. Tiene el reconocimiento al Perfil PRODEP, así como miembro del Padrón de Investigadores BUAP. Tiene publicaciones en revistas arbitradas en el área de visión computacional y aprendizaje máquina, así como en temas educativos apoyados por sistemas computacionales, en donde ha involucrado a estudiantes de licenciatura, maestría y doctorado. Correo electrónico: ivan.olmos@viep.com. ID: <https://orcid.org/0000-0003-1698-000X>.

Juan Manuel González Calleros. Profesor-Investigador en la Benemérita Universidad Autónoma de Puebla, México. Es investigador en el ámbito de la educación y la tecnología, con numerosas publicaciones en revistas indexadas y conferencias internacionales. Sus trabajos abarcan temas como la aceptación y uso de evaluaciones basadas en computadora, estrategias para encontrar recursos educativos abiertos y el desarrollo de sistemas de tutoría afectiva. Además ha investigado sobre la inclusión digital en México y las competencias digitales en instituciones de educación superior. Su enfoque interdisciplinario combina la ingeniería, la informática y la pedagogía. Correo electrónico: juanmanuel.gonzalezcalleros@correo.buap.mx. ID: <https://orcid.org/0000-0002-9661-3615>.

factores que influyen en la adopción de tecnología por parte de los usuarios. Además se ha enfatizado la importancia de aplicar la tecnología en las actividades de enseñanza y aprendizaje (Davis, 2011; Davis et al., 2011; Scherer et al., 2019). Entre estas actividades, la evaluación es uno de los elementos clave en la práctica educativa.

La evaluación es una medida utilizada para evaluar el desempeño y el progreso de un individuo (Llamas-Nistal et al., 2013). Sin embargo, como mencionan Viramontes Anaya y Viramontes Campos (2021), la evaluación basada únicamente en la medición ha sido superada en la literatura académica. Una definición frecuente y acertada, con la que coinciden Casanova (2020), De la Oliva (2020) y Miller (2020), conceptualiza la evaluación como un proceso sistemático de recolección de información sobre el avance y forma en que se da el aprendizaje. Este proceso es seguido de una valoración con la intención de mejorar la consecución de las intenciones educativas que se pretende desarrollar en el alumnado.

La evaluación electrónica *—e-assessment—*, según su definición más amplia (JISC, 2007), incluye todo uso de una computadora como parte de cualquier actividad relacionada con la evaluación, ya sea sumativa, formativa o diagnóstica. Así, su alcance abarca desde la presentación en línea de una tarea para la evaluación por un humano, la evaluación de un portafolio electrónico, la retroalimentación entregada por archivos de audio grabados en la computadora y, más comúnmente, hasta los exámenes en línea calificados por computadora.

Al-Qdah y Ababneh (2017), Bukie (2014), Jordan (2013), Kundu y Bej (2021) y Timmis et al. (2016) mencionan que la evaluación basada en computadora, evaluación digital, pruebas basadas en computadora, evaluación asistida por computadora, pruebas asistidas por computadora, pruebas administradas por computadora, evaluación mejorada por la tecnología, evaluación habilitada por la tecnología, evaluación informatizada, pruebas informatizadas, evaluación basada en la *web*, examen electrónico, pruebas electrónicas y evaluación en línea, son el conjunto de términos que generalmente se consideran sinónimos de evaluación electrónica dentro de la literatura.

Aunque la adopción de la enseñanza electrónica ha sido ampliamente estudiada, una investigación limitada se ha centrado en la evaluación basada en computadora *—Computer Based Assessment, CBA—* (Mo et al., 2022). La mayoría de las plataformas de aprendizaje electrónico de última generación brindan soporte para la evaluación en línea entre sus características y, en muchos casos, este soporte incluye la evaluación automática de pruebas. Un buen ejemplo son las pruebas de opción múltiple, en las que los estudiantes tienen que elegir la respuesta correcta entre varias posibilidades.

Sin embargo, en muchos casos los profesores prefieren exámenes abiertos o de respuesta abierta, donde los estudiantes responden libremente a las preguntas planteadas, utilizando sus propias palabras y complementándolas con gráficos, fórmulas y otros elementos. En este caso, y además de las respuestas reales proporcionadas, los profesores también pueden evaluar otros aspectos como la expresividad, la organización, la originalidad o la capacidad de resumir.

En el caso de las pruebas de respuesta abierta, las herramientas de evaluación más populares (por ejemplo, las proporcionadas por las plataformas de aprendizaje electrónico, como Moodle, Blackboard, Canvas, etc.) no admiten la evaluación automática. La evaluación automática de pruebas de respuesta abierta es un área de investigación actual que combina técnicas de *procesamiento del lenguaje natural* –PLN– e *inteligencia artificial* –IA–. Las contribuciones y resultados relevantes aún son escasos en este campo (Llamas-Nistal et al., 2013). Es importante mencionar que en este campo existe específicamente un tipo de pregunta que se conoce como *de respuesta corta* y la longitud de la respuesta debe ser aproximadamente entre una frase y un párrafo. La evaluación de las respuestas debe centrarse en el contenido en lugar del estilo de escritura (Burrows et al., 2015).

Aceptación de la evaluación basada en computadora

Si bien se han realizado muchos estudios de aceptación de tecnología en el campo de la educación, la gran mayoría de ellos han estado en la aceptación del *e-learning*. Además, la mayoría de estos estudios utilizaron el TAM como el modelo base, ampliándolo con otros constructos. Son pocos los estudios que han utilizado el TAM sin extenderlo (Imtiaz y Maarop, 2014).

