



UPC
Universidad Peruana
de Ciencias Aplicadas

REVISTA DIGITAL DE INVESTIGACIÓN EN DOCENCIA UNIVERSITARIA

2023, 17 (2)

ISSN 2223-2516

Índice

Editorial

César Merino-Soto

<https://orcid.org/0000-0001-7798-5301>

Artículos

1

Validez y confiabilidad de un instrumento para evaluación de la microenseñanza en ciencias químico biológicas

Roberto Cruz González-Mélendez; Martha Asunción Sánchez-Rodríguez; Francisca Robles-López

<https://doi.org/10.19083/ridu.2023.1581>

2

Evaluación del desempeño docente durante la pandemia: opinión de estudiantes universitarios mexicanos

Patricio Henriquez Ritchie

<https://doi.org/10.19083/ridu.2023.1780>

3

Relación de la prueba Saber 11, examen de admisión, promedio académico, prueba saber pro de estudiantes del programa de psicología de la Universidad del Magdalena

Amanda Miguel Iguarán Jiménez; María Fernanda Cabas-Manjarrés; Carmelina Paba Barbosa; Paola Diazgranados Rincones

<https://doi.org/10.19083/ridu.2023.1421>

4

Más breve puede ser mejor: Propiedades métricas de la UWES – 3 ítems para evaluar “engagement” en profesores universitarios mexicanos

Arturo Juárez-García; César Merino-Soto; Daniela Villamar Sánchez

<https://doi.org/10.19083/ridu.2023.1623>

5

Rendimiento académico, valoración del docente e interacción alumno-profesor en la docencia online frente a presencial

Alejandro M. Fernández-Castro; Roberto Sánchez-Cabrero; Yousef Hussein Eiadat

<https://doi.org/10.19083/ridu.2023.1583>

6

Cultura de evaluación para el aprendizaje según estudiantes y docentes de una universidad pública mexicana

Sofía Contreras Roldán; Patricio Henríquez Ritchie

<https://doi.org/10.19083/ridu.2023.1831>



7

Autoevaluación en educación superior: percepción de los estudiantes sobre su utilidad en el proceso de aprendizaje

María Isabel Ramírez Olivares; Katherina Edith Gallardo

<https://doi.org/10.19083/ridu.2023.1574>

8

Invarianza factorial del Maslach Burnout Inventory Student Survey (mbi-ss) en estudiantes universitarios chilenos y españoles

María Pía Valdés Castro; Kevin Rodrigo Mancilla Mancilla; Javiera Ignacia Morales Olivares; Gabriela Belén Acevedo Meza; Ricardo Jorquera Gutiérrez

<https://doi.org/10.19083/ridu.2023.1694>

9

Representaciones semánticas de la bioética por jóvenes universitarias desde la ética del cuidado

Pedro César Cantú-Martínez

<https://doi.org/10.19083/ridu.2023.1592>

10

Principles and assumptions of psychometric measurement

Gavin T. L. Brown

<https://doi.org/10.19083/ridu.2023.1834>

11

Aprendizagem baseada em projetos, portfólio coletivo e nutrição: integrando saberes no Sistema Único de Saúde (SUS)

Renato Pereira da Silva*1; Ângela Maria Campos Santana2; Ana Íris Mendes Coelho3; Ana Vlândia Bandeira Moreira4; Luiza Carla Vidigal de Castro 5; Tiago Rocha Pinto6; Mara Rejane Barroso Barcelos7

<https://doi.org/10.19083/ridu.2023.1193>

12

El COVID-19 y la valoración del proceso enseñanza - aprendizaje en estudiantes de derecho

Constanza Astudillo Meza

<https://doi.org/10.19083/ridu.2023.1279>

Editorial

César Merino-Soto

Universidad de San Martín de Porres, Lima, Perú

Tecnológico de Monterrey, Nuevo León, México

<https://orcid.org/0000-0001-7798-5301>

El contenido del título de este número especial es suficientemente descriptivo de lo que se espera: un enfoque de contenido en instrumentación y prácticas de evaluación educativa y psicológica en el contexto de la educación superior. La evaluación psicoeducativa generalmente es entendida como una acción integral enfocado en el contexto específico de la educación, con el fin último de crear intervenciones empáticas con las necesidades de los estudiantes (Isaacs et al., 2013). Sin embargo, en los procesos de evaluación en la educación superior, también se incluyen aspectos que van desde el proceso de admisión, para continuar con el ciclo de vida estudiantil (Halpern & Butler, 2013). Y temáticamente, la evaluación educativa en general, y en educación superior en particular, no puede entenderse únicamente como un conjunto de manuscritos de desarrollo y adaptación de medidas.

El objeto de estas evaluaciones tradicionalmente es la variabilidad individual de los estudiantes. Habitualmente, estas evaluaciones exploran los atributos cognitivos o no cognitivos que tienen consecuencias con los aprendizajes y rendimiento académico. Pero también, el proceso de evaluación puede dirigirse a otros componentes de la experiencia estudiantil, y por lo tanto puede, y debe, orientarse hacia aspectos del contexto, incluidos los propios procesos de educación y evaluación. Estos pueden definirse ampliamente como factores psicosociales del rendimiento académico, y que implican la interacción de la variabilidad individual con la variabilidad del contexto (Dixson et al., 2016; Tindle et al., 2022). Por lo tanto, en el ciclo de vida del estudiante, la evaluación educativa en educación superior posee un valor incomparable para medir, describir y servir de recurso extremadamente útil en la toma de decisiones. Estas decisiones estarían encaminadas hacia los que directamente se verán afectados (Halpern & Butler, 2013): los estudiantes y sus familias (e.g., obtención o mantención de becas, elección de especialidad, elección de estudios de postgrado) y la institución (e.g., caracterización de aprendizajes logrados, nuevos contratos en relación con la demanda estudiantil, cursos complementarios).

En atención a lo anterior, el presente número contiene trabajos empíricos relacionados con constructos como work engagement, cultura de evaluación, desempeño docente, autoevaluación, cohesión grupal, entre otros. Estos constructos resaltan la visión amplia de este número especial, así como los diferentes intereses de los investigadores para focalizarse en temas importantes para sus contextos institucionales, y regionales. En diferente grado de complejidad y sofisticación de sus análisis, las mediciones hechas en estos estudios también muestran datos de validez, de tal manera que se evita inducir sus propiedades desde estudios anteriores que implicaron otras condiciones de evaluación y otras muestras de participantes (Merino-Soto & Calderón-De la Cruz, 2018; Merino-Soto & Angulo-Ramos, 2021a, 2021b). Este problema, inducción de la validez de medida, ha sido resaltado en Perú (Merino-Soto & Calderón-De la Cruz, 2018; Merino-Soto et al., 2021), pero también en diferentes constructos medidos en población hispana y no hispana (Salas-Blas et al., 2022; Merino-Soto et al., 2023; Rosario-Hernández et al., 2021; Merino-Soto et al., 2022; Merino-Soto et al., 2022; Cejalvo et al., 2021).



En la perspectiva del usuario final, investigador o del gestor educativo, la comprensión de los instrumentos o medidas usadas en el proceso de evaluación psicoeducativa debe involucrar un conocimiento suficiente de conceptos psicométricos. Normas, confiabilidad, puntajes, estandarización y validez, son conceptos habituales en esta área (Isaacs et al., 2013), pero también contienen un potencial transdisciplinar inherente a su uso (Brown, 2022). Estos conceptos son altamente necesarios para comprender el uso y resultados de las evaluaciones en general, pero son especialmente sensibles cuando estos resultados influenciarán en las decisiones del estudiante, y los procesos educativos de la institución. La proliferación de medidas y estudios es un aparente termómetro de la importancia de estudios de validez, y que son consecuencias de las demandas de evaluación y de constructos del contexto (Merino-Soto, et al., 2021). Debido a la aparente inclusión de mediciones no cognitivas en las evaluaciones de educación (Halpern & Butler, 2013), una visión verificadora de las propiedades psicométricas posiblemente nunca puede ser redundante. Por lo tanto, aunque estos conceptos son conocidos por los investigadores, seguramente será siempre necesario “refrescar” este conocimiento en los tomadores de decisiones de las instituciones, y darlos a conocer claramente a los usuarios finales. Por este motivo, el primer manuscrito abre el número especial con los principles and assumptions of psychometric Measurement, una revisión lúcida y generalizable a la evaluación educativa en educación superior.

Referencias

- Brown, G. T. L. (2022). The past, present and future of educational assessment: A transdisciplinary perspective. *Front. Educ.* 7:1060633. <https://doi.org/10.3389/feduc.2022.1060633>
- Cejalvo, E., Martí-Vilar, M., Merino-Soto, C., Aguirre-Morales, M.T. (2021). Caregiving Role and Psychosocial and Individual Factors: A Systematic Review. *Healthcare*, 9, 1690. <https://doi.org/10.3390/healthcare912169>
- Dixson, D.D., Worrell, F.C., Olszewski-Kubilius, P. and Subotnik, R.F. (2016). Beyond perceived ability: the contribution of psychosocial factors to academic performance. *Ann. N.Y. Acad. Sci.*, 1377: 67-77. <https://doi.org/10.1111/nyas.13210>
- Halpern, D. F., & Butler, H. A. (2013). Assessment in higher education: Admissions and outcomes. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J.-I. C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Eds.), *APA handbook of testing and assessment in psychology*, Vol. 3. *Testing and assessment in school psychology and education* (pp. 319–336). American Psychological Association. <https://doi.org/10.1037/14049-015>
- Isaacs, T., Zara, C., Herbert, G., Coombs, S. J., & Smith, C. (2013). *Key concepts in educational assessment*. SAGE Publications Ltd, <https://doi.org/10.4135/9781473915077> (pp. 1-4)
- Merino-Soto, C., Angulo-Ramos, M., Rovira-Millán, L.V., & Rosario-Hernandez, E. (2023). Psychometric properties of the generalized anxiety disorder-7 (GAD-7) in a sample of workers. *Front. Psychiatry* 14:999242. <https://doi.org/10.3389/fpsyt.2023.999242>
- Merino-Soto, C., & Angulo-Ramos, M. (2021a). Estudios métricos del Compliance Questionnaire on Rheumatology (CQR): ¿Un caso de inducción de la validez? *Reumatología clínica*, S1699-258X(21)00093-0. <https://doi.org/10.1016/j.reuma.2021.03.004>
- Merino-Soto, C., & Angulo-Ramos, M. (2021b). Inducción de la validez: comentarios al estudio de validación del Compliance Questionnaire on Rheumatology. *Revista Colombiana de Reumatología*, 28: 312–313.
- Merino-Soto, C., & Calderón-De la Cruz, G. A. (2018). Validez de estudios peruanos sobre estrés y burnout. *Revista Peruana de Medicina Experimental y Salud Pública*, 35, 353–354, <http://dx.doi.org/10.17843/rpmesp.2018.352.3521>

- Merino-Soto, C., Chávez-Ventura, G., López-Fernández, V., Chans, G. M., & Toledano-Toledano, F. (2022). Learning Self-Regulation Questionnaire (SRQ-L): Psychometric and Measurement Invariance Evidence in Peruvian Undergraduate Students. *Sustainability*, 14(18), 11239. MDPI AG. <http://dx.doi.org/10.3390/su141811239>
- Merino-Soto, C., Lozano-Huamán, M., Lima-Mendoza, S., Calderón de la Cruz, G., Juárez-García, A., & Toledano-Toledano, F. (2022). Ultrashort Version of the Utrecht Work Engagement Scale (UWES-3): A Psychometric Assessment. *International journal of environmental research and public health*, 19(2), 890. <https://doi.org/10.3390/ijer-ph19020890>
- Merino-Soto, C., Oakland, T., Dominguez-Lara, S., & Copez-Lonzoy, A. (2021). International status of test development and use: Possible implications for Peru. *Revista Interamericana de Psicología*, 55(1), e964. <https://doi.org/10.30849/ripijp.v55i1.964>
- Rosario-Hernández E, Rovira-Millán LV and Merino-Soto, C. (2021). Review of the Internal Structure, Psychometric Properties, and Measurement Invariance of the Work-Related Rumination Scale – Spanish Version. *Frontiers in Psychology*. 12:774472. <https://doi.org/10.3389/fpsyg.2021.774472>
- Salas-Blas, E., Merino-Soto C, Pérez-Amezcu, B., & Toledano-Toledano, F. (2022) Social Networks Addiction (SNA-6) – Short: Validity of Measurement in Mexican Youths. *Frontiers in Psychology*. 12:774847. <https://doi.org/10.3389/fpsyg.2021.774847>
- Tindle, R., Abo Hamza, E. G., Helal, A. A., Ayoub, A. E. A., & Moustafa, A. A. (2022). A scoping review of the psychosocial correlates of academic performance. *Review of Education*, 10, e3371. <https://doi.org/10.1002/rev3.3371>

Validez y confiabilidad de un instrumento para evaluación de la microenseñanza en ciencias químico biológicas

Roberto Cruz González-Mélendez^{*1}; Martha Asunción Sánchez-Rodríguez²; Francisca Robles-López³

^{1,2,3} Facultad de Estudios Superiores Zaragoza, Universidad Nacional Autónoma de México, Ciudad de México, México ¹ <https://orcid.org/0000-0002-4947-347X> Chichenitza2007@gmail.com ² <https://orcid.org/0000-0002-7130-4074> masanrod@comunidad.unam.mx ³<https://orcid.org/0000-0002-8097-0255> roblesfl@hotmail.com

Citar como: Cruz González-Mélendez, R., Sánchez-Rodríguez, M., & Robles-López, F. (2023). Validez y confiabilidad de un instrumento para evaluación de la microenseñanza en ciencias químico biológicas. *Revista Digital de Investigación en Docencia Universitaria*, 17(2), e1581. <https://doi.org/10.19083/ridu.2023.1581>

Recibido: 25/11/2021. **Revisado:** 05/02/2022. **Aceptado:** 04/04/2023. **Publicado:** 30/06/2023.

Resumen

Introducción: la microenseñanza es una técnica que permite el mejoramiento de las habilidades didácticas del docente con sus estudiantes en el aula, y para evaluar la realización de esta técnica es necesario contar con instrumentos válidos y confiables.

Objetivo: determinar la validez y confiabilidad del Cuestionario de Evaluación de Microenseñanza (CEM) de docentes de la Carrera de QFB en la FES Zaragoza, UNAM. **Método:** el CEM se aplicó a una muestra no probabilística de 124 estudiantes después de una clase de 20 minutos. Resultados: se obtuvo un instrumento de 49 preguntas distribuidas en tres factores con una varianza total de 51.57%, según el Análisis Factorial Exploratorio, y alfa de Cronbach de 0.961. **Discusión:** se considera que la validez y la confiabilidad obtenidas hacen que el CEM sea un instrumento útil para que los estudiantes puedan evaluar las habilidades didácticas de los docentes en la carrera de QFB.

Palabras clave: validez; confiabilidad; microenseñanza; habilidad didáctica; análisis factorial exploratorio

Validity and reliability of an instrument for assessing microteaching in chemical biological sciences

Abstract

Introduction: Microteaching is a technique that allows the improvement of teachers' didactic skills with their students in the classroom, and in order to evaluate the performance of this technique it is necessary to have valid and reliable instruments. **Objective:** To determine the validity and reliability of the Microteaching Assessment Questionnaire (MAQ) of teachers of the QFB Career at the FES Zaragoza, UNAM. **Method:** The MAQ was applied to a non-probabilistic sample of 124 students after a 20-minute class. Results. An instrument of 49 questions distributed in three factors was obtained with a total variance of 51.57%, according to the Exploratory Factor Analysis, and Cronbach's alpha of 0.961. **Discussion:** It is considered that the validity and reliability obtained make the MAQ a useful instrument for students to evaluate the teaching skills of teachers in the QFB career.

Keywords: validity; reliability; microteaching; didactic ability; exploratory factor analysis

*Correspondencia:

Roberto Cruz González-Mélendez
roberto.sanchez@uam.es



Introducción

La microenseñanza es una técnica para el entrenamiento y perfeccionamiento de las habilidades pedagógicas, se centra en la manera en cómo los maestros enseñan (técnica) y no en el que se enseña (contenidos) ([Escoto, 2001](#)). Consiste en simplificar las complejidades de una clase normal, disminuyendo el tiempo de exposición, la extensión del contenido y el tiempo de duración durante el entrenamiento ([Escoto, 2001](#)). Esta técnica tuvo sus inicios en la Universidad de Stanford en 1963 y con algunas adaptaciones se ha difundido internacionalmente ([Dwight, 1967](#)). Fue creada para abatir muchas de las deficiencias de los programas tradicionales de la educación de maestros, y para acrecentar el entendimiento del proceso que se establece entre el profesor y la enseñanza. ("[Secretaría de Educación Pública, 2016](#)"). Es un encuentro de enseñanza reducido que se ha desarrollado para servir a 3 propósitos: (1) como experiencia preliminar y práctica en la enseñanza, (2) como un vehículo de investigación para explorar los efectos del entrenamiento bajo condiciones controladas, y (3) como un instrumento de capacitación en servicio para maestros experimentados. Se les exige que enseñen lecciones breves (5 a 25 minutos) en su materia, a un grupo pequeño de estudiantes (hasta 5). Estas breves lecciones brindan la oportunidad de una supervisión intensa, grabación de video para comentarios inmediatos y la recopilación y utilización de comentarios de los participantes ([Dwight, 1967](#)).

Para la recopilación de información de los estudiantes se han diseñado instrumentos o cuestionarios para evaluar la técnica de la microenseñanza. Entre ellos existe gran diversidad de modelos, habilidades didácticas empleadas, número de preguntas, población aplicada y diseño para la validación de cada uno de éstos, aplicados en diferentes áreas.

Al respecto, [Dwight \(1967\)](#), elaboró un cuestionario de 35 preguntas para los estudiantes universitarios (docentes de recién ingreso) dividido en 7 habilidades de enseñanza (reforzamiento, variación de estímulos, inducción, lectura y uso de medios audiovisuales, habilidad de explicación y uso

de ejemplos y preguntas iniciadas por los alumnos), en donde cada pregunta tiene siete niveles de acuerdo con una escala de Likert. No se establece una validez y confiabilidad, solo se menciona que se obtuvieron resultados estadísticamente significativos ($p= 0.001$). [Guillermo \(1997\)](#), diseñó un instrumento de 24 preguntas aplicado a docentes de la licenciatura en docencia, dividido en tres habilidades didácticas (reforzamiento, formulación de preguntas e integración), considerando 5 niveles de calificación de acuerdo con la escala de Likert, sin mencionar confiabilidad ni validez del instrumento empleado.

Por otro lado, [Capa en 2005](#), validó la versión turca del cuestionario *Teachers's Sense of Efficacy Scale* aplicado a docentes en formación en educación, que está conformado por 24 preguntas, dividido en 3 dimensiones (eficacia en la participación de estudiantes, eficacia en las estrategias de instrucción y eficacia en la gestión del aula), con 9 niveles de calificación de acuerdo con la escala de Likert. Obtuvo una alfa de Cronbach de 0.82, 0.86 y 0.84 para cada una de las habilidades y una confiabilidad del instrumento de 0.93, empleando análisis factorial exploratorio.

El cuestionario de [Watson \(2007\)](#), cuenta con 28 preguntas para docentes nuevos de la Universidad de Ciencias Aplicadas, organizadas en 4 competencias (diseña y desarrolla secuencias de clase que promueven un aprendizaje significativo, creativo y activo; elabora y utiliza medios y materiales de enseñanza y aprendizaje, haciendo uso efectivo de la tecnología educativa y de las TIC; se comunica de forma efectiva involucrando a los alumnos en el tema desarrollado y mantiene una buena relación con los alumnos durante el proceso de enseñanza y aprendizaje) y Deniz (2011), empleó un cuestionario para profesores en formación de un Programa de Desarrollo Infantil, que mide las opiniones sobre las habilidades de enseñanza, considerando 5 habilidades didácticas (planeación y preparación para la lección, gestión del aula y relación con los estudiantes, empleo de métodos de enseñanza, implementación del proceso de enseñanza y evaluación), en 31 preguntas, con cinco niveles de calificaciones de acuerdo a la escala de Likert con una confiabi-

lidad alfa del cuestionario del preexamen de .94 y post examen .91.

Otro instrumento reportado es el de las [Unidades Tecnológicas Santander \(2011\)](#), cuyo cuestionario aplicado a profesores, está conformado por 31 preguntas, dividido en 7 habilidades didácticas (inicio y motivación, comunicación verbal y no verbal, variación del estímulo, formulación de preguntas, integración de saberes, organización lógica y cierre de clase (refuerzo)), con 5 niveles de calificación en escala de Likert, aunque no se especifica la validez ni confiabilidad del instrumento. Se reportan dos cuestionarios más, el de [Remesh \(2013\)](#) que consta de 39 preguntas aplicado a profesores principiantes en la Carrera de Medicina, muy parecido al de la Universidad de Stanford ([Dwight, 1967](#)), que está dividido en 8 habilidades de enseñanza (planeación de clases, presentación y explicación, ilustrar con ejemplos, reforzamiento, variación del estímulo, preguntas de sondeo o exploración, manejo de clase y uso de medios audiovisuales), sin mencionar análisis de validez y confiabilidad; y la otra propuesta es de [Toro \(2013\)](#), que consta de 19 preguntas aplicado a estudiantes de la Especialidad en Tecnología del Vestido, dividido en tres dimensiones (planificación, conducción y evaluación), 5 niveles de calificación en una escala de Likert, con una validez por juicio de expertos de 88% y coeficiente de confiabilidad de Cronbach del instrumento de 0.96.

Recientemente [Dayanindhi \(2018\)](#), empleó un instrumento similar al utilizado en la Universidad de Stanford ([Dwight, 1967](#)), aplicado a profesores médicos en servicio, contiene seis dominios (inducción, planeación, presentación, uso de medios audiovisuales, participación de los alumnos y cierre), con 19 preguntas, aunque no se menciona la validez y confiabilidad del cuestionario. Con base en lo anterior se puede observar que la mayoría de los instrumentos no tienen evidencias de validez y confiabilidad y fueron diseñados exclusivamente para uso local.

Por otro lado, la Carrera de Química Farmacéutico Biológica (Q.F.B.) de la Facultad de Estudios Superiores (FES) Zaragoza de la Universidad Nacional Autónoma de México (UNAM) ha realizado

por varios años procesos de evaluación de la docencia, y ha implementado diferentes estrategias con la finalidad de fortalecer estas evaluaciones, como proponer cursos de actualización docente, cursos de formación pedagógica, estímulos académicos, entre otras acciones para mejorar el desempeño académico de los docentes en el aula de clases. Dentro de estas estrategias se ha propuesto la implementación de la técnica de la microenseñanza como una alternativa práctica que permita mejorar las habilidades didácticas y pedagógicas de los docentes de la carrera. Para conocer si se consigue lo anterior es necesario contar con un instrumento válido y confiable que permita evaluar las mejoras en las habilidades didácticas que presentan los profesores al aplicar la técnica de la microenseñanza, mediante la opinión de los estudiantes.

Es importante señalar que en el proceso de evaluación de la docencia se aplica un cuestionario de 22 preguntas de las cuales 9 consideran las habilidades didácticas de los docentes (propicia el interés de los alumnos por los temas, promueve la participación de los alumnos, fomenta el uso de las tecnologías de la información y comunicación (TIC), prepara las clases, relaciona la teoría con la práctica, asesora al alumno en las actividades propuestas, propicia el trabajo colaborativo entre los alumnos, genera un ambiente de confianza durante las actividades académicas y domina los contenidos del programa) y son lo que más se acerca a dichas habilidades en la técnica de la microenseñanza, pero no evalúan todos los aspectos que la conforman ([Facultad de Estudios Superiores Zaragoza, 2015](#)).

Con base en lo anterior, se consideró importante contar con un instrumento de evaluación de la microenseñanza válido y confiable que tome en cuenta la opinión de los estudiantes como parte sustancial de los procesos de enseñanza y aprendizaje, y que además sea adecuado y pertinente para ser aplicado en la carrera de Q.F.B. de la FES Zaragoza, UNAM.

Un aspecto muy importante a considerar es la validez de acuerdo en la evidencia basada en las consecuencias de las pruebas, a este respecto Martínez Rizo et al., 2015 ([citado en Taut, 2016](#)) in-

dican un conjunto de criterios de buena práctica en la validación, entre los que se pueden mencionar los siguientes: Contar con argumentos lógicos y evidencia empírica que respalde los usos y consecuencias previstas; evaluar el grado en que se producen las consecuencias previstas y/o deseables de la prueba; la difusión de los resultados en tiempos razonables, sin discriminación; informar a los usuarios sobre los propósitos y características de las pruebas; utilizar lenguaje claro y preciso sin tecnicismos; describir el perfil y características de la población de referencia; Identificar los usos o interpretaciones inadecuadas de la prueba e investigar en detalle cuando exista evidencia confiable de usos inapropiados.

Método

Diseño.

Este estudio empleo un muestreo no probabilístico por oportunidad, en donde la elección de los elementos no depende de la probabilidad sino de las características de la investigación. ([Hernández, 2010](#)).

Participantes

En este estudio participaron cinco docentes con estudios de posgrado en educación o en educación superior, con más de 20 años de experiencia en el área educativa en las carreras de la FES Zaragoza. La muestra de estudiantes fue 124, de diferentes semestres de la carrera de Q.F.B. de la FES Zaragoza UNAM (con un rango de edad entre los 18 y 24 años, con una participación del sexo femenino del 51.22% y de hombres del 48.78%, y con un nivel socioeconómico de un ingreso menor a 4 salarios mínimos). Los criterios de selección fueron: a) inscritos oficialmente en el semestre que cursaban, 2) que quisieran participar en el proyecto. El 17.7% de los estudiantes cursaban el tercer semestre, el 30.6% el quinto semestre, el 5.6% el octavo semestre, el 26.6% el noveno semestre y el 19.4% el décimo semestre. Así mismo, el 54.8% de los estudiantes se encontraban en clases teóricas (la cual consiste en la

presentación de un tema, lógicamente estructurado, en donde el recurso principal es el lenguaje oral, en donde la información debe activar los conocimientos previos, presentar los contenidos en forma activa y estructurar los contenidos ([Videla, 2005](#)) y el 45.2% en prácticas de laboratorio (la cual tiene como objetivo que los estudiantes adquieran las habilidades propias de los métodos y técnicas de trabajo y de la investigación científica; amplíen, profundicen, consoliden, generalicen y comprueben los fundamentos teóricos de la disciplina mediante la experimentación, empleando para ellos los medios necesarios ([Rodríguez, 2017](#))).

Instrumentos

Cuestionario de Evaluación de la Microenseñanza (CEM)

Para el diseño y estructuración de las habilidades didácticas y sus niveles de calificación del CEM se consultaron a autores como: [Dwight, A \(1967\)](#); [Guillermo, 1997](#); [Capa, \(2005\)](#); [Watson, \(2007\)](#); [Deniz, \(2011\)](#); [Unidades Tecnológicas de Santander \(2011\)](#); [Remesh, \(2013\)](#); [Toro, \(2013\)](#); [Dayanindhi, \(2018\)](#). Así mismo, se consideraron las nueve preguntas de la habilidad didáctica contenida en el Cuestionario de Evaluación de la Docencia aplicados en la Carrera de QFB, en la FES Zaragoza de la UNAM.

Con base en lo anterior, el CEM contempló cuatro habilidades didácticas para la evaluación de la microenseñanza, con diferentes números de preguntas, considerando 5 niveles de calificación para cada una de las preguntas (0=nunca, 1=casi nunca, 2=algunas veces, 3=frecuentemente, 4=siempre). Las definiciones de cada una de las habilidades didácticas de acuerdo con Escoto (2001), se describen en la tabla No. 1.

Instrumento de Opinión sobre el Cuestionario de Evaluación de la Microenseñanza (IOCEM).

Con la finalidad de determinar si cada una de las preguntas incluidas, así como los apartados fueron entendibles para los alumnos, se diseñó el IOCEM, con base en las preguntas de cada una de las habilidades didácticas incluidas en el CEM y la experiencia de los investigadores.

Tabla 1.
Definición de las habilidades didácticas consideradas en el CEM.

Habilidad	Definición	No. de preguntas
A) Presentación	Son aquellos comportamientos que establecen un eslabón entre la actividad del docente y la del estudiante; incluye comportamientos del profesor tales como: presentar los objetivos del tema, unidad o programa a los alumnos para dirigir su atención; mostrar a los alumnos los contenidos y actividades en forma atractiva; enlazar los contenidos con lo desconocido; presentar ordenada, sistemáticamente y con continuidad los contenidos de la clase; ajustar los recursos didácticos a los objetivos	7
B) Comunicación verbal y no verbal	Se refiere a la habilidades para transmitir información, con la finalidad de controlar, motivar, expresar emociones e informar; incluye comportamientos del docente como: utilizar un vocabulario preciso y correcto, adaptado a la capacidad del receptor; reestructurar el tema de acuerdo con las preguntas que se le formulen; partir de lo conocido a lo nuevo; favorecer la continuidad de las ideas; modular y regular el volumen y tono de voz; pronunciar claramente; controlar la velocidad del lenguaje; manejar las emociones al comunicar el mensaje; usar el enfoque verbal; eliminar el uso de muletillas; apoyar la exposición con gesticulación; mantener contacto visual con los alumnos y estimular la comunicación no verbal	18
C) Formulación de preguntas	Consisten en que el docente se organiza con la finalidad de involucrar activamente al estudiante y estimular o desencadenarles procesos mentales superiores, incluyendo conductas como: formular preguntas centradas en las ideas principales; favorecer el pensamiento convergente o divergente mediante las preguntas; formular preguntas que se involucren en la clase, tiempo para dar asesoría en la parte práctica facilitando el desarrollo de habilidades y destrezas, procurando que desarrollen las prácticas en condiciones adecuadas, resolviendo problemas prácticos con ejemplos o demostraciones y ayudándolos a comprender conceptos propios de la disciplina aplicados a la práctica	10
D) Reforzamiento	Favorecen que el estudiante pueda consolidar sus conocimientos e incluye conductas como: consolidar conceptos e ideas principales antes de iniciar un nuevo tema; volver a los puntos principales con el objeto de clasificarlos; resumir y contemplar las ideas principales tras el análisis y discusión con los estudiantes, iniciando las clases reafirmando los contenidos que son un prerrequisito para el nuevo tema; elaborar esquemas y cuadros sinópticos que condensen la información que cubre los objetivos del tema; propiciar la aplicación de los conocimientos a situaciones nuevas y analizar con los alumnos las posibilidades de uso de la información.	16

Procedimientos

El primer paso fue crear la versión uno del CEM, con base en los cuestionario de microenseñanza de diversos autores y los nueve preguntas de la habilidad didáctica contenida en el Cuestionario de Evaluación del Estudiante al Ejercicio Docente aplicado en la carrera QFB de la FES Zaragoza. El segundo paso fue llevar a cabo la validación del contenido por cinco profesores expertos en educación superior, quienes analizaron y emitieron sugerencias y observaciones para mejorar el CEM en contenido, estructura, congruencia, redacción y niveles de calificación, dando origen a la versión dos del CEM. El tercer paso fue que con base en esta versión y con la experiencia de los investigadores, se propusieron preguntas para conformar el IOCEM. Después de un análisis de cada propuesta se obtuvo el cuestionario final con 8 preguntas, que permitió recabar la opinión de los estudiantes sobre la claridad y entendimiento de la introducción, objetivos, escala de calificación, y redacción, así como, la extensión y utilidad del CEM.

El cuarto paso fue realizar la validación del constructo y confiabilidad del CEM, para ello se extendió una invitación por escrito a varios profesores de la carrera de Q.F.B. a quienes se les explicó el propósito, objetivo del estudio y las habilidades didácticas que se evaluarían con base en el CEM, solicitándoles su autorización para asistir a una clase de 20 minutos, y posteriormente llevar a cabo la aplicación del CEM y del IOCEM a sus alumnos. En el aula se explicó a los estudiantes el propósito del estudio, así como, el objetivo de su participación al evaluar la técnica de la microenseñanza del docente mediante la aplicación del CEM, y el objetivo del IOCEM. Es importante mencionar que la aplicación de los cuestionarios fue de forma anónima para evitar cualquier tipo de coerción por parte de los docentes evaluados, por lo que a los estudiantes participantes se les asignó un número conforme fueron entregando los cuestionarios resueltos. Cabe señalar que el compromiso con los profesores fue hacerles entrega por escrito de los resultados obtenidos de la evaluación de sus alumnos.

El quinto paso fue el tratamiento estadístico de los datos mediante el Análisis Factorial Exploratorio, Alpha de Cronbach, dando origen al ver-

sión tres y final del CEM, y análisis de frecuencias para el IOCEM

Es importante mencionar que cada uno de los profesores participantes firmaron un consentimiento informado, en donde se dieron a conocer el objetivo y propósito de la participación de cada uno de ellos en el proyecto, la solicitud para asistir a una clase de 20 minutos, el compromiso de mantener la confidencialidad de la información derivada de la investigación, así como el aviso de privacidad de los datos personales.

Análisis de datos

Para determinar la validez del contenido del CEM se hizo mediante el juicio de expertos (face validity) con un análisis de los comentarios sobre los apartados del instrumento, las preguntas, ubicación de estas en cada habilidad didáctica, la escala de calificación, redacción y claridad del cuestionario. Para determinar la confiabilidad del CEM se utilizó el alfa de Cronbach. Para la validez del constructo se empleó Análisis Factorial Exploratorio mediante el programa SPSS V21 (IMB SPSS statistics Armomk, NY, EU). Para el IOCEM se realizó mediante análisis de frecuencias.

De acuerdo a [Hernández, 2010](#), el Alpha de Cronbach se interpreta con los siguientes valores: 0.25 en la correlación o coeficiente, esto indica baja confiabilidad; si el resultado es 0.50, la fiabilidad es media o regular. En cambio, si supera el 0.75 es aceptable, y si es mayor a 0.90 es elevada.

