

[Cierre de edición el 31 de diciembre del 2023]

<https://doi.org/10.15359/ree.27-3.17254>
<https://www.revistas.una.ac.cr/index.php/educare>
educare@una.ac.cr

Propiedades psicométricas de un instrumento para evaluar la ejecución de una secuencia didáctica en educación superior

Psychometric Properties of an Instrument to Evaluate the Execution of a Didactic Sequence in Higher Education

Propriedades psicométricas de um instrumento para avaliar a implementação de uma sequência didática no ensino superior



Martha Gabriela Ávila-Camacho

Escuela Normal de Educación Preescolar del Estado de Coahuila
Saltillo, México

Centro Universitario CIFE
Cuernavaca, México

gabriela.avila@gmail.com

 <https://orcid.org/000-0001-8735-7579>

Luis Gibran Juárez-Hernández

Centro Universitario CIFE
Cuernavaca, México

luisgibrancife@gmail.com

 <https://orcid.org/0000-0003-0658-6818>

Recibido • Received • Recebido: 05 / 07 / 2022

Corregido • Revised • Revisado: 23 / 05 / 2023

Aceptado • Accepted • Aprovado: 22 / 06 / 2023

Resumen:

Objetivo. Analizar las propiedades psicométricas de validez de constructo y confiabilidad de un instrumento para evaluar una secuencia didáctica en educación superior. **Metodología.** Se aplicó el instrumento a 488 estudiantes de las licenciaturas en Educación Preescolar y Primaria que trabajaban los Planes de Estudio 2012. Se informó el propósito de la investigación, la confidencialidad de los resultados y se solicitó consentimiento para su uso. Para el análisis de validez de constructo se empleó un proceso de validación cruzada que involucró el análisis factorial exploratorio (AFE) y análisis factorial confirmatorio (AFC); la confiabilidad fue evaluada con el alfa de Cronbach. **Resultados.** El AFE reveló la correspondencia del modelo factorial obtenido con el modelo teórico, ya que únicamente se denotó un factor que explicó 63% de varianza, en el cual se encontraron representados la totalidad de ítems. El AFC mostró la sustentabilidad empírica del modelo factorial y óptimo ajuste (razón χ^2/gf : 2.19; GFI: 0.942; RMSAE: 0.063; RMR: 0.029; CFI: 0.977; TLI: 0.971). Finalmente se refiere la existencia



<https://doi.org/10.15359/ree.27-3.17254>

<https://www.revistas.una.ac.cr/index.php/educare>
educare@una.ac.cr

de validez convergente. La confiabilidad fue óptima (alfa de Cronbach: 0,931). **Conclusiones.** El instrumento posee adecuadas propiedades psicométricas, por lo tanto, su uso garantiza la obtención de resultados confiables en la evaluación de secuencias didácticas en nivel superior.

Palabras claves: Análisis factorial; competencias del personal docente; educación superior; evaluación; evaluación del personal docente.

Abstract:

Aim. To analyze the psychometric properties of construct validity and reliability of an instrument to assess the didactic sequence. **Methodology.** The instrument was administered to 488 students enrolled in Bachelor's Degrees in Preschool and Elementary Education; the students were working on the 2012 study plans. They were informed of the purpose of the research and the confidentiality of the results; they also were asked for their consent to use the data collected. The construct validity analysis implemented a cross-validation process involving an exploratory factor analysis (EFA) and a confirmatory factor analysis (CFA). Reliability was evaluated with Cronbach's alpha. **Results.** The EFA revealed the correspondence of the factorial model obtained with the theoretical model, since only one factor was found to explain the 63% of variance, in which all items were represented. The CFA showed the empirical sustainability of the factorial model and optimal fit (Ratio χ^2/df : 2.19; GFI: 0.942; RMSAE: 0.063; RMR: 0.029; CFI: 0.977; TLI: 0.971). The study also confirmed the presence of convergent validity. Finally, the Cronbach's alpha result was 0.931, indicating optimal reliability. **Conclusions.** The instrument has adequate psychometric properties. Therefore, its utilization ensures the acquisition of reliable results in the evaluation of didactic sequences in higher education.

Keywords: Factor analysis; teacher competencies; higher education; assessment; teacher assessment.

Resumo:

Objetivo. Analisar as propriedades psicométricas de validade de construto e confiabilidade de um instrumento para avaliar uma sequência didática no ensino superior. **Metodologia.** O instrumento foi aplicado a 488 alunos dos cursos de Bacharelado em Educação Infantil e Ensino Fundamental que estavam trabalhando nos Planos de Estudo de 2012. Os participantes foram informados sobre o objetivo da pesquisa, a confidencialidade dos resultados e foram solicitados a dar seu consentimento para o uso dos dados coletados. Um processo de validação cruzada envolvendo a análise fatorial exploratória (AFE) e a análise fatorial confirmatória (AFC) foi usado para analisar a validade do construto. A confiabilidade foi avaliada utilizando o coeficiente alfa de Cronbach. **Resultados.** A AFE revelou a correspondência do modelo fatorial obtido com o modelo teórico, já que apenas um fator foi encontrado para explicar 63% da variação, no qual todos os itens estavam representados. A AFC mostrou a sustentabilidade empírica do modelo fatorial e o ajuste ideal (Razão χ^2/df : 2,19; GFI: 0,942; RMSAE: 0,063; RMR: 0,029; CFI: 0,977; TLI: 0,971). Além disso, foi mencionada a existência de validade convergente. Finalmente, o resultado do alfa de Cronbach foi 0,931, demonstrando confiabilidade ótima. **Conclusões.** O instrumento tem propriedades psicométricas adequadas. Portanto, o seu uso garante a obtenção de resultados confiáveis na avaliação de sequências didáticas no ensino superior.

