

CAPÍTULO 36.

DISEÑO Y VALIDACIÓN DE UN INSTRUMENTO DE EVALUACIÓN DE ACTITUDES (EAM) BASADO EN EL MODELO ABC PARA CONTEXTOS DE APRENDIZAJE HÍBRIDO

Zoila Esther Cherres Lopez¹

¹: Docente.

 <https://orcid.org/0000-0001-7244-4493>

RESUMEN

Esta revisión narrativa sistematiza la evidencia internacional (1977-2022) sobre el diseño y validación de instrumentos que miden actitudes hacia el aprendizaje híbrido articulando los componentes Afectivo, Conductual y Cognitivo (modelo ABC). De 1.247 registros se analizaron 32 estudios empíricos, revelando escalas fiables ($\alpha > .85$) y estructura tridimensional confirmada solo en nueve investigaciones; la dimensión conductual presenta la menor robustez. Se detecta ausencia de un instrumento estandarizado, invariante y de libre uso en español, así como estudios longitudinales y análisis de funcionamiento diferencial por plataforma LMS. Se propone un protocolo de seis fases para desarrollar el EAM-ABC-H que permita a instituciones diagnosticar barreras actitudinales y diseñar intervenciones focalizadas para mejorar la experiencia estudiantil en contextos híbridos.

Palabras clave: aprendizaje híbrido, modelo ABC, validación de escalas, actitudes estudiantiles, invarianza métrica.

INTRODUCCIÓN

La expansión del aprendizaje híbrido—combinación de instrucción presencial y actividades mediadas por tecnología— ha intensificado la necesidad de comprender los factores psicoeducativos que inciden en su efectividad (Graham, 2019). Entre ellos, la actitud del estudiante se perfila como predictor temprano del rendimiento académico, la persistencia y la satisfacción con la experiencia formativa (Alvarez & Gómez, 2021; Bond et al., 2020). Sin embargo, la actitud no es un constructo unidimensional: implica evaluaciones afectivas hacia el entorno de aprendizaje, creencias cognitivas sobre su utilidad e intenciones conductuales de uso de las herramientas digitales (Breckler, 1984). En contextos híbridos, estas dimensiones se amplifican porque el estudiante debe coordinar recursos tecnológicos, expectativas docentes y dinámicas sociales que no siempre están alineadas (Boelens et al., 2018). Por ello, disponer de instrumentos fiables y válidos que midan la actitud se torna crítico para diagnosticar barreras de adopción y diseñar intervenciones focalizadas.

El modelo ABC (Afectivo, Conductual, Cognitivo) propuesto originalmente por Rosenberg y Hovland (1960) y reformulado por Breckler (1984) ofrece un marco teórico robusto para descomponer la actitud en sus tres componentes interrelacionados. El componente afectivo refleja emociones y sentimientos (p. ej., “Me siento ansioso cuando uso la plataforma LMS”), el cognitivo engloba creencias y conocimientos (“El aprendizaje híbrido me permite comprender mejor los contenidos”) y el conductual incluye predisposiciones a la acción (“Participaré activamente en los foros virtuales”). Esta triada ha sido validada transversalmente en ámbitos como el consumo de tecnología (Venkatesh et al., 2012) y la educación en línea (Keller & Karau, 2020), pero su aplicación específica a la modalidad híbrida aún es incipiente y dispersa.

A pesar de la creciente producción académica, no se ha identificado un instrumento estandarizado —en español o portugués— que integre

explícitamente los tres componentes ABC y que haya demostrado evidencias de validez y confiabilidad en poblaciones de educación básica y superior bajo modalidad híbrida (Reyes & Fernández, 2019; Alarcón-Urbistondo & Jornet, 2022). La mayoría de las escalas existentes se centran en la satisfacción o en la aceptación tecnológica (p. ej., TAM), dejando de lado la predisposición conductual y la carga emocional que implica alternar espacios físicos y virtuales (Parkes et al., 2015). Esta brecha limita la comparabilidad de resultados entre regiones y dificulta la implementación de mejoras basadas en evidencia.

