

Capítulo 7

Diseño y validación de las escalas básicas de autoeficacia en competencias digitales de estudiantes de educación superior

Yéssica Sandybel Garduño Espinoza

Juan Carlos Pérez Morán

José Eduardo Perezchica Vega

<https://doi.org/10.61728/AE20255435>



Resumen

El diseño y la validez de escalas para evaluar las habilidades y competencias digitales es en la actualidad una práctica bastante común en distintos sistemas educativos alrededor del mundo. El objetivo del estudio es recolectar evidencias del diseño y la validación del aspecto del contenido de las Escalas de Autoeficacia en Competencias Digitales Básicas de Estudiantes de Educación Superior (EA-CDigit-B-EES). Para ello, se integraron dos comités de expertos en el área de las competencias digitales, uno para el diseño de los ítems de las EA-CDigit-B-EES ($N = 3$) y otro para la validación de su contenido ($N = 6$). Como producto del trabajo de ambos comités, se desarrollaron y validaron 88 ítems organizados en dos subescalas asociadas con las dos primeras áreas del Marco de Competencias Digitales para la Ciudadanía actualizado (DigComp 2.2, por sus siglas en inglés): (1) Búsqueda y gestión de información y datos, y (2) Comunicación y colaboración. Para la validación de los ítems se obtuvieron distintos Indicadores de Validez de Contenido (IVC) adecuados en las dos subescalas y en la mayoría de los ítems ($PAJ = 0.84$, $V\text{-Aiken} = 0.84$, RVC de Lawshe $= 0.27$, RVC' Tristán-López $= 0.63$). Dados los resultados del estudio, se concluye que los ítems de las EA-CDigit-B-EES cumplen con los criterios de calidad preestablecidos. Para futuros estudios se recomienda corregir el diseño de los ítems con puntuaciones bajas en los IVC y realizar un piloteo empírico para probar la confiabilidad y propiedades psicométricas de las EA-CDigit-B-EES.

Introducción

Las competencias digitales son un factor clave en el desarrollo y desempeño de los individuos en ambientes donde se favorece el uso de las Tecnologías de la Información y la Comunicación (TIC). Es por ello que la evaluación de estas competencias es una práctica común en los sistemas

educativos alrededor del mundo, pues permite verificar si los estudiantes cuentan con la preparación necesaria para responder a requerimientos de aprendizaje en los contextos educativos y en el mundo laboral y social (González-Martínez et al., 2018; Mirete et al.; Vega-Hernández et al., 2018). Dicha tendencia es particularmente relevante en Instituciones de Educación Superior (IES), pues se busca contar con indicadores fiables que permitan evaluar el nivel de dichas competencias, así como medir cómo estas impactan en la competitividad, la innovación y la adaptación en el ámbito laboral y a lo largo de toda la vida.

Para el logro del Objetivo de Desarrollo Sostenible 4.4.2. “Porcentaje de jóvenes/adultos que han alcanzado al menos un nivel mínimo de competencia en habilidades de alfabetización digital”. La Comisión Europea (CE) desarrolló el Marco Común Europeo de las Competencias Digitales (DigComp 2.0). En este marco, las competencias digitales son entendidas como el conjunto de conocimientos, habilidades y actitudes que representan el uso crítico y responsable de las tecnologías digitales para aprender, laborar y participar en la sociedad (Law et al., 2018). El componente de conocimiento en el DigComp 2.0 comprende el entendimiento sobre cómo los sistemas y tecnologías digitales pueden favorecer la comunicación, la creatividad y la innovación, así como ser consciente de las oportunidades, limitaciones y riesgos que implica su uso. Por su parte, el componente de la habilidad digital se refiere a la capacidad de usar las tecnologías digitales para alcanzar objetivos específicos, mientras que el componente de actitud digital implica una disposición crítica y reflexiva al interactuar con las tecnologías, desde un enfoque ético, responsable y seguro.

Desde el 2016, se organizan en el DigComp las competencias digitales a través de cinco dimensiones competenciales: (a) Búsqueda y gestión de información y datos, (b) Comunicación y colaboración, (c) Creación de contenidos digitales, (d) Seguridad y (e) Resolución de problemas (Vuorikari et al., 2016). Las tres primeras tienen que ver con competencias observables en actividades y usos específicos, mientras que las últimas dos se consideran transversales, dado que son aplicables a cualquier actividad realizada por medios digitales. Asimismo, en la actualidad, el DigComp 2.2 incluye ocho niveles de aptitud agrupados en cuatro categorías de habilidad. Estas se consideran básicas (niveles 1 y 2), si los individuos pueden

ejecutar tareas sencillas con asistencia y la exigencia cognitiva es recordar; intermedias (niveles 3 y 4), si los individuos pueden realizar tareas y problemas bien delimitados de forma autónoma y la exigencia cognitiva es que entiendan y comprendan; avanzadas (niveles 5 y 6), si los individuos son capaces de hacer tareas complejas sin ayuda y, además, pueden asistir a otras personas mediante la evaluación de las competencias; y especializadas (niveles 7 y 8), cuando los individuos pueden resolver problemas complejos y proponer ideas, así como contar con la capacidad de crear contenidos digitales.