Palavras-chave: Análise de factores; competências dos professores; ensino superior; avaliação; avaliação dos professores.



Introducción

Una secuencia didáctica (SD) es un recurso que permite al personal docente organizar una serie de estrategias en la búsqueda del logro de ciertos aprendizajes para el estudiantado, caracterizada, entre otras cosas, por partir de los conocimientos previos base para la consolidación de mejores y mayores niveles de desempeño de la competencia. Para ello se debe utilizar todo el bagaje de saberes conceptuales, actitudinales y procedimentales en la solución de problemas del contexto propios de su área de intervención profesional, ya que de esa manera se incrementa y se percibe todo lo aprendido (Maine, 2017; Tobón Tobón et al., 2010). Para lograrlo se requiere reflexionar con base en los propósitos compartidos entre docente y estudiantes, los retos cognitivos que habrán de afrontar, pues va más allá de solo de dosificar una serie de actividades en un tiempo determinado.

A este respecto, la SD es importante, ya que por este medio el personal docente busca estratégicamente que el estudiantado movilice sus recursos actitudinales, conceptuales, procedimentales e intelectuales, para dar solución a los desafíos que se les presenten (Brousseau, 2007). Con una SD bien planeada, en la que se incorporen actividades para reconocer los saberes previos del alumnado, en que resuelvan problemáticas propias de la profesión partiendo de referentes teóricos que gestionan por sí mismo, con uso de una serie de recursos tecnológicos y científicos, además de trabajar en colaboración con otras personas, con las competencias necesarias para poder contribuir en el nivel educativo en que se desempeñaran en la formación de la nueva ciudadanía del mundo. Además, adquirirán la capacidad reflexiva de identificar sus fortalezas y áreas de oportunidad para mejorar de manera permanente (Ávila-Camacho et al., 2019).

Dada la importancia del impacto que una secuencia didáctica tiene en el proceso de adquisición y mejora de competencias, es fundamental valorar la manera en que estas se piensan y se operan en la práctica; además de su efectividad con fundamento en el desempeño docente y los resultados del estudiantado. Como señalan Darling-Hammond & McLaughlin (2003), la evaluación es una oportunidad de desarrollo profesional en la escuela, en la medida en que se parta de ella para mejorar los aspectos reconocidos como objeto de mejora; por lo tanto, se puede precisar que el hecho de trabajar con secuencias didácticas no garantiza la consecución de las metas establecidas *per se*; es necesario y responsable verificar la efectividad de todo el proceso que implica, para mejorar y adecuar de manera permanente con base en los hallazgos producto de la reflexión.

A este respecto, en la bibliografía se reporta una serie de instrumentos para la evaluación de la secuencia didáctica desde la perspectiva del personal docente (García Marimón et al., 2021; Guevara-Rodríguez & Veytia-Bucheli, 2021; Filadelfo & Guridi, 2014); y desde perspectiva de asesorías técnicas pedagógicas de preescolar, primaria y secundaria (Ibarra-Piza et al., 2018); sin embargo, son escasos los aportes en que se compartan instrumentos o evaluaciones en



<https://doi.org/10.15359/ree.27-3.17254>
<https://www.revistas.una.ac.cr/index.php/educare>
educare@una.ac.cr

los cuales el alumnado funja como evaluador. En este sentido, se considera que la evaluación hecha desde el punto de vista del estudiantado, al ser el actor central del proceso de enseñanza podría ser la más significativa debido a que son testigos de las formas de enseñanza y con base en su percepción se puede mejorar la intervención docente.

Por otro lado, [Tena & Couso \(2023\)](#) destacan la importancia de la evaluación de las secuencias didácticas diseñadas en un contexto de investigación, debido a que actualmente se ha incrementado, de manera significativa, su generación. Hacen una propuesta basada en la revisión de otros trabajos con el paradigma de la investigación basada en el diseño en que se abordan elementos, prácticas y criterios de calidad. Proponen que la evaluación se haga a partir del análisis del grado de exigencia del resultado con base la calidad determinada por: la validez, utilidad y confiabilidad de las SD.

Desde esta perspectiva, se encuentra la propuesta de [Ávila Camacho et al. \(2019\)](#) denominada *Instrumento de valoración del desempeño docente en la ejecución de una secuencia didáctica (IVDD-ESD V1)*. Este instrumento tiene el objetivo de conocer cómo el personal docente ejecuta una secuencia didáctica como medio para promover el desarrollo de competencias en el estudiantado y el momento en que este moviliza sus propias competencias docentes.

Como se precisó, el instrumento mencionado tiene por objetivo evaluar el desempeño docente a través de la ejecución de una secuencia didáctica, para lo que se consideran como referencia los tres momentos del proyecto de enseñanza de educación primaria, que fueron adecuados al nivel superior: planeación, intervención y reflexión ([Secretaría de Educación Pública \[SEP\], 2017](#)). Los aspectos que se evalúan son: "elementos curriculares, problemas del contexto, análisis de saberes previos, gestión del conocimiento, contextualización, resolución de problemas, recursos, organización de actividades, evaluación, metacognición y finalmente análisis y reflexión" ([Ávila-Camacho, 2019, pp. 2-3](#)).