El TAM se desarrolló específicamente para modelar la aceptación de las tecnologías relacionadas con la informática por parte de los usuarios (Davis, 1989; Davis et al., 1989). El TAM sugiere que la aceptación del uso de una nueva aplicación TIC está significativamente determinada por dos factores: la *utilidad percibida* –PU– y la *facilidad de uso percibida* –PEOU–. Según las definiciones de TAM, PU se define como el grado en que una persona cree que usar un sistema en particular mejorará su desempeño laboral, y PEOU se define como el grado en que una persona cree que usar un sistema en particular mejorará su desempeño laboral sin esfuerzo (Davis, 1989).

Después de su introducción, el TAM y sus modelos sucesores se han aplicado ampliamente para examinar varios tipos de factores que afectan el comportamiento de los estudiantes hacia la aplicación de las TIC en varios dominios, como la biblioteca electrónica (Jeong, 2011; Joo y Choi, 2015), el aprendizaje electrónico (Mac Callum et al., 2014; Ngai et al., 2007) y la evaluación electrónica (Nikou y Economides, 2017). El *Modelo de teoría unificada de aceptación y uso de tecnología* –UTAUT– desarrollado por Venkatesh et al. (2003) fue un sucesor importante entre los modelos de extensión del TAM. En el modelo UTAUT se aplicaron factores complementarios como la *influencia social* –SI– y la *ansiedad informática* –CA– para examinar la intención de comportamiento de los usuarios hacia una nueva aplicación TIC.

Con relación a los instrumentos sobre la aceptación de la evaluación basada en computadora, se han realizado pocos trabajos de investigación. La mayoría de estos estudios han sido llevados a cabo por Terzis et al. (2012a, 2012b, 2013), Terzis y Economides

(2011) y Nikou y Economides (2017). Terzis y Economides (2011) desarrollaron un modelo de escala con 30 ítems, conocido como el *Modelo de aceptación de evaluación basado en computadora –Computer Based Assessment Acceptance Model, CBAAM–*, para investigar la intención de los estudiantes de utilizar la *evaluación basada en computadora –EBC–*. Es importante destacar que la aplicación original de este modelo se diseñó específicamente para evaluaciones automáticas basadas en preguntas de opción múltiple. El CBAAM incluye nueve constructos que abarcan diversas percepciones del estudiante y factores externos que influyen en la intención de uso de la EBC. Las definiciones de cada constructo se presentan en la Tabla 1. El modelo fue construido sobre la base de modelos de aceptación previos, como el *Modelo de aceptación de tecnología –TAM–*, *Teoría del comportamiento planificado –TPB–* y *Teoría unificada de la aceptación y uso de la tecnología –UTAUT–*.

A pesar de su relevancia, el CBAAM tiene limitaciones significativas. Su enfoque en evaluaciones de opción múltiple deja un vacío respecto a otros formatos, como las evaluaciones de respuesta corta. Además, los estudios se centran en pocos contextos y poblaciones, limitando la generalización de los hallazgos.

Es necesario desarrollar y validar nuevos instrumentos que midan la aceptación de diversos formatos de EBC en distintos contextos educativos. La investigación futura debe explorar cómo diferentes tipos de evaluaciones basadas en computadora afectan la percepción y la intención de uso de los estudiantes, y validar estos instrumentos en diversas poblaciones para asegurar su robustez. Adaptar estos modelos a diversos contextos educativos contribuirá significativamente al conocimiento en el campo de la EBC.

En México, Madrid et al. (2016) realizaron un metaanálisis sobre la educación mediada por las TIC. A partir de la revisión en bases de datos y sitios web especializados, considerando artículos de revistas, ponencias de congresos y tesis del año 2004 al 2014 identificaron que el 49% de las investigaciones se han realizado en el nivel superior. En cuanto a los instrumentos relacionados con las TIC en México, se han enfocado al estudio de la competencia digital a un único contexto educativo, sin contemplar el acceso y hábitos de uso de las TIC (Olivares et al., 2016). Por lo tanto, se requieren instrumentos formales que puedan complementar una evaluación objetiva sobre el uso de las TIC en los estudiantes (Tirado-Lara y Roque-Hernández, 2022).

En la educación superior, los sistemas de aprendizaje y evaluación electrónica son cada vez más relevantes por su capacidad de ofrecer experiencias educativas flexibles, accesibles y personalizadas. La EBC se ha vuelto esencial, facilitando evaluaciones eficientes y proporcionando retroalimentación inmediata y adaptativa con recursos digitales, con la intención de mejorar el aprendizaje estudiantil. Dado su papel crucial, es vital investigar los factores que influyen en la actitud de los estudiantes hacia la EBC, ya que esta actitud puede afectar su aceptación y uso. Una actitud positiva fomenta un mayor compromiso y aprovechamiento de los beneficios de la EBC, mientras que una actitud negativa puede obstaculizar su implementación y éxito.

Para comprender la actitud de los estudiantes hacia la EBC es indispensable contar con instrumentos de medición robustos y fiables. Estos instrumentos deben identificar y analizar diversos factores que pueden influir en la actitud de los estudiantes. En este contexto, el objetivo de este trabajo es analizar la confiabilidad de una escala adaptada para medir la aceptación y la intención de uso de un sistema de evaluación basado en preguntas de respuesta corta con evaluación automática por parte de estudiantes universitarios mexicanos. Para ello se adaptó y tradujo al español el instrumento CBAAM de Terzis y Economides (2011), con el fin de evaluar su aplicación en un contexto y época diferentes de evaluación basada en computadora.