A partir de los marcos de DigComp en distintas partes del mundo, hay instrumentos desarrollados para medir las competencias digitales en estudiantes, docentes, trabajadores y profesionales especializados de distintos campos, y en la población en general (Litiņa y Miltuze, 2021; Nóbile y Gutiérrez, 2022; Sánchez-Canut, et al., 2023). En especial, hay una amplia variedad de instrumentos desarrollados para medir una variedad de competencias digitales de estudiantes universitarios (Litiņa y Miltuze, 2021). En Latinoamérica, solo algunos de estos instrumentos reportan evidencias de confiabilidad y validez, como el Cuestionario de Competencia Digital para Futuros Maestros (C-CD-FM) desarrollado por Cabero-Almenara et al. (2020) y el COMPDIG-PED desarrollado por Silva-Quiroz y colaboradores (2022). Además, son pocos los estudios en donde: (a) se apliquen, para el diseño y validación de los instrumentos, métodos mediante el juicio de expertos; (b) se analicen el contenido y la imparcialidad de los ítems mediante distintos criterios de validación; y (c) se recolecten evidencias de validez del aspecto de contenido por medio de indicadores empíricos (García-Ruiz et al., 2023; Sillat et al., 2021).

Con base en la problemática planteada, se estableció como objetivo del presente estudio diseñar y obtener evidencias de validez del aspecto de contenido de las Escalas de Autoeficacia en Competencias Digitales Básicas de Estudiantes de Educación Superior (EA-CDigit-B-EES) diseñadas por Garduño y colaboradores (2022). Para ello, se partió de la alineación con las primeras dos competencias del DigComp (Búsqueda y gestión de información y datos y Comunicación y colaboración) y de estándares de calidad técnica para el diseño, validación y uso de pruebas psicológicas y educativas. Con ello, se hace posible garantizar que las EA-CDigit-B-EES

representen adecuadamente el constructo teórico de competencias digitales, y justificar las interpretaciones derivadas de sus resultados. En especial, se procuró que los ítems diseñados demanden competencias propias de contextos escolares y fueran relevantes para favorecer los procesos de enseñanza-aprendizaje en las aulas universitarias.

Método

Tipo de estudio

El estudio es de tipo psicométrico y consiste en el diseño, desarrollo y validación de contenido de las EA-CDigit-B-EES. Respecto al diseño y desarrollo de dicho instrumento, se siguieron las recomendaciones de diseño y validación de ítems propuestas por Haladyna y Rodríguez (2013). En este sentido, las escalas fueron diseñadas con el propósito de detectar fortalezas y áreas de oportunidad en competencias digitales básicas de estudiantes de Educación Media Superior (EMS) y Educación Superior (ES). Para ello, se adoptó un enfoque de evaluación educativa de bajo impacto (low stakes; Martínez-Rizo, 2009; Popham, 2013; Ravela, 2013, 2020; Scriven, 1967).

Por su parte, para la validación del contenido de las EA-CDigit-B-EES, se siguieron lineamientos específicos de documentos orientados a la validación de instrumentos educativos y psicológicos. En este sentido, se atendieron los estándares 1.9 y 1.10 propuestos por la AERA et al. (2018), referentes a aspectos relacionados con las muestras y los contextos de aplicación durante la validación, y al estándar 1.11, que aborda las formas específicas de evidencia de validación en función del contenido de las pruebas. De la misma manera, se observaron los estándares 3.0 a 3.5 relacionados con el desarrollo, aplicación y procedimientos de calificación de los instrumentos y pruebas que minimizan las barreras para llegar a interpretaciones válidas de los puntajes de una variedad más amplia de individuos y grupos relevantes en los estudios. Asimismo, se siguieron las recomendaciones de la Comisión Internacional de Test (ITC, por sus siglas en inglés) referentes a: (a) el uso ético de las pruebas y buenas prácticas en el uso de los instrumentos (ITC, 2001); (b) el control de calidad de la

puntuación, el análisis de las pruebas y la comunicación de los resultados de las pruebas (ITC, 2014); y (c) el uso de pruebas y otros instrumentos de evaluación con fines de investigación (ITC, 2014b).

Participantes

El estudio contó con la participación de dos grupos de especialistas: un primer grupo de tres participantes que conformaron el Comité de Diseño y Desarrollo (CDD); y un segundo grupo de seis participantes que conformaron el Comité de Validación de Contenido (CVC). Los integrantes del CDD fueron los autores responsables del presente estudio, quienes cuentan con experiencia, además de conocimientos, en el campo de la medición educativa, y de los cuales dos tienen el grado de doctorado en ciencias de la educación y uno con licenciatura. En cuanto al CVC, lo conformaron seis especialistas con grado académico de maestro o doctor y con experiencia en la evaluación de competencias digitales o constructos similares. Además, tienen experiencia en medición de constructos latentes y el desarrollo de procesos de validación de instrumentos.