Para el AFE se realizó siguiendo cuatro fases: la primera fue determinar la matriz de correlación mediante la prueba de esfericidad de Barlett, que indica que los datos de todas las variables (preguntas) se correlacionan entre sí, y la media de la adecuación de la muestra Kaiser-Meyer-Olkin (KMO) ([Aguilar 2003](#)). Kaiser y Rice (1974) ([citado en Ferrando, 2022](#)) propusieron el índice KMO que establece que, a mayor valor del índice, más comunalidad disponible en la matriz de correlación. Se han sugerido toda una serie de valores para cualificar la cantidad de la comunalidad disponible; solo matrices con valores de KMO superiores a .75 merecen ser estudiadas mediante Análisis Factorial. La fase dos fue la estimación de factores la cual se realizó mediante un Análisis Factorial Exploratorio (AFE) utilizando el méto-

do de factorización de eje principal. La siguiente fase fue la extracción de factores, empleando para ello el Análisis Paralelo (AP) (Lloret-Segura, 2014), además, se consideró el “criterio de contraste de caída” (Hair, 1999) o “gráfico de sedimentación” (López, 2015), tomando el número de factores a extraer determinado por el cambio en la continuidad en la curva. La siguiente etapa fue la rotación de factores, empleando una rotación oblicua, mediante el método Promax, empleando como valor de saturación de los ítems de 0.35, lo que representa la correlación entre la variable original y su factor (Spector, 1992).

En la última etapa se realizó por parte de los investigadores, un análisis conceptual sobre la agrupación y congruencia de la variable para definir el significado de cada uno de los factores o habilidades didácticas, posteriormente, se reorganizaron y renombraron las habilidades didácticas, considerando que definieran al constructo de la microenseñanza.(Morales, 2011).

Resultados

Validez

Para la validación de contenido, las observaciones realizadas a la versión 1 del CEM por parte de los docentes expertos, 7 de las preguntas perma-

necieron sin cambio (15.2%), 38 preguntas (82.6%) presentaron cambios en su redacción principalmente en la conjugación de verbos, en algunas de estas preguntas se cambió la redacción para hacerlas más entendibles y sólo una pregunta (2.1.%) fue eliminada. Así mismo, fueron añadidas 6 preguntas nuevas: 4 en la habilidad de comunicación verbal y no verbal y 2 en la habilidad de formulación de preguntas. Se cambiaron los niveles de calificación de las preguntas por: 0=muy deficiente, 1= deficiente, 2= regular, 3=bueno y 4=muy bueno. Con base en lo anterior, se conformó la versión 2 del CEM con 51 preguntas, con 4 habilidades didácticas, como se observa en la tabla 2.

Para la validación de la estructura interna se aplicó el CEM y simultáneamente el IOCEM, a 6 grupos (3 grupos de actividades prácticas y de 3 de actividades teóricas) con una participación de 124 alumnos.

Los resultados del AFE del CEM son los siguientes:

- 1. En la matriz de correlación se observó una distribución normal, la prueba de esfericidad de Barlett indicó que todas las variables se correlacionan entre sí (p< 0.0001), y el KMO cumplió con el criterio para realizar un AFE con un valor de 0.897.
- 2. La estimación de factores realizado por el método de factorización de eje principal se

Tabla 2. Comparativo de las versiones del CEM.

Versión	Habilidades didácticas	Número de preguntas	Total de preguntas del CEM
1 (inicial)	Presentación	7	46
	Comunicación verbal y no verbal	14	
	Formulación de preguntas	9	
	Reforzamiento	16	
2	Presentación	7	51
	Comunicación verbal y no verbal	18	
	Formulación de preguntas	10	
	Reforzamiento	16	
3 (final)	Comunicación verbal y no verbal	29	49
	Formulación de preguntas	15	
	Empleo de medios audiovisuales y	5	
	Tics		

- observó una distribución de las preguntas en 11 factores.
3. Para la extracción de factores se empleó el Análisis Paralelo (macro de SPSS, ([Meneses, 2020](#))) y el gráfico de sedimentación (fig. 1), dando como resultados la distribución de las preguntas en tres factores. En la tabla se muestra que la varianza acumulada para los tres factores fue del 51.57% (tabla 3)
 4. Los resultados de la rotación oblicua mediante el método Promax, se observa en la matriz patrón (tabla 4).

En la tabla 4 se muestra el resultado de la rotación oblicua de la matriz patrón de tres factores para determinar el número de variables (preguntas) para cada factor. Las preguntas “Utilizó gestos para conducir al grupo /gesticulación del rostro” y “utilizó pizarrón” fueron eliminadas debidos a que tuvieron cargas factoriales menores 0.35.

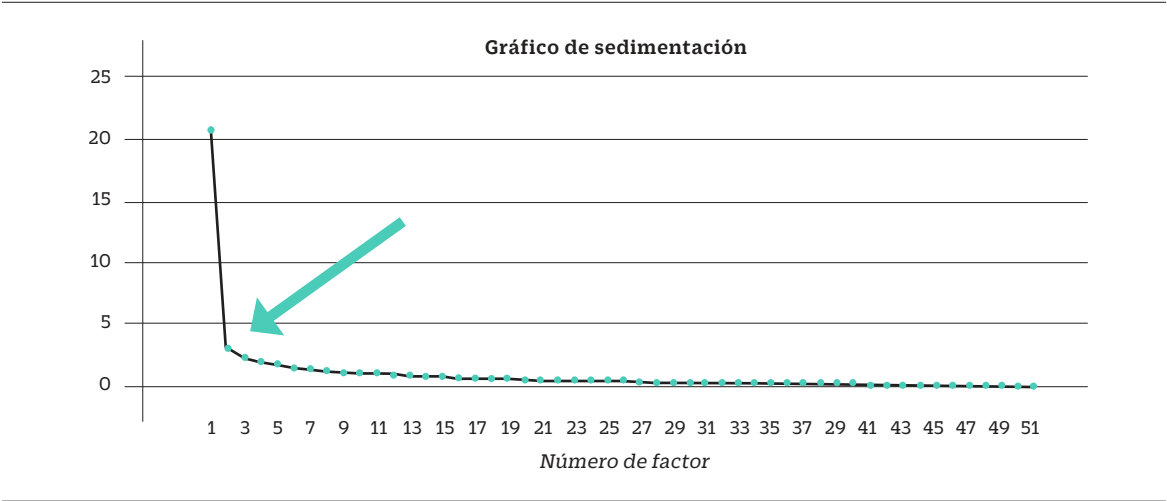
Con base en los resultados del AFE, y la experiencia de los investigadores se reagruparon y nombraron los factores o habilidades didácticas, dando origen al versión 3 de CEM con 49 preguntas, distribuidas en 3 factores: Factor 1 se denominó “Comunicación verbal y no verbal con 29 preguntas (40.83% de varianza); Factor 2 de acuerdo a los ítems que lo integraron se llamó “Formulación de preguntas y reforzamiento” integrado por 15 preguntas (6.05% de varianza) y al factor 3 de acuerdo a los preguntas que los integraron se denominó “Empleo de medios audiovisuales y las TIC” con 5 preguntas (4.67 de varianza). Lo anterior dio origen a la versión 3 y final del CEM. (Tabla 5).

El tiempo que tardaron en contestar los alumnos el CEM y el IOCEM fue en promedio de 28 minutos.

Confiabilidad

El valor de la confiabilidad del alfa de Cronbach para el CEM fue de 0.961.

Figura 1. Gráfico de sedimentación.



Nota: La flecha muestra el punto de inflexión indicando el número de factores a extraer

Tabla 3. Varianza total explicada para tres factores.

Factor	Autovalores iniciales		
	Total	% de varianza	% acumulado
1	20.826	40.835	40.835
2	3.090	6.059	46.894
3	2.385	4.676	51.570

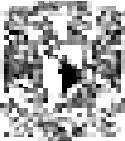
Método de extracción: factorización de eje principal

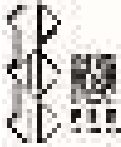
Tabla 4.
Matriz patrón de los tres factores con cargas factoriales

Variable	Matriz de patrón		
	Factor 1	Factor 2	Factor 3
p6. Mostró seguridad y dominio de los temas impartidos en clase	.812		
p13. Mostró postura adecuada en el aula	.788		
p48. Analizó ejemplos del tema estudiado	.765		
p40. Evitó expresiones faciales de molestia o agresividad ante conductas inapropiadas de los alumnos	.756		
p41. Hizo referencias a respuestas o intervenciones de los alumnos	.718		
p9. Reguló el tono de voz	.714		
p45. Enfatizó los temas importantes del contenido	.698		
p5. Presentó organización, limpieza y calidad de los materiales didácticos	.634		
p44. Evitó sentir que una pregunta o participación fuera tonta	.628		
p17. Dirigió su atención a los alumnos	.615		
p42. Asesoró a los alumnos en actividades propuestas	.610		
p36. Motivó la participación de los alumnos con las palabras “bien”, “excelente”, “buena pregunta” o “buena respuesta”	.578		
p12. Se desplazó en el aula	.576		
p10. Utilizó lenguaje claro (sin tartamudeos, ni titubeos, ni muletillas)	.574		
p14. Presentó movimiento de manos, cabeza y cuerpo, sin tics nerviosos	.561		
p8. Reguló el volumen de voz	.553		
p4. Brindó una secuencia lógica de contenido de la clase	.550		
p21. Se comportó amable y respetuosa en clase	.539		
p37. Animó a los alumnos a participar utilizando expresiones como “mm”, “sí”, “adelante”, “perfecto”, entre otras.	.525		
p3. Proporcionó los temas y objetivos relativos al programa	.519		
p49. Relacionó el tema con otros temas vistos con anterioridad o que se tratarán a futuro	.495		
p43. Dio oportunidad a que los alumnos ejemplificaran el concepto estudiado	.468		
p39.Utilizó expresiones como “piense”. “repita de nuevo”, entre otros, para ayudar a los alumnos a emitir sus respuestas correctas.	.448		
p2. Presentación personal (aseo, vestimenta, entre otros)	.430		
p51. Anunció los temas que se estudiarán en la próxima clase	.428		
p46. Sugirió cómo aplicar lo visto en clase	.426		
p35. Permitió que los alumnos formularan preguntas	.395		
p1. Motivó a los alumnos a iniciar la clase	.386		
p38. Motivó a los alumnos a participar mediante sonrisas, afirmaciones con la cabeza, escribiendo sus aportaciones en el pizarrón, mirando y escuchando amablemente.	.369		
p28. Realizó preguntas claras sobre el tema de clase		1.009	
p26. Realizó preguntas relacionadas al tema revidado en clase		.985	
p27. Realizó preguntas relevantes del tema		.984	
p29. Realizó preguntas con un orden lógico y coherente		.970	
p20. Interactuó visual y verbalmente con los alumnos		.788	
p19. Empleó pausas para obtener las respuestas a las preguntas		.746	
p31. Hizo preguntas que motivaron la participación del grupo		.623	
p18. Realizó contacto visual con los alumnos		.555	
p30. Hizo preguntas con diferentes niveles de complejidad y dificultad		.521	
p34. Motivó a los alumnos a formular preguntas		.489	
p32. Aprovechó la pregunta de un alumno para el análisis de nuevas ideas		.482	
p7. Propició confianza, trabajo colaborativo y cooperativo entre los alumnos		.428	
p33. Ayudó a los alumnos a estructurar sus respuestas correctamente		.398	
p50. Relacionó la teoría con la práctica		.383	
p11. Presentó velocidad moderada al hablar		.370	
p24. Utilizó tecnologías de información y de la comunicación			.901
p23. Utilizó medios audiovisuales			.882
p47. Realizó reforzamiento de aspectos más importantes del tema mediante cuadros sinópticos, mapas conceptuales, mapas mentales, entre otros			.623
p25. Usó medios gráficos como rotafolios, esquemas, maquetas, carteles. Modelos a escala, entre otros.			.622
p16. Enfocó su atención a objetos a objetos			.367

Método de extracción: factorización de eje principal.
Método de rotación: Promax con normalización Kaiser.^a
a. La rotación ha convergido en 5 iteraciones.

Tabla 5.
Cuestionarios de Evaluación de la Microenseñanza





Universidad Nacional Autónoma de México

Facultad de Estudios Superiores Zaragoza

Carrera de Química Farmacéutico Biológica

CUESTIONARIO PARA EVALUACIÓN DE LA MICROENSEÑANZA

Introducción

La microenseñanza, como una técnica de entrenamiento para profesores, tuvo sus inicios en la Universidad de Stanford, en 1963. La microenseñanza consiste básicamente en simplificar las condiciones de un salón de clases permitiendo que un profesor instruya durante corto tiempo a un reducido grupo de alumnos, tratando de practicar una habilidad específica al enseñar. Con esta técnica se abordan diversos aspectos del proceso pedagógico de los docentes, incluyendo la comunicación en el aula, el apoyo de recursos didácticos y la diversidad de los alumnos. Estos aspectos se trabajan con la autoobservación y la auto reflexión del profesor sobre su labor docente, recibe retroalimentación de un facilitador y de sus compañeros, desarrolla destrezas y habilidades durante la experiencia, logra motivación para una capacitación y perfeccionamiento permanente y crea lazos de apoyo entre docentes, alumnos y personal administrativo.

La microenseñanza es en la actualidad, uno de los mecanismos de formación de docentes más eficaz que existe, ya que los profesores son capaces de identificar e intervenir problemas en sus prácticas, de crear redes de apoyo en sus planteles para elevar el nivel de aprovechamiento de sus estudiantes.

Objetivo.

Evaluar las habilidades pedagógicas y didácticas de los profesores de Q.F.B con el fin de mejorar la práctica docente y elevar el nivel de aprovechamiento de los estudiantes.

Instrucciones de llenado del cuestionario.

El cuestionario consta de tres columnas: la primera describe la habilidad pedagógica y/o didáctica a evaluar de cada profesor; la segunda columna describe los atributos que se tienen que evaluar de acuerdo con cada habilidad; y la tercera columna indica la calificación que da el alumno a cada atributo del profesor, de acuerdo con la siguiente escala:

1	Calificación	2	Interpretación
3	0	4	Muy deficiente
5	1	6	Deficiente
7	2	8	Regular
9	3	10	Bueno
11	4	12	Muy bueno

Con base en lo anterior, y después de la exposición de una clase de 20 minutos por parte del profesor, el alumno colocará en el cuestionario una  en la calificación (0-4) que a su consideración describa el atributo del profesor.

Cuestionario

Nombre del profesor evaluado: _____ Fecha: _____

Habilidad pedagógica y/o didáctica	No. de atributo	Atributos del profesor a evaluar	CALIFICACION				
			0	1	2	3	4
Comunicación Verbal y no verbal	1	Mostró seguridad y dominio de los temas impartidos en clase					
	2	Mostró postura adecuada en el aula					
	3	Analizó ejemplos del tema estudiado					
	4	Evitó expresiones faciales de molestia o agresividad ante conductas inapropiadas de los alumnos					
	5	Hizo referencias a respuestas o intervenciones de los alumnos					
	6	Reguló el tono de voz					
	7	Enfatizó los temas importantes del contenido					
	8	Presentó organización, limpieza y calidad de los materiales didácticos					
	9	Evitó sentir que una pregunta o participación fuera tonta					
	10	Dirigió su atención a los alumnos					
	11	Asesoró a los alumnos en las actividades propuestas					
	12	Motivó la participación de los alumnos con las palabras “bien”, “excelente”, “buena pregunta” o “buena respuesta”					
	13	Se desplazó en el aula					
	14	Utilizó lenguaje claro (sin tartamudeos, ni titubeos, ni muletillas)					
	15	Presentó movimiento de manos, cabeza y cuerpo, sin tics nerviosos					
	16	Reguló el volumen de voz					
	17	Brindó una secuencia lógica de los contenidos de la clase					
	18	Se comportó de manera amable y respetuosa en clase					
	19	Animó a los alumnos a participar utilizando expresiones como “mm”, “sí” “adelante”, “perfecto”, etc					
	20	Proporcionó los temas y objetivos relativos al programa					
	21	Relacionó el tema con otros temas vistos con anterioridad o que se tratarán a futuro.					
	22	Dio oportunidad a que los alumnos ejemplificaran el concepto estudiado					
	23	Utilizó expresiones como “piense”, “repita de nuevo”, etc., para ayudar a los alumnos a emitir sus respuestas correctas					
	24	Presentación personal adecuada (aseo, vestimenta, etc.)					
	25	Anunció los temas que se estudiarían en la próxima clase.					
	26	Sugirió cómo aplicar lo visto en clase					
	27	Permitió que los alumnos formularan preguntas					
	28	Motivó a los alumnos al iniciar la clase					
	29	Motivó a los alumnos a participar mediante sonrisas, afirmaciones con la cabeza, escribiendo sus aportaciones en el pizarrón, mirando y escuchando amablemente					

Formulación de preguntas y reforzamiento	30	Realizó preguntas claras sobre el tema de clase.					
	31	Realizó preguntas relacionadas con el tema revisado en clase.					
	32	Realizó preguntas relevantes del tema					
	33	Realizó preguntas con un orden lógico y coherente					
	34	Interactuó visual y verbalmente con los alumnos.					
	35	Empleó pausas para obtener las respuestas a las preguntas.					
	36	Hizo preguntas que motivaron la participación del grupo					
	37	Realizó contacto visual con los alumnos					
	38	Hizo preguntas con diferentes niveles de complejidad y dificultad					
	39	Motivó a los alumnos a formular preguntas					
	40	Aprovechó la pregunta de un alumno para el análisis de nuevas ideas					
	41	Propició confianza, trabajo colaborativo y cooperativo entre los alumnos					
	42	Ayudó a los alumnos a estructurar sus respuestas correctamente					
	43	Relacionó la teoría con la práctica					
	44	Presentó velocidad moderada al hablar					
Empleo de medios audiovisuales y TIC	45	Utilizó tecnologías de la información y de la comunicación					
	46	Utilizó medios audiovisuales					
	47	Realizó un reforzamiento de los aspectos más importantes del tema, mediante cuadros sinópticos, mapas conceptuales, mapas mentales, etc.					
	48	Usó medios gráficos como: rotafolios, esquemas, maquetas, carteles, modelos a escala, etc.					
	49	Enfocó su atención principalmente a objetos y no en los alumnos					

Agradecimiento

Agradezco su participación, por contestar el presente cuestionario, sus opiniones redundarán en mejorar los procesos de enseñanza y aprendizaje en el aula, así como, tener profesores más competentes para formar mejores alumnos.

Atte.
M en C. Roberto Cruz González Meléndez
Responsable del Proyecto.

En relación con la claridad y entendimiento del IOCEM (preguntas 1, 2, 3 y 5) se obtuvo una aceptación superior al 85%, (a 1 de cada 3 alumnos le pareció extenso), del 37.1 % de los estudiantes que tuvieron duda al contestar el cuestionario, la pregunta 16 tuvo un 65.2% de duda, seguida de la pregunta 44 con un 6.52%, y se obtuvo una aceptación superior al 90% para mejorar la calidad de la enseñanza y para ser aplicado a los profesores de la FES Zaragoza. (tabla 6).

Discusión

La evaluación de la docencia es un proceso complejo en donde el papel de los estudiantes es importante y necesario, además, la inclusión de diferentes técnicas pedagógicas para impartir los contenidos de un programa académico, como la microenseñanza, es un punto dentro de algunos planes institucionales en México ([Universidad Veracruzana \[UV,2015\]](#); [Instituto Politécnico Nacional \[IPN;2018\]](#)). Particularmente, la evaluación del proceso de microenseñanza requiere que se tomen en cuenta diferentes habilidades didácticas (Dwight, 1967), considerando como una actividad importante en este momento el uso de las TIC, por lo que contar con instrumentos actuales, válidos y

confiables para ser resueltos por estudiantes, y tener parámetros de comparación y retroalimentación, tanto para el docente como para el personal encargado de la parte académico-administrativa, es un punto ineludible a nivel educativo.

En el diseño del CEM del presente trabajo se buscó que tuviera congruencia con los nueve aspectos incluidos en la habilidad didáctica en el cuestionario de la Evaluación de la Docencia aplicado en la Carrera de Q.F.B., en FES Zaragoza, y que, además, la redacción fuera clara y entendible ([Morales, 2011](#)) para que estudiantes y docentes lo comprendan de la misma manera.

El CEM diseñado quedó conformado por 49 preguntas en 3 habilidades didácticas: comunicación verbal y no verbal, formulación de preguntas y reforzamiento, y empleo de medios audiovisuales y las TIC. En este sentido, el número de habilidades didácticas en los instrumentos propuestos con anterioridad son diferentes para cada autor. [Watson \(2007\)](#) y [Toro \(2017\)](#) consideran tres habilidades, la Universidad Rafael Landívar (2010) propone 6, [Dayanindhi \(2018\)](#), [Dwight \(1967\)](#) y [Unidades Tecnológicas Santander \(2011\)](#) sugieren 7 y [Remesh \(2013\)](#) indica 9 habilidades didácticas, por lo que este instrumento está acorde con el mínimo de habilidades a evaluar según los diferentes estudios.

Analizando por habilidad evaluada, la de comunicación verbal y no verbal es considerada por

Tabla 6.
Frecuencias de la aceptación del IOCEM

Pregunta	Aceptación
	Frecuencia (porcentaje)
1.- ¿La introducción y el objetivo explican claramente el propósito del cuestionario?	123 (99.2)
2.- ¿Las instrucciones del llenado del cuestionario son claras y entendibles?	115 (92.7)
3.- ¿Consideras que la escala de calificación que utiliza el cuestionario es clara y entendible?	103 (83.1)
4.- ¿Te pareció extenso el cuestionario?	43 (34.7)
5.- ¿Consideras que la redacción de cada una de las preguntas es clara y entendible?	88 (71.0)
6.- ¿Tuviste duda en contestar en alguna de las preguntas?	46 (37.1)
7.- ¿Consideras que el cuestionario podría ayudar a mejorar la calidad de enseñanza del profesor?	121 (97.6)
8.- ¿Consideras que la Facultad debe implementar el cuestionario para todos los profesores?	117 (94.4)

Watson (2007), la Universidad Rafael Landívar (2010), y las Unidades Tecnológicas de Santander (2011); la habilidad de formulación de preguntas la proponen la Universidad Rafael Landívar (2010) y las Unidades Tecnológicas de Santander (2011); y la de reforzamiento la contemplan Dwight (1967), la Universidad Rafael y Remesh (2013) y Landívar (2010), y finalmente la habilidad de empleo de medios audiovisuales y las TIC es considerada por Remesh (2013) y Dayanindhi (2018); información que fue tomada en cuenta para el desarrollo del CEM. Cabe resaltar que esta última habilidad es de reciente inclusión debido a los avances tecnológicos y, en este momento, por la necesidad de su uso debido a la situación de contingencia mundial en la que nos encontramos debido a la pandemia de la COVID-19.

En lo referente a la extensión del cuestionario es muy variable, reportándose en un rango de entre 19 (Toro, 2013) y 38 preguntas (Unidades Tecnológicas de Santander, 2011), lo que haría que el CEM sea un poco más extenso. Sin duda, las habilidades didácticas consideradas en cada instrumento, así como el número de preguntas en cada uno, dependerán en gran medida de los objetivos, propósitos y alcances de cada investigación. Morales (2011) comenta que no hay un número óptimo de preguntas, pero a mayor número inicial de preguntas se tendrá una mayor probabilidad de encontrar en el análisis un conjunto de preguntas definitivas con una fiabilidad suficientemente alta. En cuanto al tiempo aplicación se consideró adecuado dado la extensión del CEM.

Para la validez de contenido con las aportaciones proporcionadas por los expertos se realizaron mejoras importantes para dar lugar a un cuestionario con una estructura, pertinente, bien definida, con una secuencia lógica, redacción clara y congruente con el objetivo del instrumento, con niveles de calificación adecuados para cada pregunta, logrando que fuera entendible para ambos actores del proceso de enseñanza-aprendizaje. La validez del constructo fue a partir de la opinión de los alumnos. Este estudio es coincidente con el de Toro, 2013, ya que la aplicación del instrumento fue a estudiantes propiamente a diferencia de los demás reportes que aplicaron sus instrumentos a docentes nuevos o en formación.

Para el diseño del IOCEM, se buscó que la estructura y redacción, fueran claras y entendible (Morales, 2011) para los alumnos y docentes, y permitiera recabar información sobre opinión de cada apartado (introducción, objetivo, instrucciones de llenado, niveles de calificación y preguntas) del CEM.

Es importante señalar que ninguno de los autores consultados refiere la aplicación de un instrumento como el IOCEM (Dwight 1967; Guillermo 1997; Watson 2010; Unidades Tecnológicas de Santander 2011; Remesh 2013; Toro 2013; Dayanindhi 2018), el cual recaba la opinión de los alumnos sobre un cuestionario de evaluación.

Se considera a los alumnos como las mejores fuentes de información del proceso de enseñanza y aprendizaje, y son de quienes, mediante su opinión, se obtienen resultados del desempeño del docente en el aula a través del tiempo (Elizalde, 2008). La percepción que tienen los estudiantes sobre su aprendizaje provee una medida de la acción docente, relacionada con los efectos que tiene esa acción en su experiencia de instrucción, por lo que es capaz de reconocer en que proporción la actividad del docente afecta su aprendizaje. (Acevedo, 2003).

Es importante mencionar que la metodología de obtener información en los procesos de enseñanza y aprendizaje a partir de la opinión de estudiantes no es el único, pero si es de los más utilizados y antiguos, y en el ámbito universitario es el de mayor difusión, reconocimiento y confiabilidad (Gómez, 2009).

El CEM demostró una validez y confiabilidad estadística. Las 49 preguntas tuvieron una distribución normal de acuerdo con lo que establece Castañeda, 2010, mientras que otros autores no refieren este supuesto (Dwight 1967; Guillermo 1997; Watson 2010; Unidades Tecnológicas de Santander 2011; Remesh 2013; Toro 2013; Dayanindhi 2018). Así mismo, todas las variables tienen relación entre sí, al encontrarse una alta correlación entre todas las preguntas (variables), a diferencia de lo reportado en algunas investigaciones que no refieren ningún dato sobre esta prueba (Dwight 1967; Guillermo 1997; Watson 2010, Unidades Tecnológicas de Santander 2011; Remesh 2013, Toro 2013; Dayanindhi 2018). Solamente Mergler

(2010) reporta un valor de la prueba de esfericidad de Bartlett significativa ($p < 0.001$), semejante a lo encontrado en este trabajo.

Con relación a la extracción de factores, se realizó mediante un Análisis Paralelo y el método del gráfico de sedimentación, en donde el eje X se encuentra el número de ítems y en el eje Y los autovalores o varianza explicada para cada factor obteniéndose tres factores. Es importante mencionar que el AP es el método que se recomienda para la extracción de factores (Lloret-Segura, 2014; Ferrando, 2022). En este sentido la mayoría de los estudios anteriores no refieren ningún dato o información al respecto (Dwight 1967; Guillermo 1997; Watson 2010; Unidades Tecnológicas de Santander 2011; Remesh 2013; Toro 2013; Dayanindhi 2018); sin embargo, Mergler (2010), empleando el método de componente principales, considerando 12 variables, reporta dos componentes que explicaron el 65% de la varianza, los cuales son eficacia del profesor en el aula con el 22% de la varianza y eficacia personal del mismo con el 32% de la varianza.

Es importante mencionar que en este estudio se utiliza el método de “factorización por eje principal” y se empleó la rotación oblicua, utilizando el método promax, los cuales son recomendados para el análisis factorial en el ámbito social. (Lloret-Segura, 2014). En el presente trabajo el primer factor encontrado fue la habilidad de “comunicación verbal y no verbal”, con una varianza explicada del 40.83% y 29 variables, siendo 14 preguntas al inicio del estudio. Esta habilidad fue considerada en los reportes de las Unidades Tecnológicas de Santander (2017), y de manera implícita en Watson (2007) en la competencia 3 “Se comunica de forma efectiva involucrando a los alumnos en el tema desarrollado” y la Universidad Rafael Landívar (2010) en la habilidad de “refuerzo verbal y no verbal”.

El segundo factor fue la habilidad de “formulación de preguntas y reforzamiento” con 15 variables o preguntas. Al respecto, en la versión inicial del cuestionario estas habilidades estaban separadas, de tal manera, que la formulación de preguntas estaba conformada por 9 preguntas y el de reforzamiento tenía 16 preguntas. Al respecto, esta habilidad ha sido contemplada con anterioridad como un elemento a evaluar en los estu-

dios de Dwight (1967) Guillermo (1997), Unidades Tecnológicas de Santander (2011), Remesh (2013), y en el trabajo de la Universidad Rafael Landívar (2010) que lo toma en cuenta de manera implícita en la “habilidad para hacer preguntas y conseguir respuestas orales”.

El tercer factor fue la habilidad de “empleo de medios audiovisuales y las TIC”, una nueva habilidad que surgió como resultado del análisis factorial, la cual contó con 5 variables que se ha encontrado en los estudios más recientes debido a la naturaleza de la habilidad evaluada. (Remesh 2013, y Dayanindhi 2018).

Estos tres factores fueron corroborados a través de la carga factorial ya que las preguntas incluidas en cada factor presentaron un valor de correlación mayor a 0.35 acorde con lo estipulado por Spector (1992) y semejante a la propuesta de Tabachnick y Fidel (como se cita en Costello, 2005). En este sentido, Thurstone (como se cita en De la Fuente, 2011) menciona que la matriz de cargas factoriales debe reunir tres características: cada factor debe tener unos pesos altos y los demás próximos a cero; cada variable no debe estar saturada más que en un factor; y no deben existir factores con la misma distribución, esto es, dos factores distintos deben presentar distribuciones diferentes de cargas altas y bajas. De esta manera cada factor tendrá una correlación alta con un grupo de variables y baja con el resto de las variables; de aquí que el CEM cumplió perfectamente con las características descritas por Thurstone.

Por otro lado, el cuestionario mostró un alto grado de confiabilidad respecto a las tres habilidades didácticas al ser utilizado por estudiantes, indistintamente del semestre que cursan y si están evaluando docentes de teoría o de actividad práctica, con un valor de alfa de Cronbach igual a 0.961. Toro (2013) reportó un alfa de Cronbach igual a la obtenida en el presente estudio, así como Capa (2005) indicó una confiabilidad semejante con el instrumento aplicado en su investigación. Otros estudios no mencionan ninguna información referente al valor de la confiabilidad de los instrumentos empleados (Dwight, 1967; Guillermo, 1997; Watson 2007; Universidad Rafael Landívar, 2010; Unidades Tecnológicas Santander, 2011; Remesh, 2013; Toro, 2013 y Dayanindhi, 2018).

Como se puede apreciar, aunque son varias las propuestas de instrumentos para la evaluación de los docentes que se desempeñan en el campo de la microenseñanza, pocos son los que han obtenido los parámetros de validez y confiabilidad como en el presente estudio, además, de considerar en su diseño y evaluación la opinión directa de los estudiantes. De acuerdo con los resultados obtenidos podemos señalar que el CEM es válido y confiable para su uso con estudiantes en la evaluación de docentes, siendo un instrumento que mide el constructo de la microenseñanza.

Conclusiones

Con base en lo anterior se puede decir que el Cuestionario de Evaluación de la Microenseñanza que incluye la comunicación verbal y no verbal, la formulación de preguntas y el reforzamiento, además del empleo de medios audiovisuales y las TIC es válido y confiable, y puede ser aplicable en la población estudiantil de la carrera de Q.F.B. o cualquier otra semejante. De acuerdo con la validez consecucional, el CEM, se aplicó sobre una población estudiantil con un perfil establecido, que fue informada sobre los objetivos y propósitos del CEM, con base en la opinión de los profesores expertos se diseño una prueba con lenguaje claro y preciso, no se observaron durante su aplicación usos imprevistos o inadecuados del cuestionario por lo que el CEM evaluó correctamente el constructo de la microenseñanza.