En este marco, el presente artículo de revisión narrativa se propone sistematizar la evidencia internacional sobre el diseño y validación de instrumentos de evaluación de actitudes (EAM) fundados en el modelo ABC y aplicados en contextos de aprendizaje híbrido. El objetivo general es describir cómo se han construido y validado estos instrumentos, mientras que los objetivos específicos son: (a) identificar las propiedades psicométricas reportadas (confiabilidad, validez de constructo, invarianza), (b) evaluar la calidad de la evidencia de validez mediante estándares internacionales (COSMIN; Prinsen et al., 2018) y (c) proponer un esquema de validación para el desarrollo de un EAM-ABC-H (Híbrido) que pueda ser utilizado en poblaciones hispanohablantes.

Alcanzar estos propósitos permitirá a investigadores y gestores educativos contar con un protocolo replicable para medir la actitud hacia el aprendizaje híbrido, facilitando diagnósticos precisos y estrategias de intervención orientadas a mejorar la experiencia estudiantil y, en última instancia, el éxito académico en la nueva normalidad educativa.

METODOLOGÍA

Se realizó una revisión narrativa sistematizada siguiendo los principios de Snyder (2019) y las recomendaciones de Ferrari (2015) para integrar rigor

y flexibilidad interpretativa. Este diseño permite mapear ampliamente la evidencia, identificar tendencias conceptuales y metodológicas, y formular proposiciones para el desarrollo futuro de instrumentos (Wong et al., 2021).

La pregunta orientadora fue: «¿Cómo se han construido y validado instrumentos que midan actitudes bajo el modelo ABC en contextos de educación híbrida?». Para responderla se diseñó una estrategia de búsqueda sensible y reproducible en seis bases de datos: ERIC, PsycINFO, Scopus, Web of Science, Redalyc y SciELO. Se combinó vocabulario controlado (MeSH y DeCS) y términos libres en español, inglés y portugués. El string final, adaptado a las sintaxis de cada motor, fue: («actitud*» OR «attitude scale*» OR «escala de atitute*» OR «affective-behavioral-cognitive» OR «model* ABC») AND («blended learning» OR «hybrid learning» OR «b-learning» OR «aprendizaje híbrido» OR «aprendizagem híbrida»). Se truncaron raíces para maximizar la recuperación y se utilizó el operador NEAR/3 cuando la plataforma lo permitió.

Los criterios de inclusión fueron: (a) artículos empíricos que informaran construcción o validación de escalas; (b) población escolar o universitaria; (c) modalidad híbrida explícita (combinación sincrónica o asincrónica de sesiones presenciales y en línea); (d) publicados entre enero de 1977 y diciembre de 2022. Se excluyeron estudios con escalas unidimensionales, aquellos que solo reportaron validez de contenido sin análisis factorial, y experiencias 100 % presenciales o totalmente en línea. No se impuso restricción de idioma para capturar la producción latinoamericana y lusófona.

La gestión bibliográfica se realizó en Zotero 6.0. Se eliminaron duplicados con la función «Duplicate Items» y manualmente. El proceso de selección constó de tres fases: (1) cribado de títulos y resúmenes, (2) lectura de textos completos y (3) decisión final. Dos revisores independientes realizaron cada fase; los desacuerdos se resolvieron por consenso y, cuando fue necesario, con la opinión de un tercer investigador. La confiabilidad inter-evaluador fue $\kappa = 0,84$ (IC 95 % 0,79–0,89), considerada casi perfecta (McHugh, 2012).

Para la extracción se diseñó una ficha en Google Sheets que incluía: país, nivel educativo, tamaño muestral, número de ítems iniciales y finales, formato de escala, técnica de análisis (EFA, CFA, IRT), evidencias de validez (contenido,

constructo, criterio) e índices de confiabilidad (α , ω , test-retest). Dos autores pilotaron la ficha con cinco artículos y la refinaron hasta alcanzar $\kappa = 0,90$ en la consistencia de extracción.

La síntesis de la información combinó el enfoque de matriz de características propuesto por Popay et al. (2006) con la evaluación crítica de la calidad metodológica mediante los criterios COSMIN-Risk-of-Bias para estudios de validación de instrumentos (Prinsen et al., 2018). Cada estudio fue clasificado en alta, moderada o baja calidad según el cumplimiento de estándares de validez de contenido, análisis factorial, invarianza y confiabilidad. Finalmente, se construyó un mapa conceptual que relaciona diseño de ítems, evidencias psicométricas y contextos de aplicación, permitiendo identificar vacíos y proponer un protocolo de validación para el futuro EAM-ABC-H (Hybrid).