En este sentido, uno de los requisitos para poder usar un instrumento es que se cumpla con los criterios de validez de contenido y constructo, además de la confiabilidad ([Carvajal et al., 2011](#)). Con base en ello, el instrumento referido fue sometido a la validación de facie, de contenido y constructo, además de haberlo sometido a evaluación de confiabilidad ([Ávila-Camacho et al., 2019](#)). Sin embargo, estos autores reconocen que la validación de constructo (realizada mediante el análisis factorial exploratorio) no pudo ser conclusiva por el tamaño de muestra ($n < 200$ participantes).

Por lo anterior, [Hair et al. \(2014\)](#) refieren que para tener resultados conclusivos en el AFE, el número de participantes deberá de ser mayor a 200. De esta misma forma, [Charter \(2003\)](#) indica que el cálculo de alfa de Cronbach con una número menor a 200 participantes tiende a ser inestable.

Puesto que la validez de constructo es considerada como la validez de mayor relevancia ([Messick, 1980](#)) y un instrumento se fortalece conforme se analicen sus propiedades psicométricas ([Carvajal et al., 2011](#)), el propósito de la presente investigación fue efectuar

el análisis de validez de constructo del *Instrumento de valoración del desempeño docente en la ejecución de una secuencia didáctica* (IVDD-ESD V1), con el objetivo de comprobar que los elementos que lo integran reproducen el constructo planteado teóricamente.

Metodología

Tipo de estudio

Se realizó un estudio instrumental del *Instrumento de valoración del desempeño docente en la ejecución de una secuencia didáctica*, IVDD-ESD V1, de manera específica este estudio abordó el análisis de validez de constructo y confiabilidad.

Instrumento

El IVDD-ESD V1 (Ávila-Camacho et al., 2019) es un instrumento diseñado con el formato de una rúbrica analítica con el objetivo de conocer cómo el personal docente ejecuta una secuencia didáctica al ser este el medio para promover el desarrollo de competencias en el estudiantado y el momento en que el mismo personal moviliza sus propias competencias docentes (Pimienta Prieto, 2011). El instrumento integra tres momentos, como en el proyecto de enseñanza de educación primaria: planeación, intervención y reflexión (SEP, 2017), y de acuerdo con Ávila-Camacho et al. (2019), la decisión de partir de ellos fue determinado por la dosificación que se tiene de las acciones que el personal docente habrá de realizar en cada uno de los tres momentos y que fueron adecuados al nivel superior y complementados con la propuesta de acciones de una secuencia didáctica socioformativa (Tobón Tobón et al., 2010). Los aspectos que se evalúan en la planeación son los elementos curriculares, problemas del contexto y análisis de saberes previos; en la intervención se evalúa la gestión del conocimiento, contextualización, resolución de problemas, recursos, organización de actividades, evaluación y metacognición; y finalmente en el tercer momento análisis y reflexión (Tobón, citado en Ávila-Camacho et al., 2019).

El instrumento referido fue diseñado como rúbrica analítica, la cual brinda una especificidad respecto del nivel ponderado en cada una de las dimensiones o criterios analizados (Gatica-Lara & Uribarren-Berrueta, 2013). En específico se precisa que el instrumento posee pocos ítems y para cada uno de ellos tiene opciones de respuesta que se describen como niveles de desempeño (preformal, receptivo, resolutivo, autónomo y estratégico) donde este último es el desempeño ideal (Tobón, 2017).

En el afán de que el instrumento tomara un carácter óptimo, en una primera etapa, fue sometido a la revisión de tres personas expertas, juicio de personas expertas para evaluar la validez de contenido, aplicación a un grupo piloto y a una muestra de 134 estudiantes de la Escuela Normal de Educación Preescolar para el análisis de validez de constructo. Mediante este esquema se demostró la validación de facie, la validez de contenido (V de Aiken > 0.80 e $ICI > 0.50$) y validez de constructo, y se encontró la correspondencia del modelo teórico propuesto con



<https://doi.org/10.15359/ree.27-3.17254>
<https://www.revistas.una.ac.cr/index.php/educare>
educare@una.ac.cr

el modelo factorial resultante (estructura unidimensional) y que todos los ítems estuvieron representados en el factor resultante. Respecto a la confiabilidad se obtuvo un valor óptimo (0,93) (Ávila-Camacho et al., 2019).

Participantes

El número de participantes se basó en los requerimientos de un proceso de validación cruzada que involucraría el análisis factorial exploratorio y confirmatorio. Por lo que se acordó que el un número de participantes estuviera entre 400 y 500, logrando que 488 quisieran participar en el estudio. Los sujetos participantes fueron estudiantes de los cuatro grados de las licenciaturas en Educación Preescolar y Primaria del Plan de Estudios 2012 del estado de Coahuila, a quienes se les informó sobre el propósito del instrumento, la confidencialidad de los datos personales y se les solicitó su consentimiento para el uso de los resultados con fines investigativos.

Los sujetos participantes evaluaron la manera en que el personal docente responsable de impartir los cursos del trayecto de práctica profesional ejecuta una secuencia didáctica, ya que son ellos quienes pueden valorar de primera mano a sus maestros y maestras de manera objetiva, por ser en ellos en quienes se centra la enseñanza. En seguida se incluye la caracterización de las personas evaluadoras (Tabla1).