Referencias

Acevedo, AR. (2003) Factores que inciden en la competencia docente universitaria. (Tesis doctoral). <http://webs.ucm.es/BUCM/tesis/edu/ucm-t26870.pdf>

American Educational Research Association, American Psychological Association & National Council for Measurement in Education [AERA, APA & NCME] (2014). The Standards for Educational and Psychological Testing. Washington, D.C.: AERA

Aguilar, V.M. (2003). Validación de un instrumento para evaluar

el desempeño docente en las especialidades médicas (tesis de maestría). <https://repositorio.unam.mx/contenidos/validacion-de-un-instrumento-para-evaluar-el-desempeno-docente-en-las-especializaciones-medicas-92500?c=%7B>

Capa, Y., Çakıroğlu, J., & Sarıkaya, H. (2005). The Development and Validation of a Turkish Version of the Teachers' Sense of Efficacy Scale. *Educacion and Science*, 30(137), 74-81. <https://toad.halileksi.net/sites/default/files/pdf/ogretmen-ozyeterlik-olcegi-toad.pdf>

Costello, A.B. & Osborne, J.W. (2005). Best practices in exploratory factor análisis: four recommendations for getting the most from your análisis. *Practical Assessment, Research & Evaluation*, 10(7), 1-9. doi: <https://doi.org/10.7275/jvj1-4868>

Dayanindhi, V. K., & Hegde, S. P. (2018). Effectiveness of microteaching as a method of developing teaching competence among in-service medical teachers. *Journal of Advances in Medical Education & Professionalism*, 6(4), 155-161. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6191834>

De la Fuente, F.S. (2011). Análisis factorial. España: Universidad Autónoma de Madrid. <https://www.fuenterrebollo.com/Economicas/ECONOMETRIA/MULTIVARIANTE/FACTORIAL/analisis-factorial.pdf>

Deniz, S. (2011). Implications of training student teachers of pre-schooling through micro-teaching activities for a classroom with mentally- disabled students. *Educational Research and Reviews*. 6(8), 560-569. <https://academicjournals.org/journal/ERR/article-full-text-pdf/8239E6B6323>

Dwight, A. (1967). Micro-teaching: a description. Stanford teacher education program. California: Stanford University. <https://files.eric.ed.gov/fulltext/ED019224.pdf>

Elizalde, L. y Reyes, R. (2008). Elementos clave para la evaluación del desempeño de los docentes. *Revista Electrónica de Investigación Educativa*, Especial, 10, 1-13. <https://www.redalyc.org/pdf/155/15511127004.pdf>

Escoto, P.L.M.C. (2001). Evaluación del cambio conductual de profesores expuestos al curso de microenseñanza de una institución de educación superior. (Tesis de Maestría). https://repositorio.unam.mx/contenidos/evaluacion-del-cambio-conductual-de-profesores-expuestos-al-curso-de-microensenanza-de-una-institucion-de-educacion-135421?c=r300-vQ&d=true&q=*&i=1&v=1&t=search_0&as=0

Facultad de Estudios Superiores Zaragoza (2015). Informe de

- evaluación de la docencia en las asignaturas teóricas de las especializaciones. Semestre 2014-1. México, D.F.
- Ferrando, P.J., Lorenzo-Seva, U., Hernández-Dorado A., Muñiz, J. (2022). Decálogo para el Análisis Factorial de los Ítems de un Test. *Psicothema*, 34(1), 7-17. doi: 10.7334/psicothema2021.456. <https://dialnet.unirioja.es/servlet/articulo?codigo=8259546>
- Gómez, BE y Gómez BY. (septiembre 2009). La evaluación de la docencia basada en la opinión de los alumnos: el perfil docente en la Licenciatura en Historia. Presentado en el X Congreso Nacional de Investigación Educativa. Veracruz, México
- Guillermo, GMC. (1997). Microenseñanza: ¿Una técnica vigente para el desarrollo de Habilidades Docentes? *Nueva Época* Vol. 1 (16), 59-68. <http://www.educacionyciencia.org/index.php/educacionyciencia/article/download/120/pdf>
- Hair, J., Anderson, R. Tatham, R., & Black, W. (1999). Análisis multivariante. (5ª ed). Madrid: Prentice Hall.
- Hernández R., Fernández C. & Baptista P. (2010). Metodología de la Investigación. (5ª. Ed). México: Mc Graw Hill.
- Instituto Politécnico Nacional (2018). Memoria Anual de Actividades 2018. Escuela Nacional de Ciencias Biológicas (ENCB). <https://www.ipn.mx/assets/files/se-cgeneral/docs/Memorias/Memoria2018/ns/enchb.pdf>
- Lloret-Segura, S., Ferreres-Traver, A.- Hernández-Baeza, A., Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *Anales de psicología*, 30(3), 1151-1169.
- López, R.P & Facheli, S. (2015). Metodología de la investigación social cuantitativa. Barcelona: Universidad Autónoma de Barcelona.
- Meneses, A. (08 de Junio de 2020). *Estimación análisis paralelo utilizando SPSS*. [Obtenido de archivo de video]. Recuperado de https://www.youtube.com/watch?v=AtCIR_67ljY
- Mergler, A. & Tangen, D. (2010). Using microteaching to enhance teacher efficacy in pre-service teachers. *Teaching Education*, 21 (2). 199-210 <https://doi.org/10.1080/10476210902998466>
- Morales, V. P. (2011). El análisis factorial en la construcción e interpretación de test, escalas y cuestionarios. Madrid: España: Universidad Pontificia Comillas. <http://www.upcomillas.es/personal/peter/investigacion/AnalisisFactorial.pdf>
- Remesh, A. (2013). Microteaching, an efficient technique for learning effective teaching. *Journal of Research Medical Sciences*, 18(2), 158-163. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3724377/>
- Rodríguez Fernández, Zenén, Chércoles Cazate, Lilia Esperanza, Santisteban Aguilera, Francisca de las Nieves, Ricardo Ramírez, José Manuel, & Uriarte Gómez, Manuel Esteban. (2017). La clase según lo establecido en reglamentos ministeriales cubanos. *MEDISAN*, 21(9), 2076-2085. Recuperado en 29 de abril de 2022, de http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1029-30192017000900015&lng=es&tlng=es.
- Secretaría de Educación Pública. (2016). Estrategias de microenseñanza para fortalecer la gestión del aula. Microenseñanza <http://www.cosdac.sems.gob.mx/eme/>
- Spector, P. E. (1992). Summating ratings scale construction: An introduction Sage University Papers Series. Quantitative applications in the social sciences. Iowa: Sage Publications. <https://home.ubalt.edu/tmitch/645/articles/Summated%20Rating%20Scales.pdf>
- Taut, S. & Palacios, D. (2016). Interpretaciones no intencionadas e intencionadas y usos de los resultados de PISA: Una perspectiva de validez consecuencial. *RELIEVE*, 22(1), <https://ojs.uv.es/index.php/RELIEVE/article/view/8294>
- Toro, G. (2017). La microenseñanza y el desempeño de los estudiantes de tecnología del vestido en la práctica docente continua de la Facultad de Tecnología de la Universidad Nacional de Educación Enrique Guzmán y Valle, 2013 (tesis doctoral). Universidad Nacional de Educación "Enrique Guzmán y Valle". <https://repositorio.une.edu.pe/bitstream/handle/UNE/1618/TD%20CE%201616%20T1%20-%20Toro%20Mejia.pdf?sequence=1&isAllowed=y>
- Unidades Tecnológicas de Santander (2011). Documento rector. Diplomado sobre estrategias pedagógicas en la educación superior. Modulo 3: La microenseñanza. <https://es.calameo.com/read/00007273578fc5b662de1>
- Universidad Rafael Landívar (2010). Técnicas de didáctica universitaria: microenseñanza, *Revista Electrónica Ingeniería Primero*, 15, 13-32. https://fgsalazar.net/LANDIVAR/ING-PRIMERO/boletin15/URL_15_INGO1.pdf
- Universidad Veracruzana (2015). Plan de Desarrollo de las entidades académicas (PlaDEA). Facultad de Ciencias Químicas. <https://www.uv.mx/tecnica/files/2018/01/PlaDEA-Cordoba-Orizaba-Ciencias-Quimicas.pdf>
- Videla, R. L. (2005). VIGENCIA Y DECADENCIA DE LA CLASE

TEÓRICA (PARTE I). Revista Argentina de Radiología, 69(4),319-325.[fecha de Consulta 29 de Abril de 2022]. ISSN: 1852-9992. Disponible en: <https://www.redalyc.org/articulo.oa?id=382538438009>

Watson, R. (2007). La microenseñanza en la UPC. *Revista Digital de Investigación en Docencia Universitaria*, 3 (1), 1-20. <https://doi.org/10.19083/ridu.3.24>

Evaluación del desempeño docente durante la pandemia: opinión de estudiantes universitarios mexicanos

Patricio Henriquez Ritchie*

Universidad Autónoma de Baja California, Ensenada, México <https://orcid.org/0000-0002-1026-3379> phenriquez@uabc.edu.mx

Citar como: Henriquez Ritchie, P. (2023). Evaluación del desempeño docente durante la pandemia: opinión de estudiantes universitarios mexicanos. *Revista Digital de Investigación en Docencia Universitaria*, 17(2), e1780. <https://doi.org/10.19083/ridu.2023.1780>

Recibido: 14/02/2023. **Revisado:** 20/03/2023. **Aceptado:** 30/05/2023 **Publicado:** 30/06/2023.

Resumen

Introducción: la evaluación del desempeño de docentes universitarios desde la opinión de sus estudiantes es un proceso recurrente, aunque durante la pandemia hay pocas experiencias. **Objetivo:** evaluar el desempeño de docentes universitarios desde la opinión de sus estudiantes, en contexto de pandemia. **Método:** muestra representativa de 673 estudiantes de una universidad pública mexicana, se utilizó un cuestionario de desempeño docente adaptado para clases virtuales. **Resultados:** la mayoría de los docentes se adaptó a una dinámica de trabajo virtual asincrónica, aunque un porcentaje importante (41.9%) continuó impartiendo clases de manera sincrónica (videoconferencias). Los docentes de etapas formativas iniciales promovieron principalmente el aprendizaje individual de contenidos, contrastando con aquellos de etapas avanzadas, quienes fomentaron un aprendizaje situado. Estudiantes con opiniones desfavorables respecto al dominio de contenidos y a las estrategias de evaluación, mencionaron la promoción de un aprendizaje memorístico, en desmedro de aprendizajes más constructivistas. Los contrastes fueron estadísticamente significativos.

Palabras Clave: evaluación docente; estrategias de enseñanza; objetivos de enseñanza; estrategias de evaluación del aprendizaje.

Assessment of teaching performance during the pandemic: opinion of Mexican university students

Abstract

Introduction: The evaluation of the performance of university teachers from the opinion of their students is a recurring process, although during the pandemic there are few experiences. **Objective:** To evaluate the performance of university teachers from the opinion of their students, in the context of a pandemic. **Method:** Representative sample of 673 students from a Mexican public university, a teacher performance questionnaire adapted for virtual classes was used. **Results:** Most of the teachers adapted to an asynchronous virtual work dynamic, although a significant percentage (41.9%) continued to teach classes synchronously (videoconferences). Teachers in initial formative stages mainly promoted individual content learning, contrasting with those in advanced stages, who promoted situated learning. Students with unfavorable opinions regarding content mastery and assessment strategies mentioned the promotion of rote learning, to the detriment of more constructivist learning. The contrasts were statistically significant.

*Correspondencia:

Patricio Henriquez Ritchie
phenriquez@uabc.edu.mx



Keywords: teacher evaluation; teaching strategies; teaching objectives; learning evaluation strategies.

Introducción

La evaluación del desempeño docente en el nivel educativo superior ha sido un tema de intenso debate, tanto por cuestiones teórico-conceptuales como técnico-metodológicas. Por un lado, a nivel teórico se ha discutido acerca de los aspectos y dimensiones principales que deben considerarse a la hora de analizar el desempeño docente: la manera en que organiza y planifica sus clases, el nivel de dominio o experticia en el manejo de los contenidos que enseña, las formas de interacción con el alumnado, las competencias didácticas para enseñar y lograr el tipo de aprendizaje que se propone con sus alumnos, las estrategias y las funciones que le asigna a la evaluación del aprendizaje, entre otros ([UNESCO, 2006](#); [García Cabrero, Loredó Enríquez & Carranza Peña, 2008](#); [Cordero Arroyo, Luna Serrano & Patiño Alonso, 2013](#)). Además, mucho se ha debatido ([Santos del Real, 2012](#); [Martínez, 2016](#); [Vázquez Cruz, Cordero Arroyo & Leyva Barajas, 2014](#)) acerca de la finalidad detrás de la evaluación del profesorado, de manera que ante la pregunta ¿Para qué evaluar a los docentes?, las respuestas se han bifurcado en direcciones apegadas a una perspectiva sumativa (de certificación, de ascensos laborales, de estímulos económicos) o formativa (retroalimentación, propuestas de mejoramiento, identificación de áreas de oportunidad).

Por otro lado, a nivel metodológico la discusión se ha centrado en los instrumentos (cuestionarios cerrados, entrevistas y observación de aula, principalmente) y los sujetos (estudiantes, colegas, directores, expertos externos) más idóneos para obtener información confiable, válida, suficiente y pertinente con base en la cual evaluar el desempeño de los docentes ([Cruz Ávila, 2007](#); [Romero Díaz & Martínez Gimeno, 2017](#); [Zamora Serrano, 2021](#)). Más allá de esta discusión, la evaluación del desempeño docente desde la opinión de sus estudiantes a través de cuestionarios cerrados sigue siendo una forma de acercamiento recu-

rrente e institucionalizada en el nivel educativo superior, así como también válida, confiable y legítima cuando se aborda con los requerimientos teóricos y metodológicos adecuados.

El presente artículo presenta los resultados del proceso de evaluación del desempeño docente según la opinión de los estudiantes de la Facultad de Ciencias Administrativas y Sociales (FCAYS), Universidad Autónoma de Baja California (UABC), México. Sus objetivos principales son: describir la opinión de los estudiantes en torno al desempeño de sus docentes a partir de cuatro dimensiones (planificación de la enseñanza, dominio de contenidos, estrategias de enseñanza y evaluación del aprendizaje); y analizar la relación entre los objetivos de enseñanza de los docentes y variables académicas de los estudiantes (etapa formativa, opinión acerca del dominio de contenidos y de las estrategias de evaluación del aprendizaje), en contexto de clases virtuales por contingencia sanitaria debido al COVID-19. La evaluación del desempeño docente desde la opinión de estudiantes universitarios se vio trastocada por la contingencia sanitaria producto de la propagación del virus COVID-19 durante 2020-21.

En efecto, los sistemas educativos de todos los países y en todos los niveles tuvieron que trabajar bajo modalidad virtual a distancia durante este periodo. Lo anterior trajo consigo un conjunto de desafíos tanto a nivel tecnológico como pedagógico: por una parte, a nivel tecnológico se tuvo que enfrentar un complicado escenario en el cual docentes y alumnos se vieron obligados a utilizar una serie de recursos digitales para acceder a los contenidos educativos y para establecer canales comunicativos. Sin embargo, como lo destacan algunos autores ([Vialart Vidal, 2020](#); [Mendoza Castillo, 2020](#); [Cencia Crispín, Carreño Colchado, Eche Querevalú, Barrantes Morales & Cárdenas Baldeón, 2021](#); [Villarreal, Pérez, Rojas-Barahona & García, 2021](#); [Moreno Tapia, Quintero López & Cáceres Mesa, 2021](#)) se presentaron dificultades tanto en términos de dominio tecnológico (conocimiento y manejo de recursos digitales) como en términos socioeconómicos (acceso y posesión de recursos digitales y conexión a la red en el hogar; espacios idóneos para la realización de tareas y trabajos).

Por otra parte, a nivel pedagógico los desafíos que presentó la pandemia no fueron menores: el profesorado se vio en la necesidad de modificar sus estrategias didácticas para enseñar los contenidos de sus materias en línea, modalidad que muchos no dominaban y a la cual tuvieron que ir adaptándose en el camino, de manera autodidacta y forzada ([Castro, Paz & Cela, 2020](#); [Valenzuela & Barrios, 2020](#); [Vialart Vidal, 2020](#); [Moreno Tapia et al., 2021](#); [Mamani-Cori, Padilla, Cervantes, Caballero & Sucari, 2022](#)). En este sentido, se observó una problemática preocupante relacionada con la incapacidad del profesorado universitario para transformar sus estrategias de enseñanza según las necesidades del contexto pandémico y adaptarse a los requerimientos de innovación didáctica que exige la educación a distancia, mediada por recursos digitales. La elaboración de diseños instruccionales adecuados, la propuesta de actividades que incluyan situaciones de aprendizaje en modalidad virtual, la incorporación de estrategias de evaluación del aprendizaje innovadoras a partir de la modalidad a distancia, entre otros elementos, salieron a luz y pusieron de manifiesto la necesidad de capacitación y preparación del profesorado al respecto.

En este contexto, resulta trascendental analizar los objetivos de enseñanza de los docentes bajo la modalidad virtual, considerando la relación entre estos objetivos y las teorías psicológicas del aprendizaje que han permeado esta discusión desde hace décadas ([Zapata-Ros, 2015](#)). Así, mientras tradicionalmente algunos docentes han preferido el uso de estrategias de enseñanza apuntadas al fomento de un aprendizaje memorístico, basado en una lógica de estímulos-respuestas y posterior suministro de retroalimentación; otros se han enfocado en la aplicación de estrategias de enseñanza que promueven un aprendizaje significativo y activo, orientado al procesamiento cognitivo de los contenidos y la construcción de significados por parte de los estudiantes. De esta manera, en función de los objetivos del presente estudio, se plantea la siguiente interrogante: ¿Cómo se desarrolló esta dinámica estratégica en el contexto de clases virtuales por el COVID-19 y cómo enfocaron los objetivos de enseñanza los docentes universitarios en dicho contexto?

Frente a las problemáticas anteriormente expuestas, el presente artículo expone los resultados del proceso de evaluación del desempeño docente desde la opinión de los estudiantes de la FCAYS de la UABC durante el período 2021-2, en que se implementaron clases virtuales a distancia a raíz de la contingencia sanitaria por el COVID-19. Se indagó la perspectiva del estudiantado respecto al desempeño de sus profesores según la planificación de clases, el dominio de contenidos, las estrategias de enseñanza y la evaluación del aprendizaje. Los resultados muestran tendencias interesantes en función de los desafíos mencionados en los párrafos anteriores y posibilitan vislumbrar elementos, tanto teóricos como metodológicos, que abonan a la discusión en torno a la evaluación del desempeño docente en el nivel educativo superior.

Método

Diseño

El enfoque metodológico del presente estudio es cuantitativo, con un diseño descriptivo, relacional y no experimental. Las variables consideradas en torno a la opinión de los estudiantes se engloban dentro de cuatro dimensiones del desempeño docente: planificación de la enseñanza, dominio de contenidos, estrategias de enseñanza y evaluación del aprendizaje (ver tabla 2). De estas variables, se consideraron los objetivos de enseñanza (variable dependiente), así como el dominio de contenidos y evaluación del aprendizaje (variables independientes) para los análisis inferenciales realizados (chi-cuadrada, ANOVA).

Participantes

El estudio consideró una muestra representativa de estudiantes de la FCAYS de la UABC, campus Ensenada. Esta facultad está integrada por ocho programas de licenciatura, que abarcan tres áreas de conocimiento: Derecho (ciencias jurídicas); Psicología, Ciencias de la Comunicación, Sociología y Ciencias de la Educación (ciencias sociales); Informática, Administración de Empresas y Contaduría (ciencias administrativas).

Durante el período 2021-2, la población fue de 4,551 estudiantes, según datos oficiales de la universidad¹. Utilizando el algoritmo propuesto por Cuesta y Herrero (2010, citado en: Organista-Sandoval y Serrano-Santoyo, 2014), con un nivel de confianza del 95% ($Z=1.96$), una frecuencia/probabilidad del factor a estudiar del 50% ($P=0.5$) y una estimación del error máximo del 0.05, se calculó que el tamaño mínimo de la muestra para que fuera representativa era de 354 estudiantes. En el caso de este estudio, la muestra alcanzó los 673 participantes (14.8% de la población estudiantil), de los cuales la mayoría se encontraba cursando la etapa disciplinar (semestre tercero al sexto: 52.6%) de su programa de estudios (tabla 1). Por su parte, menos de un quinto (18,6%) de los participantes se encontraban en la etapa básica (semestres 1 y 2) de sus licenciaturas.

Tabla 1
Frecuencia y porcentaje de participantes por etapa formativa

Etapas de formación	n	%
Básica	124	18,6
Disciplinar	351	52,6
Terminal	192	28,8
Total	673	100,0

Fuente: elaboración propia

Instrumentos

En la FCAYS de la UABC se aplica semestralmente un cuestionario de evaluación del desempeño docente a los estudiantes para conocer su opinión en torno a cuatro dimensiones (Henríquez, Arámburo & Dávila, 2017; Henríquez, Arámburo & Boroel, 2018; Henríquez & Arámburo, 2021): i) *planeación de la enseñanza*: referida aspectos relacionados con la organización del curso por parte del docente (entrega, explicación y cumplimiento de plan de clases, duración de sesiones, carga de tareas, entre otros); ii) *dominio de contenidos*:

asociada a la experticia en torno a los contenidos enseñados por parte del docente (claridad, manejo, actualización, asociación con otras materias, relación con situaciones de la vida real, entre otros); iii) *estrategias de enseñanza*: dimensión que apunta a explorar las estrategias didácticas utilizadas por el profesor, con relación a las estrategias de enseñanza en torno a los contenidos y los objetivos de aprendizaje detrás de su acción docente (estrategias tradicionales apuntadas a objetivos de enseñanza y aprendizaje transmisivos y memorísticos; estrategias constructivistas asociadas a objetivos de enseñanza y aprendizaje significativos y socio-culturales); y iv) *evaluación del aprendizaje*: referida a los instrumentos de evaluación que utiliza el docente en su práctica de enseñanza y las actitudes de los estudiantes respecto a las estrategias, métodos y recursos de evaluación del aprendizaje.

Este cuestionario tiene el propósito de aportar información valiosa, desde la perspectiva del estudiantado de la FCAYS, sobre el desempeño de sus docentes para favorecer su formación continua, identificar áreas de oportunidad y fortalecerlas a través de programas de capacitación. Como fundamento teórico-conceptual, este instrumento se basa en modelos multidimensionales construidos, validados y reportados en el marco de investigaciones internacionales respecto a la evaluación de la enseñanza desde la opinión de los estudiantes (Arreola, 2007; Bazán-Ramírez, Pérez-Morán & Bernal-Baldenebro, 2021; Fink, 2008). De esta manera, se atienden las recomendaciones López-Aguado & Gutiérrez-Provecho (2019), quienes destacan la trascendencia de considerar los conocimientos teóricos del ámbito de estudio como primer paso en el proceso de construcción de escalas: de esta manera, la selección de variables son base en esta orientación teórica evita el levantamiento de futuras estructuras factoriales artificiales².

¹ <http://cgsege.uabc.mx/documents/10845/65446/Poblaci%C3%B3n%20Estudiantil%202021-2>

² Como se mencionó, el instrumento acá presentado fue una adaptación del cuestionario original para aplicarse en contexto de contingencia sanitaria por el COVID-19 y la obligatoriedad de clases a distancia. El instrumento original fue retomado al finalizar el confinamiento pandémico y sus análisis psicométricos se encuentran en proceso de publicación.

En su primera versión ([Henríquez, Arámburo & Dávila, 2017](#)), el cuestionario se compuso de 16 reactivos a través de los cuales se pretendía describir aspectos relacionados con la impartición presencial de clases por parte de la planta docente de la FCAYS. En el presente documento, se presentan los resultados obtenidos a partir de la adaptación de este instrumento, producto de la contingencia sanitaria causada por la propagación del COVID-19 y la impartición obligatoria de clases virtuales durante el período 2021-2. De

esta manera, algunos reactivos debieron ser modificados y otros incorporados para indagar la perspectiva del alumnado acerca del desempeño de sus docentes en torno a las mismas dimensiones, pero en condiciones de clases virtuales mediadas por plataformas y medios digitales. En la tabla 2 se muestra la distribución de las dimensiones, variables, escalas de medición, número de reactivos y los índices de confiabilidad a través del coeficiente de Alfa, de cada dimensión del instrumento:

Tabla 2
Dimensiones, variables, escalas y coeficientes de confiabilidad del cuestionario desempeño docente para estudiantes de la FCAYS, UABC, 2021-2

Dimensiones	Variables	Escala	Número de reactivos	Índice de confiabilidad
Planeación de la enseñanza	Entrega y explicación del programa al inicio del semestre	Nominal	8	.426
	Atención a lo establecido en el programa	Ordinal		
	Duración de sesiones virtuales	Ordinal		
	Carga de tareas durante sesiones virtuales	Ordinal		
	Dinámica de trabajo durante sesiones virtuales	Nominal		
	Medios/plataformas de apoyo	Nominal		
	Retroalimentación posterior a metas alcanzadas	Ordinal		
	Idoneidad de recursos utilizados	Nominal		
Dominio de contenidos	Claridad	Ordinal	6	.858
	Manejo y actualización	Ordinal		
	Conexión con situaciones cotidianas (situado)	Ordinal		
	Conexión con otras materias	Ordinal		
	Fomento de la participación	Ordinal		
	Flexibilidad	Ordinal		
Estrategias de Enseñanza	Estrategias de trabajo en torno a contenidos	Nominal	2	.310
	Objetivos de aprendizaje	Nominal		
Evaluación del aprendizaje	Instrumentos de evaluación del aprendizaje	Nominal	8	.811
	Actitudes en torno a las estrategias de evaluación	Ordinal		

Fuente: elaboración propia

Así, con base en esta nueva versión, el cuestionario de evaluación del desempeño docente en contexto de pandemia constó de 4 dimensiones, 18 variables y 24 reactivos (además de la dimensión de información sociodemográfica y académica de los participantes: edad, licenciatura, materia/asignatura, semestre que cursa al momento de contestar el cuestionario; y la dimensión de estrategias de retroalimentación, que no se consideraron en este reporte). Como se observa en la tabla anterior, dos de las dimensiones teóricas consideradas en el instrumento (dominio de contenidos y evaluación del aprendizaje) arrojaron índices de confiabilidad aceptables ($>.80$). La dimensión de planeación de la enseñanza arrojó un índice bajo ($<.60$). Mientras que la dimensión de estrategias de enseñanza mostró un índice por debajo de los aceptable ($<.40$). Lo anterior se puede explicar, como lo argumentan [Barrios & Cosculluela \(2013\)](#) y [Oviedo & Campo-Aria \(2005\)](#), debido a la poca cantidad de ítems que conforman esta dimensión. Al respecto, se asevera que la cantidad de ítems afecta directamente el valor del alfa de una escala, así como también una mayor cantidad de individuos que respondan un instrumento.

Procedimiento

Dadas las condiciones de contingencia sanitarias producto de la propagación del COVID-19, durante el período 2021-2 se trabajó obligatoriamente a distancia mediante clases virtuales, mediadas por plataformas digitales y medios electrónicos para la entrega de contenidos e interacción entre docentes y estudiantes. Las condiciones anteriores forzaron a que el cuestionario de evaluación del desempeño docente fuera aplicado en línea al estudiantado de la FCYS de la UABC, mediante su acceso en la plataforma abierta de formularios de Google, durante el mes de octubre de 2021. De esta manera, con el apoyo de las coordinaciones de cada licenciatura, se envió el vínculo electrónico a los jefes de grupo de cada semestre dentro de todos los programas educativos, a quienes se les pidió compartirlo con sus compañeros, indicándoles que el formulario estaría 30 días disponible para ser respondido. Al cabo de un mes, se cerró la captura de respuestas y se procedió a la configuración de los archivos de procesamiento de datos.

Análisis de datos

En función de los objetivos de la investigación y de la naturaleza de los datos recolectados, se procedieron a ejecutar análisis estadísticos descriptivos e inferenciales con apoyo del paquete estadístico para ciencias sociales SPSS®, versión 25.0. Por un lado, se obtuvieron estadísticos descriptivos básicos según el tipo de variable: distribuciones de frecuencia, indicadores de tendencia central e índices de dispersión. Por otro lado, de acuerdo a las escalas con que se midieron algunas variables, se ejecutaron dos tipos de análisis inferenciales: paramétricos (contraste de medias, ANOVA) y no paramétricos (contraste de hipótesis de independencia, chi-cuadrada), en la opinión de los estudiantes de la FCYS acerca de los objetivos de enseñanza de sus docentes y variables académicas (etapa formativa, opinión acerca del dominio de contenidos y de las estrategias de evaluación del aprendizaje).

Resultados

Análisis descriptivos de la opinión de los estudiantes

Como se observa en la tabla 3, para indagar la dinámica de trabajo que los docentes privilegiaron en el desarrollo de sus clases durante la contingencia sanitaria (variable dentro de la dimensión Planeación de la enseñanza, ver tabla 2), se presentaron dos opciones a los estudiantes: i) clases asincrónicas: referidas a actividades planificadas con anticipación, objetivos y tareas planeadas con fechas preestablecidas, entrega de productos en plataformas digitales, clases virtuales de corta duración para esclarecer dudas y aclarar contenidos; y ii) clases sincrónicas: en las que el docente siguió impartiendo la materia como normalmente lo hace de manera presencial, sin variar sus estrategias (clases en los horarios normales, exposición de contenidos por videoconferencia y duración de clases similar a las presenciales). Aunque los porcentajes de opinión entre ambas opciones se mostraron equilibrados, la mayoría de estudiantes (58,1%) aseveró que sus docentes privilegiaron dinámicas asincrónicas de trabajo. Para esto, se apoyaron principalmente en la

plataforma Blackboard (92,3%), la cual permite administrar los contenidos de los cursos en diversas modalidades (presencial, semi-presencial, virtual) y ofrece diversos medios de comunicación (chat, mensajería, videoconferencias) entre docentes y estudiantes. No obstante lo anterior, llama la atención que más del 40% de los participantes aseveraron que sus docentes continuaron con una dinámica de trabajo sincrónica durante las clases virtuales en contexto de COVID.

Respecto al dominio de contenidos por parte de los docentes durante la contingencia sanitaria, se indagó a través seis ítems (ver tabla 2) en los que los participantes debían aseverar la frecuencia baja (nunca, algunas veces) o alta (casi siempre, siempre) con la que el profesorado enseñó los contenidos de manera clara, con un buen dominio y conocimiento actualizado, conectándolos con situaciones o problemas de la vida cotidiana (situado), relacionándolos con contenidos de otras materias, fomentando la participación activa de los estudiantes y con flexibilidad frente a situaciones que requerían de su atención especial. Como se puede observar en la tabla 4, en todos los ítems la opinión de los estudiantes arrojó porcentajes de frecuencia alta (sobre el 80%) respecto al manejo de contenidos por parte de sus docentes; excepto la conexión que establecen los docentes entre su materia y otras asignaturas del plan de estudios. En este aspecto, cerca de la mitad de los estudiantes (47,1%) aseveró una frecuencia baja de ejecución por parte del profesorado. Cabe destacar se ejecutaron análisis estadísticos para contrastar las frecuencias observadas y las teóricamente esperadas (chi-cuadrado, X^2) en todas las categorías, arrojando valores estadísticamente significativos (sig.=.000; 95% de confianza) para todos los casos; a excepción del manejo situado de contenidos (asociándolos con situaciones de la

vida real) por parte de los docentes, en el cual las diferencias entre las frecuencias no fueron significativas (sig.=.133). Con el propósito de complementar los resultados anteriores, se ejecutaron análisis para comparar las proporciones de cada categoría (estimador de tamaño del efecto, ETE) por medio del indicador H de Cohen. En todos los casos, los valores obtenidos se encuentran entre <0.2 y >0.5, lo que denota un grado de tamaño del efecto intermedio (Khalilzadeh & Tasci, 2017). Destaca la categoría que remite a la relación que establecen los docentes entre su materia y los contenidos de otras asignaturas, en que el indicador arrojó un valor >0.2 (efecto mínimo).

A su vez, se indagó la opinión de los estudiantes acerca de los objetivos de enseñanza que persiguieron sus docentes a través de la impartición de materias en modalidad virtual (tabla 5). Se propusieron cinco categorías de respuesta: i) memorizar y reforzar la retención de los contenidos de la materia (memorístico); ii) asociar los conocimientos previos de los estudiantes con los contenidos nuevos de la materia (significativo); iii) promover el aprendizaje individual a través de tareas y trabajos individuales de los estudiantes (individual); iv) contextualizar el aprendizaje en situaciones reales, relacionadas con el contexto laboral o la comunidad local (situado); y v) promover el aprendizaje colaborativo a través de tareas y trabajos en grupos de estudiantes (grupal). En este caso, la categoría con mayor porcentaje de respuesta fue el aprendizaje situado (32,1%), seguido del fomento del aprendizaje individual (22,4%) por parte de los docentes. Se ejecutó la prueba de contraste entre las frecuencias observadas y esperadas (chi cuadrado X^2) para cada categoría relacionada con objetivos de enseñanza: el resultado arrojó un índice significativo (sig.=.000), con un 95% de confianza.

Tabla 3
Distribuciones de frecuencia en torno a la dinámica de trabajo y medios

Dinámica de trabajo en clases virtuales	n	%	Medios	n	%
Clases asincrónicas	391	58,1	Blackboard	621	92,3
			Videoconferencias	43	6,4
Clases sincrónicas	282	41,9	Otros	9	1,3
Total	673	100,0	Total	673	100,0

Fuente: elaboración propia

Tabla 4
Distribuciones de frecuencia en torno al dominio de contenidos

Manejo de contenidos	Baja		Alta		Chi cuadrado	H de Cohen
	n	%	n	%	Sig.	
Claridad	107	15,9	566	84,1	.000	,341
Dominio	56	8,3	616	91,7	.000	,417
Situado	120	17,9	553	82,1	.000	,322
Otras materias	317	47,1	356	52,9	.133	,029
Participación	117	17,4	556	82,6	.000	,326
Flexibilidad	105	15,6	568	84,4	.000	,344

Fuente: elaboración propia

Tabla 5
Distribuciones de frecuencia en torno a los objetivos de enseñanza

Estrategias	n	%
Aprendizaje memorístico	88	13,1
Aprendizaje significativo	107	15,9
Aprendizaje individual	151	22,4
Aprendizaje situado	216	32,1
Aprendizaje grupal	111	16,5
Total	673	100,0
	Valor	Sig.
Chi cuadrado	77.156	.000

Fuente: elaboración propia

Por su parte, se preguntó acerca de los principales instrumentos de evaluación del aprendizaje utilizados por la planta docente de la FCAYS en contexto de clases virtuales por pandemia. Al respecto, cerca de la mitad de los estudiantes (44,5%) aseveró que sus docentes utilizaron los exámenes de opción múltiple individuales para evaluarlos, seguidos de un quinto de la muestra (20,9%) que mencionó los exámenes tipo ensayo individuales como principal instrumento utilizado. El resto mencionó otras herramientas de evaluación, tales como trabajos de investigación y reportes de lectura.