Procedimiento

El procedimiento se dividió en dos etapas: (I) Diseño y desarrollo; y (II) Validación del contenido del instrumento. La primera etapa constó de tres actividades: (a) integración y capacitación del CDD; (b) elaboración del diseño y especificaciones del instrumento; y (c) desarrollo del contenido del instrumento. Durante las primeras actividades realizadas por los integrantes del CDD, se llevó a cabo la revisión de textos referentes a la conceptualización y medición de competencias digitales. Después, dicho comité elaboró el marco general del instrumento, es decir, se definieron: (a) los usos e interpretaciones previstos; (b) la teoría sustantiva; (c) las variables que pretende medir; (d) los usuarios y la población objetivo; (e) el contexto y circunstancias de aplicación; (f) el impacto de los resultados; y (g) otras cuestiones técnicas (Muñiz y Fonseca-Pedrero, 2019). Asimismo, se elaboraron las especificaciones, donde se delimitaron aspectos como el formato del instrumento, las guías e indicadores para redactar ítems,

las pautas para su administración y las dimensiones cognitivas manifiestas al responder el instrumento (Haladyna y Rodriguez, 2013; Muñiz y Fonseca-Pedrero, 2019). También, el CDD generó las instrucciones, los enunciados de los ítems y las opciones de respuesta disponibles para los sustentantes y la documentación del instrumento. Una vez hecho esto, se elaboró la documentación técnica, la cual comprende la ficha técnica del instrumento y la versión de aplicación en formato papel y lápiz.

Por su parte, en la segunda etapa se realizaron dos actividades: (a) integración y capacitación del CVC y (b) validación de los reactivos del instrumento por parte de los integrantes del CVC. En la primera actividad se reclutó y capacitó a especialistas para realizar actividades relacionadas con el juicio de contenido e imparcialidad del instrumento. Para ello, se enviaron solicitudes por medio de correo electrónico en donde se explicaban los propósitos y objetivos del estudio. A aquellos especialistas que aceptaron fungir como jueces se les proporcionó acceso a una carpeta electrónica que alberga los materiales necesarios para su capacitación y evaluación.

Posteriormente, en la segunda actividad, cada especialista del CVC evaluó los ítems del instrumento en función de cuatro criterios de validez de contenido (Relevancia, Precisión, Claridad y Simplicidad) (Aiken, 1985) y siete elementos (Definición precisa de constructo; Inclusión poblacional; Accesibilidad e imparcialidad; Acomodación flexible de los contenidos; Procedimientos e instrucciones simples, claras e intuitivas; Comprensibilidad; y Máxima legibilidad) del Modelo de Diseño Universal (MDU) descritos por Thompson et al. (2002) y adaptados por Pérez-Morán (2014). Estos fueron evaluados a través de una escala tipo Likert que presenta cuatro opciones de respuesta, en donde 1 es la valoración más baja y 4 la valoración más alta. Asimismo, cada especialista realizó sugerencias para mejorar la calidad de las especificaciones y contenidos de los ítems que integran a la EA-CDigit-B-EES.

Análisis de datos

Una vez obtenidas las valoraciones de los especialistas que integraron el CVC, se procedió al cálculo y análisis de distintos Índices de Validez de Contenido (IVC): la obtención de los porcentajes de acuerdo entre jueces

(PAJ), el coeficiente V de Aiken (1985) y la razón de validez de contenido (RVC) desarrollada por Lawshe (1975), así como su versión corregida propuesta por Tristán-López (2008; RVC'). Para ello, se transformaron las valoraciones de los especialistas y se calculó el PAJ y el coeficiente V de Aiken. Primero se aplicó una regla de proporción donde las puntuaciones con un valor de 4 se recodificaron a 1, las de 3 a 0.66, las de 2 a 0.33 y las de 1 a 0.0. Para el cálculo de RVC y RVC', las valoraciones de 4 se convirtieron a 1 y las demás opciones de respuesta a 0.0. Para el cálculo del coeficiente V de Aiken se empleó la fórmula:

$$V = \frac{\bar{x} - l}{k} \quad (1)$$

En donde la variable V es establece como el coeficiente de Aiken, $\bar{x}$ se refiere al promedio de valoración de los especialistas para cada ítem, l a la calificación que es más baja posible con la que se puede evaluar un ítem y k es el rango de opciones de respuesta que presenta la escala. Para el presente estudio, el criterio de corte establecido para el Coeficiente V de Aiken fue de 0.69 (Aiken, 1985). Por su parte, para el cálculo de la RVC de Lawshe se consideró la fórmula:

$$RVC = \frac{n_e - \frac{N}{2}}{\frac{N}{2}} \quad (2)$$

En donde n se asigna a la cantidad de jueces que calificaron el ítem con base en los criterios esencial o relevante, y N el número total de jueces que validaron el contenido de la prueba. En el presente estudio, se calculó la RVC de cada ítem y subescala. En este estudio se fijó un criterio de RVC > 0.56 para denotar evidencias de validez adecuada. Asimismo, se calculó la RVC'; se obtuvo a través de la ecuación:

$$RVC' = \frac{RVC + 1}{2} \quad (3)$$

En donde RVC es la razón de validez de contenido de cada ítem. Una vez calculados los coeficientes y razones de validez de contenido, se hicieron adecuaciones basadas en las valoraciones y comentarios de los especialistas. En este sentido, se modificarán los contenidos que potencialmente limitan la validez del instrumento. Es decir, aquellos que incumplan con los criterios establecidos de acuerdo con las recomendaciones de especialistas en procesos de validación del contenido (AERA et al., 2018).

Instrumentos y materiales

En el presente estudio se emplearon distintos instrumentos y materiales durante el proceso de validación del contenido. Entre los instrumentos se encuentra la versión preliminar de las Escalas Básicas de Autoeficacia en Competencias Digitales de Estudiantes de Educación Superior (EA-CDigit-B-EES). También se empleó el protocolo de validación de contenido, que sirve como guía para evaluar los contenidos de interés según los criterios de validez de contenido, así como los elementos del Modelo de Diseño Universal (MDU) propuestos por Thompson et al. (2002) y adaptados por Perez-Moran (2014). El formato sigue las recomendaciones de Aiken (1985), en donde se califica a distintos criterios a través de cuatro opciones de valoración ordinales tipo Likert (pe. 1 = nada preciso, 2 = poco preciso, 3 = preciso, 4 = muy preciso) para evaluar los criterios del contenido. Además de estos, se emplearon materiales como un acuerdo de confidencialidad y consentimiento informado, el manual donde se describen los procedimientos para el proceso de validación de contenido y la ficha técnica de las EA-CDigit-B-EES.

Resultados

Diseño de las EA-CDigit-B-EES

Como resultado, se desarrolló la primera versión de las EA-CDigit-B-EES, que cuenta con 88 ítems, cada uno con cuatro opciones de respuesta ordinal (1 = Muy difícil; 2 = Difícil; 3 = Fácil, y 4 = Muy fácil), que miden la autoeficacia en competencias digitales de los estudiantes. Asimismo, se

compone de dos dimensiones: Búsqueda y gestión de información y datos (BYG) e interacción a través de tecnologías de la información (ITI). La dimensión BYG incluye tres subescalas: (1.1) Navegar, buscar y filtrar datos, información y contenidos digitales, (1.2) Evaluar datos, información y contenidos digitales, y (1.3) Gestión de datos, información y contenidos digitales. La primera, 1.1. Navegar, buscar y filtrar datos, información y contenidos digitales se refiere a articular las necesidades de información, buscar datos, información y contenidos en entornos digitales, así como acceder y navegar entre ellos; además, crear y actualizar estrategias de búsqueda personales. La segunda, 1.2. Evaluar datos, información y contenidos digitales implica analizar, comparar y evaluar críticamente la fiabilidad y seriedad de los recursos de datos, información y contenido digital, así como interpretar y evaluar de manera crítica los datos y contenidos digitales. Finalmente, la tercera, 1.3. Gestión de datos, información y contenidos digitales: se entiende como la habilidad de poder organizar, almacenar y recuperar información, datos y contenidos en entornos digitales, así como de estructurar, organizar y procesarlos en distintos entornos.

Por su parte, la dimensión ITI está formada por seis escalas: (2.1) Interactuar a través de tecnologías digitales, (2.2) Compartir a través de tecnologías digitales, (2.3) Participación ciudadana a través de tecnologías digitales, (2.4) Colaboración a través de tecnologías digitales, (2.5) Comportamiento en la red y (2.6) Gestión de la identidad digital. La primera, 2.1. Interactuar a través de tecnologías digitales se refiere a interactuar mediante diferentes tecnologías digitales y entender los medios de comunicación digitales apropiados para un contexto determinado. La segunda, 2.2. Compartir a través de tecnologías digitales implica compartir datos, información y contenidos digitales con otros a través de las tecnologías adecuadas, actuando como intermediario y conociendo las prácticas de referencia y atribución. La tercera, 2.3. La participación ciudadana a través de las tecnologías digitales se centra en participar en la sociedad mediante el uso de servicios digitales públicos y privados, buscando oportunidades de empoderamiento y promoviendo una ciudadanía participativa a través de tecnologías digitales adecuadas. La cuarta, 2.4. Colaboración a través de las tecnologías digitales: trata del uso de herramientas y tecnologías digitales en procesos colaborativos para la co-construcción y co-creación

de datos, recursos y conocimiento. La quinta, 2.5. Comportamiento en la red: Implica estar al tanto de las normas de comportamiento y del saber cómo en el uso de tecnologías y en la interacción en entornos digitales, adaptando las estrategias de comunicación a una audiencia específica y teniendo en cuenta la diversidad cultural y generacional de los entornos digitales. Finalmente, la sexta escala, 2.6. Gestión de la identidad digital: Se refiere a crear y gestionar una o varias identidades digitales para proteger la propia reputación y manejar los datos que uno produce a través de diversas herramientas, entornos y servicios digitales. En la Tabla 3 se observa la distribución de los ítems.