Tabla1: Caracterización sociodemográfica de las personas participantes

Indicador	Descripción
Sexo (%)	Mujeres 90,16%
	Hombres 9,8%
Edad (media ± DE)	20.54±2.91
Zona de residencia (%)	Zona urbana 89,9%
	Zona rural 4,9%
	Zona semi-urbana 5,1%
Grado satisfacción de la carrera (%)	Excelente 29,7%
	Bueno 44%
	Aceptable 21,5%
	Bajo 3,6%
	Muy Bajo 1,02%
Grado de satisfacción desempeño docentes	Excelente 13,3%
	Bueno 47,3%
	Aceptable 32,5%
	Bajo 5,9%
	Muy Bajo ,8%

Nota: Elaboración propia.

Análisis factorial exploratorio y confirmatorio

Inicialmente se realizó un análisis de ajuste de los ítems a la distribución normal mediante el cálculo de la asimetría y curtosis, considerando que si algún ítem presentaba un valor superior a ± 2 (Bollen & Long, 1993), era sujeto de eliminación. Posteriormente se analizó la correlación ítem-prueba y si algún ítem presentaba un valor superior a 0,90 o menor a 0,20 se consideraría eliminarlo (Tabachnick & Fidell, 2001).

Mediante un proceso de aleatorización la muestra total se particionó en dos submuestras de igual tamaño, con el objetivo de efectuar una validación cruzada, el cual se corresponde con la recomendación clásica y de mayor pertinencia (Brown, 2015). Con la primer submuestra se efectuó el análisis factorial exploratorio (AFE) y la pertinencia de los datos a este análisis se efectuó a través de la observación de la matriz de correlaciones, la prueba de Kaiser Meyer Olkin y de Barlett (Juárez-Hernández, 2018). Una vez verificados los supuestos del método, se procedió con el análisis factorial con el empleo del método de estimación de máxima verosimilitud, si se comprobaba la normalidad, o bien el método de mínimos cuadrados no ponderados cuando no se cumplió el supuesto normalidad (Juárez-Hernández, 2018; Thompson, 2004). Para la elección del número de factores a retener se consideró la regla de Gutman-Kaiser y porcentaje de la varianza explicada (Hair et al., 2014; Thompson, 2004). Para determinar la significancia de las cargas factoriales acorde al tamaño de muestra se consultó lo indicado por Hair et al. (2014). Acorde con diversos estudios como los de Thompson (2004), Hair et al. (2014) y Juárez-Hernández (2018), cuando los ítems presentan cargas factoriales significativas a más de un factor, se efectúa la rotación matricial.

Con la segunda submuestra se realizó el análisis factorial confirmatorio (AFC) y se empleó el método de estimación de máxima verosimilitud, y se evaluó la bondad de ajuste del modelo mediante diversos indicadores (razón chi-cuadrada/grados de libertad (χ^2/df) índice de bondad de ajuste (GFI); error cuadrático medio de aproximación (RMSAE); raíz cuadrada residual (RMR); índice comparativo de ajuste (CFI); índice de Tucker-Lewis (TLI), se consideraron los valores de referencia establecidos por Yuan (2005). Finalmente se calculó la varianza media extraída y la confiabilidad compuesta, se aceptó el valor umbral para la primera mayor a 0.50 (Fornell & Larcker, 1981) y para la segunda de 0.70 (Hair et al., 2014). Así mismo se realizó el cálculo de la confiabilidad el alfa de Cronbach (Cronbach, 1951). El cálculo de la asimetría, curtosis, correlación ítem-prueba, confiabilidad y análisis factorial exploratorio, confirmatorio se realizó con el programa JASP versión 0.11.1 (JASP Team, 2019).



<https://doi.org/10.15359/ree.27-3.17254>
<https://www.revistas.una.ac.cr/index.php/educare>
educare@una.ac.cr

Resultados

Ninguno de los ítems violó los criterios establecidos de asimetría y curtosis (Tabla 2), por lo que se puede asumir que los datos se presentan una distribución normal. Acorde con el indicador de ítem-test-correlación, todos los ítems contribuyen al valor total, por lo cual no existió necesidad de eliminar alguno.

Tabla 2: Análisis descriptivo de los ítems

Ítem	Media	Asimetría	Curtosis	Test correlación
1	3.701	-0.422	0.337	0.720
2	3.753	-0.345	0.125	0.671
3	3.497	-0.516	-0.225	0.740
4	3.644	-0.31	-0.056	0.774
5	3.671	-0.427	-0.0003307	0.743
6	3.661	-0.308	-0.115	0.758
7	3.648	-0.784	-0.536	0.705
8	3.612	-0.683	0.549	0.789
9	3.602	-0.426	-0.242	0.758
10	3.566	-0.702	0.251	0.790
11	3.526	-0.488	-0.233	0.773

Nota: Elaboración propia.

Los datos mostraron pertinencia para ser analizados mediante el análisis factorial exploratorio, ya que se presentaron correlaciones significativas entre ítems ($r > 0.50$; $p < 0.05$) (Tabla 3), así como se evidenció mediante las pruebas de pertinencia efectuadas (KMO: 0.963; Prueba de Barlet: 3511.798, gl 55, $p < 0.001$). El 100% de los ítems estuvo representado dentro del modelo factorial en donde un solo factor presentó un autovalor mayor a uno, el cual explicó más del 63% de la varianza, y en el cual estuvieron representados todos los ítems con cargas factoriales significativas (Tabla 4). Finalmente, el instrumento presentó un valor óptimo de confiabilidad (Alfa de Cronbach: 0.931; IC 95%: 0.921 ± 0.940).