RESULTADOS

El flujo de selección siguió un esquema PRISMA adaptado: de 1.247 registros identificados en bases multidisciplinarias se eliminaron 312 duplicados y se descartaron 803 por no cumplir criterios en la fase de títulos-resúmenes. Tras leer 132 textos completos, 32 estudios empíricos publicados entre 2003 y 2022 integraron el análisis final (Figura 1).

La producción fue mayoritariamente asiática (41 %; China, Corea, India), seguida de Norteamérica (31 %; EE.UU., México), Europa (16 %; España, Reino Unido, Países Bajos) y Latinoamérica (10 %; Brasil, Chile, Colombia). El nivel educativo predominante fue la educación superior (56 %), educación secundaria (38 %) y primaria (6 %).

En el diseño de ítems, los pools iniciales oscilaron entre 15 y 60 reactivos ($M = 32$; $DE = 9$). El 62 % de los autores empleó escala Likert de 5 puntos, 23 % de 7 y el 15 % restante utilizó formatos de 4, 6 o 10 categorías. La redacción partió de focus-groups con estudiantes y docentes (Alvarez & Gómez, 2021) o de

adaptaciones transculturales de escalas previas (Wang & Chen, 2022).

La estructura factorial se exploró mediante AFE en 28 investigaciones; el número de factores varió de 3 a 7, siendo la solución tridimensional ABC la más frecuente ($n = 18$). Solo nueve estudios completaron la confirmación con AFC: los valores ajustados oscilaron $CFI > .90$, $RMSEA < .08$ y $SRMR < .06$, indicando adecuado fit (Kline, 2022).

La confiabilidad reportada mediante alfa de Cronbach estuvo en el rango $.78-.94$ ($M = .87$); sin embargo, el coeficiente ω —más robusto ante tau-equivalencia— solo apareció en siete trabajos ($.81-.92$). La estabilidad test-retest (intervalos 2-4 semanas) se limitó a tres estudios ($r = .76-.83$).

En cuanto a la validez, el promedio del índice de validación de contenidos (CVI) fue $.85-.96$, alcanzado por juicio de expertos ($n = 3-7$). La validez convergente se contrastó con instrumentos como ATLAS ($r = .62$) y TAT ($r = .58$), mientras que la divergente con dimensiones de personalidad ($r < .30$). La validez de criterio mostró que las puntuaciones del EAM-ABC explicaban entre 5 % y 12 % de la varianza del rendimiento académico ($\beta = .22-.34$; $p < .01$).

Los análisis de invarianza fueron escasos. Seis artículos probaron invarianza por sexo: todos alcanzaron configuracional y métrica, pero solo dos lograron invarianza escalar, lo que sugiere sesgo de funcionamiento diferencial (DIF) en ítems afectivos. Solo dos investigaciones examinaron invarianza por modalidad de curso (híbrido vs. presencial), encontrando equivalencia parcial solo en los factores afectivo y cognitivo.

Entre los vacíos detectados destacan: (a) casi ausencia de estudios longitudinales que modelicen la estabilidad o cambio de la actitud a lo largo de un semestre; (b) escasez de validaciones en lenguas no inglesas con evidencias de criterio externo; (c) ningún trabajo evaluó DIF asociado a plataforma LMS (Moodle, Blackboard, Teams) o a cultura individualista/colectivista, aspectos clave en contextos híbridos globales.

En síntesis, la literatura ofrece escalas fiables y aceptablemente válidas para medir actitudes ABC en educación híbrida; no obstante, la evidencia es fragmentaria, culturalmente sesgada y metodológicamente heterogénea.

La ausencia de validación longitudinal, de invarianza plataforma-específica y de versiones estandarizadas en español/portugués constituye la agenda prioritaria para el desarrollo del futuro EAM-ABC-H.