Tabla 3. Estructura de las EA-CDigit-B-EES

Dimensión	Escala	ítems
1. Búsqueda y gestión de información y datos	1.1. Navegar, buscar y filtrar datos, información y contenidos digitales	k = 11 (1.1.1 a 1.1.11)
	1.2. Evaluar datos, información y contenidos digitales	k = 2 (1.2.1 a 1.2.12)
	1.3. Gestión de datos, información y contenidos digitales	k = 10 (1.3.1 a 1.3.10)
2. Comunicación y colaboración	2.1. Interactuar a través de tecnologías digitales	k = 9 (2.1.1 a 2.1.9)
	2.2. Compartir a través de tecnologías digitales	k = 8 (2.2.1 y 2.2.8)
	2.3. Participación ciudadana a través de las tecnologías digitales	k = 10 (2.3.1 a 2.3.10)
	2.4. Colaborar a través de las tecnologías digitales	k = 9 (2.4.1 a 2.4.9)
	2.5. Comportamiento en la red	k = 9 (2.5.1 a 2.5.9)
	2.6. Gestionar la identidad digital	k = 10 (2.6.1 a 2.6.10)

Validación del contenido

Como resultado del cálculo de los IVC, se obtuvo un PAJ con un valor medio de 0.84 global y con una valoración promedio de 0.81 en los ítems de la dimensión BYG y 0.87 en los ítems de la dimensión ITI. También se encontró que la subescala 1.1. NBFD no cumple con el criterio de PAJ superior a .80, mientras que las demás escalas sí superan dicho criterio (ver Tabla 4).

Tabla 4. *Valores medios de IVC de las EA-CDigit-B-EES por escalas*

Dimensión	Escala	PAJ	v de Aiken	RVC de Lawshe	RVC' de Tristán-López
BYG	1.1. NBFD	0.76*	0.76	0.11*	0.56
	1.2. EDIC	0.80	0.80	0.26*	0.63
	1.3. GDIC	0.87	0.87	0.50*	0.75
ITI	2.1. ITD	0.87	0.87	0.23*	0.62
	2.2. CTD	0.88	0.88	0.28*	0.62
	2.3. PCTD	0.89	0.89	0.33*	0.66
	2.4. COTD	0.85	0.85	0.17*	0.58
	2.5. CR	0.86	0.86	0.20*	0.60
	2.6. GID	0.88	0.88	0.23*	0.62
	Global	0.84	0.84	0.27*	0.63

*No cumple con el criterio establecido a priori (PAJ > 0.80, coeficiente V > 0.69, RVC > 0.56. RVC' > 0.56).

De manera específica, se reporta que 66 ítems (75 %) de los 88 ítems de la EA-CDigit-B-EES superan los criterios de corte. Nótese que 22 ítems (25 %) incumplen alguno de los criterios de validez evaluados (relevancia, precisión, claridad y simplicidad) (ver Tabla 5).

Tabla 5. *Ítems que incumplen los criterios y puntos de corte de los IVC*

Ítems	PAJ	V de Aiken	RVC de Lawshe	RVC' de Tristan-López
1.1.4, 1.1.10, 1.1.11, 1.2.5	x	x	x	x
1.1.2, 1.1.3, 1.1.5, 1.1.8, 1.2.1, 1.2.10, 1.3.2, 2.1.2, 2.1.4, 2.3.3, 2.4.3, 2.4.7, 2.5.3, 2.5.8, 2.6.4	x		x	x
1.2.6, 1.2.7, y 1.2.12	x		x	
2.1.1, 2.2.8, 2.4.1, 2.4.5, 2.4.6, 2.5.1, 2.5.2, 2.5.4, 2.6.1, 2.6.7, 2.6.8			x	x
1.1.1, 1.1.9, 1.2.2, 1.2.3, 1.2.4, 1.2.8, 1.2.9, 1.2.11, 1.3.1, 1.3.7, 1.3.8, 1.3.10, 2.1.5, 2.1.6, 2.1.7, 2.1.8, 2.1.9, 2.2.1, 2.2.2, 2.2.3, 2.2.4, 2.2.5, 2.2.6, 2.2.7, 2.3.1, 2.3.2, 2.3.4, 2.3.5, 2.3.6, 2.3.7, 2.3.8, 2.3.9, 2.3.10, 2.4.2, 2.4.4, 2.4.8, 2.4.9, 2.5.5, 2.5.6, 2.5.7, 2.5.9, 2.6.2, 2.6.3, 2.6.5, 2.6.6, 2.6.9, 2.6.10			x	