Tabla 3: Matriz de correlaciones entre los ítems

	Ítem 1 EC	Ítem 2 PC	Ítem 3 ASP	Ítem 4 GC	Ítem 5 CON	Ítem 6 RP	Ítem 7 REC	Ítem 8 OA	Ítem 9 EV	Ítem 10 META	Ítem 11 AYR
Ítem 1 EC	1.000										
Ítem 2 PC	.610*	1.000									
Ítem 3 ASP	.575*	.566*	1.000								
Ítem 4 GC	.595*	.595*	.620*	1.000							
Ítem 5 CON	.599*	.635*	.608*	.629*	1.000						
Ítem 6 RP	.550*	.575*	.587*	.613*	.606*	1.000					
Ítem 7 REC	.503*	.500*	.566*	.558*	.544*	.557*	1.000				
Ítem 8 OA	.587*	.563*	.576*	.622*	.585*	.624*	.598*	1.000*			
Ítem 9 EV	.597*	.537*	.586*	.602*	.612*	.602*	.598*	.650*	1.000		
Ítem 10 META	.563*	.573*	.606*	.610*	.642*	.571*	.592*	.696*	.685*	1.000	
Ítem 11 AYR	.561*	.557*	.622*	.571*	.561*	.584*	.614*	.651*	.643*	.687*	1.000

* correlaciones significativas $p < 0.05$

Nota: Elaboración propia.



<https://doi.org/10.15359/ree.27-3.17254>
<https://www.revistas.una.ac.cr/index.php/educare>
educare@una.ac.cr

Tabla 4: Comunalidad y carga factorial por ítem

Ítem	Comunalidad	Carga Factorial
Ítem 1. Elementos curriculares (EC)	0.5488	0.7408
Ítem 2. Problemas del contexto (PC)	0.5422	0.7364
Ítem 3. Análisis de saberes previos (ASP)	0.5858	0.7654
Ítem 4. Gestión del conocimiento (GC)	0.6089	0.7803
Ítem 5. Contextualización (CON)	0.6098	0.7809
Ítem 6. Resolución de problemas (RP)	0.5770	0.7596
Ítem 7. Recursos (REC)	0.5276	0.7264
Ítem 8. Organización de actividades (OA)	0.6416	0.8010
Ítem 9. Evaluación (EV)	0.6324	0.7953
Ítem 10. Metacognición (META)	0.6598	0.8123
Ítem 11. Análisis y reflexión (AYR)	0.6185	0.7864

Nota: Elaboración propia.

El análisis factorial confirmatorio reveló un buen ajuste del modelo (Tabla 5). Específicamente el valor de la razón χ^2/df los índices de ajuste absoluto (GFI, SRMR y RMSA) e índices incrementales (CFI, NFI, IFI) mostraron un valor óptimo (Tabla 5). Finalmente, el cálculo de la varianza extractada media fue de 0,51 y la confiabilidad compuesta fue de 0,93 (Tabla 6) y se refiere que cada ítem presentó una CFE mayor a 0,50 (Tabla 6 y Figura 1).

Tabla 5: Ajuste del modelo factorial

Index	Valor esperado (Yuan, 2005)	Valor obtenido
Razón chi-cuadrada/grados de libertad (χ^2/df)	Menor a 3	2.19
Índice de bondad de ajuste (GFI)	Mayor a 0.90	0.942
Error cuadrático medio de aproximación (RMSAE)	.050 a .080	0.063 (0.046 a 0.080)
Raíz cuadrada residual (RMR)	Menor a .050	0.029
Índice comparativo de ajuste (CFI)	Mayor a 0.95	0.977
Índice de Tucker-Lewis (TLI)	Mayor a .90	0.971

Nota: Elaboración propia.

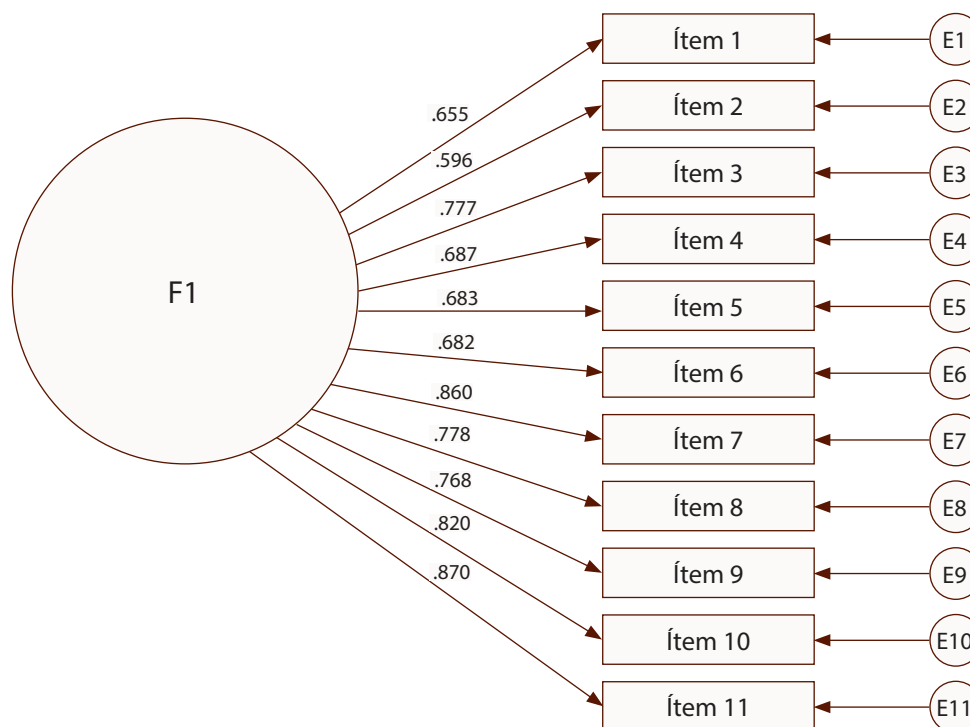


Tabla 6: Resumen del modelo

Constructo	# de ítems	Confiabilidad compuesta	Alfa de Cronbach	VEM	Cargas factoriales estandarizadas
Secuencia didáctica	11	.932	.931	.56	I1 (.655), I2 (.596), I3 (.77), I4 (.687), V5 (.683), V6 (.682), V7 (.86), V8 (.778), V9 (.768), V10 (.82), V11 (0.87)

Nota: Elaboración propia.