Respecto a las estrategias de evaluación del aprendizaje utilizadas por sus docentes, a nivel general se observó una opinión positiva por parte del estudiantado de FCAYS (tabla 6): más de la

mitad de los participantes estuvieron de acuerdo en el gusto por dichas estrategias (61,8%) y más de dos tercios coincidieron en que aprendieron de mejor manera, que sus docentes los apoyaron de manera paralela a la aplicación de instrumentos y que aplicaron estrategias relacionadas con los principios de la evaluación auténtica. No obstante lo anterior, llama la atención que cerca de la mitad de los estudiantes (44,9%) opinaron que sus docentes deben incluir otras estrategias de evaluación del aprendizaje, aparte de las utilizadas. En este caso, también se ejecutaron análisis de contraste entre las frecuencias observadas y las teóricamente esperadas (chi cuadrado, X^2) en todas las categorías, arrojando índices estadísticamente significativos para todos los casos (95%

de confianza). Nuevamente, con el propósito de complementar los resultados anteriores, se ejecutaron análisis para comparar las proporciones de cada categoría (estimador de tamaño del efecto, ETE) por medio del indicador H de Cohen. En todos los casos, los valores obtenidos se encuentran cercanos o bajo 0.2, lo que denota un grado de tamaño del efecto mínimo ([Khalilzadeh & Tasci, 2017](#)).

Análisis inferenciales en torno a la opinión de los estudiantes

En la tabla 7, se observan los resultados de un análisis estadístico no paramétrico (chi-cuadrada) realizado con base en la opinión de los estudiantes acerca de los objetivos de enseñanza según su etapa formativa dentro de sus respectivos planes de estudio. Lo anterior con el propósito de evidenciar eventuales diferencias significativas en

la opinión de los estudiantes a partir de esta última variable, contrastando la hipótesis de independencia entre ambas. Mientras los estudiantes de la etapa básica identificaron el aprendizaje individual como el objetivo de enseñanza mayormente promovido por sus docentes (27,4%), aquellos que se encontraban cursando las etapas formativas disciplinar y terminal mencionaron el aprendizaje situado como objetivo de enseñanza principalmente fomentado por sus profesores (33,9% y 33,3%, respectivamente). Como se puede observar, los resultados del análisis inferencial realizado indicaron diferencias significativas, con un índice de sig=.032 (95% de confianza). Sin embargo, el tamaño del efecto (V de Cramer) arroja un valor igual a .112, el cual denota una asociación débil (aunque significativa) entre las variables objetivos de enseñanza y etapa formativa de los estudiantes.

Tabla 6
Distribuciones de frecuencia de opinión en torno a las estrategias de evaluación

Estrategias de evaluación	Desacuerdo		Acuerdo		Chi cuadrado	H de Cohen
	n	%	n	%	Sig.	
Me gustaron	257	38,2	416	61,8	.000	,118
Aprendí mejor	214	31,8	459	68,2	.000	,182
Apoyo paralelo	208	30,9	465	69,1	.000	,191
Auténtica	194	28,8	479	71,2	.000	,212
Diversidad	255	37,9	418	62,1	.000	,121
Incluir otras	302	44,9	371	55,1	.008	,051

Fuente: elaboración propia

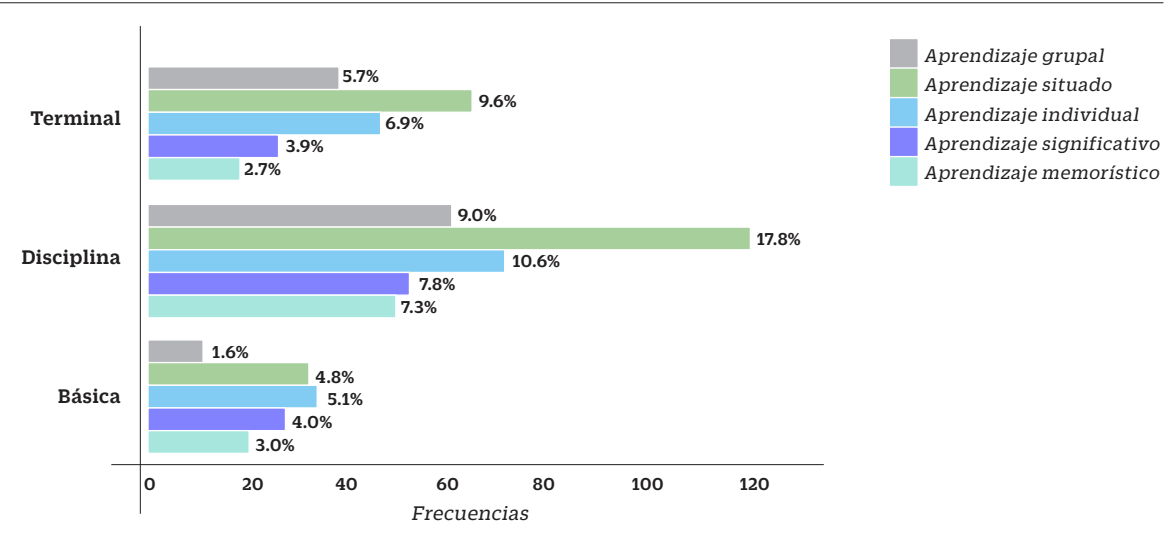
Tabla 7
Tabla de contingencia y chi-cuadrada de objetivos de enseñanza según etapa formativa

	Básica		Disciplinar		Terminal	
	n	%	n	%	n	%
Aprendizaje memorístico	20	16,1	49	14,0	18	9,4
Aprendizaje significativo	27	21,8	52	14,8	26	13,5
Aprendizaje individual	34	27,4	71	20,2	46	24,0
Aprendizaje situado	32	25,8	119	33,9	64	33,3
Aprendizaje grupal	11	8,9	60	17,1	38	19,8
			Valor		Sig.	
Chi-cuadrada*			16,781		.032	
V de Cramer			.112		.032	

*95% de confianza

Fuente: elaboración propia

Figura 1
Gráfico de barras de objetivos de enseñanza según etapa formativa



Fuente: elaboración propia

Tabla 8
ANOVA del índice de dominio de contenidos según objetivos de enseñanza

Estrategias	n	Media	D.E.	Suma de cuadrados	Sig.
Aprendizaje memorístico	88	1,8	0,8	40.4	.000
Aprendizaje significativo	107	2,3	0,7		
Aprendizaje individual	151	2,3	0,6		
Aprendizaje situado	216	2,6	0,4		
Aprendizaje grupal	111	2,4	0,6		

*95% de confianza; indicador de tamaño de efecto (ETA cuadrado)=0.147 (sig.=.000)

Fuente: elaboración propia

En la figura 1 se pueden observar de manera más clara las diferencias significativas de la opinión de los estudiantes respecto a los objetivos de enseñanza según su etapa formativa: a nivel general, en aquellos participantes que se encuentran cursando los primeros semestres de sus programas de estudio predominó el aprendizaje individual, mientras que en los alumnos que se encuentran en la etapa formativa intermedia (disciplinar) y final (terminal), predominó el aprendizaje situado como principal objetivo de enseñanza fomentado por sus docentes.

Por su parte, se obtuvo el promedio global de la opinión de los estudiantes acerca del dominio de los contenidos por parte de sus docentes, con el propósito de compararlo según los objetivos de enseñanza. El conjunto de ítems de esta dimensión (seis en total) fue medido con una escala de Likert de cuatro pasos: nunca (0), algunas veces (1), casi siempre (2) y siempre (3). Sobre la base de esta codificación, se obtuvo la media global del total de puntuaciones. Como se observa en la tabla 8, el objetivo relacionado con la promoción del aprendizaje memorístico fue el único

que arrojó un índice de frecuencia bajo (1,8) en torno al manejo de contenidos. Para analizar las diferencias significativas de esta variable según los objetivos de enseñanza, se aplicó un análisis paramétrico (ANOVA) de comparación de medias: el resultado de este análisis arrojó un índice de significatividad de .000, del cual se infieren diferencias estadísticamente significativas en el dominio de contenidos a partir de los objetivos de enseñanza. Además, el indicador de tamaño del efecto (ETA cuadrado) arrojó un valor de 0.147, el cual se considera bajo (sig.=.000), con relación a las diferencias entre los valores de la media del índice de dominio de contenidos obtenida para cada objetivo de enseñanza ([Cárdenas Castro & Arancibia Martini, 2014](#)).

Por último, un análisis similar se realizó con base en la opinión del estudiantado acerca de las estrategias de evaluación del aprendizaje que utilizan sus docentes: se obtuvo la media global para compararla según los objetivos de enseñanza que persiguieron a través de sus clases virtuales en contexto de pandemia. Igualmente, el conjunto de ítems (seis en total) se midieron por medio de una escala de Likert de cuatro pasos: totalmente en desacuerdo (0), en desacuerdo (1), de acuerdo (2) y totalmente de acuerdo (3). Con base en esto, se obtuvo el promedio del total de puntuaciones en torno a la opinión acerca de las estrategias de evaluación del aprendizaje. Como se observa en la tabla 9, nuevamente el objetivo de enseñanza relacionado con la promoción del aprendizaje

memorístico fue el que arrojó la media más baja dentro de la opinión del estudiantado (2,4). Se ejecutó un análisis paramétrico de comparación de medias (ANOVA) para analizar las diferencias estadísticas: el resultado arrojó un índice de significatividad de .003, del cual se infieren diferencias estadísticamente significativas respecto al resto de los objetivos de enseñanza. En este caso, el indicador de tamaño del efecto (ETA cuadrado) arrojó un valor de 0.024, el cual se considera bajo (sig.=.003), con relación a las diferencias entre los valores de la media del índice de opinión acerca de las estrategias de evaluación obtenida para cada objetivo de enseñanza ([Cárdenas Castro & Arancibia Martini, 2014](#)).

Discusión

Los análisis realizados a partir de la aplicación del cuestionario de evaluación del desempeño de los docentes de la FCAYS de la UABC, período 2021-2, muestran que la variable objetivos de enseñanza es determinante en la opinión del estudiantado a partir de su cruce con algunas variables académicas, tales como la etapa formativa, el dominio de contenidos y las estrategias de evaluación del aprendizaje utilizadas por el profesorado. En efecto, los resultados exponen que los estudiantes que se encuentran cursando la etapa inicial de las licenciaturas aseveraron que los objetivos de

Tabla 9
ANOVA de índice de estrategias de evaluación según objetivos de enseñanza

Estrategias	n	Media	D.E.	Suma de cuadrados	Sig.
Aprendizaje memorístico	88	2,4	0,9	15.3	.003
Aprendizaje significativo	107	2,8	0,9		
Aprendizaje individual	151	2,9	0,9		
Aprendizaje situado	215	2,9	1,1		
Aprendizaje grupal	111	2,8	1,0		

*95% de confianza; indicador de tamaño de efecto (ETA cuadrado)=0.024 (sig.=.003)
Fuente: elaboración propia

enseñanza de sus docentes apuntan a promover el aprendizaje individual, contrastando significativamente con aquellos de etapas disciplinar y terminal, que mencionaron el aprendizaje situado como el principal propósito de enseñanza. De la misma manera, los estudiantes que mostraron una opinión negativa acerca del dominio de contenidos y de las estrategias de evaluación del aprendizaje utilizadas por sus docentes, destacaron que éstos últimos promueven principalmente el aprendizaje memorístico, en desmedro de otros objetivos de enseñanza. Mientras que una opinión positiva al respecto por parte del estudiantado, se asoció con objetivos de enseñanza de índole constructivistas, tales como la promoción del aprendizaje significativo y situado.

Los hallazgos anteriores adquieren especial relevancia considerando el contexto de pandemia que imperó a nivel mundial durante el período 2020-21 y afectó a todos los niveles educativos, obligando al trabajo a distancia por medio de recursos digitales para implementar los procesos de enseñanza y aprendizaje en todos los niveles educativos. En efecto, esta virtualización forzosa trastocó insoslayablemente las estrategias didácticas en torno a los contenidos educativos y, por ende, los objetivos tanto implícitos como explícitos, así como los modelos de enseñanza y aprendizaje a nivel de educación superior, en todas las áreas de conocimiento ([Castro, Paz & Cela, 2020](#); [Valenzuela & Barrios, 2020](#); [Miguel Román, 2020](#); [Cencia Crispín et al., 2021](#); [Moreno Tapia, Quintero López & Cáceres Mesa, 2021](#); [Mamani-Cori, Padilla, Cervantes, Caballero & Sucari, 2022](#)).

Estas transformaciones didácticas afectaron de sobremanera a aquellos docentes neófitos en el dominio de estrategias soportadas por medios digitales, dificultando su incorporación a este acelerado cambio de modelo. Lo anterior quedó en evidencia dentro de la opinión de los estudiantes de la FCAYS de la UABC durante el período 2021-2, ya que más del 40% aseveró que sus docentes continuaron trabajando sus clases a través de una dinámica sincrónica durante el contexto de COVID-19, lo que implicó la continuación de formas de trabajo similares a las utilizadas durante clases presenciales. Al respecto, algunos autores ([Valenzuela & Barrios, 2020](#); [Moreno Tapia et](#)

[al., 2021](#)) destacan la necesidad de diseñar e implementar de manera sistemática capacitaciones teórico-prácticas a docentes universitarios en torno al dominio de metodologías activas, interactivas y variadas en el marco de procesos educativos virtuales, fomentando la motivación, el aprendizaje autónomo, colaborativo y la metacognición en los estudiantes.

Lo anterior se justifica a partir de los resultados expuestos en el presente artículo, en el cual se evidencia que aquellos estudiantes de la FCAYS de la UABC que se encuentran cursando la etapa formativa inicial de sus licenciaturas y que tuvieron una opinión negativa acerca del dominio de contenidos, así como de las estrategias de evaluación del aprendizaje utilizadas por sus docentes en contexto de COVID-19, destacaron el fomento de estrategias de aprendizaje individuales y memorísticas principalmente por parte de sus docentes, en desmedro de estrategias constructivistas. A su vez, como lo aseveran Cencia Crispín et al. (2021), un error recurrente en muchos docentes a nivel universitario es implementar procesos de enseñanza-aprendizaje en línea y virtuales bajo estrategias didácticas similares a aquellas utilizadas en contextos de presencialidad dentro del aula, ignorando sus diferencias sustanciales. En este sentido, resulta preocupante que los docentes sigan utilizando en la educación virtual las mismas estrategias e instrumentos que utilizan en la enseñanza presencial ([Villarroel et al., 2021](#)).

A su vez, como lo destacan algunos autores ([Vialart Vidal, 2020](#); [Mendoza Castillo, 2020](#); [Cencia Crispín et al., 2021](#); [Villarroel et al., 2021](#)), estas modificaciones y transformaciones que trajo consigo la virtualización de la educación afectó principalmente a aquellos sujetos carentes de conocimientos tecnológicos, de equipamiento o recursos tecnológicos y de disponibilidad o acceso a redes desde su hogar. Y no sólo en lo referente a carencias tecnológicas: como también lo destacan [Moreno Tapia et al. \(2021\)](#), la contingencia sanitaria afectó negativamente a muchos estudiantes que no contaban con espacios ideales de concentración y atención para la realización de sus obligaciones académicas, lo cual sumado a limitaciones de disponibilidad de recursos digitales, provocó serios obstáculos en sus procesos de

aprendizaje y en la efectividad de la enseñanza. Por lo tanto, también se debe considerar un factor que va más allá de lo tecnológico y que se relaciona con las condiciones socioeconómicas que afectan los contextos educativos virtuales mediados por recursos digitales. Y dado que la contingencia sanitaria a causa de la propagación del COVID-19 obligó a virtualizar los procesos de enseñanza-aprendizaje en todos los niveles educativos, no se debe soslayar este elemento a la hora de analizar los factores determinantes que afectaron dichos procesos en este contexto.

A su vez, hay que destacar que el presente estudio se vio afectado por algunas limitaciones de índole metodológicas, a saber: el tipo de muestreo no probabilístico que no permite generalizar los resultados a toda la población de estudiantes de la FCAYS de la UABC; los bajos índices de confiabilidad que arrojaron algunas dimensiones del instrumento, sobre las cuales se trazan algunas interpretaciones; y la limitada cantidad de participantes, que impidió realizar una limpieza de datos relacionados con sesgos de respuesta. Sin embargo, respecto a estas limitaciones, se debe recordar que el cuestionario utilizado representa una adaptación circunstancial para evaluar el desempeño docente en contexto de pandemia. El instrumento original se volvió a aplicar una vez terminado el confinamiento forzoso y se encuentra en proceso de retroalimentación, validación y aplicación para procesos de enseñanza y aprendizaje presenciales.

En conclusión, el presente estudio aporta elementos para analizar, desde la perspectiva del estudiantado universitario, algunos aspectos en torno al desempeño de sus docentes con contexto de clases virtuales por la pandemia COVID-19. El análisis mostró que, según la opinión del estudiantado, los objetivos de enseñanza de los docentes se bifurcan en fines individuales y memorísticos, por un lado, o en propósitos de índole más constructivistas (significativo, situado, grupal), por otro lado. Lo anterior se relacionó de manera significativa con ciertas variables académicas de los estudiantes (etapa formativa, opinión acerca del dominio de contenidos y de las estrategias de evaluación del aprendizaje utilizadas por sus docentes), a partir de lo cual queda abierto el

camino para seguir explorando estas tendencias. Esto principalmente en función de las modalidades educativas en que se imparten clases a nivel de educación superior, considerando que el contexto de educación virtual obligatoria ya fue superado y que actualmente la modalidad de educación presencial ha vuelto a retomarse en todos los países y niveles educativos.

Referencias

- Arreola, R.A. (2007). *Developing a comprehensive faculty evaluation system: A guide to designing, building, and operating large-scale faculty evaluation systems* (3rd. Ed.). Bolton, MA: Anker.
- Barrios, M. & Cosculluela, A. (2013). *Fiabilidad*. Barcelona: Universitat Oberta de Catalunya. https://openaccess.uoc.edu/bitstream/10609/69325/3/Psicometria-C3%ADa_M%C3%B3dulo%202_%20Fiabilidad.pdf
- Bazán-Ramírez, A., Pérez-Morán, J. C., and Bernal-Baldenegro, B. (2021). Criteria for Teaching Performance in Psychology: Invariance According to Age, Sex, and Academic Stage of Peruvian Students. *Front. Psychol.* 12:764081. doi:10.3389/fpsyg.2021.764081
- Cárdenas Castro, M. & Arancibia Martini, H. (2014). Potencia estadística y cálculo del tamaño del efecto G*Power: complementos a las pruebas de significación estadística y su aplicación en psicología. *Salud & Sociedad*, 5(2), 210-224. <https://dialnet.unirioja.es/descarga/articulo/4945415.pdf>
- Castro, M., Paz, M. & Cela, E. (2020). Aprendiendo a enseñar en tiempos de pandemia COVID-19: nuestra experiencia en una universidad pública de Argentina. *Revista Digital de Investigación en Docencia Universitaria*, 14(2), 1-11. http://www.scielo.org.pe/scielo.php?pid=S2223-25162020000200006&script=sci_abstract
- Cencia Crispín, O., Carreño Colchado, M., Eche Querevalú, P., Barrantes Morales, I. & Cárdenas Baldeón, G. (2021). Estrategias docentes de profesores universitarios en tiempos de COVID-19. *Horizonte de la Ciencia*, 11(21), 347-360. <https://www.redalyc.org/journal/5709/570967307025/html/>
- Cordero Arroyo, G., Luna Serrano, E. y Patiño Alonso, N.X. (2013). La evaluación docente en educación básica en México: panorama y agenda pendiente. *Revista Electrónica Sinéctica*, número 41, 1-19. <http://www.>

- redalyc.org/pdf/998/99828325007.pdf
- Cruz Ávila, M. (2007). *Una propuesta para la evaluación del profesorado universitario*. Trabajo de investigación para la obtención del grado de doctor, Facultad de Ciencias de la Educación, Departamento de Pedagogía Sistemática y Social, Universitat Autònoma de Barcelona. <https://www.tdx.cat/bitstream/handle/10803/5285/mca1de1.pdf>
- De Vet, H., Mokkink, L., Mosmuller, D. & Terwee, C. (2017). Spearman–Brown prophecy formula and Cronbach's alpha: different faces of reliability and opportunities for new applications. *Journal of Clinical Epidemiology*, 85, 45-49. <https://doi.org/10.1016/j.jclinepi.2017.01.013>
- Fink, D. (2008). "Evaluating teaching: A new approach to an old problem" in To improve the academy: Resources for faculty, instructional, and organizational development. eds. S. Chadwick-Blossey and D. R. Robertson (Jossey-Bass)
- García Cabrero, B., Loredó Enríquez, J. y Carranza Peña, G. (2008). Análisis de la práctica educativa de los docentes: pensamiento, interacción y reflexión. *Revista Electrónica de Investigación Educativa, Especial*. <http://redie.uabc.mx/redie/article/viewFile/200/345>
- Henríquez, P., Arámburo, V. & Dávila, E. (2017). Percepción de los estudiantes universitarios acerca de las estrategias pedagógicas y de evaluación del aprendizaje utilizadas por sus profesores: el caso de la FCAYS de la UABC. *Memorias del XIV Congreso Nacional de Investigación Educativa, COMIE*. <https://www.comie.org.mx/congreso/memoriaelectronica/v14/seccion4.htm>
- Henríquez, P., Arámburo, V., & Boroel, B. (2018). Análisis de las estrategias de enseñanza según áreas de conocimiento en el nivel educativo superior: Percepciones de estudiantes de la Facultad de Ciencias Administrativas y Sociales (FCAYS), Universidad Autónoma de Baja California (UABC), México. *Compendio Investigativo de Academic Journals Celaya 2018*, tomo 14, 2320-2325. <https://www.academiajournals.com/pub-celaya-2018>
- Henríquez, P. & Arámburo, V. (2021). Evaluación del desempeño docente por áreas de conocimiento: el caso de la Facultad de Ciencias Administrativas y Sociales de la Universidad Autónoma de Baja California, México. *Revista Actualidades Investigativas en Educación*, 21(3), 1-20. <https://revistas.ucr.ac.cr/index.php/aie/article/view/46294>
- Khalilzadeh, J. & Tasci, A.D. (2017). Large sample size, significance level, and the effect size: Solutions to perils of using big data for academic research. *Tourism Management*, 62, 89-96. <https://doi.org/10.1016/j.tourman.2017.03.026>
- López-Aguado, M. & Gutiérrez-Provecho, L. (2019). Cómo realizar e interpretar un análisis factorial exploratorio utilizando SPSS. *REIRE, Revista d'Innovació i Recerca en Educació*, 12(2), 1-14. <http://doi.org/10.1344/reire2019.12.227057>
- Mamani-Cori, V., Padilla, T., Cervantes, Sh., Caballero, L. & sucari, W. (2022). Estrategias y recursos didácticos empleados en la enseñanza/aprendizaje virtual en estudiantes universitarios en el contexto de la COVID-19. *Revista Innova Educación*, 4(1), 78-91. <https://revistainnovaeducacion.com/index.php/rie/article/view/465>
- Martínez, F. (2016). *La evaluación de los maestros y la calidad educativa [Versión electrónica]*. En *La evaluación de docentes de educación básica. Una revisión de la experiencia internacional*. México: Instituto Nacional para la Evaluación de la Educación. <https://www.inee.edu.mx/wp-content/uploads/2018/12/P1C233.pdf>
- Mendoza Castillo, L. (2020). Lo que la pandemia nos enseñó sobre la educación a distancia. *Revista Latinoamericana de Estudios Educativos*, 50(número especial), 343-352. <https://rlee.iberomx/index.php/rlee/article/view/119/492>
- Miguel Román, J. (2020). La educación superior en tiempos de pandemia: una visión desde dentro del proceso formativo. *Revista Latinoamericana de Estudios Educativos*, 50(número especial), 13-40. <https://www.redalyc.org/jatsRepo/270/27063237017/html/index.html>
- Moreno Tapia, J., Quintero López, I., & Cáceres Mesa, L. (2021). Estrategias de aprendizaje de estudiantes universitarios durante la pandemia. *Revista de Educación, Cooperación y Bienestar Social*, número 19, 179-183. <https://revistadecooperacion.com/numero19/19-15.pdf>
- Organista Sandoval, J. & Serrano Santoyo, A. (2014). Aspectos de posesión, permisos y usos educativos de dispositivos portátiles durante el trayecto de primaria a universidad. *Apertura* 6(2), 1-11. <https://www.redalyc.org/articulo.oa?id=68835725004>
- Oviedo, H. & Campo-Aria, A. (2005). Aproximación al uso del coeficiente alfa de Cronbach. *Revista Colombiana de Psiquiatría*, 34(4), 572-580. <https://www.redalyc.org/articulo.oa?id=80634409>
- Romero Díaz, T. & Martínez Gimeno, A. (2017). Construcción de

- instrumentos de evaluación del desempeño docente universitario desde una perspectiva cualitativa. *Revista Universitaria del Caribe*, 18(1), 34-42. https://www.researchgate.net/publication/320210350_Construccion_de_instrumentos_de_evaluacion_del_desempeno_docente_universitario_desde_una_perspectiva_cualitativa
- Santos del Real, A. (2012). Evaluación docente. *Educacion Química*, 23(2), 200-204. <https://www.scielo.org.mx/pdf/eq/v23n2/v23n2a5.pdf>
- UNESCO (2006). *Evaluación del desempeño y carrera profesional docente. Un estudio comparado entre 50 países de América y Europa*. Oficina Regional de Educación para América Latina y el Caribe, Santiago, Chile. <http://unesdoc.unesco.org/imagenes/0015/001529/152934s.pdf>
- Valenzuela, L. & Barrios, C. (2020). Caracterización de de las estrategias de la enseñanza universitaria y la actitud del profesorado hacia la innovación en tiempos de pandemia. *Ciencia Latina Revista Científica Multidisciplinar*, 4(2), 843-859. <https://ciencialatina.org/index.php/cienciala/article/view/121>
- Vázquez Cruz, M. A., Cordero Arroyo, G. y Leyva Barajas, Y. E. (2014, septiembre-diciembre). Análisis comparativo de criterios de desempeño profesional para la enseñanza en cuatro países de América [Versión electrónica]. *Revista Electrónica Actualidades Investigativas en Educación*, 14(3), 1-20. <https://www.redalyc.org/articulo.oa?id=44732048018>
- Vialart Vidal, M. (2020). Estrategias didácticas para la virtualización del proceso de enseñanza aprendizaje en tiempos de COVID-19. *Educación Médica Superior*, 34(3), 1-10. http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21412020000300015
- Villarroel, V., Pérez, C., Rojas-Barahona, C. & García, R. (2021). Educación remota en contexto de pandemia: caracterización del proceso educativo en las universidades chilenas. *Formación Universitaria*, 14(6), 65-76. https://www.scielo.cl/scielo.php?pid=S0718-50062021000600065&script=sci_arttext
- Zamora Serrano, E. (2021). La evaluación del desempeño docente mediante cuestionarios en la universidad: su legitimidad según la literatura y los requerimientos para que sea efectiva. *Revista Actualidades Investigativas en Educación*, 21(3), 1-23. <https://www.scielo.sa.cr/pdf/aie/v21n3/1409-4703-aie-21-03-00726.pdf>
- Zapata-Ros, M. (2015). Teorías y modelos sobre el aprendizaje en entornos conectados y ubicuos. Bases para un nuevo modelo teórico a partir de una visión crítica del conectivismo. *Education in the Knowledge Society*, 16(1), 69-102. <https://www.redalyc.org/pdf/5355/535554757006.pdf>

Relación de la prueba Saber 11, examen de admisión, promedio académico, prueba saber pro de estudiantes del programa de psicología de la Universidad del Magdalena

Amanda Miguel Iguarán Jiménez^{*1}; María Fernanda Cabas-Manjarrés²; Carmelina Paba Barbosa³; Paola Diazgranados Rincones⁴

^{1, 2, 3, 4}Universidad del Magdalena, Santa Marta, Colombia^a ¹<https://orcid.org/0000-0002-5658-6560> amanjosemiguel0222@hotmail.com
²<http://orcid.org/0000-0002-0359-6669> mcabasm@unimagdalena.edu.co ³<https://orcid.org/0000-0002-9928-2970> carmelinapaba@gmail.com ⁴<https://orcid.org/0000-0003-2499-1235> paoladiazgranadosrin@gmail.com

Citar como: Iguarán Jiménez, A., Cabas-Manjarrés, M., Paba Barbosa, C., & Diazgranados Rincones, P. (2023). Relación de la prueba Saber 11, examen de admisión, promedio académico, prueba saber pro de estudiantes del programa de psicología de la Universidad del Magdalena. *Revista Digital de Investigación en Docencia Universitaria*, 17(2), e1421. <https://doi.org/10.19083/ridu.2023.1421>

Recibido: 29/01/2021 **Revisado:** 12/04/2021. **Aceptado:** 02/11/2021. **Publicado:** 30/06/2023.

Resumen

Introducción: la calidad de la educación superior en Colombia se mide a través de los indicadores que arrojan las Pruebas Saber-Pro, este examen evalúa competencias genéricas y específicas del futuro profesional en psicología, por ello, en esta investigación se realizó el análisis de la relación entre el promedio académico, los resultados obtenidos en el examen de admisión, las Pruebas Saber 11, las Pruebas Saber Pro y promedio académico de los estudiantes de Psicología de la Universidad del Magdalena. **Métodos:** se seleccionaron todos los estudiantes que fueron admitidos entre el año 2012 y el 2016, siendo 246 del género femenino y 86 del género masculino, se utilizó análisis multivariado. **Objetivo:** establecer la relación con el rendimiento en la prueba estandarizada SABER PRO para psicología. **Resultados:** el análisis que se realizó fue con la cohorte de ingreso 2012 por Ji-cuadrado del Saber-Pro y las variables de los puntajes y las sociodemográficas, encontrando una asociación estadísticamente significativa entre los cuartiles de Saber Pro y los cuartiles de los puntajes de admisión.

Palabras clave: Educación superior, Saber Pro, Saber 11, Colombia, desempeño académico.

List of the Saber 11 Test, admission exam, academic average, saber protest of students of the psychology program of the University of Magdalena

Abstract

Introduction: The quality of higher education in Colombia is measured through the indicators that the Saber-Pro Tests show, this exam evaluates generic and specific competencies of the future professional in psychology, therefore, in this research the analysis of the relationship between the academic average, the results obtained in the admission exam, the Saber 11 Tests, the Saber Pro Tests and the academic average of Psychology students at the University of Magdalena. **Methods:** All students who were admitted between 2012 and 2016 were selected, being 246 female and 86 male, multivariate analysis was used. **Objective:** to establish the relationship with performance in the SABER PRO standardized test for psychology. **Results:** The

*Correspondencia:

Amanda Miguel Iguarán Jiménez
amanjosemiguel0222@hotmail.com



analysis that was carried out was with the 2012 admission cohort by Chi-square of Saber-Pro and the variables of the scores and sociodemographics, finding a statistically significant association between the Saber Pro quartiles and the quartiles of the scores of admission.

Keywords: Higher education, Saber Pro, Saber 11, Colombia, academic performance.

Introducción

En Colombia, el gobierno central ha establecido que las instituciones universitarias deben esforzarse por lograr una educación superior de alta calidad mediante la aplicación de los exámenes Sabre-Pro a los estudiantes. El propósito de estos exámenes estatales es medir la calidad de la educación superior para tratar de evaluar la caad de los servicios de educación pública y su función en la inspección y seguimiento ([Delgado, 2013](#)). Esta herramienta de evaluación propuesta por el Ministerio de Educación (MEN) es utilizada para evaluar habilidades y desempeños de los futuros profesionales en Colombia, por ello a partir del 2002 el MEN comienza aplicar esta prueba por primera vez con estudiantes de las carreras de Ingeniería Mecánica, derecho y medicina, estas carreras realizaron las pruebas en el segundo año correspondiente, y los resultados obtenidos ayudaron a conocer las ventajas y desventajas de los diferentes cursos de grado que participan en la evaluación, el ICFES ([Instituto Colombiano para el Fomento de la Educación Superior](#)) convoca a las universidades de todo el país, para analizar las facultades y sus profesionales y demás miembros de la comunidad académica, para que en conjunto reformularán una prueba que fuera aplicada a todas las áreas del conocimiento, entre ellas: las ciencias de la salud; arquitectura, urbanismo, ciencias sociales y humanas.

En el segundo año posterior a este encuentro se sumaron otros programas académicos, en 2004 participaron en esta evaluación 42 programas de diferentes universidades de Colombia. En la actualidad, 55 carreras de pregrado de diferentes universidades cumplen con los requisitos del examen estatal. Hasta el momento, el Instituto Colombiano para la Promoción de la Educación Superior ([ICFES, 2012, 2016a, 2016b y 2016c](#)) sigue aplicando el examen normalmente a 55 programas. Con el fin de

buscar el cumplimiento regular del cronograma establecido desde 2003. En el proceso, el examen ha sido revisado muchas veces por diferentes instituciones educativas del país.

De las revisiones, el MEN, en el 2009 realizó replanteamientos en la estructura de la prueba. En la Actualidad el ICFES cuenta con una nueva estructura gracias al trámite legal que pudo realizar el Congreso de la República, creando una nueva ley que regulara el examen de estado. Por ello el Ministerio de Educación Nacional expide nuevos estándares de educación superior para que, con base en ellos, se evalúe la estructura de los ECAES ([Delgado, 2013](#) y [Colombia Aprende, 2017](#)).

Por lo tanto, el gobierno reformó el ECAES (Exámenes de Calidad de Educación Superior) en el examen Sabre-Pro (Saber Profesional). Con esta nueva reorganización ha podido lograr el MEN (Ministerio de Educación Nacional), la mejor forma de evaluar la formación profesional que han adquirido los estudiantes a lo largo del proceso de aprendizaje, y de enfatizar los resultados de ECAES como un aporte para reajustar la política educativa y mejorar la calidad de la educación superior y promover la toma de decisiones en estos temas.

Estos resultados no son más que unos indicadores de desempeño por competencias, y es por ello que [Torres, et al. \(2008\)](#), expresan su preocupación al elevar las voces de algunos estudiantes, que consideran que la evaluación debe tomar más relevancia en las expresiones educativas y pedagógicas, los autores creen que *investigar en evaluación es comprender una cultura, una forma de leer y de proyectar el mundo* (p. 117), esto es solo una expresión del surgimiento del sistema de evaluación, el cual permitirá orientar a las universidades sobre la importancia de la evaluación, puesto que estas no tienen idea sobre la esencia de lo que se evalúa y para qué se evalúa en educación superior.