Cabe destacar que al revisar la literatura no se encontraron instrumentos con reportes de confiabilidad para medir la aceptación y la intención de uso por parte de estudiantes universitarios mexicanos de un sistema de evaluación basada en computadora –EBC–.

METODOLOGÍA

El presente estudio se enmarcó en el proceso de desarrollo de un instrumento educativo que se puede utilizar para evaluar la comprensión de los estudiantes de conceptos básicos en física, con la característica de que contiene preguntas en formato de respuesta corta y su correspondiente calificación automática. La evaluación de este tipo de preguntas usando computadora sigue siendo un gran desafío y una importante área de investigación en el campo de la evaluación electrónica. En este contexto, se considera muy relevante estudiar la confiabilidad de la versión en español de las escalas del Modelo de aceptación de evaluación basada en computadora –CBAAM– en una población de estudiantes mexicanos para evaluar la aceptación de esta tecnología.

Participantes

Se trabajó en una universidad privada ubicada en una zona urbana en la ciudad de Puebla. Para el ciclo escolar 2022, dicha institución tenía una matrícula total de 4,398 estudiantes distribuidos en nivel posgrado, profesional y preparatoria. La población objetivo para el estudio realizado fueron los estudiantes de primer semestre de la Escuela de Ingeniería y Ciencias, que contaba con una matrícula de 283 nuevos ingresos.

Siguiendo el protocolo de ética en la investigación institucional, se informó a los estudiantes sobre los objetivos del estudio. Un total de 84 estudiantes asistieron de manera voluntaria a las sesiones para realizar el examen basado en computadora y posteriormente completaron la información del instrumento CBAAM. El muestreo fue no probabilístico y los criterios de inclusión fueron: ser estudiante del primer año de Ingeniería y estar inscrito en el curso de Física aplicada a la ingeniería. La lectura y aceptación del consentimiento informado fueron requisitos indispensables para la

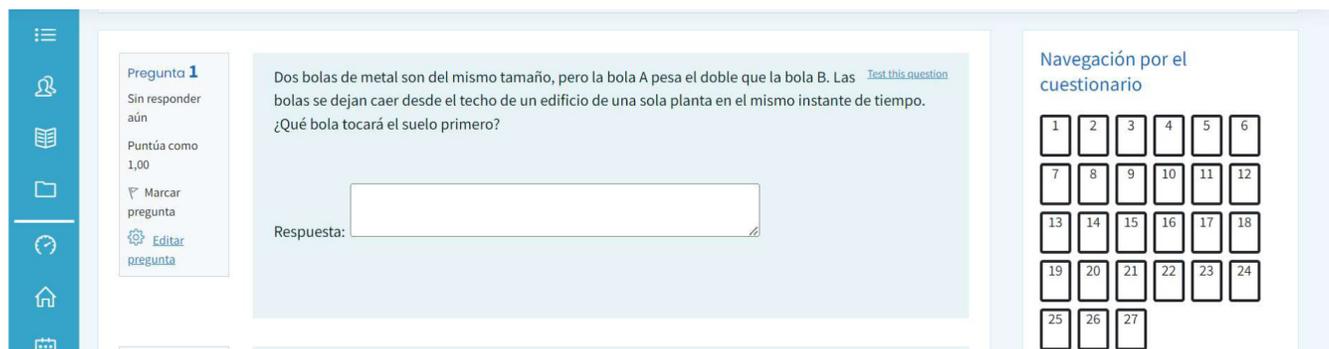
participación. La edad media fue de 18.39 años ($DE = 0.94$) con un rango de 17 a 22 años; por sexo respondieron 26 (30%) mujeres, 54 (65%) hombres y 4 (5%) no lo indicaron. Promedio escolar se ubicó en 93.043 ($DE = 5.33$). El procedimiento de recolección de datos se estructuró en dos fases principales, la primera fase consistió en la administración de un examen basado en computadora, mientras que la segunda implicó la recopilación de las respuestas al instrumento CBAAM por parte de los estudiantes.

Fase 1. Aplicación del instrumento de evaluación

El uso del sistema de evaluación basado en preguntas de respuesta corta con evaluación automática fue muy simple: el estudiante solo tenía que escribir en sus propias palabras la respuesta correcta, con una longitud no mayor a 20 palabras. La duración del examen fue de 30 minutos para 27 preguntas de contenido exclusivamente conceptual, basado en el *inventario de conceptos de fuerza –Force Concept Inventory, FCI–* propuesto por Hestenes et al. (1992). Este instrumento de diagnóstico fue desarrollado con la idea de proporcionar un esquema más sistemático y completo de los diversos conceptos erróneos de mecánica newtoniana de los estudiantes, y cuenta con más de 50 publicaciones que respaldan su validez. En su forma actual, la mayoría de los inventarios de conceptos hacen uso del tipo de pregunta de opción múltiple, por lo que es muy relevante destacar el uso del formato de pregunta de respuesta corta con evaluación automática en el presente trabajo.

La Figura 1 muestra la interfaz del sistema de evaluación a través de una pregunta de muestra. El examen se presenta completo en una sola página. Las instrucciones y el texto estaban en español, y no se dio ninguna otra instrucción especial al principio. La única pregunta que hicieron los estudiantes fue si tenían que contar las palabras, lo cual no era necesario, ya que el propio sistema emitía una advertencia para cada pregunta al exceder las 20 palabras.

Figura 1
Pregunta muestra



Fuente: Adaptado del inventario de conceptos de Hestenes et al., 1992.

Para el desarrollo de esta investigación se eligió utilizar una tecnología desarrollada en la comunidad de *software* libre, la cual se integra al sistema gestor de aprendizaje Moodle a través de un *plugin* llamado “Pattern Match”; con esta integración es posible construir preguntas en formato de respuesta corta y calificarlas de forma automática.