Nota: x = Ítems que no superan el criterio de corte en alguno de los IVC

De manera similar, el valor promedio del coeficiente V de Aiken fue de 0.84, siendo que se reportó un valor promedio de 0.81 en los ítems que integran la dimensión BYG y de 0.87 para aquellos de la dimensión ITI. También se obtuvo que, en todas las subescalas, los valores del coeficiente V de Aiken superan el criterio establecido a priori (coeficiente $V > .69$; ver Tabla 4). Por su parte, 84 ítems (95.45%) superan el valor establecido en el criterio de corte. De manera que cuatro ítems (4.55%; 1.1.4, 1.1.10, 1.1.11 y 1.2.5) presentan valores bajos en alguno de los criterios revisados (Tabla 5).

En cuanto al RVC de Lawshe, se obtuvo un valor promedio de 0.27, en donde los ítems de la dimensión BYG obtuvieron una media de 0.29 y de la dimensión ITI de 0.24. Los valores promedio de RVC en función de las subescalas fueron bajos, de modo que ninguna subescala cumplió con los criterios establecidos (ver Tabla 4). En un sentido más granulado, se encontró que ocho ítems obtuvieron un RVC promedio superior a 0.56 (9.09%; 1.1.6, 1.1.7, 1.3.3, 1.3.4, 1.3.5, 1.3.6, 1.3.9, 2.1.3), mientras que el resto incumplen los criterios (Tabla 5).

Por último, para el RVC de Tristan-López, la puntuación media fue de 0.63; los ítems de la dimensión BYG presentaron un valor promedio de 0.65, mientras que los ítems de la dimensión ITI fue de 0.62. Asimismo, todas las subescalas tienen un valor de RVC' mayor al criterio establecido por Tristán-López (2008; $RVC' > 0.56$) para evaluar el cumplimiento de los criterios revisados. De manera más detallada, 58 ítems (65.91 %) presentan un RVC' superior a 0.56, mientras que 30 (34.09 %; 1.1.2, 1.1.3, 1.1.4, 1.1.5, 1.1.8, 1.1.10, 1.1.11, 1.2.1, 1.2.5, 1.2.10, 1.3.2, 2.1.1, 2.1.2, 2.1.4, 2.2.8, 2.3.3, 2.4.1, 2.4.3, 2.4.5, 2.4.6, 2.4.7, 2.5.1, 2.5.2, 2.5.3, 2.5.4, 2.5.8, 2.6.1, 2.6.4, 2.6.7, 2.6.8) incumplen alguno de los criterios. En la Tabla 5 se observan los ítems que no cumplen con los criterios establecidos para al menos uno de los índices de validez de contenido.

En cuanto al resultado del análisis de imparcialidad mediante los elementos del MDU, se recibieron 16 comentarios distribuidos en 12 ítems (13.63 %; 1.1.4, 1.1.9, 1.1.10, 1.2.1, 1.2.5, 1.2.12, 2.1.1, 2.1.7, 2.3.1, 2.3.2, 2.4.6 y 2.5.3). El ítem con más comentarios fue el 1.1.4, con un total de cuatro comentarios relacionados con la Definición Precisa del Constructo (DPC), Inclusión Poblacional (IP) y Procedimientos e Instrucciones claras (PI). En especial, se señalaron recomendaciones para los ítems 1.1.4, 1.1.9, 1.1.10 y 1.2.1 sobre mejorar la alineación de los ítems con sus descriptores; los ítems 1.1.4, 1.2.12, 2.1.1 y 2.5.3 recibieron comentarios sobre que pueden ser poco accesibles para algunos estudiantes; respecto a la Accesibilidad e Imparcialidad (AI), en el ítem 1.2.5 se sugiere una contextualización del término “deep-fake”; los ítems 1.1.4 y 2.4.6 requieren mayor precisión y el uso de términos completos antes de los acrónimos; en términos de Compresibilidad (CO), se indica que los ítems 2.1.7, 2.3.1 y 2.3.2 deben ajustarse e incluir los signos de puntuación requeridos; y con referencia a la Máxima Legibilidad (ML), se recomienda el uso de términos en mayúsculas para el ítem 1.1.4. En ninguno de los ítems se presentaron comentarios respecto a la Acomodación Flexible de los contenidos (AF). En la Tabla 6 se observan los ítems que tuvieron algún comentario.