Específicamente, la evaluación a gran escala de la educación superior comenzó en 2003, primero aplicando ECAES a 27 programas diferentes y luego a 55 programas en 2008, por lo tanto, los resultados despertaron el interés de la gente por el aprendizaje. La primera etapa, así como los nuevos requisitos legales, motivaron la reestructuración de estos exámenes, que desde 2009 se denominan SABER PRO y son requisitos obligatorios para las carreras de

grado. Esta es la primera vez que el ICFES (2016a) ha proporcionado resultados comparables para toda la educación superior ([ICFES, 2016b; Oficina asesora de comunicaciones, 2013 y ICFES, 2016c](#))

En este sentido, el [Tirado \(2017\)](#) junto con el ICFES, unen esfuerzos para posibilitar la construcción de ciudadanía, haciendo hincapié en las competencias genéricas y específicas, por medio de los conocimientos y habilidades que se enmarcan en la comprensión del entorno y en la coexistencia inclusiva dentro del marco que propone la constitución política de Colombia, por tanto, el ejercicio de la ciudadanía debe basarse en la participación activa de los ciudadanos y sus comunidades, pues un ciudadano capaz es una persona que comprende su entorno social y político; comprende sus derechos y obligaciones; además de interesarse por los asuntos comunitarios, desarrolla la capacidad de reflexionar sobre los problemas sociales; participar en la búsqueda de soluciones a los problemas sociales y buscar el bienestar de su comunidad.

Es aquí, donde lo expuesto, cobra cada vez más relevancia, atendiendo la necesidad de mejorar cada vez más, los procesos de evaluación de las competencias específicas de los estudiantes de pregrado del programa de psicología, esto se hace, a través del análisis de problemáticas psicológicas, lo cual el [MEN \(2010\)](#), explica, que la propuesta consiste en identificar los niveles de comprensión que tienen los futuros psicólogos, teniendo en cuenta los factores biológicos y socioculturales, y el nivel de comprensión del proceso de evaluación e intervención de problemas psicológicos en los entornos social, clínico y sanitario, educativo y organizativo.

Evaluar estas habilidades mediante el diseño y la realización de preguntas de intervención psicológica, se necesita que los estudiantes identifiquen problemas relacionados a situaciones vulnerables psicológicas, definir el propósito de la intervención, definir estrategias de implementación y evaluación y comunicar los resultados ([ICFES, 2016](#)). Por tanto, los nuevos exámenes están diseñados para evaluar dos tipos de competencias: las genéricas y las específicas (rigen desde 2009 para los Exámenes de Estado de la Educación Superior – Ley 1323 y Decreto 3963– que señala la necesidad de adecuar el diseño de los instrumentos a los parámetros establecidos en la Ley).

Hasta aquí la tarea sigue siendo ardua, ya que es importante evaluar los factores intrínsecos de la evaluación, pero el ICFES, el cual propone una alternativa para el diseño de un examen. En este asunto el ASCOFAPSI (Asociación Colombiana de Facultades de Psicología) asume el reto y elabora un documento inicial sobre las habilidades básicas, en un 75% de los estudiantes en cada curso académico, el cual se puede analizar desde la perspectiva de las habilidades generales que han desarrollado hasta ahora en el examen y desde otros campos o cursos de formación (Texto de la convocatoria que realizó el ICFES en 2010 por ASCOFAPSI, 2012, citado por [Tirado, 2017](#)). Por lo tanto, ASCOFAPSI es la entidad seleccionada por el ICFES, para adelantar una propuesta sobre las competencias fundamentales de los estudiantes, convocando un equipo de académico que elabore una propuesta para ser sometida a discusión por parte del ICFES.

De esta propuesta surgen unas reformas en los exámenes Saber Pro y son: que deben ser obligatorios, evaluar competencias genéricas como la lectura crítica, el razonamiento cuantitativo, la cultura ciudadana y el entendimiento del entorno, la redacción y el inglés y competencias específicas relacionadas con los problemas comunes a distintos programas, más que contenidos curriculares específicos los cuales usan para evaluar estudiantes de distintos programas, incluso de diversas áreas del conocimiento comunes a grupos de programas, horario de aplicación de la prueba y módulos que serán obligatorios para estudiantes, que cumplan con las características mínimas y con la presentación de resultados.

La prueba ha venido evolucionando a nivel conceptual y práctico y sigue siendo un tema de interés por la deficiencia en los resultados que varían cada año, aunque existen estudios que fundamentan el impacto de revisar la relación del puntaje del ICFES, puntaje de admisión, promedio académico y pruebas Saber Pro, surgen interrogantes que se relacionan con los factores que inciden en los bajos desempeños de la prueba a nivel nacional.

Otro aspecto de la evaluación son los esfuerzos del Ministerio de Educación para recomendar becas entre universidades del departamento elegido por Colciencias para jóvenes investigadores. Por ello, el ICFES sigue posicionado la prueba

Sabre Pro como una herramienta para medir la calidad de la educación superior y el nivel de habilidad de los estudiantes bajo estándares internacionales ([MEN, 2010](#)).

Estos intentos evaluativos por parte del Estado, han propiciado diversas investigaciones, para [León, et al. \(2012\)](#), la relación entre comprensión lectora, inteligencia y desempeño en pruebas Saber Pro en una muestra de estudiantes universitarios del Programa de Psicología de la Universidad de la Costa - CUC, no existe una correlación significativa entre los puntajes obtenidos en comprensión lectora de la prueba Saber Pro y los resultados de las pruebas que evalúan memoria operativa (PHC-Comprensión lectora y RIAS) e inteligencia verbal y general.

Además de la evaluación de la memoria y la comprensión lectora, otros autores ([Gil, Rodríguez, Sepúlveda, Rondón y Gómez, 2013](#)), consideran los resultados en la prueba nacional de calidad de la educación superior (Saber Pro). Los autores utilizaron la información de 4.498 estudiantes de medicina evaluados en la prueba Saber Pro, en 2009, utilizando un modelo jerárquico, que evaluaba 40 facultades de medicina, la correlación entre las características de estudiantes y universidades y los puntajes obtenidos en la prueba, encontraron puntajes promedio por universidades de la prueba Saber Pro 2009 para estudiantes de medicina con un índice de 100,4, en un rango entre 84,3 y 110,8 puntos y la variabilidad de los puntajes en la prueba fue explicada en un 29% por las diferencias entre universidades, por lo que concluyen que las universidades oficiales y las facultades de medicina con propios hospitales universitarios obtuvieron mejores resultados en promedio, y también que la provisión de cursos profesionales médicos estaba asociada con puntajes más bajos.

[Bahamón y Reyes \(2014\)](#), realizaron una caracterización de la capacidad intelectual, factores sociodemográficos y académicos de estudiantes con alto y bajo desempeño en los exámenes Saber Pro (2012) en la Universidad Simón Bolívar, las autoras destacaron que los puntajes reportados por el Instituto Colombiano para el Fomento de la Educación Superior (ICFES) de los periodos comprendidos entre el I y II de 2012 y los datos reportados por el programa de psicología de la Universidad

en una muestra de 68 estudiantes que cursaban noveno y décimo semestre en edades entre los 21 y 52 años, aplicando estadígrafos descriptivos y prueba T-Student, determinaron la capacidad intelectual de los estudiante y la mayoría de ellos se clasificaron con puntajes correspondientes al término medio (36.8%) y superior al medio (30.9%) mientras que los menores puntajes los ubicaron en los términos de deficiente (4.4%) y superior (1.5%), en cuanto al análisis de los datos en función del desempeño en las pruebas Saber Pro, los estudiantes con mejor rendimiento se ubicaron en término superior al medio (60%) y medio (30%), a diferencia de los datos de estudiantes con menor desempeño, esta distribución se ubicó en la categoría término medio e inferior al término medio (31%), relacionado con características socio-demográficas, las autoras concluyeron que los datos en función del desempeño en las pruebas Saber Pro con estudiantes con mejor rendimiento están en su mayoría en término superior al medio (60%) y medio (30%), por tanto, los estudiantes con mejor desempeño tienen características sociodemográficas específicas, capacidad intelectual superior, buen desempeño académico, intereses profesionales relacionados con las áreas de la carrera y adecuados hábitos y técnicas de estudio.

Lo anterior, podría indicar que los estudiantes que tienen mejor desempeño académico obtienen mejores puntajes en la prueba de estado en educación superior, por ello [Isáziga, et al \(2014\)](#), a través del apoyo del Instituto Colombiano para la Evaluación de la Educación (ICFES), creen que la intervención académica es la construcción de una sociedad con calidad, ellos proponen realizar un análisis del valor agregado en el proceso formativo colombiano, los autores tuvieron en cuenta el análisis de los resultados de la evaluación del examen Saber11 y SaberPRO por regiones colombianas (Corpes), basándose en la metodología del valor agregado (VA), este estudio evidenció que en el componente de comprensión lectora y lenguaje en todas las regiones tenían valores por debajo de 0,05 (grado de significancia), lo que indica que la correlación es no cero (existe asociación), con un 95% de nivel de confiabilidad, los autores analizaron una asociación significativa entre el %efecSaber11 y su posterior %efecSaberPRO; lo que resalta además la

inexistencia de una correlación muy grande entre los dos desempeños, en las regiones Corpes que analizaron existe un notable mejoramiento del porcentaje de efectividad máxima antes y después de la formación universitaria (%efecSaberPRO fue superior a %efecSaber11), por tanto surge el interés de seguir realizando este tipo de estudios.

Estos dos puntos coyunturales, establecen un punto de encuentro para comprender el porqué de la evaluación de los desempeños o las deficiencias de los puntajes, por ello [Ramírez \(2014\)](#), analiza otros factores sociodemográficos (socio-económicos, educativo, nivel de formación y género) que podrían estar asociados con el desempeño académico de los estudiantes que presentan la prueba Saber Pro, la autora para identificar la asociación, utilizó las bases de datos suministradas por el ICFES, de las cuales conformó una base de datos de trabajo con 4.031 observaciones, 2.145 son mujeres (53.2%), 1.886 son hombres (46.8%) y el 95% de ellos son solteros, la edad promedio fue de 21 años, el tipo de prueba presentada fueron prueba Saber Pro 2009 y prueba Saber 11 en 2005 y 2006, la metodología que utilizó fue de tipo descriptivo utilizando distribuciones de frecuencia y un análisis inferencial y pruebas de hipótesis, para verificar la independencia entre variables, además implementó un análisis de regresión lineal múltiple para determinar los factores asociados al desempeño académico, según género y nivel de formación, de los estudiantes evaluados con la prueba Saber Pro 2009.

En los resultados la autora encontró que según el nivel de formación, el 15% son estudiantes de programas técnicos profesionales, 63% son de programas tecnológicos y 22% son de programas profesionales, otro dato relevante del estudio realizado por Ramírez fue que el 89% de la población colombiana pertenece a los estratos uno, dos y tres ([Conpes 3386, 2005, citado por Ramírez, 2014](#)), mientras que en la población que ella utilizó fueron 84% estudiantes que pertenecen a esos estratos y el porcentaje restante al 15% que se distribuye en los tres estratos más altos, el nivel de educación de las madres es comparativamente mayor que la de los padres, ya que el 69% de los padres de los estudiantes no han realizado estudios de nivel superior o no tiene ningún tipo de estudios o tiene algún tipo de estudios

primarios o secundarios; mientras que el porcentaje para las madres es de 74%, en conclusión este estudio expone que sin importar el nivel de formación o el género de los estudiantes, el desempeño académico previo, medido en el resultado de las prueba Saber 11 en las áreas de química, física, matemáticas, lenguaje, filosofía, biología e idiomas, está fuertemente asociado con el desempeño académico de los estudiantes que presentaron la prueba Saber 2009 y que el desempeño académico previo medido en los puntajes obtenidos en las pruebas Saber 11 está asociado fuerte y positivamente con el de desempeño de los estudiantes sin importar su género o nivel de formación y en las variables socio-económicas y educativas en comparación con el desempeño previo, resultaron estar débilmente asociadas con el desempeño en la educación superior.

También, [Timarán, et al \(2016\)](#), indagaron sobre las tendencias del desempeño académico en competencias genéricas-pruebas Sabe Pro, efectuando un análisis estadístico preliminar de los datos con el fin de obtener las tendencias de desempeño académico de los estudiantes, en las cuatro competencias genéricas de las pruebas Saber Pro 2011-2, en cuanto a las variables sociodemográficas, económicas, académicas e institucionales de los estudiantes, el análisis de medidas estadísticas arrojó una comparación de los puntajes en habilidades generales. El resultado mostró una alta correlación en la lectura crítica del razonamiento cuantitativo y la lectura crítica del inglés ($r > 0.5$), la lectura escrita y la lectura crítica del razonamiento cuantitativo en inglés muestran una correlación moderada ($0.3 < r \leq 0.5$), y la correlación entre el razonamiento cuantitativo y las habilidades de expresión escrita de la comunicación escrita en inglés es muy baja ($r \leq 0$; 3), en términos de género y desempeño académico en habilidades generales, los hombres superan a las mujeres en todas las habilidades, excepto en la habilidad para escribir (A: alto), en términos de estado civil, las mujeres se desempeñan mejor que los hombres en habilidades generales; los solteros obtienen mejores resultados en todas las pruebas (A), el desempeño en casados, separados o divorciados es bajo (B), entre las cuatro habilidades, los viudos y los estudiantes que están en convivencia mostraron un desempeño B en lectu-

ra y escritura críticas, mientras que en el módulo de razonamiento cuantitativo, los estudiantes solteros, estaban en el grupo intermedio B y en el último grupo C.

A parte, [Betancourt y Frías \(2015\)](#), cambian de perspectiva y se enfocan en las competencias específicas por área de desempeño, tal es el caso de las competencias argumentativas de los estudiantes de derecho en el marco de las pruebas Saber-Pro en la Universidad Sergio Arboleda de la ciudad de Bogotá, los investigadores encontraron que en las competencias genéricas y las competencias ciudadanas tuvieron un promedio de III (representa un nivel medio en una escala de I a V), en la escritura 5 (en una escala de 1 a 8; se aprecia un buen uso del lenguaje, aunque pueden encontrarse errores en la aplicación de algunas reglas de ortografía), en inglés A1 (es capaz de comprender y utilizar expresiones cotidianas) y en lectura crítica III, en contraste con lo anterior los autores compararon los resultados con el periodo 2012-2 con otros estudiantes (81 participantes) del Programa de Derecho y en el análisis de las competencias los autores resaltaron las competencias ciudadanas, los estudiantes obtuvieron un promedio de II (representa un nivel bajo en una escala de I a V), en la escritura un promedio de 5, en inglés un promedio A1, en lectura crítica un promedio de II y en razonamiento cuantitativo un promedio quintil II – 40 %–, esto indica que los puntajes en lectura crítica bajaron y los demás se mantuvieron, en conclusión los estudiantes necesitan una mejor formación lingüística, y esto se da mejorando el currículo, reestructurar los modelos evaluativos del estado y las políticas de cualificación estudiantil y mejorar los referentes teóricos de la carrera, ya que según ellos esto conllevará a que la orientación pedagógica posibilite, rescatar y validar el criterio de un educando como ciudadano y como profesional en formación.

La Universidad de Nariño, evaluó el valor agregado relacionado con las pruebas saber 11 y saber pro-2010-2014, y encontraron que la universidad generó un valor agregado de un 10% a los estudiantes en el periodo estudiado, esto indica la efectividad educativa y el impacto de la enseñanza de los maestros que han tenido durante el progreso de la carrera universitaria ([Burgos, Ruales, Bastidas y Ortiz, 2019](#)).

De lo anterior, en esta investigación se propuso interpretar la relación de los resultados obtenidos en el examen de admisión, acumulado del 1er semestre, y Saber Pro con el promedio académico de los estudiantes de los años 2012 al 2016, con el fin de analizar el valor que tienen algunas variables que se analizaron en el estudiante antes de iniciar sus estudios universitarios, permitiendo establecer la relación con el rendimiento en la prueba estandarizada SABER PRO para psicología. Este análisis podría resultar de interés para seleccionar candidatos en los procesos de admisión dadas las características tanto académicas como personales de entrada a la universidad.

Método

Diseño

Es de tipo cuantitativo ya que se fundamenta en cuantificar los fenómenos sociales que se presentan para luego crear un marco conceptual acertado con el problema detallado y por lo tanto, se analizan una serie de hipótesis que tienen relaciones entre las variables observadas en la deducción. Este método suele tender a generalizar y normalizar los resultados. El tipo de diseño no es experimental. El diseño recolecta datos de forma pasiva sin introducir cambios o tratamientos, es decir, no hay manipulación deliberada de las variables, y es transversal porque los datos solo se recolectaron en un momento único de la investigación ([Bernal, 2010](#)).

Participantes

La unidad de análisis estuvo conformada por 332 estudiantes admitidos y matriculados durante un periodo de 5 años en el Programa de Psicología comprendidos entre los años 2012 al 2016, de diferentes sexos, edad, estado civil, estrato socioeconómico, lugar de origen y colegio de procedencia.

Instrumentos

Se recolectó la información del total de 1020 estudiantes admitidos por medio de las bases de datos existentes en el Departamento de admisiones y las plataformas virtuales que ofrece el ICFES.

Los datos fueron analizados a partir de la información suministrada por el departamento de Admisiones de la Universidad del Magdalena. Se realizó la caracterización sociodemográfica de los estudiantes de psicología de las cohortes del 2008 al 2012 con las variables: sexo, tipo de colegio, departamento de procedencia y estrato socioeconómico (ver tabla 1 y tabla 2).

Procedimientos

Fase 1. Solicitud y Revisión de la información
Durante este momento de la investigación se realizaron las solicitudes oficiales para el suministro de la información a la dependencia de Admisiones y Registro Académico de la Universidad del Magdalena y se revisaron las plataformas virtuales del ICFES existentes en la página web. Posteriormente se revisó la información de tal manera que se comprobó e que es la solicitada y está completa.

Fase 2. Procesamiento de la Información
Se elaboraron las bases de datos por cada una de las variables de estudio, la cual permitió una organización de la información año por año y por materia de mayor interés para el proyecto, lo cual permitió establecer una base de datos ordenada y clara.

Fase 3. Análisis de Datos
Para el análisis de los datos se aplicaron estadígrafos de tendencia central y para la comparación entre variables planteadas en el proyecto se aplicará el índice de correlación de Pearson, la

cual es una prueba estadística para analizar la relación entre dos variables medidas en un nivel de intervalo (Bernal, 2010).

Análisis de datos

Se realizó correlación de Pearson con las variables continuas, entre los puntajes de Saber Pro y los puntajes de admisión, acumulado de 1er semestre, y promedio acumulado general.
Se realizó análisis por Ji-cuadrado de las variables sociodemográficas con las variables continuas (los puntajes) pero obteniendo rangos entre ellas, basados en los cuartiles de cada puntaje.

Resultados

Los siguientes resultados se presentan, realizando en primer lugar un análisis descriptivo y finalmente un análisis de las correlaciones entre las variables propuestas en el estudio. El primer objetivo corresponde a las características sociodemográficas relacionadas con sexo, departamento de procedencia, colegio de procedencia y nivel socioeconómico de los estudiantes del Programa de Psicología de los años 2008 al 2016. Se observa que el sexo femenino es la mayoría de la población con cerca del 74%, la mayoría procede de un colegio público (72%), la mayoría procede del departamento del Magdalena (66%), y la mayoría

Tabla 1
Operacionalización de las variables cuantitativas y cualitativas

	Variable	Categorización o escala
Cuantitativa	Puntajes Saber pro-2008 al 2011	0 a 13
	Saber Pro del 2016	0 a 300
	Puntaje de admisión	0 a 1000
	Acumulado de 1er semestre	0 a 500
	Promedio acumulado general	0 a 500
	Sexo	Femenino o masculino
Cualitativa	Tipo de colegio	Público o privado
	Estrato socioeconómico	1 a 5
	Departamento de procedencia del estudiante	Magdalena o fuera del departamento del Magdalena

Nota: Esta tabla muestra la organización de las variables cualitativas y cuantitativas que se analizaron en el estudio.

pertenece a los estratos socioeconómicos 1 y 2 (40 y 43% respectivamente) (ver tabla 2).

En cuanto a los resultados del examen de admisión. Con respecto a esta variable se encontró que durante el año 2008 no se obtuvo reporte de datos de puntajes de admisión. En promedio los puntajes de ingreso estuvieron entre 608 puntos (año 2013) y 651 puntos (año 2016). El box-plot (Ver figura 1) nos indica que efectivamente el promedio y la mediana más alta se presentaron en los años 2012 y 2016.

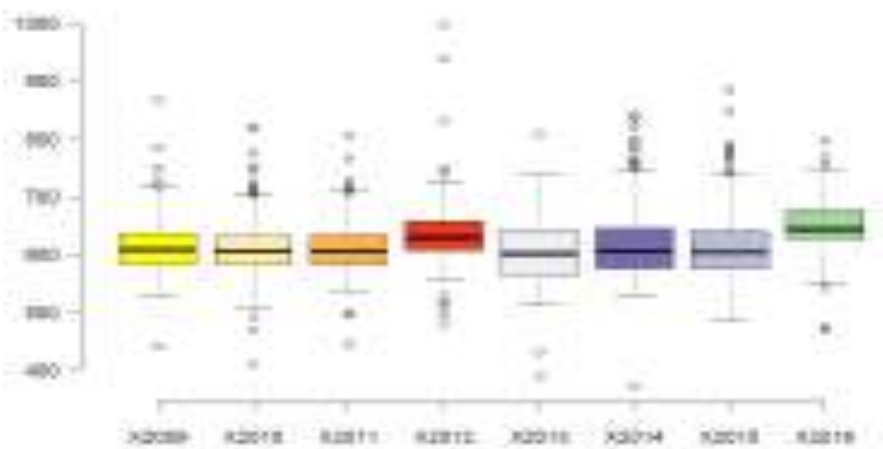
En la figura 2, se aprecia en promedio el acumulado general y se observó entre 341 (año 2013) y 401 puntos (año 2011). Esto indicó que desde el año 2013 al 2016 se ha presentado una mediana más baja al igual que el 50% del acumulado, es decir, los datos que se encuentran entre el 1er cuartil y el 3er cuartil. Esto concuerda con el promedio acumulado entre los años 2008 y 2012 los cuales estuvieron en un rango entre 389 y 401 puntos, siendo los años con mayor promedio.

Tabla 2
Características sociodemográficas estudiantes Programa de Psicología cohortes de ingreso 2008-2016

Variables		Frecuencia	Porcentaje
Sexo	Femenino	246	74,10%
	Masculino	86	25,90%
Tipo de colegio	Privado	93	28,01%
	Público	239	71,99%
Procedencia del estudiante	Fuera del Magdalena	113	34,04%
	Magdalena	219	65,96%
Estrato socio - económico	1	133	40,06%
	2	143	43,07%
	3	50	15,06%
	4	4	1,20%
	5	2	0,60%

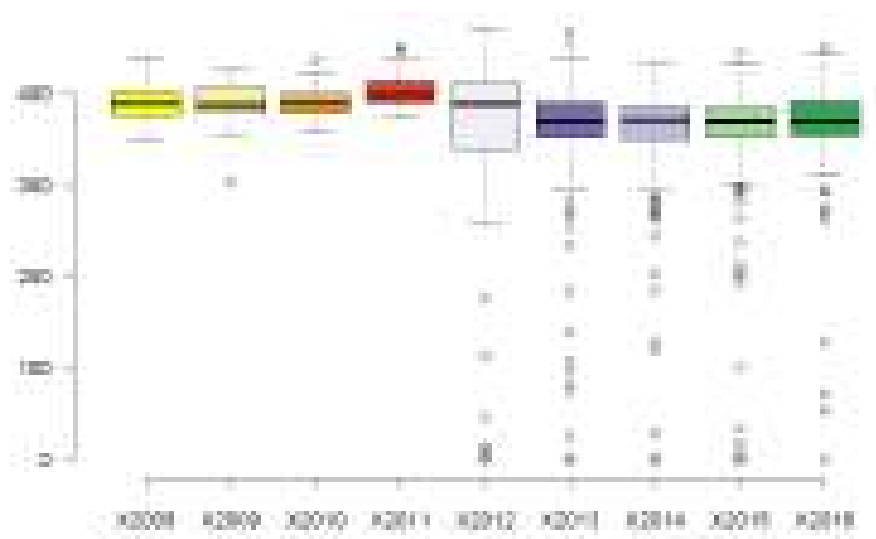
Nota: Esta tabla presenta la caracterización de los datos sociodemográficos que se utilizaron para el análisis de los puntajes de las cohortes de ingres desde el 2008 hasta el 2016.

Figura 1.
box-plots puntajes de admisión 2009-2016



En la tabla 3 y figura 2, en cuanto al promedio acumulado del primer semestre se observó entre 330 puntos (años 2012 y 2013) y 365 puntos (año 2009). Esto indica que en el año 2009 se presentó una mediana (365 puntos) ligeramente más alta con relación a los otros años, y con un 50% de los datos entre 339 (1er cuartil) y 392 puntos (3er cuartil); pero en general los datos entre el 1er cuartil y el 3er cuartil son muy homogéneos.

Figura 2.
Promedio acumulado 2008-2016



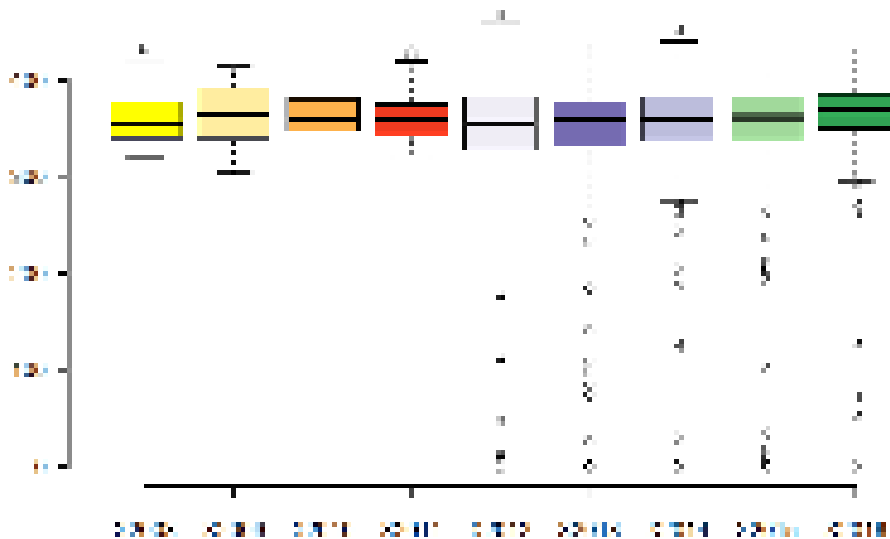
Fuente: Elaboración Propia.

Tabla 3
Cuartiles de acumulado general entre los años 2008-2016

	2008	2009	2010	2011	2012	2013	2014	2015	2016
Upper whisker	440.00	425.00	423.00	437.00	471.00	439.00	432.00	434.00	445.00
3rd quartile	401.50	406.50	400.00	412.00	411.00	390.00	386.00	385.00	391.00
Median	389.00	387.00	389.00	398.00	388.00	370.00	370.50	369.00	369.50
1st quartile	375.50	377.50	378.00	390.00	340.00	351.00	349.00	351.00	353.00
Lower whisker	350.00	354.00	361.00	374.00	257.00	293.00	297.00	300.00	310.00
Nr. of data points	71.00	55.00	55.00	57.00	139.00	185.00	268.00	229.00	138.00
Mean	389.63	389.51	390.40	401.14	350.17	341.12	350.08	349.79	362.86

Nota: en la tabla se señalan los cuartiles acumulados en general, entre los años 2008 y 2016.

Figura 3.
Promedio acumulado primer semestre 2008-2016



Fuente: Elaboración Propia.

Tabla 4
Cuartiles acumulados primer semestre 2008-2016

	2008	2009	2010	2011	2012	2013	2014	2015	2016
Upper whisker	418.00	414.00	426.00	419.00	460.00	438.00	439.00	438.00	431.00
3rd quartile	377.00	392.00	379.50	374.00	383.00	378.00	382.00	383.00	385.00
Median	356.00	365.00	358.00	360.00	356.00	358.00	360.50	362.00	368.00
1st quartile	340.00	339.50	346.50	343.00	328.00	333.00	338.00	338.00	349.00
Lower whisker	320.00	304.00	327.00	321.00	257.00	267.00	275.00	281.00	295.00
Nr. of data points	71.00	55.00	55.00	57.00	139.00	185.00	268.00	229.00	138.00

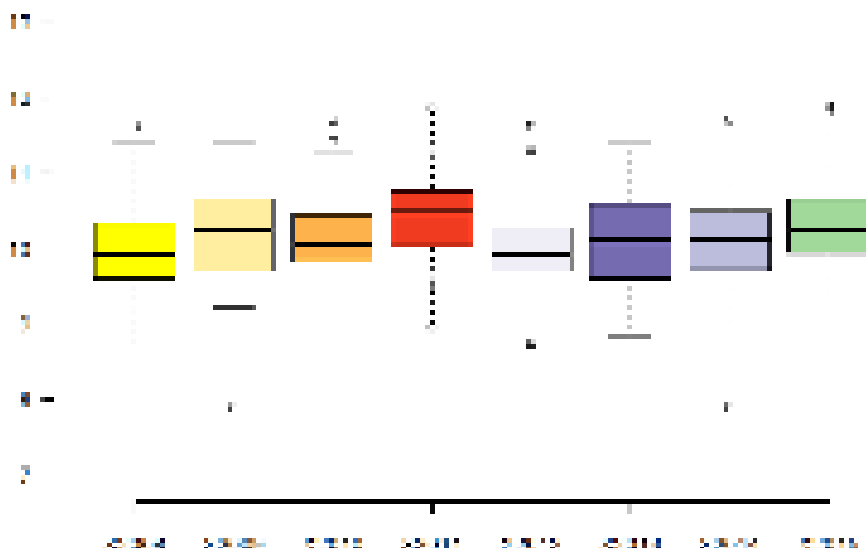
Nota: En esta tabla se especifica el análisis del tercer cuartil acumulado en el primer semestre entre los periodos del 2008 hasta el 2016.

En cuanto al análisis de los resultados obtenidos por los estudiantes de Psicología de los años 2012 al 2016 de la Prueba Saber Pro. Es de aclarar que se presentan en primera instancia los resultados de los años comprendidos de los años 2008 al 2015 ya que el MEN a través del IC-FES realiza un cambio de valoración en la escala de la Prueba para el año 2016, es así como presentamos los resultados del año 2016 de manera

independiente (ver tabla 4 y figura 3).

Los resultados de los promedios de Saber-Pro estuvieron en un rango entre 9,96 y 10,46. En el año 2011 se presentó el promedio más alto (10,46) e igualmente la mediana más alta (10,48); igualmente se observó que en el año 2011 el 50% de los datos fueron los más altos de todos los comparados, en un rango entre 10,04 (1er cuartil) y 10,76 (3er cuartil).

Figura 4.
Promedios Saber Pro2008-2015



Fuente: Elaboración Propia.

Tabla 5
Cuartiles de los promedios Saber Pro2008-2015

	2008	2009	2010	2011	2012	2013	2014	2015
Upper whisker	11.40	11.40	11.28	11.82	10.94	11.40	11.42	11.64
3rd quartile	10.33	10.67	10.44	10.76	10.30	10.58	10.50	10.64
Median	9.92	10.26	10.06	10.48	9.93	10.10	10.11	10.26
1st quartile	9.60	9.69	9.84	10.04	9.70	9.60	9.72	9.94
Lower whisker	8.74	9.22	9.36	9.00	8.92	8.84	8.92	8.94
Nr. of data points	71.00	55.00	55.00	57.00	70.00	74.00	90.00	70.00

Nota: En esta tabla se especifica el análisis del primer cuartil acumulado en el promedio de las pruebas Saber Pro entre los periodos del 2008 hasta el 2015.

Tabla 6
Datos descriptivos de los promedios Saber Pro2016

Año	Obs	Total	M	S	DE	Min	25%	Me	75%	Max	Mo
2016	94	14673	156,09	363,01	19,05	92	144	156	168	205	151

Nota: En la tabla se aprecia los datos de los promedios de las pruebas Saber Pro del año 2016.

En el año 2016 los puntajes de Saber Pro se presentaron en una escala entre 0 y 300 (ver tabla 9 y 10). Se presentaron 94 estudiantes, se obtuvo un

mínimo de 92 y un máximo de 205 puntos. La media de fue de 156 puntos. Se observa que el 50% de los puntajes estuvo entre 144 y 168 puntos.

En la interpretación de los resultados obtenidos en el examen de admisión, el promedio académico acumulado, el promedio académico del primer semestre de la carrera y la Prueba Saber Pro de los años 2012 al 2016. Para el análisis de las

variables se realizó una correlación de Pearson y regresión lineal múltiple y logística se realizaron desde las cohortes 2009 al 2011, dado que para el año 2008 no se suministraron los puntajes del examen de admisión.

Figura 5.
Puntajes Saber Pro2016



Fuente: Elaboración Propia.

Tabla 7
Cuartiles de los puntajes Saber Pro2016

Min	25%	Median	75%	Max
92	144	156	168	205

Nota: En esta tabla se muestra los cuartiles de los puntajes entre el 25% y 75% correspondiente a la prueba Saber Pro del 2016.

Tabla 8
Análisis correlación de Pearson entre puntajes Saber Pro- cohortes ingreso 2009 y 2011, puntajes examen de admisión, acumulado primer semestre y acumulado general.