Fase 2. Aplicación del instrumento CBAAM

El proceso de traducción, adaptación y validación intercultural del instrumento para su uso requiere un proceso metodológico riguroso con el objetivo de lograr la equivalencia entre el instrumento original y su traducción (Muñiz et al., 2013; Sousa y Rojjanasrirat, 2011).

El proceso consistió en los siguientes pasos:

1. Traducción del instrumento original al idioma de destino (traducción hacia adelante o de un solo sentido), realizada por el autor con el apoyo de dos colegas expertos en el área de adopción tecnológica.
2. Retrotraducción ciega (traducción hacia atrás o traducción doble ciego) de la versión preliminar traducida del instrumento, realizada por dos investigadores con amplia experiencia en el campo de adopción tecnológica.
3. Comparación de las versiones retrotraducidas del instrumento, con el objetivo de encontrar la equivalencia conceptual, semántica y de contenido. Durante este proceso también se consideraron las traducciones publicadas en el idioma de destino de instrumentos similares aplicados en el campo de la adopción tecnológica.
4. Pruebas de confiabilidad de la versión final del instrumento traducido en una muestra de la población objetivo mediante el análisis descriptivo, utilizando el *software* Minitab 21.3.1.

Previo a la implementación del instrumento CBAAM se llevó a cabo un estudio piloto con 55 estudiantes de primer año de Ingeniería. El objetivo de este estudio fue obtener su opinión sobre la claridad de los reactivos y corregir cualquier aspecto, tanto de forma como de fondo, que dificultara la comprensión. A estos participantes se les informó sobre el propósito de la prueba y se les solicitó su consentimiento informado a través de un formulario electrónico.

El instrumento final fue administrado a través de Formularios de Google, asegurando su disponibilidad inmediata después de que los alumnos completaran el sistema de evaluación basado en preguntas de respuesta corta con evaluación automática. El instrumento contenía los 30 ítems con las siete opciones de respuesta (1 - Absolutamente en desacuerdo, 2 - Fuertemente en desacuerdo, 3 - En desacuerdo, 4 - Ni en acuerdo ni desacuerdo, 5 - De acuerdo, 6 - Fuertemente en acuerdo, 7 - Absolutamente de acuerdo) (Tabla 1).

Tabla 1

Traducción del Modelo de aceptación de evaluación basada en computadora (CBAAM)

Constructo	Definición operacional	Ítems
Utilidad percibida	Se define como el grado en que una persona cree que el uso de un sistema en particular mejorará su rendimiento	1. El uso de la evaluación basada en computadora (EBC) mejorará mi trabajo 2. El uso de la evaluación basada en computadora (EBC) mejorará mi eficacia 3. El uso de la evaluación basada en computadora (EBC) aumentará mi productividad
Facilidad de uso percibida	Se define como el grado en que una persona cree que el uso del sistema estaría libre de esfuerzo	4. Mi interacción con el sistema es clara y comprensible 5. Es fácil para mí llegar a ser hábil en el uso del sistema 6. El sistema es fácil de usar
Autoeficacia informática	Se define como las percepciones del individuo sobre su capacidad para utilizar computadoras	7. Podría completar un trabajo o tarea usando la computadora 8. Podría completar un trabajo o una tarea usando la computadora si alguien me mostrara cómo hacerlo primero 9. Puedo navegar fácilmente por la <i>web</i> para encontrar cualquier información que necesite 10. Era completamente capaz de usar la computadora e Internet antes de comenzar a usar la evaluación basada en computadora (EBC)
Influencia social	El efecto de la opinión de otras personas, la influencia de los superiores y la influencia de los compañeros	11. Las personas que influyen en mi comportamiento piensan que debería usar la EBC 12. Las personas que son importantes para mí piensan que debería usar la EBC 13. Los estudiantes de último año de mi universidad han sido de ayuda en el uso de la EBC 14. En general, mi universidad ha apoyado el uso de la EBC
Condiciones facilitadoras	Son factores que influyen en la creencia de un individuo sobre su capacidad para realizar un procedimiento	15. Cuando necesito ayuda para usar la EBC, alguien está ahí para ayudarme 16. Cuando necesito ayuda para aprender a usar la EBC, el soporte de ayuda está ahí para enseñarme
Contenido	Es uno de los determinantes de la satisfacción del alumno al usar sistemas de aprendizaje	17. Las preguntas de la EBC fueron claras y comprensibles 18. Las preguntas de la EBC fueron fáciles de responder 19. Las preguntas de la EBC eran relativas al temario del curso 20. Las preguntas de la EBC fueron útiles para mi curso
Expectativa de meta	Es una variable que influye en la creencia de un individuo de estar adecuadamente preparado para usar la EBC	21. La preparación de los cursos fue suficiente para la EBC 22. Mi preparación personal fue suficiente para la EBC 23. Mis expectativas de desempeño fueron acordes a los resultados de la EBC
Percepción lúdica	Se define por tres dimensiones: concentración, curiosidad y disfrute del usuario al interactuar con el sistema	24. Usar EBC me mantiene feliz para realizar mi tarea 25. Usar EBC me hace disfrutar de mi aprendizaje 26. Usando EBC, mi curiosidad se estimula 27. El uso de EBC conducirá a mi exploración
Intención de utilizar la EBC	Se refiere al grado en que una persona tiene la intención de utilizar una tecnología en particular	28. Tengo la intención de utilizar EBC en el futuro 29. Predigo que usaría EBC en el futuro 30. Planeo usar EBC en el futuro

Fuente: Adaptado de Terzis y Economides, 2011.