Figura 1: Representación del modelo factorial



Nota: Elaboración propia.

Discusión

Las secuencias didácticas son un recurso a la mano del personal docente que permite organizar una serie de acciones y que resulta relevante en la promoción de competencias, en la medida en que se planifique considerando el propósito y enfoque de los saberes que

<https://doi.org/10.15359/ree.27-3.17254>
<https://www.revistas.una.ac.cr/index.php/educare>
educare@una.ac.cr

habrá de promover, parte de los conocimientos previos del estudiantado para que sean la base de aquellos que habrán de construir; que las actividades estén relacionadas con los saberes conceptuales y con problemáticas propias de la profesión, de tal forma que el alumnado, con el apoyo del personal docente, les encuentre aplicación para solucionar los retos con el uso estratégico de recursos y gracias a un ambiente propicio para la reflexión, en el que exista la confianza de compartir experiencias desafiantes para tomar decisiones que permitan tanto al alumnado como al personal docente mejorar su desempeño (Ávila-Camacho et al., 2019; Tobón Tobón et al., 2010).

Dada la relevancia que una secuencia didáctica tiene en la promoción de competencias en las futuras generaciones profesionales de la educación, resulta preponderante que sea evaluada, con la finalidad de identificar las fortalezas que una persona docente podría estar teniendo y de igual manera aspectos en que podría mejorar (Guevara-Rodríguez & Veytia-Bucheli, 2021). Sucede que en ocasiones se buscan razones externas para explicar el bajo desempeño de sus estudiantes, sin considerar que la razón podría estar en la práctica docente; para identificar su origen se requiere, además de tener mente abierta y voluntad para mejorar, un instrumento confiable que dé soporte a los resultados obtenidos.

Por lo anterior, es que la propuesta del instrumento *IVDD-ESD V1* (Ávila-Camacho et al., 2019) es relevante, pues da respuesta a la necesidad de evaluar la ejecución de secuencias didácticas por parte del estudiantado. En este orden, conviene señalar que aun cuando ha sido pensado originalmente el instrumento para que el estudiantado evalúe el desempeño docente (heteroevaluación), resulta ser un recurso que puede ser aplicado con fines de autoevaluación, a iniciativa del personal docente interesado en reflexionar sobre su desempeño; así como coevaluación en pares para retroalimentarse por algún colega, ya sea a iniciativa propia o como parte de los procesos de evaluación instruccional. Entre otra de sus bondades se encuentra que consta de pocos ítems (11) y, al contener descriptores de desempeño, la interpretación de los resultados es mucho más rápida, precisa y completa (Ávila-Camacho et al., 2019).

Un aspecto para señalar del instrumento es que fue sujeto al análisis de propiedades psicométricas (validez de facie, validez de contenido, validez de constructo, confiabilidad) (Ávila-Camacho et al., 2019); sin embargo, si bien se refiere que el análisis de validez de constructo efectuado mediante el análisis factorial exploratorio es factible con muestras de 100 participantes, se recomienda una muestra mayor (Hair et al., 2014) con el objetivo de que los resultados sean conclusivos. Por lo anterior, en el presente estudio se volvió a efectuar el análisis de validez de constructo con una muestra mayor (488), la cual se considera como óptima (Hair et al., 2014) y con un procedimiento de validación cruzado, referido como el de mayor conveniencia y robustez (Brown, 2015).

En este orden, del análisis de validez de constructo mediante el AFE se indica la correspondencia con el modelo teórico, al encontrar la representación de todos los ítems con cargas factoriales significativas a un solo factor, similar a lo reportado por [Ávila-Camacho et al. \(2019\)](#). A este respecto, como se indicó, en ambos casos se ha revelado la existencia de un solo factor, con lo que se denota la unidimensionalidad del instrumento y que desde la perspectiva teórica significa la existencia de un único rasgo subyace a las respuestas de los sujetos a los ítems ([Hattie, 1985, citado en Burga León, 2006](#)) e indica la presencia de consistencia interna, lo cual indica que los 11 ítems del instrumento efectivamente miden una sola variable: la secuencia didáctica, coincidiendo con el primer análisis.

A este respecto, en el presente el único factor encontrado explicó más del 63% de la varianza e integró a todos los ítems, destacando que todos presentaron cargas factoriales superiores al mínimo establecido ($CFE > 0.50$) ([Hair et al., 2014](#)), lo anterior revela la validez de constructo del instrumento, enfatizando que los ítems propuestos reproducen el constructo planteado teóricamente. Es importante señalar este punto, ya que la validez de constructo es el principal tipo de validez, pues resulta importante determinar la relevancia de los ítems para el objetivo que ha sido diseñado y que cada uno de ellos representa al constructo evaluado ([Messick, 1980](#)). Ello denota la importancia de la fase que antecede a la presente, es decir, del juicio experto, ya que como se indica, la validez de contenido es un componente esencial de la validez de constructo.