	Saber Pro 2009-2011
Saber Pro	1
Puntaje de admisión	0,40 (p < 0,001)
Acumulado de 1er semestre	0,42 (p < 0,001)
Promedio acumulado general	0,48 (p < 0,001)

Nota: En esta tabla se muestra los resultados del análisis de la correlación de Pearson entre puntajes Saber Pro- cohortes ingreso 2009 y 2011, puntajes examen de admisión, acumulado primer semestre y acumulado general (1 y p < 0,001).

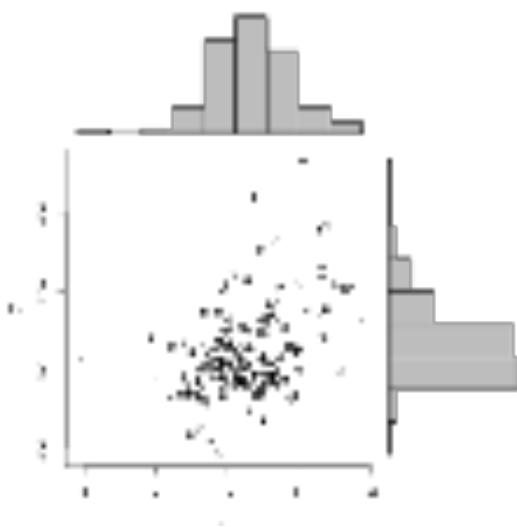


Figura 2. Correlación de Pearson entre los puntajes de Saber Pro, cohortes de ingreso 2009-2011, eje X y los puntajes de admisión (eje Y). $r = 0,40$ ($p < 0,001$).

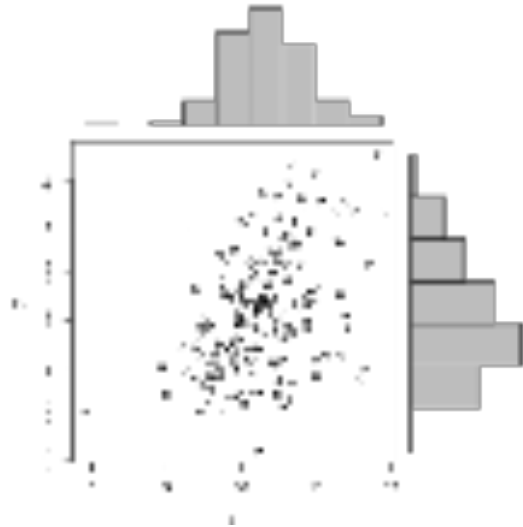


Figura 3. Correlación de Pearson entre los puntajes de Saber Pro, cohortes de ingreso 2009-2011, eje X y el acumulado del 1er semestre (eje Y). $r = 0,42$ ($p < 0,001$).

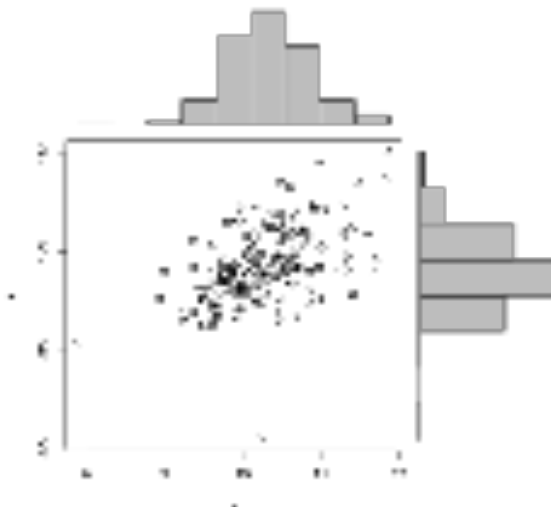


Figura 4. Correlación de Pearson entre puntajes Saber Pro, cohortes de ingreso 2009-2011, eje X y el promedio acumulado general (eje Y). $r=048$, ($p< 0,001$).

Fuente: Elaboración Propia.

Las tres variables tuvieron una correlación estadísticamente significativa lo que estaría indi-

cando que existe una correlación directamente proporcional entre las variables.

Tabla 9
Análisis de correlación de Pearson entre los puntajes de Saber Pro presentados en el 2016 de la Cohorte 2012 y los puntajes de admisión, acumulado de 1er semestre y el promedio acumulado general de la carrera.

Puntaje Saber Pro	1
Puntaje de admisión	0,46 (p < 0,001)
Acumulado 1er semestre	0,378 (p < 0,001)
Promedio acumulado general	0,38 (p < 0,001)

Nota: En esta tabla se muestra la correlación de Pearson entre los puntajes de Saber Pro con un valor de 1, los puntajes de admisión 0,46 (p < 0,001), acumulado de 1er semestre 0,378 (p < 0,001) y el promedio acumulado general de la carrera 0,38 (p < 0,001).

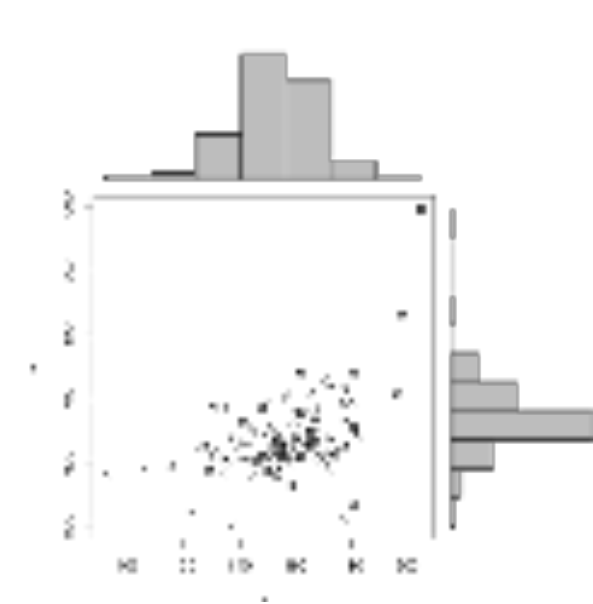


Figura 5. Correlación de Pearson entre los puntajes de Saber Pro (año 2016), cohorte 2012, eje X y el puntaje de admisión (eje Y).

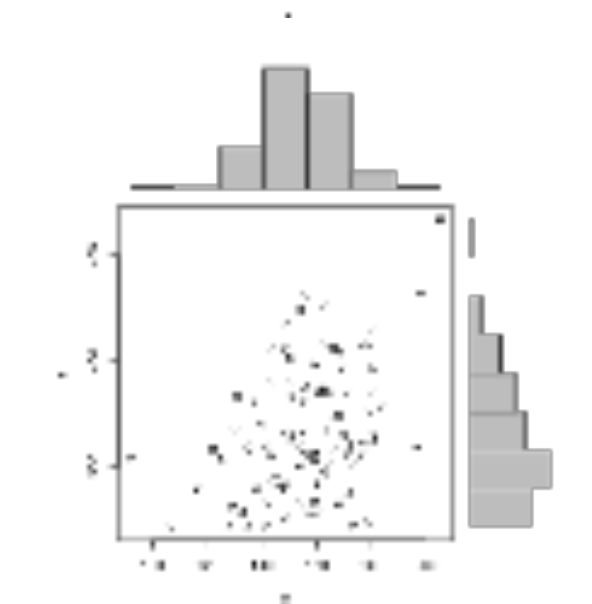


Figura 6. Correlación de Pearson entre los puntajes de Saber Pro (año 2016), cohorte 2012, eje X y el acumulado de Primer semestre (eje Y).

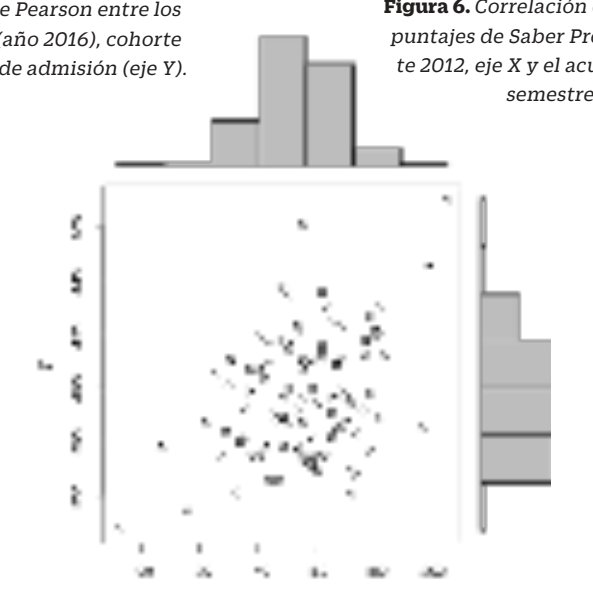


Figura 7. Correlación de Pearson entre los puntajes de Saber Pro (año 216), cohorte 2012, eje X y el promedio acumulado general (eje Y).

Fuente: Elaboración Propia.

Las tres variables tuvieron una correlación estadísticamente significativa, lo está indicando que existe una relación directamente proporcional entre las variables.

Dado que se presentaron correlaciones estadísticas significativas entre las variables continuas analizadas, se procedió a realizar los respectivos análisis de regresión lineal múltiple entre las mismas variables. Se realizó la regresión lineal entre los resultados de las pruebas Saber Pro y los puntajes de admisión, el acumulado del 1er semestre y el acumulado general y en este modelo de regresión lineal múltiple, tomando como variable dependiente (Y) los resultados de las pruebas Saber-pro-2016 y las variables independientes o predictoras: puntaje de admisión, el acumulado del 1er semestre y el promedio acumulado general, se observa un coeficiente de determinación $R^2 = 0,26$, indicando que este modelo con las tres variables predictoras explicaría el 26% de los resultados de las pruebas Saber-Pro.

La validación de este modelo se realizó con el análisis de la varianza donde se obtuvo una $F = 10,79$ ($p < 0,05$) siendo estadísticamente significativo, donde se observó que solo la variable independiente (X) puntaje de admisión fue estadísticamente significativamente relacionada con la variable Y = prueba Saber- pro.

Discusión

En la caracterización sociodemográfica de los estudiantes de psicología de la Universidad del Magdalena se observó que cerca del 70% de los estudiantes son mujeres y proceden de colegios públicos, cerca de un 60% proceden del departamento del Magdalena y cerca del 80% pertenecen a los estratos 1 y 2. A nivel de los puntajes de admisión la media máxima fue de 640, en el año 2012; en el acumulado de 1er semestre la media máxima fue de 366, el año 2012; en el acumulado general la media máxima fue de 401, en el año 2011; y en las pruebas Saber-Pro la media máxima fue de 10,45, pruebas presentadas entre 2012 y 2015; y la media de las pruebas Saber-Pro presentadas en el año 2016 fue de 156.

Cuando se realizaron los análisis multivariados teniendo como variable respuesta los resultados de las pruebas Saber-Pro, para las cohortes de ingreso 2009-2011, las variables predictoras estadísticamente significativas fueron puntaje de admisión y promedio del acumulado general, donde estas dos variables explicarían el 31% de los resultados de Saber-Pro.

En los análisis multivariados para la cohorte de ingreso 2012 con pruebas Saber-Pro presentadas en el año 2016, evidenciaron dos modelos estadísticamente significativos, así: el 1er modelo fue entre puntajes Saber-Pro y los puntajes de admisión y promedio de acumulado general; y el 2do modelo fue entre los puntajes Saber-Pro y los puntajes de admisión y el promedio del 1er semestre, en ambos modelos las variables predictoras explicarían cerca del 25% de los resultados de las pruebas Saber-Pro.

Se puede concluir que las variables puntajes de admisión y promedio acumulado general explicarían entre el 26 y 31% de los resultados de las pruebas Saber-Pro, en las cohortes de ingreso desde el 2009 al 2012.

En los análisis multivariados para la cohorte de ingreso 2012 con pruebas Saber-Pro presentadas en el año 2016, evidenciaron dos modelos estadísticamente significativos, así: el 1er modelo fue entre puntajes Saber-Pro y los puntajes de admisión y promedio de acumulado general (ver tabla 3, 4, 5 y 6) ; y el 2do modelo fue entre los puntajes Saber-Pro y los puntajes de admisión y el promedio del 1er semestre (ver tabla 3), en ambos modelos las variables explicarían cerca del 25% de los resultados de las pruebas Saber-Pro.

En los análisis de la cohorte de ingreso 2012 por Ji-cuadrado entre los resultados de Saber-Pro y las variables de los puntajes y las sociodemográficas y solo se observó asociación estadísticamente significativa entre los cuartiles de Saber Pro y los cuartiles de los puntajes de admisión, por el motivo anterior no se realizó análisis por regresión logística, ya que solo se observó una sola variable (X) predictora relacionada estadísticamente con la variable explicativa (Y), Saber-Pro.

En Colombia se han realizado trabajos en relación con el rendimiento académico de los universitarios teniendo como referencias las variables pruebas saber 11, puntaje de admisión, promedio

académico acumulado y pruebas Saber-Pro; así mismo se ha tenido en cuenta una gran diversidad de variables sociodemográficas incluidas sexo, estrato socioeconómico y procedencia del colegio. Sin embargo, los análisis por encontrar un solo modelo predictivo son diversos, esto se podría explicar bien por el tamaño de la muestra y por factores intrínsecos de las universidades y/o de las regiones. A continuación, se presentan una relación entre nuestros resultados y los de otros estudios en Colombia ([Osma y cols., 2014](#), [Tapasco y cols., 2016](#); [Melo y cols., 2017](#)).

En el año 2014 se conoció sobre los factores asociados a los resultados de las pruebas Saber Pro 2010 de los estudiantes de ingeniería civil de toda Colombia, donde se analizaron un total de 15364 estudiantes y se concluyó que existe asociación entre los puntajes Saber Pro superiores a la media y ser hombre, ser estudiante de ingeniería de institución pública, tener padres con estudios universitarios y ser de estratos medios y altos (Osma y et al, 2014), esta información se contrasta con la encontrada en esta investigación donde se indicó que el promedio y la mediana en general de los datos encontrados entre el 1er cuartil y el 3er cuartil son muy homogéneos, teniendo en cuenta que más del 70%, lo conforma el género femenino lo que no guarda asociación con el puntaje, puesto que al analizar los datos por cada año que presentaron las pruebas los estudiantes, los mismos guardaron homogeneidad teniendo en cuenta la variable de género (ver tabla 3 y figura 1).

En el año de 2016 el trabajo realizado por la Universidad de Caldas sobre la relación entre los puntajes de las pruebas Saber 11 y el rendimiento académico de los estudiantes universitarios de las cohortes 2008-2012, reveló que, en las variables de género, estrato, tipo de colegio, procedencia y puntaje promedio de todos los programas, se encontró que el puntaje de las pruebas Saber 11 tuvo una correlación débil con el desempeño académico. Y por análisis de regresión logística se encontraron asociaciones significativas entre Pruebas Saber Pro y ser mujer, estrato socioeconómico alto, puntaje de admisión superior y una edad joven ([Tapasco y cols., 2016](#)), en comparación con los datos de esta investigación, en la

tabla 4 y 5, se aprecia un promedio acumulado general y observándose puntajes entre 341 (año 2013) y 401 puntos (año 2011), lo que explica que desde el año 2013 al 2016 se ha presentado una mediana más baja al igual que el 50% del acumulado, es decir, los datos que se encuentran entre el 1er cuartil y el 3er cuartil, concuerda con el promedio acumulado entre los años 2008 y 2012 los cuales estuvieron en un rango entre 389 y 401 puntos, siendo los años con mayor promedio. Por ende, los análisis por regresión logística de la cohorte de ingreso 2009-2011 entre las variables se observó un único modelo estadísticamente significativo entre las pruebas Saber-Pro y las variables predictoras acumulado del 1er semestre y el promedio acumulado en general.

También se analizó en este modelo que los estudiantes que presentaron un acumulado de 1er semestre igual o mayor a 359 puntos tenían 2,67 veces de alcanzar un promedio de pruebas Saber-Pro mayor a 10,15, y los estudiantes que presentaron un promedio acumulado general igual o mayor a 390 puntos tenían 3,24 veces de alcanzar un promedio de pruebas Saber-Pro mayor a 10,15.

Recientemente, en el año 2017, un análisis nacional en Colombia sobre el desempeño universitario observado con diferentes variables destacó que el estrato socioeconómico y el nivel de educación de los padres, se refleja en el rendimiento académico de los estudiantes de los diferentes programas académicos ([Melo y cols., 2017](#)). Esto estaría sugiriendo que el rendimiento de los estudiantes en sus programas académicos universitarios y en las pruebas Saber Pro estaría restringido a los factores que rodean al estudiante como son el estrato socioeconómico y el nivel de educación de sus padres. Además de otros factores que si son inherentes a las universidades como serían: criterios de admisión de los docentes, la investigación y la administración.

Se concluye que los resultados hallados en el programa de Psicología están analizados como procesos internos, por ello al realizar el análisis de las variables entre el periodo 2009-2011, estas tuvieron una correlación estadísticamente significativa con Saber Pro 1; Puntaje de admisión 0,40 ($p < 0,001$); Acumulado de 1er semestre 0,42 ($p < 0,001$) y Promedio acumulado general 0,48 ($p <$

0,001) (ver tabla 11) y así mismo, se realizó el análisis con el periodo del 2012 -2016 con el Puntaje Saber Pro 1 junto con el Puntaje de admisión 0,46 ($p < 0,001$); el Acumulado 1er semestre 0,378 ($p < 0,001$) y Promedio acumulado general 0,38 ($p < 0,001$) (ver tabla 8 y 9), esto indica una relación directamente proporcional entre las variables que se analizaron, así como el tipo y grado de participación organizacional, y responsabilidades entre docentes y estudiantes y en el logro de las metas, lo que da indicios de la calidad en que se basan los objetivos de la organización. Además, que invita a diseñar y ajustar planes de mejora con base a los resultados de la evaluación.

Este análisis convoca al personal directivo, a los docentes y estudiantes a seguir participando activamente en la construcción de una educación de calidad. A su vez, las actitudes de docentes y alumnos hacia la evaluación indican motivación, ya sea para expresarse y planes de vida, la autonomía para gestionar su propio progreso, e indicadores de las competencias docentes de los docentes, esto ayuda a encontrar una metodología que se aproxime a un estado de medición ideal y a una valoración del capital intelectual adquirido a través de un proceso de formación, que lleve a revisar dos puntos que son la culminación de los estudios de nivel medio (Saber11) y la finalización de los estudios de nivel superior (Saber Pro).

Se recomienda que las instituciones consideren los resultados de la evaluación de Saber Pro en el comité curricular para su análisis. El proceso de formación del curso no es para preparar el examen, sino para evaluar a los estudiantes enseñados y su estilo de aprendizaje. La docencia y la forma de evaluar los conocimientos a partir de la formación y el entrenamiento de la habilidad especial y la habilidad y perfil general a evaluar.

Referencias

- Bahamón, M. & Reyes, R. (2014). Caracterización de la capacidad intelectual, factores sociodemográficos y académicos de estudiantes con alto y bajo desempeño en los exámenes Saber Pro - año 2012. *Avances en Psicología Latinoamericana*, 32(3), 459-476.
- Barahona, P. (2014). Factores determinantes del rendimiento académico de los estudiantes de la Universidad de Atacama. *Estudios pedagógicos*. 11(1), 25-39, 2014. <http://dx.doi.org/10.4067/S0718-07052014000100002>
- Bernal, César A. (2010). *Metodología de la Investigación*. Tercera edición. Pearson Educación:
- Betancourt, R. y Frías, L. (2015). Competencias argumentativas de los estudiantes de derecho en el marco de las pruebas Saber-Pro. *Civilizar. Ciencias Sociales y Humanas*, 15(28), 213-228. <https://www.redalyc.org/articulo.oa?id=100241608015>
- Burgos, M., Ruales, K., Bastidas, Y. y Ortiz, E. (2019). Cálculo del valor agregado generado por la Universidad de Nariño en relación con las pruebas Saber 11- Saber Pro-2010-2014. *Revista de la Facultad de Ciencias Económicas y Administrativas*. 20 (2), 203 -226. ISSN-E 2539-0554. <https://vlex.com.co/vid/calculo-valor-agregado-generado-843966065>
- Colombia Aprende. (2017). Historia de los ECAES. http://www.colombiaaprende.edu.co/html/home/1592/article-156083.html#h2_1
- Delgado, M. (2013). Examen de Estado de la Calidad de la Educación Superior: SABER PRO. ¿Qué indican sus resultados? *Revista Colombiana de Anestesiología*, 41(3):177-178. <https://www.redalyc.org/articulo.oa?id=195128413001>
- Gil, F., Rodríguez, V., Rondón, M. y Gómez, C. (2013). Impacto de las facultades de medicina y de los estudiantes sobre los resultados en la prueba nacional de calidad de la educación superior (SABER PRO). *Revista Colombiana de Anestesiología*. 41(3): 196-204.
- Instituto Colombiano para la Evaluación de la Educación (2012). Estudios sobre la calidad de la educación en Colombia. Bogotá, Colombia: ICFES.
- Instituto Colombiano para la Evaluación de la Educación. (2016a). Módulo de Análisis de problemáticas psicológicas Saber Pro-2016-2.
- Instituto Colombiano para la Evaluación de la Educación. (2016b). Pruebas Saber Pro-Psicología. Bogotá, Colombia: ICFES.
- Instituto Colombiano para la Evaluación de la Educación. (2016c). Módulo de Competencias ciudadanas Saber Pro-2016-2.
- Isáziga, C.H., Gabalán, J. y Vásquez, F. (2014). La intervención académica en la construcción de una sociedad con calidad: análisis del valor agregado en el proceso for-

- mativo colombiano. Hallazgos, 11(22):359-384.
- León, A., Amaya, s. & Orozco, D. (2012). Relación entre comprensión lectora, inteligencia y desempeño en pruebas saber pro en una muestra de estudiantes universitarios. Cultura, Educación y Sociedad. 3(1): 187-204.
- Melo, L., Ramos, J. Y Hernández, P. (2017). La educación superior en Colombia: situación actual y análisis de eficiencia. Revista Desarrollo y Sociedad. (78): 59-111. doi: <https://doi.org/10.13043/DYS.78.2>
- MEN (Ministerio de Educación Nacional). (2019). Módulo de Análisis de problemáticas psicológicas. Las Pruebas Saber Pro. Publicación del Instituto Colombiano para la Evaluación de la Educación. (Icfes). Bogotá. ISBN de la versión digital: 978-958-11-0839-8.
- Oficina Asesora de Comunicaciones MEN. (2013). DEL ECAES A SABER PRO. https://www.mineducacion.gov.co/1621/articles-299905_recurso_1.pdf
- Osma, Wa., Mojica, A. y Rivera, T. (2014). Factores asociados al rendimiento en las pruebas saber pro en estudiantes de ingeniería civil en una universidad colombiana. *Innovaciencia Facultad Ciencias Exactas Física y Naturales*. 2(1): 17 – 24.
- Ramírez, C. (2014). Factores asociados al desempeño académico según nivel de formación pregrado y género de los estudiantes de educación superior Colombia. *Revista Colombiana de Educación*. (66): 204-224.
- Samaniego, A. y Buenahora, M. (2016). Variables relacionadas con ansiedad social en adolescentes: un modelo de regresión lineal múltiple. *Interacciones: Lima, Perú*. 2 (2): 109-122. <https://dialnet.unirioja.es/servlet/articulo?codigo=5765069>
- Tapasco Alzate, O.A.; Ruiz Ortega, F.J. & Osorio García, D. (2016). Estudio del poder predictivo del puntaje de admisión sobre el desempeño académico universitario. *Revista Latinoamericana de Estudios Educativos*, 12(2): 148-165. <https://www.redalyc.org/articulo.oa?id=134149931008>
- Timarán, S. R., Hernández, I., Caicedo, S. J., Hidalgo, A. y Alvarado, J. C. (2016). Tendencias de desempeño académico en competencias genéricas - pruebas Saber Pro. En Descubrimiento de patrones de desempeño académico con árboles de decisión en las competencias genéricas de la formación profesional. Ediciones Universidad Cooperativa de Colombia. 87-99. <http://dx.doi.org/10.16925/9789587600490>
- Tirado, M. (2017). Aporte a los exámenes de estado, saber pro, al desarrollo curricular y la calidad en dos programas de psicología de Bogotá, D.C. (2017). Trabajo realizado en el marco de la maestría en psicología de la Universidad Católica de Colombia. Repositorio. <https://repository.ucatolica.edu.co/server/api/core/bitstreams/437e82cd-d52b-47b0-8cc5-2698597b7e2a/content>
- Torres, G., Marulanda, J., Barreto, J., Quiroga, M., Centeno, M., Mesa, S. y Ramírez, S. (2008). El surgimiento del Servicio Nacional de Pruebas del ICFES en las voces de sus protagonistas. *Magistro*. 2 (3). 115-134; <https://dialnet.unirioja.es/servlet/articulo?codigo=4037231>

Más breve puede ser mejor: Propiedades métricas de la UWES – 3 ítems para evaluar “engagement” en profesores universitarios mexicanos

Arturo Juárez-García ^{*1}; César Merino-Soto²; Daniela Villamar Sánchez³

¹Centro de Investigación Transdisciplinar en Psicología, Universidad Autónoma del Estado de Morelos, México <https://orcid.org/0000-0003-3264-679X> arturojuarezg@hotmail.com ²Instituto de Investigación de Psicología, Universidad de San Martín de Porres, Lima, Perú, <https://orcid.org/0000-0002-1407-8306> sikayax@yahoo.com.ar ³Instituto de Ciencias de la Educación, Universidad Autónoma del Estado de Morelos, México <https://orcid.org/0000-0002-5843-9628> danielavillamarsanchez@gmail.com

Citar como: Juárez-García, A.; Merino-Soto, C.; Villamar Sánchez, D. (2023). Más breve puede ser mejor: Propiedades métricas de la UWES – 3 ítems para evaluar “engagement” en profesores universitarios mexicanos. *Revista Digital de Investigación en Docencia Universitaria*, 17(2), e1623. <https://doi.org/10.19083/ridu.2023.1623>

Recibido: 21/04/2022. **Revisado:** 29/04/2022 **Aceptado:** 02/02/2023. **Publicado:** 30/06-2023.

Resumen

Introducción: el *engagement* laboral es un efecto psicológico positivo relevante en la literatura, y prácticamente son inexistentes estudios de validez de la versión ultra-breve de la Utrecht Work Engagement Scale en Latinoamérica. **Objetivo:** identificar las propiedades psicométricas de la versión ultra-breve de la Utrecht *Work Engagement Scale* (UWES-3) y analizar su equivalencia con la versión UWES-9 en una muestra de docentes de una Universidad Pública de México. **Método:** la muestra final fue de N=247 profesores universitarios, quienes contestaron la UWES en sus versiones de 9 y 3 ítems, así como el Cuestionario de Evaluación de Síndrome de Quemarse por el Trabajo para personal de educación (CESQT-PE) con el propósito de estimar su validez discriminante. **Resultados:** los resultados confirmaron correlaciones muy altas entre ambas versiones del UWES (3-9) ($r > .82-.93$), una estructura unidimensional del UWES- 3 con mayores cargas factoriales ($> .75$), valores de consistencia interna satisfactorios ($> .70$) y correlaciones altas en la dirección esperada y similares con el CESQT-PE ($q = .02-.08$). **Conclusiones:** se concluye que existe equivalencia del UWES-3 y el UWES-9 y se discuten los beneficios prácticos de la versión ultra-breve para la medición del *engagement* laboral en docentes universitarios..

Palabras clave: entusiasmo laboral; UWES; Docentes universitarios; Escalas de medición

Shorter maybe better: Psychometric properties of the UWES-3 items for measuring engagement in university Mexican teachers

Abstract

Introduction: Work engagement is a relevant positive psychological effect in the literature, and validity studies of the ultra-brief version of the Utrecht Work Engagement Scale are practically nonexistent in Latin America. **Objective:** to identify the psychometric properties of the ultra-brief version of the Utrecht Work Engagement Scale (UWES-3) and to analyze its equivalence with the UWES-9 version in a sample of teachers from a public university in Mexico. **Method:** the final sample consisted of

*Correspondencia:

Arturo Juárez García
arturojuarezg@hotmail.com



N=247 university teachers, who answered the 9-item and 3-item versions of the UWES and the Burnout Evaluation Questionnaire for Education Personnel (CESQT-PE) to estimate its discriminant validity. **Results:** The results confirmed very high correlations between both versions of the UWES (3-9) ($r=>.82-.93$), a unidimensional structure of the UWES- 3 with higher factor loadings ($>.75$), satisfactory internal consistency values ($>.70$) and high and similar correlations in the expected direction between both versions and the CESQT-PE ($q=.02-.08$). **Conclusions:** We conclude that there is equivalence of the UWES-3 and the UWES-9. The practical benefits of the ultra-brief version for the measurement of work engagement in university teachers are discussed.

Keywords: Engagement; UWES; University teachers; Measurement Scales.

Introducción

El *engagement* laboral es quizás el efecto psicológico laboral con enfoque positivo más popular en últimas las décadas. Aunque existen intentos de su traducción al castellano en conceptos tales como “entusiasmo laboral” (Juárez-García, 2015), su acepción en inglés sigue siendo la más común, tanto en ámbitos de consultoría como en contextos académicos, quizás para no confundirse con otros fenómenos psicológicos. Maslach y Leiter (1997) lo posicionaron justo como un opuesto directo al síndrome de *burnout*: esto es, sentirse energético en contraparte al síntoma de agotamiento emocional, involucrarse psicológicamente en la actividad laboral en contraparte a la despersonalización del *burnout*; y finalmente, sentirse auto eficaz profesionalmente en oposición a la ausencia de tal certeza en las capacidades cuando hay desgaste psicológico. Sin embargo, este paradigma padece de un enfoque restringido a dos polos opuestos mutuamente excluyentes, pues su lógica básica es que bajos niveles de *burnout* representan en automático altos niveles de *engagement* y por ende ambos pueden medirse con el mismo instrumento (Maslach Burnout Inventory -MBI-) (Maslach, Jackson, Leiter, Schaufeli & Schwab, 1986), lo que evidentemente ha recibido críticas (González-Roma, Schaufeli, Bakker & Lloret, 2006).

De acuerdo con Schaufeli, Salanova, González-Roma y Bakker (2002), el *engagement* laboral

es un fenómeno correlacionado negativamente al *burnout*, pero al mismo tiempo independiente. Lo definen como un efecto motivacional positivo que se da por el trabajo y se caracteriza por tres componentes: 1) Vigor, que se caracteriza por los altos niveles de energía y resiliencia mental mientras se trabaja, voluntad de esforzarse y persistir en situaciones difíciles en el trabajo; 2) Dedicación, que se refiere a estar fuertemente motivado en el trabajo y experimentar un sentido de significancia, inspiración, orgullo y desafío; y 3) Absorción, que se refiere al hecho estar completamente concentrado e inmerso en el trabajo, con la sensación de que el tiempo “pasa volando”, en una especie de vinculación plena con el trabajo (Salanova & Schaufeli, 2009). El *engagement* ha demostrado su relación con la motivación, la satisfacción, la felicidad, la iniciativa y el buen desempeño laboral entre diversos efectos individuales y organizacionales favorables en diversos contextos (Salanova & Schaufeli, 2009). Incluso, se han encontrado asociaciones del *engagement* con la eficacia productiva de los países, el producto interno bruto, la presencia de integridad y ausencia de corrupción, así como con culturas con alto individualismo, alto disfrute de la vida y baja distancia de poder (Schaufeli, 2018).

El *engagement* laboral en la tarea docente ha sido poco investigado, sin embargo, se han demostrado ya las relaciones positivas de este fenómeno con la satisfacción de los profesores (Han, Yin, Wang, & Zhang, 2020), la presencia de más emociones positivas mientras trabajan (Burić & Macuka, 2018), mayor eficacia en su actividad laboral (Wyatt & Dikilitaş, 2016), e incluso, con el mejoramiento de desempeño cognitivo de sus estudiantes (Zhu, 2001), entre muchas otras consecuencias positivas.

Un parteaguas en el desarrollo del constructo se instaló en los trabajos de Schaufeli y colaboradores (Schaufeli et al., 2002), quienes no solo demostraron que el constructo era independiente al *burnout* (Gonzalez-Roma et al., 2006), sino que lo definieron y diferenciaron conceptualmente de otros constructos con base en evidencia empírica. Además, propusieron una medida con satisfactorias propiedades psicométricas en distintas muestras y regiones: el *Utrecht Engagement Scale*

(UWES) ([Salanova, Schaufeli, Llorens, Peiro & Grau, 2000](#); [Schaufeli et al., 2002](#); [Schaufeli, 2017](#)).

La escala UWES ha sido validada en diversos idiomas en más de 30 países en los cinco continentes ([Schaufeli & Bakker, 2010](#)). Su versión inicial de 24 ítems se validó en una propuesta de 17 ítems funcionales, con la estructura trifactorial esperada e índices de consistencia interna (α) superiores a .70 en poblaciones de estudiantes y trabajadores ([Schaufeli et al., 2002](#)). Posteriormente [Salanova et al., \(2000\)](#), probaron una versión de la UWES de 15 ítems en una muestra de trabajadores que utilizan las tecnologías de la información, a la que dedicaron cinco preguntas para vigor, cinco para dedicación y cinco para absorción en una escala de siete opciones de respuesta de frecuencia Likert que va desde "0=Nunca", hasta "6=Todos los días". Los resultados confirmaron la estructura tridimensional e índices de consistencia interna (α) superiores a .73 en todas las subescalas, así como la independencia con el constructo de burnout y relaciones estadísticamente significativas en las direcciones esperadas con indicadores organizacionales tales como el compromiso organizacional, la satisfacción o el gozo laboral.

En una siguiente etapa, [Schaufeli, Bakker y Salanova \(2006\)](#) validaron una versión de la UWES de 9 ítems (3 para cada componente) en 10 países europeos de diferentes grupos ocupacionales, la consistencia interna fue satisfactoria ($\alpha > .70$) y aunque la estructura trifactorial se comprobó, los mejores índices fueron para una estructura unidimensional. Dadas las altas correlaciones entre los tres componentes, los autores sugirieron el uso del puntaje total en una medida única global de *engagement*.