Tabla 2*Media y desviación estándar de las puntuaciones totales por ítem*

Constructo	Ítems	Media	D.E.	Asimetría	Curtosis
Utilidad percibida	UP1:	5.890	0.943	-0.68	0.11
	UP2:	5.927	0.927	-0.62	-0.37
	UP3:	5.890	1.155	-1.61	3.91
Facilidad de uso percibida	FU4:	6.073	1.225	-1.59	2.87
	FU5:	6.207	1.086	-1.73	3.81
	FU6:	6.451	1.032	-2.87	10.38
Autoeficacia unformática	AI7:	6.5732	0.8753	-3.80	20.24
	AI8:	6.463	1.239	-3.07	9.90
	AI9:	6.5122	0.8783	-2.78	10.23
	AI10:	6.476	1.091	-2.62	7.02
Influencia social	IS11:	5.220	1.548	-0.56	-0.14
	IS12:	5.061	1.485	-0.50	-0.03
	IS13:	5.049	1.742	-0.77	-0.11
	IS14:	5.878	1.299	-1.26	1.17
Condiciones facilitadoras	CF15:	5.610	1.377	-1.12	1.23
	CF16:	5.854	1.268	-1.21	1.34
Contenido	C17:	5.829	1.368	-1.31	1.64
	C18:	5.354	1.290	-0.38	-0.58
	C19:	5.805	1.401	-1.27	1.34
	C20:	5.841	1.271	-1.06	0.63
Expectativa de meta	EM21:	5.634	1.436	-0.79	-0.11
	EM22:	4.988	1.544	-0.43	-0.37
	EM23:	4.890	1.771	-0.61	-0.50
Percepción lúdica	PL24:	5.354	1.364	-0.91	1.19
	PL25:	5.488	1.317	-1.17	2.01
	PL26:	5.793	1.264	-1.03	1.07
	PL27:	5.927	1.086	-1.38	3.80
Intención de utilizar la EBC	IU28:	5.939	1.280	-1.48	2.35
	IU29:	6.159	1.262	-1.82	3.59
	IU30:	5.915	1.363	-1.70	3.19

Fuente: Elaboración propia con los resultados obtenidos en Minitab.

RESULTADOS

Los 30 ítems fueron evaluados por sus medias, desviaciones estándar, asimetría y curtosis. Las estadísticas descriptivas presentadas en la Tabla 2 indican una disposición positiva hacia los ítems. Mientras que la desviación estándar (D.E.) osciló entre 0.8753 y 1.742, estos valores indican una dispersión estrecha alrededor de la media. Para los análisis de normalidad univariada se utilizaron los indicadores de asimetría y curtosis de las variables. Curran et al. (1996) establecen los límites de asimetría y curtosis en valores absolutos. Valores de hasta 2 para la asimetría y 7 para la curtosis son considerados normales; valores comprendidos entre 2 y 3 para la asimetría y entre 7 y 21 para la curtosis son considerados moderadamente normales, y valores superiores a 7 en la asimetría y a 21 en la curtosis son considerados no-normales. Posteriormente, para comprobar la normalidad multivariada se realizó un análisis de las puntuaciones obtenidas de todo el instrumento y el resultado mostró un valor $p = 0.656$, lo que nos sugiere una muestra paramétrica.

Para evaluar la confiabilidad del instrumento –CBAAM– se obtuvo un coeficiente de consistencia interna con la prueba de alfa de Cronbach, que resultó en 0.9190, un valor alto. El coeficiente puede oscilar entre 0 y 1, donde 0 significa nula confiabilidad y 1 confiabilidad total.

Se utilizó el *software* Minitab 21.3.1 para efectuar el cálculo; los coeficientes de consistencia interna se encontraron en un rango de 0.6838 a 0.9325. Únicamente la subescala de expectativa de meta está por debajo de 0.70, mínimo que se considera adecuado. En la Tabla 3 se muestran los puntajes en cada subescala del instrumento.

Tabla 3
Confiabilidad de la versión en español de las subescalas del Modelo de aceptación de evaluación basada en computadora

Subescala	CBAAM Español Coeficiente alpha de Cronbach
1. Utilidad percibida	0.7184
2. Facilidad de uso percibida	0.8501
3. Autoeficacia informática	0.8076
4. Influencia social	0.8065
5. Condiciones facilitadoras	0.7619
6. Contenido	0.7386
7. Expectativa de meta	0.6838
8. Percepción lúdica	0.8763
9. Intención de utilizar la EBC	0.9325

Fuente: Elaboración propia con los resultados obtenidos en Minitab.

DISCUSIÓN

En el presente trabajo se realizó un estudio de la versión en español del instrumento asociado al Modelo de aceptación de evaluación basada en computadora –CBAAM–. Los resultados obtenidos sugieren que las subescalas y, en general, la totalidad de la versión en español del instrumento muestran niveles adecuados de confiabilidad para su aplicación en estudiantes de una universidad privada en México. Los índices de consistencia interna de la versión en español presentados en este estudio son congruentes con los reportados para la versión original en inglés de Terzis y Economidis (2011), lo que indica un alto grado de homogeneidad y coherencia entre los ítems de la escala. El coeficiente de consistencia del total de la versión en español del CBAAM fue alto y estadísticamente significativo, al igual que el de las subescalas, con la excepción de la expectativa de meta.