Referente a la evaluación del modelo mediante el AFC, se precisa que brindó elementos de comprobación del ajuste del modelo a los datos, ya que para los diferentes indicadores empleados (razón X^2/df , índices de ajuste absoluto e índices incrementales) mostraron valores óptimos. Se destaca que un aspecto importante y aportado por este análisis es el valor de las cargas factoriales estandarizadas ($CFE > 0,50$) y la confiabilidad compuesta ($CC: 0,891$) así como la varianza extraída media ($VEM: 0,51$). Mediante estos indicadores se puede referir la sustentabilidad empírica del modelo propuesto y que los indicadores propuestos miden dicho factor ([Cheung & Wang, 2017; Fornell & Larcker, 1981](#)). A este respecto, es a través del valor de las cargas factoriales estandarizadas; la confiabilidad compuesta y la varianza extraída media que se puede indicar la existencia de validez convergente ([Fornell & Larcker, 1981; Hair et al., 2014](#)), esto es: el constructo es adecuadamente medido por los indicadores propuestos ([Cheung & Wang, 2017](#)).

Acorde con el valor de la confiabilidad obtenida (0,931), se puede considerar como óptima, esto significa, en primera instancia, la capacidad del instrumento de obtener mediciones con un error mínimo y, en segunda instancia, demuestra la correlación entre el concepto abordado e ítems de este ([Jabrayilov et al., 2016](#)).

Se puede concluir que el proceso metodológico efectuado de manera inicial ([Ávila-Camacho et al., 2019](#)), así como en el presente para el instrumento *IVDD-ESD V1*, permiten



<https://doi.org/10.15359/ree.27-3.17254>
<https://www.revistas.una.ac.cr/index.php/educare>
educare@una.ac.cr

asegurar que el instrumento posee validez de contenido y constructo, así como es confiable, traduciéndose en que los elementos que lo integran son pertenecientes, pertinentes, relevantes y representativos el constructo planteado teóricamente (Haynes et al., 1995). Por lo anterior, la demostración de las propiedades psicométricas analizadas en el presente permite asegurar que el instrumento podrá entregar evidencias válidas y confiables (Kerlinger & Lee, 2002).

Finalmente, de este estudio se puede generar una línea de investigación que identifique los aspectos que destacan de los resultados de las evaluaciones realizadas con el instrumento IVDD-ESD V1, ya sea porque requieren ser reforzados o porque implican fortalezas en la ejecución de secuencias didácticas en educación superior.

Declaración de contribuciones

Las personas autoras declaran que han contribuido en los siguientes roles: **M. G. A. C.** contribuyó en la escritura del artículo, la supervisión y administración del proceso investigativo, la obtención de recursos, y la conceptualización, metodología y curación de la investigación. **L. G. J. H.** contribuyó en la revisión-edición y primer borrador del artículo, la validación del proceso investigativo, y la metodología, conducción, curación y análisis de la investigación.

Declaración de material complementario

Este artículo tiene disponible, como material complementario:

-La versión preprint del artículo en <https://doi.org/10.35542/osf.io/p9fqz>

Referencias

- Ávila-Camacho, G. (2019). Diagnóstico del desempeño docente en la ejecución de proyectos formativos. *4to. Congreso Nacional de Investigación sobre Educación Normal*. CONISEN. <https://conisen.mx/memorias2019/memorias/2/P267.pdf>
- Ávila-Camacho, M. G., Juárez-Hernández L. G., Arreola-González, A. L., & Palmares Villarreal, O. G. (2019). Construcción y validación de un instrumento de valoración del desempeño docente en la ejecución de una secuencia didáctica. *Revista de Investigación en Educación*, 17(2), 122-142. <https://revistas.uvigo.es/index.php/reined/article/view/2150/2190>
- Bollen, K. A. & Long, J. S. (Editores). (1993). *Testing structural equation models*. Sage
- Brousseau, G. (2007). *Iniciación al estudio de la teoría de las situaciones didácticas*. Zorzal.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford Publications.