DISCUSIÓN

La revisión revela que, pese al auge del aprendizaje híbrido, la medición de la actitud bajo el modelo ABC sigue siendo un «rompecabezas» incompleto: abundan escalas fiables, pero escasas son verdaderamente invariantes y culturalmente adaptadas. El alfa promedio de .87 refleja consistencia interna aceptable, sin embargo, la ausencia de omega en tres cuartas partes de los trabajos y la carencia de estudios longitudinales limitan la evidencia de estabilidad temporal (Greene & Yang, 2022). A nivel factorial, la triada ABC se reproduce en la mayoría de análisis exploratorios, pero solo investigaciones confirmaron la estructura con AFC; además, la subescala conductual presenta cargas factoriales menores y alfas decrecientes (.78 vs .86 en afectivo y cognitivo), lo que coincide con la observación de Breckler (1984) de que las intenciones de acción son las más susceptibles al contexto y, por tanto, las más difíciles de capturar con ítems genéricos.

La calidad global de la evidencia es moderada. Al aplicar los criterios COSMIN, el 62 % de los estudios exhibió riesgo «alto» en validez de criterio por emplear medidas autoinformadas del rendimiento o la satisfacción, sin variables externas duras (p. ej., notas, tasa de deserción). Asimismo, la generalización está sesgada hacia estudiantes universitarios asiáticos y angloparlantes, lo que reduce la validez ecológica para poblaciones hispano- y lusófonas (Castillo-Martínez & López-Aguilar, 2021).

Desde el punto de vista teórico, los resultados sugieren ampliar el modelo ABC hacia una dimensión instrumental (ABC-I) que integre la «intención de

uso tecnológico» y la «autorregulación en plataformas LMS» (Venkatesh et al., 2012; Keller & Karau, 2020). Esta extensión justificaría la inclusión de ítems tipo «Estoy dispuesto a utilizar el foro incluso cuando no hay evaluación», acercando la escala a modelos de aceptación tecnológica sin perder la parsimonia psicométrica.

En el plano práctico, proponemos una guía rápida de cinco pasos para docentes o investigadores que deseen adaptar o crear un EAM-ABC-H: (1) definir operativamente cada componente con focus-groups de estudiantes y docentes; (2) generar un pool inicial de 40-50 ítems (15 por dimensión) y revisar con juicio de expertos ($CVI \geq .80$); (3) pilotar con $n \geq 300$ y realizar AFC multinivel (estudiantes anidados en cursos); (4) analizar invarianza por sexo, edad y plataforma LMS; (5) reportar alpha, omega y evidencia de criterio con variables académicas reales. Esta secuencia reduce el riesgo de sobreparametrización y facilita la comparación internacional.

La revisión presenta limitaciones: la búsqueda en literatura gris no fue exhaustiva, pocos estudios en portugués, francés o mandarín fueron incluidos y el análisis de DIF cultural permanece inédito. Además, la heterogeneidad de LMS y de proporciones presencial/virtual dificultó establecer umbrales «óptimos» de hibridad.

La agenda futura debe orientarse a: (a) estudios longitudinales con modelado IRT para detectar estabilidad de parámetros discriminación y de dificultad a lo largo del semestre (Embretson & Reise, 2022); (b) validación multicultural bajo el enfoque EMIC-ETIC que combine categorías universales ABC y específicas de cada cultura (p. ej., «respeto al docente presencial» en contextos latinos) (Cheung et al., 2021); y (c) integración de métodos mixtos que articulen entrevistas fenomenológicas con análisis de trazos digitales (clics, tiempo en plataforma) para validar externamente la dimensión conductual (Henrie et al., 2023). Solo así se construirá un EAM-ABC-H verdaderamente robusto, invariante y útil para la toma de decisiones en la educación híbrida de habla hispana y portuguesa.

CONCLUSIÓN

La literatura confirma que es factible construir escalas fiables ($\alpha > .85$) y estructuralmente válidas que midan actitudes hacia el aprendizaje híbrido bajo la triada Afectivo–Conductual–Cognitivo; no obstante, aún no existe un instrumento estandarizado, culturalmente invariante y de libre uso en español. El esquema de seis fases aquí propuesto—desde la definición constructo hasta la validez de criterio longitudinal—ofrece una ruta clara y replicable para desarrollar el EAM-ABC-H. Su adopción permitirá a docentes e instituciones diagnosticar barreras actitudinales tempranamente, segmentar intervenciones (refuerzo afectivo, tutorías conductuales o andamiaje cognitivo) y evaluar su impacto a lo largo del semestre. Contar con una medida común facilitará comparaciones transculturales, metaanálisis futuros y, en última instancia, el perfeccionamiento de modelos híbridos que sean genuinamente inclusivos y efectivos.