Un aspecto relevante y distintivo de este estudio es su enfoque comparativo, el cual examina la adaptación de una escala de medición para evaluar la aceptación de la evaluación basada en computadora utilizando un formato de evaluación diferente basado en preguntas de respuesta corta. Mientras que todas las investigaciones previas se han centrado en un único formato de evaluación basado en preguntas de opción múltiple, este estudio introduce un enfoque dual que permite una comparación directa de cómo el formato de las preguntas puede influir en la percepción y aceptación estudiantil de la EBC.

Este enfoque comparativo representa una innovación metodológica significativa, ya que proporciona evidencia empírica de que el modelo CBAAM se adapta bien para estudiar los factores de aceptación de evaluaciones basadas en preguntas de respuesta corta. Esta adaptación metodológica no solo amplía el alcance del modelo CBAAM sino que también abre nuevas vías para la investigación en la evaluación educativa, subrayando la importancia de considerar diferentes formatos de evaluación en el estudio de la aceptación de tecnologías educativas.

A pesar de los hallazgos significativos de este estudio, es crucial reconocer sus limitaciones y proponer áreas para futuras investigaciones que puedan superarlas. Una de las principales limitaciones radica en la generalización de los resultados, ya que la muestra se centra en estudiantes de una institución de educación superior específica y en el contexto de la enseñanza de conceptos básicos de física. Además se identificaron dos sesgos importantes: por un lado tenemos que, de acuerdo con la caracterización que se hizo de la muestra de alumnos que participaron en el estudio, se observó que el 75% tiene un promedio académico alto situado en el rango de 100-90, la literatura reporta que los efectos de la evaluación basada en computadora con retroalimentación inmediata son más notorios en estudiantes de rendimiento académico medio-bajo (Jordan y Mitchell, 2009). De manera similar tenemos que de los estudiantes que participaron en el estudio 68% fueron hombres y 32% fueron

mujeres, la literatura reporta que los estudios sobre la aceptación de la tecnología no son iguales para los hombres y para las mujeres (Venkatesh et al., 2003).

El siguiente paso consiste en efectuar pruebas de validez que permitan ampliar y consolidar el conocimiento sobre las propiedades del instrumento en el contexto mexicano. Para futuras investigaciones se recomienda continuar con el proceso de corroborar los resultados sobre las propiedades del CBAAM de este trabajo y resolver las limitaciones que presenta. Una forma de hacerlo implicaría la revisión de posibles errores en la traducción que pudieron afectar la comprensión de algunos reactivos.

En general, se cree que la EBC es útil para los estudiantes, los docentes y las instituciones, ya que brinda flexibilidad en el lugar y el tiempo de la evaluación, mejora los resultados de aprendizaje y facilita una retroalimentación inmediata y de mayor calidad por parte de los docentes (Alruwais et al., 2018). Por otra parte, la abundante información y los datos recopilados por la EBC, como el desempeño integrado de los estudiantes en diferentes evaluaciones relacionadas (por ejemplo, tareas, exámenes y otras actividades educativas) en un periodo específico o durante todo el programa de estudios, sería un activo valioso para los maestros y universidades para analizar de manera sistemática e individual el rendimiento y los resultados de aprendizaje de los estudiantes (Gañán et al., 2017).

CONCLUSIONES

Este estudio tiene implicaciones prácticas ya que aporta un instrumento con evidencia de confiabilidad para la medición del uso y aceptación de la evaluación basada en computadora en el contexto educativo mexicano. Adicionalmente permite a los investigadores conocer las formas en que los estudiantes perciben el uso de las TIC específicamente aplicadas al proceso de evaluación.

Por otra parte, desde el punto de vista teórico, al tratarse de una adaptación de un instrumento ya publicado en otro idioma y contexto, permite la comparación de los constructos e indicadores observables. Se puede concluir que la presente escala presenta suficientes evidencias de confiabilidad para continuar con estudios de validación correspondientes.

A partir de lo anterior se subraya la importancia de contar con instrumentos confiables que aporten datos para el proceso de evaluación de las TIC para un uso específico.

Finalmente, es importante mencionar que, al momento de redactar este artículo, se está llevando a cabo el análisis del CBAAM en este mismo contexto educativo mediante el enfoque de modelado de ecuaciones estructurales. El objetivo es probar tanto el modelo de medición como el modelo estructural del CBAAM, y estudiar las relaciones significativas entre las múltiples variables consideradas en este modelo que influyen en la adopción y el uso exitoso de esta tecnología por parte de los estudiantes.

Agradecimientos

Este trabajo fue realizado con el apoyo de la beca CONAHCYT 769631.