Más recientemente, [Schaufeli, Shimazu, Hakonen, Salanova & De Witte \(2019\)](#), validaron una versión ultra-breve de solo tres ítems (UWES-3) en muestras de varios países, la cual mantuvo buenos índices psicométricos, compartió hasta un 92% de varianza con la versión de nueve ítems (UWES-9) y no presentó diferencias en los patrones de correlación con otras variables. Los autores sugieren el uso del UWES-3 en estudios futuros e incorporarlo en encuestas nacionales, por todas las ventajas prácticas que representa.

Aunque en Latinoamérica existen algunas validaciones de la escala de *engagement* laboral

(UWES) en muestras de docentes ([Flores, Fernández, Juárez, Merino y Guimet, 2015](#); [Álvarez Garzón & Peña Fuentes, 2019](#); [Martins & Mendonça, 2019](#)), estos se han realizado con las versiones de 17, 15 ó 9 ítems y solo se identificó una publicación muy reciente que validó la versión ultra-breve UWES-3 en estudiantes peruanos. Se trata del estudio peruano de [Domínguez, Fernández y Seperak \(2021\)](#) quienes concluyeron que las propiedades psicométricas de esta versión breve no diferían a las otras versiones del UWES. No obstante lo anterior, siguen pendientes estudios de validación de esta versión ultra-breve en poblaciones de trabajadores académicos en poblaciones latinas, en particular en México, donde no hay estudios al respecto. Las implicaciones prácticas y de investigación de contar con una medida ultra-breve para el tamizaje del fenómeno de *engagement* en docentes, representa múltiples ventajas asociadas a una detección sencilla, rápida y con una relación lo más positiva posible en términos de efectividad/costo y de eficiencia/tiempo en la evaluación de este fenómeno en profesores.

Dado lo anterior, el objetivo del presente estudio es analizar las propiedades psicométricas de la versión ultra-breve UWES-3 y analizar su equivalencia con la UWES-9 en una muestra de docentes universitarios mexicanos.

Método

Participantes

Considerando una población total de 1,825 profesores e investigadores en los centros temáticos y regionales de la Universidad de Guadalajara, se envió un correo electrónico a todos ellos con la invitación a participar y un vínculo de acceso a la encuesta a través de la plataforma Survey-Monkey®. Para ingresar a contestar el cuestionario fue necesaria la aceptación de la carta de consentimiento. Aceptaron participar 259 académicos, por lo que la tasa de respuesta fue del 14.19%. Los criterios de inclusión fueron: realizar actividades de investigación y docencia, ser un trabajador activo en la Universidad al momento del estudio, tener un contrato de tiempo com-

pleto, aceptar participar en el estudio y firmar la carta de consentimiento informado. Los criterios de exclusión fueron: ser docentes en sabático, ser profesores con contrato distinto al tiempo completo y no participar en algún programa estímulo. El criterio de eliminación fue no contestar la batería de instrumentos o contestarla parcialmente. Finalmente, conforme a estos criterios se obtuvo una muestra de 247 académicos.

La muestra final estuvo distribuida de la siguiente manera: por 122 hombres (49.4%) y 125 mujeres (50.6%), con una edad promedio de 49.28 años (Desviación estándar = 9.677). Respecto al estado civil, 149 fueron casados (60.3%), 53 solteros (21.5%), 23 viven en unión libre (9.3%), 18 están divorciados (7.3%) y 4 son viudos (1.6%). En nivel máximo de estudios se observa que 38 tienen maestría (15.4%), 179 tienen doctorado (72.5%), y 30 tienen postdoctorado (12.1%). Respecto al centro universitario de adscripción 205 pertenecen a un Centro Temático (82.9%) y 42 a un Centro Regional (17.1%).

Conforme al nombramiento oficial, 240 participantes señalaron ser Profesor-Investigador (97.2%) y 7 señalaron ser Técnico Académico (2.8%). Sin embargo, todos se dedican a realizar actividades de investigación y docencia. El promedio de antigüedad en la institución fue de 19.32 años (Desviación estándar= 10.983), aunque el promedio en el puesto actual fue de 8.81 años (Desviación estándar = 7.832).

Consideraciones éticas

Se formuló una carta de consentimiento informado. A través de este documento se solicitó la participación voluntaria de los académicos, se garantizó la confidencialidad de su identidad y se les confirió la libertad de abandono de la encuesta en el momento deseado. No hubo identificadores personales, ni algún procedimiento invasivo. Se siguieron todas las recomendaciones de la declaración de Helsinki ([World Medical Association \[AMM\], 2013](#))

Instrumentos

En la misma lógica de los estudios disponibles sobre la validez y equivalencia de la versión corta UWES-3, se seleccionó como criterio concurrente la versión de 9 ítems por ser la versión mejorada de

versiones previas y la que posee mejores capacidades psicométricas que la versión de 17 ó 15 ítems (Schaufeli et al., 2006). Asimismo, como criterio externo para valorar los patrones de equivalencia de correlaciones entre ambas versiones con un criterio externo, se consideró la edad y el sexo, así como una medida de *burnout* o desgaste emocional con la que se espera un tipo de validez o correlación divergente con las medidas del UWES, como sucedió en estudios iniciales. En este estudio se utilizó el CESQT, un instrumento que denomina síndrome de quemarse por el trabajo al *burnout*. Tales instrumentos se desglosan a continuación:

Utrecht Work Engagement Scale (UWES) en sus versiones de 9 y 3 ítems ([Schaufeli et al., 2017](#)).

Es una medida de evaluación del *engagement*, dentro del mismo modelo en que se crearon sus versiones anteriores (i.e., UWES-15 y UWES-9). La versión ultra-breve UWES-3 está compuesta por tres ítems que representan a cada uno de los componentes originales, referidos a sentir energía (ítem 1, Vigor), entusiasmo (ítem 2, Dedicación) y e inmersión (ítem 8, Absorción). Sus opciones de respuesta están escaladas ordinalmente de 0 (*Nunca-Ninguna vez*) a 6 (*Siempre-Todos los días*). La confiabilidad de consistencia interna (α) en el estudio original osciló entre un de .77 y .85.

Cuestionario de Evaluación de Síndrome de Quemarse por el Trabajo para personal de educación (CEQST-PE).

Contiene 20 ítems distribuidos en 4 dimensiones ([Gil-Monte, 2011](#)). La Ilusión por el trabajo (1) implica el deseo del individuo de alcanzar las metas laborales porque suponen una fuente de satisfacción personal. El Desgaste psíquico (2), se refiere al Agotamiento emocional y físico debido a que en el trabajo se tiene que tratar a diario con personas que presentan o causan problemas. La Indolencia (3) representa la presencia de actitudes negativas de indiferencia y cinismo hacia los clientes de la organización (pacientes, alumnos, etc.). La Culpa (4), involucra sentimientos de culpa que la persona tiene por el comportamiento y las actitudes negativas desarrolladas en el trabajo, en especial hacia las personas con las que se establecen relaciones laborales. Las opciones de respuesta incluyen una escala de frecuencia de cinco grados que van de 0 (nunca) a 4 (Muy frecuentemente; todos los días) ([Gil-Mon-](#)

te & Noyola, 2011). El CESQT-PE ha mostrado validez y confiabilidad satisfactoria en muestras mexicanas y latinas (Gil-Monte, Unda & Sandoval, 2009) y con eficacia superior al *Maslach Burnout Inventory - Human Services Survey* (MBI-HSS) (Juárez-García, 2015, Calderón-de la Cruz, Merino-Soto, Juárez-García, Dominguez-Lara & Fernández-Arata, 2020), por lo que la escala y sus dimensiones, son un criterio idóneo en la validez discriminante del UWES-3.

Procedimiento

Análisis. Consistió en dos procedimientos, el primero preliminar para prevenir el efecto de distorsiones no intencionadas de respuesta (Wetzel, Böhnke, & Brown, 2016) y para describir las respuestas a los ítems, y el segundo enfocado en el objetivo principal del estudio.

Análisis preliminar. En el primero, se escanearon los datos para detectar posibles sesgos de respuestas expresados por valores atípicos multivariados (Meade, & Craig, 2012). Considerando su eficiencia (Meade, & Craig, 2012), se utilizó la distancia D_2 (Mahalanobis, 1936) con el programa *normtest* (DeCarlo, 1997). Los participantes detectados como valores atípicos fueron removidos de la base de datos. Luego en el nivel de los ítems se obtuvieron estadísticos descriptivos y distribucionales, y correlacionales (igualdad de las correlaciones inter-ítem; Steiger 1980), en los programas R *psych* (Revelle, 2020), y *rcompanion* (Mangiafico, 2021)

Análisis psicométrico. Para el objetivo principal, el análisis de las propiedades psicométricas fue secuencialmente realizado mediante la evaluación de la dimensionalidad (i.e., número de factores latentes), la escalabilidad de los puntajes e ítems observados, las características estructurales, la equivalencia entre grupos, y la correlación con otras variables.

Dimensionalidad. Debido que existe discrepancia entre la representación de constructo más apropiada con el UWES, primero se exploró el número de variables latentes mediante la evaluación del número de factores latentes. Se implementó un procedimiento de consenso entre 9 métodos de extracción de factores disponibles (e.g., análisis paralelo, coordenadas óptimas,

MAP-Velicer, etc.), mediante el programa *psycho* (Makowski, 2018). También se utilizó el método de *Comparación de Datos* (CD; Ruscio, & Roche, 2012; programa *RGenData*, Ruscio, 2018) debido que tiende a mostrar adecuada precisión (Auerwald, & Moshagen, 2019).

Propiedades estructurales. Se aplicaron dos enfoques, uno no paramétrico apropiado para medidas con pocos ítems (Van Schuur, 2003), y otro paramétrico, analizados en ese orden. Respecto al modelo no paramétrico, se corroboró la dimensionalidad predominante hallada en el análisis anterior mediante el *Mokken Scale Analysis* (MSA; Mokken, 1971). Como un proxy del número de dimensiones latentes (Straat, Van der Ark, & Sijtsma, 2013), se aplicó el algoritmo Automated Item Selection Procedure (AISP; Straat, Van der Ark, & Sijtsma, 2013) para seleccionar ítems escalables dentro de grupos homogéneos, e identificar así el número de escalas. Se utilizó el rango de .30 hasta .70, en pasos de .10 unidades del coeficiente de escalabilidad H para la detección de grupos homogéneos de ítems. Una vez completado esto, en el siguiente paso se corroboraron las propiedades de los puntajes observados, que son precursoras para el modelamiento paramétrico más complejo (e.g., basados en SEM o IRT). De este modo, se verificó el cumplimiento del *modelo de homogeneidad monotonica* (MHM; Sijtsma, & Van der Ark, 2017), mediante la evaluación de: a) la escalabilidad con el coeficiente H (Loevinger, 1948) para el puntaje y los ítems (H_i); coeficientes mayores a .40 son aceptables; Sijtsma, & Van der Ark, 2017). Luego, b) la independencia local, mediante los coeficientes $W^{(1)}$ y $W^{(2)}$ para las relaciones inter-ítem, y $W^{(2)}$ para cada ítem (Straat, Van der Ark, & Sijtsma, 2016); finalmente, c) la monotonidad de los ítems (i.e., la relación monotonica incremental entre los ítems y el puntaje de la escala), evaluado mediante gráficos de curvas características de los ítems. Finalmente, desde el MSA la confiabilidad se estimó con el coeficiente MS . Para el análisis MSA, se usó el programa *mokken* (Van der Ark, 2012).

En el segundo enfoque, mediante el modelamiento paramétrico lineal, se utilizó el modelamiento de ecuaciones estructurales (SEM), con el estimador WLSMV para variables categóricas,

e índices prácticos de ajuste, como CFI (≥ 95), TLI (≥ 95) y SRMR ($\leq .05$). Se evaluaron tres modelos: dimensiones correlacionadas y unidimensional en el UWES-9, y unidimensional en el UWES-3. Se utilizó el programa lavaan (Rosseel, 2012) y sem-Tools (Jorgensen, Pornprasertmanit, Schoemann, & Rosseel, 2020).

Confiabilidad. Se examinó la confiabilidad de consistencia interna, mediante el coeficiente MS (Molenaar & Sijtsma, 1984, 1988) desde el modelamiento Mokken Scaling Analysis; desde el modelamiento lineal, se estimaron los coeficientes alfa (α) y omega (ω), con intervalos de confianza en 95%.

Equivalencia. La equivalencia entre las versiones (UWES-9 y UWES-3) fue evaluada mediante: a) la correlación ajustada entre sus puntajes directos (Levy, 1967); y el coeficiente AC1 (Gwet, 2008, 2019), para el acuerdo en la clasificación de los participantes en cuartiles. Finalmente, la equivalencia de la validez de constructo se hizo mediante la comparación de correlaciones dependientes (Zou, 2007), entre el UWES-9 y UWES-3 con las escalas de CES-QT y variables demográficas (sexo y edad). Se usó el programa cocor (Diedenhofen & Musch, 2015).

Resultados

Análisis preliminar.

Información descriptiva de los ítems. La tendencia de respuesta del UWES (Tabla 1) se ubicó entre las opciones 3 (“regularmente: algunas veces al mes”) y 4 (“frecuentemente: una vez por semana”), mientras que la primera opción (“esporádicamente: pocas veces al año o menos”) fue predominantemente escasa y hubo una pobre frecuencia de respuesta; esto fue más claro en los ítems 7 y 8, donde no hubo respuestas en la primera opción. Las diferencias en el ranking promedio de cada ítem fueron estadísticamente significativas (Friedman - $\chi^2 = 278$, gl = 8, $p < .001$), con tamaño de efecto Kendall - $W = .523$ (IC 95% = .514, .262). Respecto a las correlaciones inter-ítem (no mostrada en la Tabla 1), la identidad correlacional no fue mantenida, $\chi^2 = 2601.53$, gl = 36 ($p < .01$), indicando que existen diferencias en la magnitud entre las correlaciones inter-ítem.

Detección de sesgo de respuesta. Mediante el procedimiento Mahalanobis - $D2$, el valor crítico establecidos en la distribución $F(9, 249)$ fue $D2 > 30.67$

Tabla 1. Frecuencia de opciones de respuesta del UWES-9 y resultados correlacionales

	Correlaciones Spearman					Frecuencia de respuesta				M	DE
	Sexo	Il. Trab	Indol.	Desg.	Culpa	1	2	3	4		
uwes1 a	-.114	.455**	-.193**	-.485**	-.220**	2	44	137	70	3.09	.690
uwes2	-.120	.361**	-.177**	-.393**	-.239**	1	37	152	63	3.09	.635
uwes5	-.140*	.514**	-.185**	-.366**	-.144*	3	26	136	88	3.22	.671
uwes3 a	-.139*	.565**	-.240**	-.363**	-.183**	1	16	139	97	3.31	.605
uwes4	-.098	.565**	-.307**	-.328**	-.266**	1	13	122	117	3.40	.607
uwes7	-.018	.410**	-.258**	-.174**	-.174**	0	5	90	158	3.60	.529
uwes6	-.028	.475**	-.189**	-.202**	-.200**	1	22	102	128	3.41	.664
uwes8 a	-.008	.407**	-.188**	-.073	-.101	0	7	125	121	3.45	.551
uwes9	.058	.301**	-.103	-.045	-.018	6	47	123	77	3.07	.763

Nota. Il. Trab.: Ilusión por el Trabajo. Indol.: Indolencia. Desg.: Desgaste. a Ítems de la versión ultra-breve UWES-3

(nivel .05 con ajuste Bonferroni); de este modo, se removieron de la base de datos 6 participantes con valores *D2* entre 31.41 y 52.97. El examen visual del patrón de respuesta de estos sujetos mostró inconsistencia de respuesta en todos los ítems del UWES.

Dimensionalidad

La dimensionalidad del UWES-9 en un solo factor fue respaldada por cinco métodos (55.5%; óptimas coordenadas, factor de aceleración, análisis paralelo, MAP-Velicer y VSS complejidad 1), mientras que tres métodos (33.3%) respaldaron dos factores, y solo un método (BIC ajustado por tamaño muestral) indicó tres factores. El método de CD señaló dos posibles factores. La mayor convergencia ocurrió para la unidimensionalidad con los métodos más recomendados en la práctica de investigación (Auerswald, & Moshagen, 2019), y por lo tanto el análisis factorial se orientó hacia la estimación de los parámetros de un modelo unidimensional general de la versión de 9 ítems. Desde otro enfoque, con el enfoque no paramétrico MSA (ver siguiente sección), se procedió a realizar la verificación de la unidimensionalidad.

Mokken Scale Analysis (MSA)

Identificación de escalas. En primer lugar, el procedimiento AISP arrojó con más claridad que se puede obtener una sola escala en el rango de *H* entre

.30 a .50 (Tabla 2, lado izquierdo). En el nivel más alto (*H* > .50) no se identificó una configuración de tres escalas, lo que converge con el análisis paramétrico del párrafo anterior, en que los ítems convergen en una sola dimensión del UWES-9.

Escalabilidad. El coeficiente *H* para el puntaje total (Tabla 2, parte central) fue alto (*H* = .554, e.e. = .033, IC 95%: .489, .618); en varones la escalabilidad fue moderadamente mayor (*H* = .604, e.e. = .044, IC 95% = .517, .690) comparado con las mujeres (*H* = .492, e.e. = .048, IC 95% = .397, .586). La escalabilidad inter-ítem (*H_{ij}*) varió entre .869 (e.e. = .059) y .029 (e.e. = .098), con el ítem 9 mostrando los coeficientes más bajos con el resto de los ítems. Finalmente, la escalabilidad para cada ítem (*H_i*) fue superior a .540, excepto para el ítem 9. Todos los coeficientes *H* fueron estadísticamente significativos (*z* > 2.50).

Dependencia local. *W*(²) varió entre 27.82 y 41.80, *W*(¹) entre .370 y 7.565, y *W*(³) entre 8.33 y 9.248; ninguno fue identificado como evidencia sustancial de dependencia local entre los ítems.

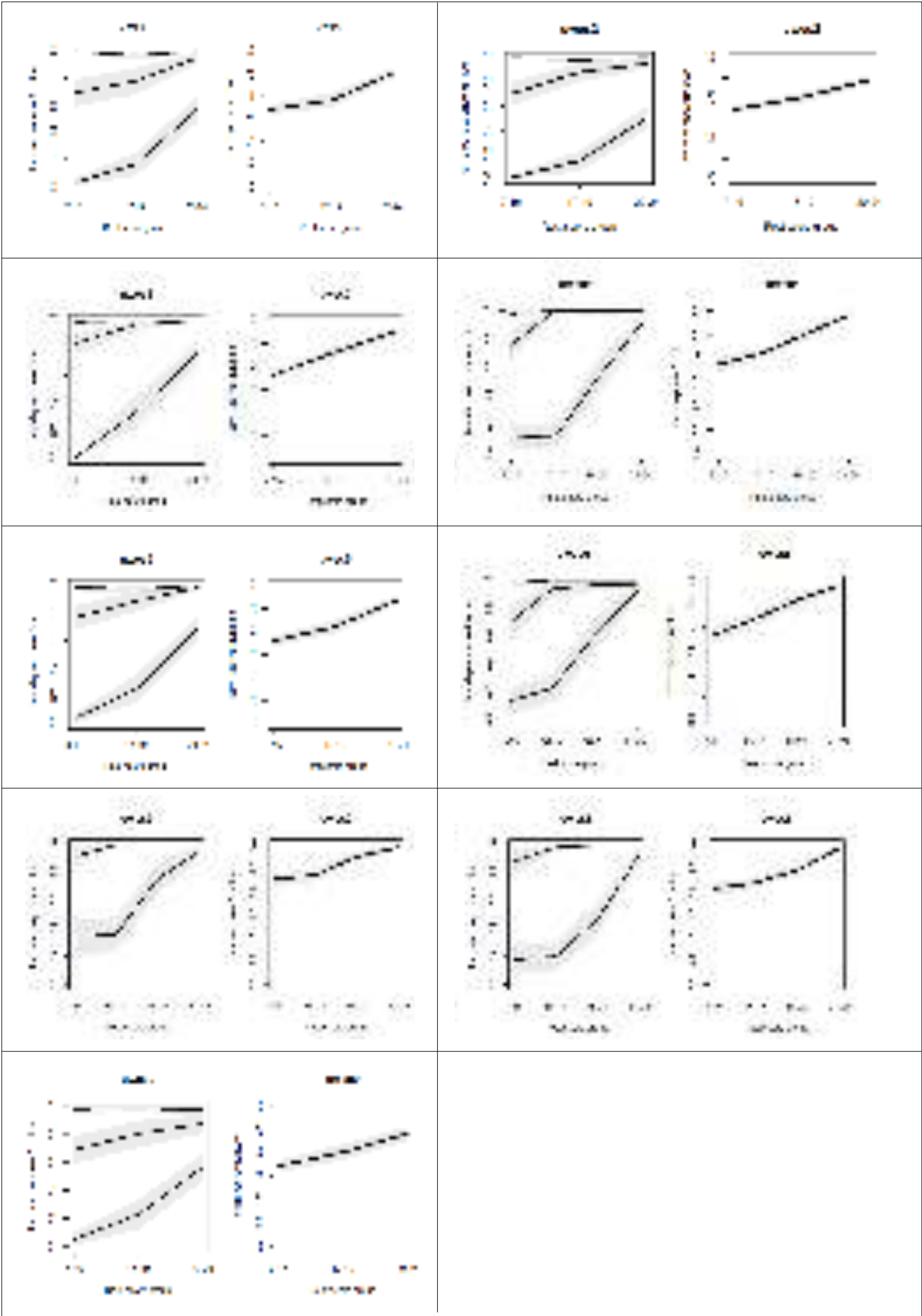
Monotonicidad. En la Figura 1 se muestra el patrón incremental de las opciones de respuesta, y del puntaje de los ítems. Las curvas menos diferenciadas con las correspondientes a las primeras opciones de respuesta, y todos los ítems mantienen un incremento constante y prácticamente lineal.

Tabla 2. Escalamiento Mokken: Procedimiento automatizado de selección de ítems (AISP)

	AISP (método normal)					Escalamiento			Homogeneidad		
						(Hi)			monotónica (MHM)		
	.30	.40	.50	.60	.70	Total	Hom	Muj	#ac	vi	CRIT
uwes1 ^a	1	1	1	1	1	.542	.629	.444	7	0	0
uwes2	1	1	1	1	1	.542	.581	.484	7	0	0
uwes5	1	1	1	1	0	.576	.611	.529	7	0	0
uwes3 ^a	1	1	1	1	1	.651	.709	.581	2	0	0
uwes4	1	1	1	1	2	.630	.631	.623	8	0	0
uwes7	1	1	1	1	2	.603	.683	.493	6	0	0
uwes6	1	1	1	0	0	.564	.586	.537	12	0	0
uwes8 ^a	1	1	1	2	0	.549	.582	.508	7	0	0
uwes9	1	0	0	2	0	.377	.454	.305	9	0	0
Total	-	-	-	-	-	.554	.604	.492	-	-	-

Nota. AISP: identificación de escalas derivadas de los ítems; 0 significa que el ítem no se agrupa en ninguna. Hi: escalabilidad para cada ítem. ^a Ítems de la versión abreviada UWES-3

Figura 1.
Monotonicidad de los ítems: opciones de respuesta y puntaje del ítem.



Modelamiento lineal (SEM)

En la versión UWES-9 (Tabla 3), el modelo de tres factores correlacionados se ajustó satisfactoriamente: WLSMV- $\chi^2 = 122.636$, $gl = 2$ ($p < .01$); CFI = .989, TLI = .984, SRMR = .076; sin embargo, las correlaciones interfactoriales fueron elevadas, sugiriendo falta de discriminación conceptual entre ellas. El modelo unidimensional también fue satisfactorio, WLSMV- $\chi^2 = 316.665$, $gl = 27$ ($p < .01$); CFI = .968, TLI = .958, SRMR = .122, aunque ligeramente bajo. El ajuste del UWES-3 fue completamente satisfactorio: WLSMV- $\chi^2 = 638.132$, $gl = 3$ ($p < .01$); CFI = 1.000, TLI = 1.000, SRMR = .000. Las cargas factoriales en todos los modelos fueron elevados ($> .50$) y estadísticamente significa-

tivos. El ítem 3 en el modelo del UWES-3 mostró carga superior a 1.0, pero no fue necesariamente un caso *Heywood* debido a su magnitud. No se detectaron modificaciones al modelo unidimensional del UWES-9, y no fueron explorados. Se observa también que los ítems correspondientes al UWES-3 constantemente fueron los de mayor carga factorial comparados con el resto de los ítems, en los modelos de tres dimensiones y en el de una dimensión (Tabla 3). Los umbrales (la separación de las opciones en el continuo latente del puntaje del UWES-3) fueron irregulares en su separación, y el ítem 8 mostró que la última opción de respuesta no fue funcional debido a la ausencia de respuestas, y el umbral no fue estimado.

Tabla 3. Modelamiento SEM del UWES-9 y UWES-3

	UWES – 9						UWES – 3			
	3 dimensiones			Unidimensional		FT	Umbrales			
	VI	DED	AB	R ²	FT		R ²	τ_1	τ_2	τ_3
uwes1a	.918			.842	.869	.755	.668	-2.143	-.908	.593
uwes2	.880			.775	.844	.712	-	-	-	-
uwes5	.891			.794	.789	.622	-	-	-	-
uwes3 a		.959		.919	.912	.832	1.126	-2.656	-1.497	.297
uwes4		.905		.819	.876	.768	-	-	-	-
uwes7		.812		.660	.783	.614	-	-	-	-
uwes6			.875	.765	.747	.559	-	-	-	-
uwes8 a			.866	.750	.756	.571	.521	-1.916	.05	-
uwes9			.617	.381	.529	.280				
Correlaciones										
VI	1.0				-	-	-	-	-	-
DED	.717	1.0			-	-	-	-	-	-
AB	.470	.707	1.0		-	-	-	-	-	-
Confiabilidad										
α					.877		.703			
(IC 95%)					(.84, .89)		(.62, .76)			
ω					.871		.754			
(IC 95%)					(.83, .89)		(.70, .81)			

Nota. VI: Vigor. DED: Dedicación. AB: Absorción. FT: factor total. Ítems de la versión abreviada del UWES. α y ω : coeficientes de confiabilidad. a Ítems de la versión abreviada UWES-3

Confiabilidad

Desde el modelo MSA, la confiabilidad MS para el UWES-3 fue .892, un nivel que puede considerarse alto. Asimismo, las confiabilidades alfa (α) y omega (ω) fueron moderadamente altas ($\geq .70$), y se observa leve discrepancia entre ambas (Tabla 3, parte inferior).

Equivalencia entre versiones

La correlación no ajustada ($r = .931$, $p < .01$) y ajustada ($r = .824$, $p < .01$) entre UWES-9 y UWES-3 fue elevada, indicando muy similar ordenamiento de sujetos desde ambos puntajes. El acuerdo clasificatorio fue moderadamente alto, $AC_1 = .611$ (IC 95% = .537, .687), indicando la tendencia a diferenciar sujetos cuando son clasificados en cuartiles. Respecto a equivalencia de la validez o discriminación de constructo con el síndrome de quemarse por el trabajo, las correlaciones del UWES-9 (puntaje único basado en el modelo unidimensional;

ver párrafos arriba) y UWES-3 obtenidas mostraron coherencia teórica (Tabla 4). La correlación con la dimensión de Ilusión por el Trabajo del CESQT-PE fue positiva y alta, mientras que con el resto de los puntajes fue de dirección negativa y bajas o moderadamente bajas, como se esperaba. Respecto la diferencia correlacional, ambas versiones produjeron magnitudes correlacionales similares; la magnitud de estas diferencias estimadas fue pequeña, porque: a) entre 8% y 11% más bajo respecto a la versión UWES-9, b) los intervalos de confianza para las diferencias correlacionales (Δ_r) incluyeron el cero (excepto en Ilusión por el trabajo), y c) la diferencia estandarizada (q) fue entre .020 y .083, es decir, de magnitud insustancial. Considerando la edad y el sexo de los participantes, los puntajes de ambas versiones correlacionaron cero en la población, y la diferencia correlacional fue insustancial (q alrededor de cero) (Tabla 4).

Tabla 4.
Correlaciones entre versiones del UWES con CESQT

	Correlaciones		Diferencia correlacional		M	DE
	UWES-9	UWES-3	Δ_r (IC 95%)	q		
Sexo	-.082	-.109	(-.018, .072)	.027	-	-
Edad	.068	.035	(-.012, .078)	.033	-	-
CESQT						
Ilusión por el trabajo	.618**	.564**	(.017, .095)*	.083	3.33	.653
Indolencia	-.245**	-.226**	(-.064, .025)	.020	.67	.555
Desgaste	-.384**	-.413**	(-.013, .072)	.034	1.75	.976
Culpa	-.188**	-.168**	(-.065, .025)	.020	.62	.620
M	29.66	9.85	-	-	-	-
DE	4.07	1.47	-	-	-	-

Nota. Δ_r : diferencia de correlaciones dependientes. q : magnitud de la diferencia. CESQT: dimensiones del Cuestionario para la Evaluación del Síndrome de Quemarse por el Trabajo. * $p < .05$; ** $p < .01$

Discusión

El objetivo del presente estudio fue analizar las propiedades psicométricas de la versión ultra-breve UWES-3 y analizar su equivalencia con la UWES-9 en una muestra de docentes universitarios mexicanos. La estrategia metodológica comprendió una exploración de la dimensionalidad y un examen de las propiedades estructurales, mediante estimaciones paramétricas y no paramétricas, así como un análisis de equivalencia de los patrones de correlaciones directos y ajustados, tanto entre ambas versiones (UWES-3 y UWES-9) como con las dimensiones del CESQT, el sexo y la edad.

Dicha estrategia metodológica elegida tiene ventajas sobre estudios previos, específicamente en la inclusión de estimaciones de correlación ajustadas, omitidas en el estudio de [Schaufeli et al. \(2017\)](#) y la estimación de resultados no paramétricos junto con los resultados paramétricos, omitidas en el estudio de [Domínguez-Lara et al. \(2021\)](#). Respecto a las correlaciones ajustadas, el control de la varianza de error mediante la corrección de [Levy \(1968\)](#) aplicada a la correlación entre ambas versiones, es una práctica que se recomienda para obtener conclusiones válidas de equivalencia ([Merino-Soto, & Angulo-Ramos, 2013](#); [Smith, McCarthy, & Anderson, 2000](#)), debido a que se tienden a obtener tamaños correlacionales espurios sin esta corrección. Por otro lado, la aplicación del modelamiento no paramétrico generalmente es considerado como un precursor de modelos paramétricos, debido que no requiere una conocida distribución del constructo (eg, distribución normal) ([Mokken, 1971](#); [Van der Ark, 2012](#)).

Los análisis dimensionales y estructurales paramétricos y no paramétricos coincidieron en la confirmación de una estructura unidimensional del UWES-9, donde justo los 3 ítems de la versión breve presentaron las cargas factoriales más altas y un ajuste paramétrico totalmente satisfactorio. Esto confirma la unidimensionalidad esencial del UWES-3, con alta validez de los ítems en relación con su constructo. Por otro lado, la consistencia interna fue aceptable para ambas versiones, aunque superior para el UWES-9. El decremento de la confiabilidad de los puntajes del UWES-3 es una consecuencia del número de ítems reducido y su

consecuente restricción de la varianza del constructo, pero no es necesariamente un problema en una visión relativa, porque la magnitud apropiada o razonable de la confiabilidad de un puntaje está condicionado por el uso de este puntaje y el contexto de aplicación (American Educational Research Association (AERA), American Psychological Association (APA) & National Council for Measurement in Education (NCME), 2014). Por lo tanto, con una magnitud de confiabilidad alrededor de .75 de los puntajes del UWES-3, su uso parece adecuarse a evaluaciones de tamizaje, monitoreo, vigilancia psicosocial y descripciones de grupos. Para usos adicionales que requieran mayor precisión del puntaje, la versión de UWES-9 puede estar mejor indicada. Como nota adicional, la diferencia entre el coeficiente alfa y omega del UWES-3 es consecuencia de la distancia de la carga factorial del ítem 3 respecto a los otros dos ítems. Una implicación es que, si esto se mantiene en otros estudios, entonces el coeficiente adecuado para estimar la confiabilidad es ω , no α .

Uno de los hallazgos del estudio fue la magnitud de los coeficientes de escalabilidad H en el grupo de mujeres, comparado con los varones. Los coeficientes de escalamiento H son interpretables como coeficientes de discriminación de un ítem respecto al puntaje de la media ([Van Abswoude, Van der Ark, & Sijsma, 2004](#)), y parcialmente se refiere también a la calidad de medición de ítem. Esto implica que las propiedades de escalamiento y de calidad general del UWES-3 en mujeres, no son tan fuertes como en el caso de los varones. Resultados como los obtenidos aquí, en que las diferencias de grupo son aparentes en la calidad de medición, requieren más investigación ([Wind, 2017](#)). En el funcionamiento de diferencial de ítems, dentro del marco *Mokken Scale Analysis* (MSA), la intersección de curvas características de las opciones de respuesta puede estar detrás de estas diferencias, demostrando una menor capacidad de discriminación de diferencias individuales en mujeres. Es probable que estas diferencias sean idiosincrásicas a la muestra de participantes, y por lo tanto no replicables, sin embargo, esto alerta sobre la necesidad de incluir otros indicadores en una condición de mayor tamaño muestral.