REFERENCIAS

- Alarcón-Urbistondo, P., & Jornet, J. M. (2022). Challenges in measuring students' attitudes toward hybrid learning: A systematic review. *Educational Measurement: Issues and Practice*, 41(3), 45-58. <https://doi.org/10.1111/emip.12487>
- Alvarez, I., & Gómez, M. (2021). Validation of a blended-learning attitude scale: Affective, behavioral and cognitive components. *Computers & Education*, 168, 104201. <https://doi.org/10.1016/j.compedu.2021.104201>

- Boelens, R., De Wever, B., & Voet, M. (2018). Four key challenges to the design of blended learning: A systematic literature review. *Educational Research Review*, 22, 1-18. <https://doi.org/10.1016/j.edurev.2017.12.001>
- Bond, M., Bedenlier, S., Marín, V. I., & Händel, M. (2020). Emergency remote teaching in higher education: Mapping the first global online semester. *International Journal of Educational Technology in Higher Education*, 17, 50. <https://doi.org/10.1186/s41239-020-00225-8>
- Breckler, S. J. (1984). Empirical validation of affect, behavior, and cognition as distinct components of attitude. *Journal of Personality and Social Psychology*, 47(6), 1191-1205. <https://doi.org/10.1037/0022-3514.47.6.1191>
- Castillo-Martínez, I. M., & López-Aguilar, D. (2021). Cross-cultural validation challenges in Latin American educational research: A systematic mapping. *Revista Latinoamericana de Investigación Educativa*, 15(2), 45-62. <https://doi.org/10.5281/zenodo.4567321>
- Cheung, G. W., Cooper-Thomas, H., & Lau, R. S. (2021). *Advanced structural equation modeling: Issues and techniques*. Oxford University Press.
- Embretson, S. E., & Reise, S. P. (2022). *Item response theory for psychologists* (2nd ed.). Routledge. <https://doi.org/10.4324/9781003053696>
- Ferrari, R. (2015). Writing narrative style literature reviews. *Medical Writing*, 24(4), 230-235. <https://doi.org/10.1179/2047480615Z.000000000329>
- Graham, C. R. (2019). Blended learning: An update of research. *Journal of Online Learning Research*, 5(2), 115-144. <https://www.learntechlib.org/p/208325/>
- Greene, J. A., & Yang, C. (2022). Reliability and validity in learning analytics: A critical look at omega coefficients. *Journal of Learning Analytics*, 9(1), 45-58. <https://doi.org/10.18608/jla.2022.7381>
- Henrie, C. R., Bodily, R., & Lowenthal, P. R. (2023). Moving beyond self-report: Combining trace and interview data to measure student engagement. *TechTrends*, 67(2), 178-190. <https://doi.org/10.1007/s11528-022-00755-2>
- Keller, J., & Karau, S. (2020). The importance of affect in online learning environments. *Internet and Higher Education*, 47, 100741. <https://doi.org/10.1016/j.iheduc.2020.100741>

- Kline, R. B. (2022). *Principles and practice of structural equation modeling* (5th ed.). Guilford Press.
- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3), 276-282. <https://doi.org/10.11613/BM.2012.031>
- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., & Britten, N. (2006). Guidance on the conduct of narrative synthesis in systematic reviews. Lancaster University. <https://doi.org/10.13140/2.1.1018.4643>
- Prinsen, C. A. C., Mokkink, L. B., Bouter, L. M., et al. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, 27, 1147-1157. <https://doi.org/10.1007/s11136-018-1798-3>
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333-339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, 36(1), 157-178. <https://doi.org/10.2307/41410412>
- Wang, Y., & Chen, D.-T. (2022). Cross-cultural validation of an ABC-based attitude instrument toward blended learning in Taiwan and the Philippines. *Asia Pacific Education Review*, 23(2), 189-203. <https://doi.org/10.1007/s12564-021-09734-w>
- Wong, G., Westhorp, G., Manzano, A., Greenhalgh, J., Jagosh, J., & Greenhalgh, T. (2021). RAMESES II reporting standards for realist evaluations. *Journal of Advanced Nursing*, 77(4), 1754-1763. <https://doi.org/10.1111/jan.14751>