Tabla 6. *Ítems con comentarios para la evaluación del modelo de evaluación del diseño universal*

Ítems	DPC	IP	AI	PI	CO	ML	AF
1.1.4	x	x				x	
1.1.9, 1.1.10 y 1.2.1	x						
1.2.12, 2.1.1 y 2.5.3		x					
1.2.5			x				
2.4.6				x			
2.1.7, 2.3.1 y 2.3.2					x		

Nota. DPC = Definición precisa de constructo; IP = Inclusión poblacional; AI = Accesibilidad e imparcialidad; PI = Procedimientos e instrucciones simples, claras e intuitivas; CO = Comprensibilidad; ML = Máxima legibilidad; AF = Acomodación flexible de los contenidos.

Conclusiones y discusión

En el presente estudio se diseñó, desarrolló y obtuvieron evidencias de validez del contenido de las Escalas de Autoeficacia en Competencias Digitales Básicas de Estudiantes de Educación Superior (EA-CDigit-B-EES). Para ello, se diseñaron y validaron 88 ítems distribuidos en dos dimensiones alineadas con el DigitComp 2.2, una para medir la competencia de 1. Búsqueda y gestión de información y datos ($k = 33$), y otra para la competencia 2. Comunicación y colaboración ($k = 55$). Su diseño permite una representación multidimensional de las competencias digitales básicas basado en el marco de la DigiComp 2.2. Adaptado al contexto de la educación superior en México.

Los resultados obtenidos del análisis de la validación del contenido muestran que el contenido de las EA-CDigit-B-EES es relevante, preciso, claro y simple. Se cumplieron los criterios de los IVC a nivel general: PAJ promedio de 0.84, coeficiente V de Aiken de 0.84 y RVC' promedio de 0.63. En total, 84 ítems (95.45 %) cumplieron con el coeficiente V, y 66 ítems (75 %) con el PAJ.

No obstante, al margen de los resultados favorables, se identifican diversas limitaciones. Primero, una parte importante de los ítems no cumplió con los criterios de validez. Esto indica que algunos ítems podrían no ser del todo comprensibles para los estudiantes. Además, la alineación de

algunos ítems con los descriptores no fue adecuada, lo que implica que se puede generar confusión o interpretaciones incorrectas de los resultados. En cuanto a la imparcialidad del instrumento, los comentarios se orientaron hacia la necesidad de mejorar la definición de los constructos y hacer los términos más accesibles.

Para finalizar, las EA-CDigit-B-EES se consideran un instrumento válido en función de la representación del contenido de las competencias digitales, como se describe en la DigiComp 2.2. Este instrumento es un avance hacia una evaluación contextualizada de las competencias digitales en el ámbito de educación superior. Se recomienda la obtención de evidencias de validez que justifiquen las interpretaciones del instrumento. En especial, el análisis de la estructura interna subyacente a las escalas y el estudio de las relaciones estadísticamente significativas con otras variables concurrentes y predictivas.

Referencias

- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and psychological measurement*, 45(1), 131–142. <https://doi.org/10.1177/0013164485451012>
- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (2018). *Standards for Educational and Psychological Testing*. American Educational Research Association. <https://doi.org/10.2307/j.ctvr43hg2>
- Cabero Almenara, J., Barroso Osuna, J. M., Gutiérrez Castillo, J. J., y Palacios-Rodríguez, A. de P. (2020). Validación del cuestionario de competencia digital para futuros maestros mediante ecuaciones estructurales. *Bordón*, 72 (2), 45-63. <https://doi.org/10.13042/Bordon.2020.73436>
- García-Ruiz, R., Buenestado Fernández, M., y Ramírez-Montoya, M. S. (2023). Evaluación de la Competencia Digital Docente: Instrumentos, resultados y propuestas. Revisión sistemática de la literatura. *Educación XX1*, 26(1), 273-301. <https://doi.org/10.5944/educxx1.33520>
- Garduño, Y. S., Pérez-Morán, J. C., Ángeles, J. F. (2022). *Ficha Técnica del diseño de la Escalas de Autoevaluación en Competencias Digitales Básicas de Estudiantes de Educación Superior (EA-CDigit-B-EES v0.1)*. [Documento interno].