- Burga León, A. (2006). La unidimensionalidad de un instrumento de medición: Perspectiva factorial. *Revista de Psicología*, 24(1), 53-80. <https://doi.org/10.18800/psico.200601.003>
- Carvajal, A., Centeno, C., Watson, R., Martínez, M., & Sanz Rubiales, Á. (2011). ¿Cómo validar un instrumento de medida de la salud? *Anales del Sistema Sanitario de Navarra*, 34(1), 63-72. <https://doi.org/10.4321/S1137-66272011000100007>
- Charter, R. A. (2003). A breakdown of reliability coefficients by test type and reliability method, and the clinical implications of low reliability. *The Journal of General Psychology*, 130(3), 290-304. <https://doi.org/10.1080/00221300309601160>
- Cheung, G. W. & Wang, C. (2017). Current approaches for assessing convergent and discriminant validity with SEM: Issues and solutions. *Academy of Management Proceedings*, (1), 12706. <https://doi.org/10.5465/AMBPP.2017.12706abstract>
- Cronbach, L. J. (1951). Coefficient Alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334. <https://doi.org/10.1007/BF02310555>
- Darling-Hammond, L. & McLaughlin, M. W. (2003). *El desarrollo profesional de los maestros. Nuevas estrategias y políticas de apoyo* (Cuadernillos de discusión). SEP. <https://educacionbasica.sep.gob.mx/multimedia/RSC/BASICA/Documento/201611/201611-3-RSC-mtL7leHGBY-cds09.pdf>
- Filadelfo, C. & Guridi, V. (2014). Evaluación de una secuencia didáctica basada en actividades lúdicas para la enseñanza de alimentación y nutrición en los primeros años de la enseñanza secundaria. *Revista Enseñanza de la Física*, 6(Extra1), 309-320. <https://dialnet.unirioja.es/servlet/articulo?codigo=5192248>
- Fornell, C. & Larcker, D. F. (1981). Evaluating structural equation Models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50. <https://doi.org/10.2307/3151312>
- García Marimón, O., Diez-Palomar, J., Morales Maure, L., & Durán González, R. E. (2021, agosto). Evaluación de secuencias de aprendizaje de matemáticas usando la herramienta de los criterios de idoneidad didáctica. *Boletim de Educação Matemática*, 35(70), 1047-1072. <https://doi.org/10.1590/1980-4415v35n70a23>
- Gatica-Lara, F. & Uribarren-Berrueta, T. del N. J. (2013). ¿Cómo elaborar una rúbrica? *Investigación en educación médica*, 2(5), 61-65. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2007-50572013000100010
- Guevara-Rodríguez, G. & Veytia-Bucheli, M. G. (2021). Validez de contenido de una rúbrica analítica del diseño de secuencias didácticas como mejora de la práctica pedagógica del equipo docente desde el enfoque de la socioformación. *Revista Electrónica Educare*, 25(1), 1-20. <https://doi.org/10.15359/ree.25-1.20>



<https://doi.org/10.15359/ree.27-3.17254>
<https://www.revistas.una.ac.cr/index.php/educare>
educare@una.ac.cr

- Hair, Jr., J. F., Black, W. C., Babin, B. H., & Anderson, R. E. (2014). *Multivariate data analysis*. Pearson.
- Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, 7(3), 238-247. <https://doi.org/10.1037/1040-3590.7.3.238>
- Ibarra-Piza, S., Segredo-Santamaría, S., Juárez-Hernández, L. G., & Tobón, S. (2018). Estudio de validez de contenido y confiabilidad de un instrumento para evaluar la metodología socioformativa en el diseño de cursos. *Revista espacios*, 39(53), 24-33. <https://revistaespacios.com/cited2017/cited2017-24.pdf>
- Jabrayilov, R., Emons, W. H., & Sijtsma, K. (2016). Comparison of classical test theory and item response theory in individual change assessment. *Applied Psychological Measurement*, 40(8), 559-572. <https://doi.org/10.1177/0146621616664046>
- JASP Team. (2019). *JASP* (Version 0.11.1) [Computer Software]. <https://jasp-stats.org/>
- Juárez-Hernández, L. G. (2018). *Manual práctico de estadística básica para la investigación*. Kresearch. <https://doi.org/10.24944/isbn.978-1-945721-24-3>
- Kerlinger, F. N. & Lee, H. B. (2002). *Investigación del comportamiento: Métodos de investigación en ciencias sociales*. McGraw-Hill.
- Maine, C. A. (2017). La secuencia didáctica una oportunidad para pensar, planear, vivir, mirar y sentir la clase de ciencias. *Anuario Digital de Investigación Educativa*, (1), 501-517. <https://revistas.bibdigital.uccor.edu.ar/index.php/adiv/article/view/3250/1962>
- Messick, S. (1980). Test Validity and Ethics of Assessment. *American Psychologist*, 35(11), 1012-1027. <https://doi.org/10.1037/0003-066X.35.11.1012>
- Pimienta Prieto, J. H. (2011). Secuencias didácticas: Aprendizaje y evaluación de competencias en educación superior. *Revista de Pedagogía*, 63(1), 77-92. <https://recyt.fecyt.es/index.php/BORDON/article/view/28906>
- Secretaría de Educación Pública (SEP). (2017). *Guía académica del sustentante para la evaluación del desempeño del personal docente tercer grupo 2017. Educación primaria. Educación básica*. https://www.academia.edu/34451803/Gu%C3%ADa_Acad%C3%A9mica_del_sustentante_para_la_Evaluaci%C3%B3n_del_Desempe%C3%B1o_del_Personal_Docente_Tercer_Grupo_2017_Educaci%C3%B3n Primaria
- Tabachnick, B. G. & Fidell, L. S. (2001). *Using multivariate statistics*. Allyn y Bacon.

- Tena, È. & Couso, D. (2023). ¿Cómo se que mi secuencia didáctica es de calidad? Propuesta de un marco de evaluación desde la perspectiva de investigación basada en diseño? *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 20(2), 280101-280115. https://doi.org/10.25267/rev_eureka_ensen_divulg_cienc.2023.v20.i2.2801
- Thompson, B. (2004). *Exploratory and confirmatory. Factor analysis: Understanding concepts and applications*. American Psychological Association. <https://doi.org/10.1037/10694-000>
- Tobón Tobón, S., Pimienta Prieto, J. H., & García-Fraile, J. A. (2010). *Secuencias didácticas: Aprendizaje y evaluación de competencias*. Pearson.
- Tobón, S. (2017). *Metodología de la elaboración de una rúbrica socioformativa*. CIFE. https://issuu.com/cife/docs/metodologia_rubrica_socioformativa
- Yuan, K.-H. (2005). Fit indices versus test statistics. *Multivariate Behavioral Research*, 40(1), 15-148. https://doi.org/10.1207/s15327906mbr4001_5