REFERENCIAS

- Alexiou-Ray, J., y Bentley, C. C. (2015). Faculty professional development for quality online teaching. *Online Journal of Distance Learning Administration*, 18(4), 1-7. https://ojdla.com/archive/winter184/ray_bentley184.pdf
- Al-Qdah, M., y Ababneh, I. (2017). Comparing online and paper exams: Performances and perceptions of Saudi students. *International Journal of Information and Education Technology*, 7(2), 106-109. <https://doi.org/10.18178/ijiet.2017.7.2.850>
- Alruwais, N., Wills, G., y Wald, M. (2018). Advantages and challenges of using e-assessment. *International Journal of Information and Education Technology*, 8(1), 34-37. <https://doi.org/10.18178/ijiet.2018.8.1.1008>
- Bralić, A., y Divjak, B. (2018). Integrating MOOCs in traditionally taught courses: Achieving learning outcomes with blended learning. *International Journal of Educational Technology in Higher Education*, 15(1). <https://doi.org/10.1186/s41239-017-0085-7>
- Bukie, O. F. (2014). Understanding technologies for e-Assessment: A systematic review approach. *Journal of Emerging Trends in Computing and Information Sciences*, 5(12). <http://www.cisjournal.org>
- Burrows, S., Gurevych, I., y Stein, B. (2015). The eras and trends of automatic short answer grading. *International Journal of Artificial Intelligence in Education*, 25, 60-117. <https://doi.org/10.1007/s40593-014-0026-8>
- Casanova, M. A. (2020). Evaluación formativa e inclusiva en línea: ¿Cómo evaluar el aprendizaje ante los retos de la pandemia por el Covid 19?. *Memorias del Quinto Congreso Internacional de Evaluación Socioformativa (Valora-2020)*. Centro Universitario CIFE, Cuernavaca, México. <https://cife.edu.mx/recursos>
- Chau, K. Y., Law, K. M., y Tang, Y. M. (2021). Impact of self-directed learning and educational technology readiness on synchronous e-Learning. *Journal of Organizational and End User Computing (JOEUC)*, 33(6), 1-20. <https://doi.org/10.4018/JOEUC.20211101.0a26>
- Curran, P. J., West, S. G., y Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1(1), 16-29. <https://doi.org/10.1037/1082-989X.1.1.16>
- Davis, F. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340. <https://doi.org/10.2307/249008>
- Davis, F. D. (2011). *Foreword in technology acceptance in education: Research and issues*. Sense Publishers.
- Davis, F. D., Bagozzi, R. P., y Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35(8), 982-1003. <https://doi.org/10.1287/mnsc.35.8.982>
- Davis, M. M., Spohrer, J. C., y Maglio, P. P. (2011). Guest editorial: How technology is changing the design and delivery of services. *Operations Management Research*, 4, 1-5. <https://doi.org/10.1007/s12063-011-0046-6>
- De la Oliva, D. (2020). La autoevaluación en las escuelas de educación primaria para la mejora de la calidad educativa y de la inclusión social. *Memorias del Quinto Congreso Internacional de Evaluación Socioformativa (Valora-2020)*. <https://cife.edu.mx/recursos/>
- Gañán, D., Caballé, S., Clarisó, R., Conesa, J., y Bañeres, D. (2017). ICT-FLAG: A web-based e-assessment platform featuring learning analytics and gamification. *International Journal of Web Information Systems*, 13(1), 25-54. <https://doi.org/10.1108/IJWIS-12-2016-0074>
- Granic, A., y Marangunic, N. (2019). Technology acceptance model in educational context: A systematic literature review. *British Journal of Educational Technology*, 50(5), 2572-2593. <https://doi.org/10.1111/bjet.12864>
- Hestenes, D., Wells, M., y Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, 30(3), 141-158. <https://doi.org/10.1119/1.2343497>
- Imtiaz, A., y Maarop, N. (2014). A review of technology acceptance studies in the field of education.

- Jurnal Teknologi*, 69(2), 2180-3722. <https://doi.org/doi:10.11113/jt.v69.3101>
- Jeong, H. (2011). An investigation of user perceptions and behavioral intentions towards the e-library. *Library Collections, Acquisitions, and Technical Services*, 35(2-3), 45-60. <https://doi.org/10.1080/14649055.2011.10766298>
- JISC [Joint Information Systems Committee] (2007). *Effective practice with e-assessment: An overview of technologies, policies and practice in further and higher education*. <https://people.cs.vt.edu/shaffer/cs6604/Papers/eAssessment.pdf>
- Joo, S., y Choi, N. (2015). Factors affecting undergraduates' selection of online library resources in academic tasks usefulness, ease-of-use, resource quality, and individual differences. *Library Hi Tech*, 33(2), 272-291. <https://doi.org/10.1108/LHT-01-2015-0008>
- Jordan, S. (2013). E-assessment: Past, present and future. *New Directions*, 9(1), 87-106. <https://oro.open.ac.uk/38536/>
- Jordan, S., y Mitchell, T. (2009). e-Assessment for learning? The potential of short-answer free-text questions with tailored feedback. *British Journal of Educational Technology*, 40(2), 371-385. <https://doi.org/10.1111/j.1467-8535.2008.00928.x>
- Kundu, A., y Bej, T. (2021). Experiencing e-assessment during COVID-19: an analysis of Indian students' perception. *Higher Education Evaluation and Development*, 15(2), 114-134. <https://doi.org/10.1108/heed-03-2021-0032>
- Llamas-Nistal, M., Fernández-Iglesias, M. J., González-Tato, J., y Mikic-Fonte, F. A. (2013). Blended e-assessment: Migrating classical exams to the digital world. *Computers & Education*, 62, 72-87. <https://doi.org/10.1016/j.compedu.2012.10.021>
- Mac Callum, K., Jeffrey, L., y Kinshuk (2014). Comparing the role of ICT literacy and anxiety in the adoption of mobile learning. *Computers in Human Behavior*, 39, 8-19. <https://doi.org/10.1016/j.chb.2014.05.024>
- Madrid, G., Angulo, J., García, L., y Olivares, C. (2016). *Estado del conocimiento en la educación mediada por TIC en México*. CIATA.
- Marangunic, N., y Granic, A. (2015). Technology acceptance model: A literature review from 1986 to 2013. *Universal Access in the Information Society*, 14, 81-95. <https://doi.org/10.1007/s10209-014-0348-1>
- Miller, A. (2020, abr. 7). Formative assessment in distance learning. *Edutopia*. <https://www.edutopia.org/article/formative-assessment-distance-learning>
- Mo, D. Y., Tang, Y. M., Wu, E. Y., y Tang, V. (2022). Theoretical model of investigating determinants for a successful Electronic Assessment System (EAS) in higher education. *Education and Information Technologies*, 27(9), 12543-12566. <https://doi.org/10.1007/s10639-022-11098-1>
- Mohammadyari, S., y Singh, H. (2015). Understanding the effect of e-learning on individual performance: The role of digital literacy. *Computers & Education*, 82, 11-25. <https://doi.org/10.1016/j.compedu.2014.10.025>
- Muñiz, J., Elosua, P., y Hambleton, R. K. (2013). Directrices para la traducción y adaptación de los tests: segunda edición. *Psicothema*, 25(2), 151-157. <https://doi.org/10.7334/psicothema2013.24>
- Ngai, E. W., Poon, J. K. L., y Chan, Y. H. (2007). Empirical examination of the adoption of WebCT using TAM. *Computers & Education*, 48(2), 250-267. <https://doi.org/10.1016/j.compedu.2004.11.007>
- Nikou, S. A., y Economides, A. A. (2017). Mobile-based assessment: Investigating the factors that influence behavioral intention to use. *Computers & Education*, 109, 56-73. <https://doi.org/10.1016/j.compedu.2017.02.005>
- Olivares, K., Angulo, A., Torres, G., y Madrid, G. (2016). Validación de un modelo de medida para la competencia digital en estudiantes universitarios. En M. E. Prieto y S. J. Pech (eds.), *La tecnología como instrumento para potenciar el aprendizaje* (pp. 72-78). CIATA-UCLM. <https://doi.org/10.13140/RG.2.1.3802.3920>
- Scherer, R., Siddiq, F., y Tondeur, J. (2019). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education*, 128, 13-35. <https://doi.org/10.1016/j.compedu.2018.09.009>
- Sousa, V. D., y Rojjanasrirat, W. (2011). Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: A clear and user-friendly guideline. *Journal of Evaluation in Clinical Practice*, 17(2), 268-274. <https://doi.org/10.1111/j.1365-2753.2010.01434.x>