- González-Martínez, J., Esteve-Mon, F. M., Rada, V. L., Vidal, C. E., y Cervera, M. G. (2018). INCOTIC 2.0. Una nueva herramienta para la autoevaluación de la competencia digital del alumnado universitario. Profesorado. *Revista de currículum y formación del profesorado*, 22(4), 133-152. <https://doi.org/10.30827/profesorado.v22i4.8401>
- Haladyna, T. M., y Rodríguez, M. C. (2013). *Developing and Validating Test Items*. Routledge.
- International Test Commission. (2001). *International Guidelines for Test Use*. *International Journal of Testing*, 1(2), 93-114. https://doi.org/10.1207/S15327574IJT0102_1
- International Test Commission. (2014a). ITC Guidelines on Quality Control in Scoring, Test Analysis, and Reporting of Test Scores. *International Journal of Testing*, 14(3), 195-217. <https://doi.org/10.1080/15305058.2014.918040>
- International Test Commission. (2014b). *ITC Statement On the Use of Tests and Other Assessment Instruments for Research Purposes*. International Test Commission. https://www.intestcom.org/files/statement_using_tests_for_research.pdf
- Law, N. W. Y., Woo, D. J., De la Torre, J., and Wong, K. W. G. (2018). *A global framework of reference on digital literacy skills for indicator 4.4.2*. United Nations Educational, Scientific and Cultural Organization.
- Lawshe, C. H. (1975). *A Quantitative Approach to Content Validity*. Personnel psychology/Berrett-Koehler Publishers.
- Litiņa, S., & Miltuze, A. (2021). Students' digital competence: A scoping review of measuring instruments. En L. Daniela (Ed.), *Human, Technologies and Quality of Education: Proceedings of Scientific Papers*, (pp. 288-308). <https://doi.org/10.22364/htqe.2021.21>
- Martínez-Rizo, F. M. (2009). Evaluación formativa en aula y evaluación a gran escala: Hacia un sistema más equilibrado. *Revista Electrónica de Investigación Educativa*, 11(2). <https://redie.uabc.mx/redie/article/view/231>
- Mirete, A.; García, F. y Hernández, F. (2015). Cuestionario para el estudio de la actitud, el conocimiento y el uso de TIC (ACUTIC) en Educación Superior: Estudio de fiabilidad y validez. *Revista Interuniversitaria de Formación del Profesorado*, 83, 75-89, <https://www.redalyc.org/articulo.oa?id=274436590006>

- Muñiz, J., y Fonseca-Pedrero, E. (2019). Diez pasos para la construcción de un test. *Psicothema*, 31(1), 7–16. <https://doi.org/10.7334/psicothema2018.291>
- Pérez Morán, J. C. (2014). *Análisis del aspecto sustantivo de la validez de constructo de una prueba de habilidades cuantitativas* [Doctorado, Universidad Autónoma de Baja California]. http://iide.ens.uabc.mx/documentos/divulgacion/tesis/DCE/2010/Juan_Carlos_Perez_Moran.pdf
- Popham, W. J. (2013). On Serving Two Masters: Formative and Summative Teacher Evaluation. *Principal Leadership*, 13(7), 18-22.
- Ravela, P. (2013). La evaluación de y en la educación. *Boletín del Instituto Nacional de Evaluación Educativa*, 1. <https://www.ineed.edu.uy/images/pdf/Aristas-La-evaluacin-de-y-en-la-educacin-I.pdf>
- Sánchez-Canut, S., Usart-Rodríguez, M., Grimalt-Álvaro, C., Martínez-Requejo, S., & Lores-Gómez, B. (2023). Professional digital competence: definition, frameworks, measurement, and gender differences: A systematic literature review. *Human Behavior and Emerging Technologies*, 2023(1), 8897227. <https://doi.org/10.1155/2023/8897227>
- Scriven, M. (1967). The Methodology of Evaluation. In R. W. Tyler, R. M. Gagne, and M. Scriven (Eds.), *Perspectives of Curriculum Evaluation Chicago* (pp. 39-83). Rand McNally.
- Sillat, L. H., Tammets, K., and Laanpere, M. (2021). Digital competence assessment methods in higher education: A systematic literature review. *Education Sciences*, 11(8), 402. <https://doi.org/10.3390/educsci11080402>
- Silva-Quiroz, J. E., Abricot Marchant, N., Aranda Fáunderz, G., y Rioseco París, M. (2022). Diseño y Validación de un instrumento para evaluar competencia digital en estudiantes de primer año de las carreras de educación de tres universidades públicas de Chile. *Edutec, Revista Electrónica De Tecnología Educativa*, (79), 319–335. <https://doi.org/10.21556/edutec.2022.79.2333>
- Thompson, S. J., Johnstone, C. J., and Thurlow, M. L. (2002). Universal design applied to large scale assessments. *Synthesis Report*, 44. Minnesota: NCEO (National Center on Educational Outcomes). <https://nceo.umn.edu/docs/OnlinePubs/Synth44.pdf>

- Tristán-López, A. (2008). Modificación al modelo de Lawshe para el dictamen cuantitativo de la validez de contenido de un instrumento objetivo. *Avances en medición*, 6(1), 37-48.
- Vega-Hernández, M. C., Patino-Alonso, M. C., and Galindo-Villardón, M. P. (2018). *Multivariate characterization of university students using the ICT for learning*. *Computers & Education*, 121, 124-130. <https://doi.org/10.1016/j.compedu.2018.03.004>
- Vuorikari, R., Punie, Y., Gomez, S. C., and Van Den Brande, G. (2016). *DigComp 2.0: The digital competence framework for citizens. Update phase 1: The conceptual reference model* (No. JRC101254). Joint Research Centre.