- Tang, Y. M., Au, K. M., Lau, H. C. W., Ho, G. T. S., y Wu, C. H. (2020). Evaluating the effectiveness of learning design with mixed reality (MR) in higher education. *Virtual Reality*, 24(4), 797-807. <https://doi.org/10.1007/s10055-020-00427-9>
- Tang, Y. M., Chen, P. C., Law, K. M. Y., Wu, C. H., Lau, Y.-y., Guan, J., He, D., y Ho, G. T. S. (2021). Comparative analysis of Student's live online learning readiness during the coronavirus (COVID-19) pandemic in the higher education sector. *Computers & Education*, 168, 104211. <https://doi.org/10.1016/j.compedu.2021.104211>
- Tang, Y. M., y Yu, K. M. (2018). Development and evaluation of a mobile platform for teaching mathematics of CAD subjects. *Computer-Aided Design and Applications*, 15(2), 164-169. <https://doi.org/10.1080/16864360.2017.1375665>
- Terzis, V., y Economides, A. A. (2011). The acceptance and use of computer based assessment. *Computers & Education*, 56(4), 1032-1044. <https://doi.org/10.1016/j.compedu.2010.11.017>
- Terzis, V., Moridis, C. N., y Economides, A. A. (2012a). How student's personality traits affect Computer Based Assessment Acceptance: Integrating BFI with CBAAM. *Computers in Human Behavior*, 28(5), 1985-1996. <https://doi.org/10.1016/j.chb.2012.05.019>
- Terzis, V., Moridis, C. N., y Economides, A. A. (2012b). The effect of emotional feedback on behavioral intention to use computer based assessment. *Computers & Education*, 59(2), 710-721. <https://doi.org/10.1016/j.compedu.2012.03.003>
- Terzis, V., Moridis, C. N., y Economides, A. A. (2013). Continuance acceptance of computer based assessment through the integration of user's expectations and perceptions. *Computers & Education*, 62, 50-61. <https://doi.org/10.1016/j.compedu.2012.10.018>
- Timmis, S., Broadfoot, P., Sutherland, R., y Oldfield, A. (2016). Rethinking assessment in a digital age: Opportunities, challenges and risks. *British Educational Research Journal*, 42(3), 454-476. <https://doi.org/10.1002/berj.3215>
- Tirado-Lara, P. J., y Roque-Hernández, M. del P. (2022). Validación de la escala Uso y función de las TIC en contextos educativos para estudiantes de educación superior. *RELATEC Revista Latinoamericana de Tecnología Educativa*, 21(1), 9-25. <https://doi.org/10.17398/1695-288X.21.1.9>
- Venkatesh, V., Morris, M. G., Davis, G. B., y Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425-478. <https://doi.org/10.2307/30036540>
- Viramontes Anaya, E., y Viramontes Campos, O. (2021). Evaluación en educación a distancia por confinamiento. *IE Revista de Investigación Educativa de la REDIECH*, 12, e1216. https://doi.org/10.33010/ie_rie_rediech.v12i0.1216

Cómo citar este artículo:

Torres Kauffman, J. G., Olmos Pineda, I., y González Calleros, J. M. (2024). Uso y aceptación de la evaluación automática: estudio de traducción y confiabilidad de un instrumento para medirla. *IE Revista de Investigación Educativa de la REDIECH*, 15, e1923. https://doi.org/10.33010/ie_rie_rediech.v15i0.1923



Todos los contenidos de *IE Revista de Investigación Educativa de la REDIECH* se publican bajo una licencia de Creative Commons Reconocimiento-NoComercial 4.0 Internacional, y pueden ser usados gratuitamente para fines no comerciales, dando los créditos a los autores y a la revista, como lo establece la licencia.