En cuanto a la equivalencia de las versiones UWES-3 y UWES-9, las correlaciones entre ambas versiones presentaron coeficientes superiores a $r = .82$, con un acuerdo clasificatorio moderadamente alto y los patrones de correlaciones con el CESQT y las variables demográficas (sexo, edad) fueron prácticamente similares tanto teórica como empíricamente, en la misma dirección y con diferencias de tamaño estadísticamente insustanciales. Ello que lleva a la conclusión que el UWES-3 es psicométricamente equivalente al UWES-9 para la evaluación del engagement laboral en profesores universitarios.

Resulta relevante señalar que la versión ultra-breve de 3 ítems (UWES-3) representa en cada ítem los tres mismos aspectos conceptuales importantes del engagement que fueron acuñados desde su origen: vigor, dedicación y absorción, por lo que no se pone en riesgo la validez de contenido, ni se pierde representación teórica construida en sus raíces. En general, no hay pérdidas significativas de información o varianza con respecto a las versiones del UWES con más ítems, lo que implica beneficios prácticos implícitos.

Respecto a ello, actualmente existe una tendencia de investigar la viabilidad de medidas ultra-breves debido a los beneficios prácticos que esto tiene, las ventajas son múltiples e incluyen entre otras: tomar menos tiempo de participantes e instituciones en el llenado de cuestionarios, disminuir las tasas de deserción por cuestionarios largos, evitar la fatiga o monotonía de los respondientes (lo que al mismo tiempo asegura la validez de las mediciones), y no menos importante, se facilita su inclusión en estudios epidemiológicos grandes o encuestas nacionales midiendo los mismos constructos sin muchos ítems. Estos beneficios pueden hacerse extensivos a la versión breve UWES-3 analizada en este estudio y la opción de una evaluación más práctica y sencilla del engagement laboral en profesores universitarios, aunque parecen necesarios más estudios que confirmen la tendencia encontrada hasta ahora.

Las limitaciones del estudio pudieran identificarse en el tamaño y la estrategia no aleatoria de la muestra, la elección de pocas variables para la discriminación (dimensiones CESQT, edad, sexo) y la implícita falta de independencia del uso del

UWES-9 y UWES-3 en el mismo estudio, lo que dificulta entender su verdadera confiabilidad y validez de forma verdaderamente independiente (Schaufeli et al., 2017). Sin embargo, las estimaciones no paramétricas y robustas utilizadas pueden atenuar en alguna medida los problemas de muestreo. Asimismo, la elección del burnout o síndrome de quemarse por el trabajo, medido mediante el CESQT, como criterio de análisis de equivalencia pudo ser suficiente dada la divergencia teórica esperada entre ambos constructos, considerando que el engagement nació y cobró identidad como justo un opuesto del Burnout. Corresponde a futuros estudios analizar el patrón de correlaciones con otras variables, en otras muestras, utilizando de forma conjunta e independiente el UWES-9 y el UWES-3 para confirmar tal equivalencia en distintos escenarios. Entretanto, sirva este estudio como una aportación inicial en este sentido, en beneficio de la medición del engagement en la fuerza laboral docente latinoamericana.

Referencias

- Álvarez Garón, S. M., & Peña Fuentes, J. N. (2019). Adaptación del Utrecht Work Engagement Scale–UWES en profesores universitarios de Villavicencio (Tesis de pregrado) *Universidad Cooperativa de Colombia, Villavicencio*. <http://hdl.handle.net/20.500.12494/12393>
- American Educational Research Association, American Psychological Association & National Council for Measurement in Education [AERA, APA & NCME] (2014). *The Standards for Educational and Psychological Testing*. Washington, D.C.: AERA
- Auerswald, M., & Moshagen, M. (2019). How to determine the number of factors to retain in exploratory factor analysis: A comparison of extraction methods under realistic conditions. *Psychological Methods*, 24(4), 468–491. <https://doi.org/10.1037/met0000200>
- Burić, I., & Macuka, I. (2018). Self-efficacy, emotions and work engagement among teachers: A two wave cross-lagged analysis. *Journal of Happiness Studies*, 19(7), 1917–1933. <https://doi.org/10.1007/s10902-017-9903-9>

- Calderón-de la Cruz, G. A., Merino-Soto, C., Juárez-García, A., Domínguez-Lara, S., & Fernández-Arata, M. (2020). ¿Es replicable la estructura factorial del Maslach Burnout Inventory Human Service Survey (MBI-HSS) en la profesión de enfermera del Perú?: un estudio nacional. *Enfermería Clínica*, 30(5), 340-348. <https://doi.org/10.1016/j.enfcli.2019.12.013>
- DeCarlo, L. T. (1997). On the meaning and use of kurtosis. *Psychological Methods*, 2(3), 292-307. <https://doi.org/10.1037/1082-989X.2.3.292>
- Diedenhofen, B. & Musch, J. (2015). A Comprehensive Solution for the Statistical Comparison of Correlations. *PLoS ONE*, 10(4), e0121945. <http://doi.org/10.1371/journal.pone.0121945>
- Domínguez-Lara, S. A., Fernández-Arata, M., & Seperak-Viera, R. (2021). Análisis psicométrico de una medida ultra-breve para el engagement académico: UWES-3S: UWES-3S en estudiantes universitarios. *Revista Argentina de Ciencias del Comportamiento*, 13(1), 25-37. <https://doi.org/10.32348/1852.4206.v13.n1.27780>
- Flores Jiménez, C., Fernández Arata, M., Juárez García, A., Merino Soto, C., & Guimet Castro, M. (2015). Entusiasmo por el trabajo (engagement): un estudio de validez en profesionales de la docencia en Lima, Perú. *Liberabit*, 21(2), 195-206. http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1729-48272015000200003
- Gil-Monte, P. R. (2011). Evaluación del cuestionario CESQT. Madrid.
- Gil-Monte, P. R., & Noyola, C. V. (2011). Estructura factorial del "Cuestionario para la Evaluación del Síndrome de Quemarse por el Trabajo" en maestros mexicanos de educación primaria. *Revista Mexicana de Psicología*, 28(1), 75-84. <http://www.redalyc.org/articulo.oa?id=243029630007>
- Gil-Monte, P. R., Unda, S. R., & Sandoval, O. J. I. (2009). Validez factorial del « Cuestionario para la Evaluación del Síndrome de Quemarse por el Trabajo » (CESQT) en una muestra de maestros mexicanos. *Salud Mental*, 32(3), 205-214. http://www.revistasaludmental.mx/index.php/salud_mental/article/view/1285
- González-Romá, V., Schaufeli, W.B., Bakker, A. & Lloret, S. (2006). Burnout and engagement: Independent factors or opposite poles? *Journal of Vocational Behavior*, 68, 165-174. <https://doi.org/10.1016/j.jvb.2005.01.003>
- Gwet, K. L. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology*, 61, 29-48. <https://doi.org/10.1348/000711006X126600>
- Gwet, K. L. (2019). irrCAC: Computing Chance-Corrected Agreement Coefficients (CAC). R package version 1.0. <https://CRAN.R-project.org/package=irrCAC>
- Han, J., Yin, H., Wang, J., & Zhang, J. (2020). Job demands and resources as antecedents of university teachers' exhaustion, engagement and job satisfaction. *Educational Psychology*, 40(3), 318-335. <https://doi.org/10.1080/01443410.2019.1674249>
- Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2020). semTools: Useful tools for structural equation modeling. R package version 0.5-3. <https://CRAN.R-project.org/package=semTools>
- Juárez-García, A. (2015). Investigaciones psicométricas de escalas psicosociales en trabajadores mexicanos. México, DF: Plaza y Valdés.
- Levy, P. (1967) The correction for spurious correlation in the evaluation of short-form tests. *Journal of Clinical Psychology*, 23, 84-86. [https://doi.org/10.1002/1097-4679\(196701\)23:1<84::aid-jclp2270230123>3.0.co;2-2](https://doi.org/10.1002/1097-4679(196701)23:1<84::aid-jclp2270230123>3.0.co;2-2)
- Loevinger, J. (1948). The technic of homogeneous tests compared with some aspects of scale analysis and factor analysis. *Psychological Bulletin*, 45(6), 507-529. <https://doi.org/10.1037/h0055827>
- Mahalanobis, P. C. (1936). On the generalized distance in statistics. *Proceedings of the National Institute of Science of India*, 12(1), 49-55. <http://hdl.handle.net/10263/6765>
- Makowski, (2018). The psycho Package: an efficient and publishing-oriented workflow for psychological science. *Journal of Open Source Software*, 3(22), 470. <https://doi.org/10.21105/joss.0047>
- Mangiafico, S. (2021). rcompanion: Functions to support extension education program evaluation. R package version 2.4.1. <https://CRAN.R-project.org/package=rcompanion>
- Martins, E. L. D. S., & Mendonça, H. (2019). Evidencias de validez de la escala de compromiso de docentes de educación básica. *Psicología em Pesquisa*, 13(3), 159-172. <http://dx.doi.org/10.34019/1982-1247.2019.v13.27110>
- Maslach, C., & Leiter, M. P. (1997). The truth about burnout. San Francisco, CA: Jossey-Bass
- Maslach, C., Jackson, S.E., Leiter, M.P., Schaufeli, W.B. and Schwab, R.L. (1986) Maslach burnout inventory instruments and scoring guides forms: General,

- human services, & educators. *Health and Quality of life Outcomes*, 7, 31. <http://www.mindgarden.com>
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, 17(3), 437–455. <https://doi.org/10.1037/a0028085>
- Merino-Soto, C., & Angulo-Ramos, M. (2013). Validación en Chile de la escala de sobrecarga del cuidador de Zarit en sus versiones original y abreviada: corrección. *Revista Médica de Chile*, 141, 1083-1084. <http://dx.doi.org/10.4067/S0034-98872013000800019>
- Mokken, R. J. (1971) *A Theory and Procedure of Scale Analysis*. Berlin, Germany: De Gruyter.
- Molenaar, I. W., & Sijtsma, K. (1984). Internal consistency and reliability in Mokken's nonparametric item response model. *Tijdschrift voor onderwijsresearch*, 9, 257–268.
- Molenaar, I. W., & Sijtsma, K. (1988). Mokken's approach to reliability estimation extended to multicategory items. *Kwantitatieve methoden*, 9(28), 115-126. <https://research.tilburguniversity.edu/en/publications/mokkens-approach-to-reliability-estimation-extended-to-multicateg>
- Revelle, W. (2020) *psych: Procedures for Personality and Psychological Research*, Northwestern University, Evanston, Illinois. R package version 2.0.12. <https://CRAN.R-project.org/package=psych>
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1-36. URL <http://www.jstatsoft.org/v48/i02/>.
- Ruscio, J. (2018). RGenData: Generates Multivariate Nonnormal Data and Determines How Many Factors to Retain. R package version 1.0. <https://CRAN.R-project.org/package=RGenData>
- Ruscio, J., & Roche, B. (2012). Determining the number of factors to retain in an exploratory factor analysis using comparison data of known factorial structure. *Psychological Assessment*, 24(2), 282–292. <https://doi.org/10.1037/a0025697>
- Salanova, M., & Schaufeli, W. (2009). *El engagement en el trabajo: cuando el trabajo se convierte en pasión*. Alianza Editorial; Madrid.
- Salanova, M., Schaufeli, W.B., Llorens, S., Peiro, J.M. & Grau, R. (2000). Desde el ‘burnout’ al ‘engagement’: ¿Una nueva perspectiva? [From ‘burnout’ to ‘engagement’: A new perspective?] *Revista de Psicología del Trabajo y las Organizaciones*, 16, 117-134. <https://journals.copmadrid.org/jwop/art/7c590f-01490190db0ed02a5070e20f01>
- Schaufeli, W. B., & Bakker, A. B. (2010). Defining and measuring work engagement: Bringing clarity to the concept. In A. B. Bakker & M. P. Leiter (Eds.), *Work engagement: A handbook of essential theory and research* (pp.10-24). New York: Psychology Press
- Schaufeli, W. B., & Bakker, A. B., & Salanova, M. (2006). The measurement of work engagement with a short questionnaire: A cross-national study. *Educational and Psychological Measurement*, 66, 701-716. <https://doi.org/10.1177/0013164405282471>
- Schaufeli, W. B., Salanova, M., González-Romá, V., & Bakker, A. B. (2002). The measurement of engagement and burnout: A two sample confirmatory factor analytic approach. *Journal of Happiness studies*, 3(1), 71-92.
- Schaufeli, W. B., Shimazu, A., Hakanen, J., Salanova, M., & De Witte, H. (2019). An ultra-short measure for work engagement: The UWES-3 validation across five countries. *European Journal of Psychological Assessment*, 35(4), 577–591. <https://doi.org/10.1027/1015-5759/a000430>
- Schaufeli, W.B. (2017). General engagement: Its conceptualization and measurement. *Journal of Well-Being Assessment*, 1, 9-24.
- Schaufeli, W.B. (2018). *Burnout in Europe: Relations with national economy, governance, and culture*. Research Unit Occupational & Organizational Psychology and Professional Learning (internal report). KU Leuven: Belgium.
- Sijtsma, K., & Van der Ark, L. A. (2017). A tutorial on how to do a Mokken scale analysis on your test and questionnaire data. *British Journal of Mathematical and Statistical Psychology*, 70(1), 137-158. <https://doi.org/10.1111/bmsp.12078>
- Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment*, 12(1), 102–111. <https://doi.org/10.1037/1040-3590.12.1.102>
- Steiger, J. H. (1980). Testing pattern hypotheses on correlation matrices: alternative statistics and some empirical results. *Multivariate Behavioral Research*, 15, 335-352. https://doi.org/10.1207/s15327906mbr1503_7
- Straat J. H., Van der Ark L. A., Sijtsma K. (2016). Using conditional association to identify locally independent item sets. *Methodology* 12 117–123. <https://doi.org/10.1027/1614-2241/a000115>

- Straat, J. H., Van der Ark, L. A., & Sijtsma, K. (2013). Comparing optimization algorithms for item selection in Mokken scale analysis. *Journal of Classification*, 30, 72-99. <https://doi.org/10.1007/s00357-013-9122-y>
- Straat, J. H., Van der Ark, L. A., and Sijtsma, K. (2013). Comparing optimization algorithms for item selection in Mokken scale analysis. *Journal of Classification*, 30, 72-99. <https://doi.org/10.1007/s00357-013-9122-y>
- Van Abswoude, A.A.H., Van der Ark, L.A., & Sijtsma, K. (2004). A comparative study of test dimensionality assessment procedures under nonparametric IRT models. *Applied Psychological Measurement*, 28, 3-4. <http://dx.doi.org/10.1177/0146621603259277>
- Van der Ark, L. A. (2012). New developments in Mokken Scale Analysis in R. *Journal of Statistical Software*, 48(5), 1-27. <http://www.jstatsoft.org/v48/i05/>
- Van Schuur, W.H. (2003). Mokken Scale Analysis: between the Guttman scale and parametric item response theory. *Political Analysis*, 11, 139-163. <https://doi.org/10.1093/pan/mpg002>
- Wetzel, E., Böhneke, J. R., & Brown, A. (2016). Response biases. In F. T. L. Leong, D. Bartram, F. M. Cheung, K. F. Geisinger, & D. Iliescu (Eds.), *The ITC international handbook of testing and assessment* (p. 349-363). Oxford University Press.
- Wind, S.A. (2017). An Instructional Module on Mokken Scale Analysis. *Educational Measurement: Issues and Practice*, 36: 50-66. <https://doi.org/10.1111/emip.12153>
- World Medical Association (AMM). (2013). Declaración de Helsinki de la AMM - Principios éticos para las investigaciones médicas en seres humanos. *World Medical Association, Inc.* <https://doi.org/10.1001/jama.2013.281053>
- Wyatt, M., & Dikilitaş, K. (2016). English language teachers becoming more efficacious through research engagement at their Turkish university. *Educational Action Research*, 24(4), 550-570. <http://dx.doi.org/10.1080/09650792.2015.1076731>
- Zhu, N. Q. (2001). *The effects of teachers' flow experiences on the cognitive engagement of students*. (Tesis Doctoral). Universidad de San Diego.
- Zou, G. Y. (2007). Toward using confidence intervals to compare correlations. *Psychological Methods*, 12, 399-413. <http://doi.org/10.1037/1082-989X.12.4.399>

Rendimiento académico, valoración del docente e interacción alumno-profesor en la docencia online frente a presencial

Alejandro M. Fernández-Castro *¹; Roberto Sánchez-Cabrero*²; Yousef Hussein Eiadat*³

¹School of Business, University College Dublin, Belfield, Dublin, Ireland <https://orcid.org/0000-0002-3484-0185> alejand.fernandez@ucd.ie ²Interfacultative Department of Evolutionary and Educational Psychology, Faculty of Teacher Training, Autonomous University of Madrid, Spain, <https://orcid.org/0000-0002-1978-7531> roberto.sanchez@uam.es ³School of Business, University College Dublin, Belfield, Dublin, Ireland <https://orcid.org/0000-0002-2126-2275> yousef.husein@ucd.ie

Citar como: Fernández-Castro, A., Sánchez-Cabrero, R., Hussein, Y. (2023). Rendimiento académico, valoración del docente e interacción alumno-profesor en la docencia online frente a presencial. *Revista Digital de Investigación en Docencia Universitaria*, 17(2), e1583. <https://doi.org/10.19083/ridu.2023.1583>

Recibido: 13/12/21. **Revisado:** 07/02/22. **Aceptado:** 21/08/2022. **Publicado:** 30/06/2023.

Resumen

Introducción: el cambio de entorno de aprendizaje, de presencial a online, provocado por la pandemia del coronavirus SARS-COV2 en el ámbito universitario de Singapur tuvo impacto en el rendimiento académico, la valoración del docente y la interacción alumno-profesor. **Objetivo:** evaluar el impacto del cambio de entorno de aprendizaje en el rendimiento académico, la valoración del docente y la interacción alumno-profesor, **Método:** se compararon las calificaciones y las encuestas de evaluación de 282 estudiantes de dos cohortes que en el año 2020 cursaron de forma consecutiva formación universitaria presencial y online. **Resultados:** los resultados muestran diferencias en el rendimiento académico, pero no así la valoración del docente y la interacción alumno-profesor. **Discusión:** es posible mantener una calidad educativa óptima sin alterar la valoración del docente y la interacción alumno-profesor, tanto en la modalidad online como presencial, pero es necesario presentar un método de evaluación coherente con la modalidad utilizada.

Palabras clave: educación a distancia; evaluación online; COVID-19; valoración del docente; interacción alumno-profesor; rendimiento académico.

Academic performance, teacher assessment and student-teacher interaction in online versus face-to-face teaching

Abstract

Introduction: The change in the learning environment, from face-to-face to online, caused by the SARS-COV2 coronavirus pandemic in the Singapore university environment had an impact on academic performance, teacher assessment and student-teacher interaction. **Objective:** To evaluate the impact of the change in the learning environment on academic performance, teacher assessment and student-teacher interaction. **Method:** The grades and evaluation surveys of 282 students from two cohorts who studied in 2020 were compared consecutively face-to-face and online university training. **Results:** The results show differences in academic performance, but not the teacher's assessment and student-teacher interaction. **Discussion:** It is possible to maintain an optimal educational quality without altering the teacher's assessment

*Correspondencia:

Roberto Sánchez-Cabrero
roberto.sanchez@uam.es



and the student-teacher interaction, both in the online and face-to-face modes, but it is necessary to present an evaluation method consistent with the modality used.

Keywords: long distance education; online evaluation; COVID-19; teacher assessment; student-teacher interaction; academic performance.

Introducción

En la localidad de Wuhan (China) a finales del año 2019 hizo su aparición mundial el coronavirus SARS-COV2, causante de la sintomatología patológica categorizada como COVID-19, altamente contagiosa y para la cual el sistema inmunológico humano no estaba convenientemente preparado. La rápida propagación del coronavirus SARS-COV2 en todo el mundo derivó en el estallido de la pandemia mundial que ha afectado profundamente a todos los países del mundo a lo largo de los dos siguientes años. Esta pandemia, derivada por el coronavirus SARS-COV2, ha tenido enormes consecuencias sociales en todos los ámbitos y niveles ([Hosseinzadeh et al., 2022](#); [Szcześniak et al., 2021](#)). Uno de los ámbitos más afectados ha sido el ámbito educativo, puesto que el estallido de la pandemia cerró de forma prolongada todos los centros educativos en todos los niveles escolares e impidió la modalidad de educación presencial durante dos cursos escolares en la mayoría de los países. Además, las consecuencias de la pandemia ocasionaron innumerables y variadas medidas que afectaron de manera profunda en el desarrollo normal del proceso educativo, sobre todo entre los más vulnerables ([de Oliveira Araújo et al., 2020](#); [Vlachos et al., 2021](#)).

[Pérez y Hernández \(2020\)](#) señalan que, en el caso de España, una vez los centros educativos migraron su actividad al nuevo entorno online durante el confinamiento, el principal problema no fue la falta de equipamiento, sino la escasa formación en el uso de las herramientas tecnológicas, tanto de estudiantes como de profesores. Si bien las posibles carencias de equipamiento fueron debidamente resueltas por los centros y las administraciones, no lo fueron las deficiencias

formativas y de competencia digital, contribuyendo a aumentar la desigualdad. En esta misma dirección apunta [Alharthi \(2020\)](#), quien señala que la mayoría de los que han recibido formación en un entorno virtual, incluso en niveles educativos superiores, consideran fundamental recibir previamente formación específica acerca de la tecnología a emplear en dicho entorno, lo cual no resultó posible durante el confinamiento ([García-Martín y García-Martín, 2021](#); [Sánchez-Cabrero et al., 2021b](#)). Hay quien incluso habla de un cambio de paradigma en la educación tras la pandemia, más allá de los cambios coyunturales experimentados, basado en la combinación de las Tecnologías de la Información y la Comunicación (TIC) y la neuroeducación ([Espino-Díaz et al., 2021](#)).

Entre otros estudios, [Baczek et al. \(2021\)](#) estudiaron los efectos del repentino cambio de entorno de aprendizaje (de presencial a online) motivado por los confinamientos. A partir de las respuestas de 804 estudiantes a su cuestionario, llegaron a la conclusión de que éstos perciben ambos entornos como igualmente válidos para facilitar su aprendizaje, y que el 73% considera el entorno virtual tan ameno como el presencial. Ampliando esta comparativa, respecto a la vinculación del cambio repentino del entorno de aprendizaje con la afectación emocional de los estudiantes, el reciente estudio de [Sánchez-Cabrero et al. \(2022\)](#) observaron cómo, a pesar del cambio tan drástico y las múltiples consecuencias psicológicas y sociales en la población, la inteligencia emocional de los estudiantes no se veía afectaba, lo que permitía mantener un funcionamiento educativo adecuado durante la pandemia sin alterar de manera visible el rendimiento académico o la ansiedad ante los exámenes. No obstante, el análisis comparativo de los entornos de aprendizaje presencial y virtual ya había sido ampliamente tratado con anterioridad a la pandemia. [Paechter y Maier \(2010\)](#), tras analizar la información de 2196 estudiantes acerca de su experiencia y preferencias por la enseñanza presencial o virtual (online), destacan como ventajas de la enseñanza virtual la claridad y coherencia de los materiales didácticos, el aprendizaje autorregulado y la distribución de información, mientras que como aspectos positivos de la enseñanza presen-

cial se señalan la comunicación y las relaciones interpersonales. En [Alonso y Blázquez \(2009\)](#) se analizan las funciones de los docentes en entornos presenciales y virtuales (online) a partir del estudio de cuatro aspectos (contenido teórico, contenido práctico, interacción con el estudiante, y diseño de las actividades), concluyendo que no se aprecian diferencias estadísticamente significativas que se expliquen por el entorno en el que se desempeña la actividad docente. Por su parte, [Johnson et al. \(2000\)](#) realizaron un estudio comparativo de una misma asignatura impartida en un entorno presencial y en uno online. Para ello se analizaron, entre otros, tres aspectos del proceso enseñanza-aprendizaje: los resultados del aprendizaje, la valoración del docente por parte de los alumnos y la interacción alumno-profesor. En este estudio se evidencia una percepción ligeramente más favorable del docente en el entorno presencial, que podría explicarse por una mayor proximidad emocional. Los resultados de aprendizaje resultaron ser muy similares en ambos entornos.

En lo que a la interacción alumno-profesor se refiere, los resultados indican una percepción más favorable en el entorno presencial. En esta misma línea, Paul y Jefferson (2019), tras analizar las calificaciones de 548 alumnos de una asignatura a lo largo de 8 años consecutivos, cursada en formato presencial (401) y online (147), concluyen que no se aprecian diferencias estadísticamente significativas en los resultados alcanzados por los estudiantes, conclusión coincidente por la presentada por [Toyne et al. \(2019\)](#).

Poco antes del inicio la pandemia ocasionada por el coronavirus SARS-COV2, [Bengtsson \(2019\)](#) concluyó en su estudio que los exámenes no presenciales eran la mejor forma de evaluación para los niveles educativos más altos porque facilitan un aprendizaje constructivo, permiten la reflexión y el desarrollo de capacidades cognitivas superiores, y al mismo tiempo convierten la evaluación en un proceso de aprendizaje. Advierte el autor de que estos exámenes no son recomendables en niveles inferiores por el riesgo evidente de conductas no éticas de los estudiantes.

[Myvry y Joutsenvirta \(2015\)](#) estudiaron la realización de exámenes de 'libro abierto', como alternativa idónea para una evaluación sumativa y formativa,

evitando la negativa influencia que puede ejercer el plagio. Concluyen que el tiempo empleado en la preparación es similar en ambos tipos de examen, pero que para la mitad de los estudiantes la realización y el aprendizaje en los exámenes abiertos supuso una inversión mucho mayor de tiempo.

En otro estudio previo a la pandemia, [Gaytan \(2005\)](#) estudió los elementos necesarios para que las técnicas de evaluación sean efectivas en un entorno online y señala la necesidad de que éstas contemplen aspectos tan diversos como los requisitos tecnológicos, el estilo de enseñanza, el enfoque pedagógico y los resultados del aprendizaje. Con posterioridad [Gaytan y McEwen \(2007\)](#) estudiaron las estrategias de evaluación online y las técnicas efectivas. Su trabajo se lleva a cabo a partir de dos cuestionarios enviados a 85 docentes y 1963 estudiantes estadounidenses que participaban en entornos online, del que obtuvieron 361 respuestas (29 profesores y 332 alumnos). El cuestionario remitido a los docentes versaba sobre su percepción acerca de si la calidad de la docencia se veía afectada al desempeñarse en un entorno online, y sobre la efectividad de las estrategias de evaluación en dicho entorno. El enviado a los estudiantes trataba de medir la percepción de éstos sobre la efectividad de Internet como entorno de aprendizaje, y sobre las estrategias de evaluación más efectivas. Tras el análisis de los resultados, los autores concluyen que las estrategias deben basarse en una amplia variedad de tareas a realizar con regularidad, a partir de las cuales el docente pueda aportar, de manera puntual, un *feedback* significativo para su aprendizaje. Establecen que las técnicas efectivas son los proyectos, los portafolios, las autoevaluaciones y co-evaluaciones, y las tareas semanales con *feedback* inmediato. Para [Deeley \(2018\)](#) la evaluación sumativa puede utilizarse para el aprendizaje, de la mano de la tecnología adecuada. Esta transformación de la evaluación 'del aprendizaje' en una evaluación 'para el aprendizaje' debe producirse en un entorno en el que docentes y estudiantes confíen y se sientan cómodos y cercanos, y requiere de flexibilidad y progresividad en su implementación.

Siempre que el contexto lo permita, la preferencia de los estudiantes sobre la evaluación online y presencial debe ser tenida en cuenta ([Hewson, 2012](#)).

Asimismo, [Hewson \(2012\)](#) estudia si esta preferencia condiciona los resultados obtenidos. Para ello, analiza dos cohortes consecutivas de 33 y 41 estudiantes, respectivamente, que cursaron la misma asignatura con los mismos docentes. Todos los estudiantes realizaron el mismo examen, pero los de la primera cohorte lo hicieron presencialmente y los de la segunda cohorte lo hicieron online. En su trabajo concluye que la preferencia no influye en los resultados, pero que ello no significa poder obviar la reticencia hacia las pruebas online, especialmente importante en los estudiantes de ciencias sociales y humanidades. En tiempos de pandemia del coronavirus SARS-COV2 cabría esperar un cambio en estas preferencias, con una mejor predisposición hacia la evaluación online, dado que las prioridades y preocupaciones de los estudiantes se sitúan en dimensiones no consideradas en el citado trabajo, como es la seguridad individual.

Autores como [Fuller et al. \(2020\)](#) Y [Sánchez-Cabrero et al. \(2021a\)](#) plantean el contexto actual como la oportunidad definitiva para ampliar y mejorar la forma en la que tradicionalmente se viene realizando la evaluación. El reciente estudio de [Sánchez-Cabrero et al. \(2021a\)](#) mostró cómo “la evaluación en línea supervisada brinda las mismas garantías de justicia e igualdad que los exámenes presenciales, incluyendo además las propias ventajas de la no presencialidad, que respaldan que se siga usando en el futuro, especialmente en títulos con muchos estudiantes que pueden provenir de muchas ubicaciones diferentes” (p. 1). No obstante, los autores también apreciaron cómo el medio de evaluación (online/presencial) generaba diferencias en el diseño de los exámenes por parte de los profesores, lo que repercutía en la dificultad de la prueba. En su estudio observaron que la inexperiencia de los profesores en el diseño de exámenes online hizo que diseñaran exámenes más fáciles, lo que hizo mejorar las calificaciones, aunque la distribución de puntuaciones según nivel se mantuvo inalterable, lo que permite concluir que no hay diferencias reales en el medio de evaluación, pero que los miedos o expectativas de los profesores sí que pueden afectar a los resultados, puesto que pueden incluir medidas compensatorias que pueden alterarlos ([Sánchez-Cabrero et al., 2021a](#)).

Por todo lo expuesto, y debido a la oportunidad única que ha supuesto la lamentable situación de pandemia mundial, este estudio se plantea evaluar el impacto del cambio de entorno de aprendizaje implementado en una asignatura de grado en el ámbito universitario de Singapur a lo largo del año 2020, como consecuencia de las restricciones a la movilidad y al contacto social impuestas durante la pandemia ocasionada por el coronavirus SARS-COV2. La investigación se diseña con dos objetivos de investigación. En primer lugar, (1) se pretende comparar el rendimiento académico de la última cohorte con entorno de aprendizaje totalmente presencial con la primera cohorte con entorno de aprendizaje totalmente online, teniendo en cuenta, no solo la evaluación final global, sino también los distintos componentes evaluativos que la componen. En segundo lugar, (2) se pretende evaluar si el cambio de entorno de aprendizaje ha variado la valoración del profesor y la interacción docente-alumno, y si estos factores han influido en el rendimiento académico.

Basándonos en la literatura previa sobre este ámbito de estudio, expuesta previamente, podemos afirmar que el cambio repentino y las carencias formativas de los docentes podrían afectar negativamente, tanto al rendimiento académico como a la relación docente-alumno ([Alharti, 2020](#); [de Oliveira Araújo et al., 2020](#); [Pérez y Hernández, 2020](#)). No obstante, otro tipo de estudios destacan la capacidad de adaptación y resiliencia de los estudiantes a los cambios educativos, por lo que podemos hipotetizar que esas características de los alumnos modularán y reducirán los efectos adversos y el grado de afectación del cambio del entorno educativo, por lo que seguramente el grado de afectación sea escaso, o incluso imperceptible ([Bączek et al., 2021](#); [Sánchez-Cabrero et al., 2021a](#); [Sánchez-Cabrero et al., 2022](#)).

Método

Diseño

El estudio propuesto se trata de un diseño *ex post facto* con carácter descriptivo, dado que parte de datos relativos a distintas calificaciones de los es-

tudiantes y de sus encuestas de valoración. Es correlacional, pues analiza la correlación entre variables, como las calificaciones, y secuencial transversal, dado que estudia dos cohortes consecutivas de una materia.

Participantes

En este estudio se pretendió utilizar una muestra representativa de la población universitaria de Singapur, pero realizando el estudio bajo el prisma de la estructura educativa occidental, así que, mediante muestreo por conglomerados, se seleccionó los estudiantes que cursaron la titulación oficial *Bachelor of Business Studies programme* de la prestigiosa *University College Dublin* en su vertiente internacional, con las dos cohortes más cercanas respecto al estallido de la Pandemia derivada del coronavirus SARS-COV2. La *University College Dublin* es una de las pocas universidades de todo el mundo que posee la acreditación 'Triple Corona' de EQUIS (Europa), AACSB (Estados Unidos) y AMBA (Reino Unido), lo que garantiza su perfil internacional, una gran excelencia educativa, transparencia y calidad en sus procesos educativos y en la fiabilidad y validez de sus procesos de evaluación. La muestra, estuvo compuesta por 282 estudiantes universitarios de grado, distribuidos en dos cohortes de 125 y 157 alumnos, con un rango de edad de entre 18 y 21 años. Todos los participantes se matricularon en el mismo módulo en ambas cohortes y completaron la nota y el formulario de evaluación estándar de la universidad, que garantizaba el anonimato de sus datos, la protección de sus derechos y su consentimiento informado al uso anónimo de sus resultados académicos y evaluaciones educativas con ánimo exclusivamente científico. Ambas cohortes presentan una muy elevada diversidad cultural, y en ningún caso se aprecian diferencias estadísticamente significativas en función del género. Se desconoce la distribución concreta de los participantes según sexo y edad por restricciones de protección a la intimidad impuestas por la universidad para garantizar el anonimato de los estudiantes. No obstante, la distribución equitativa y aleatoria según género y el limitado rango de edad evaluado (18-21 años) garantizan que ambas variables son variables controladas que no interfirieron en los resultados.

El sistema educativo en el que se enmarcan las enseñanzas objeto del análisis es el de Singapur, si bien la titulación también pertenece al *Espacio Europeo de Educación Superior (EEES)*, que exige para su validación altos requisitos de fiabilidad y validez en todos sus procesos, lo que incluye un exquisito y escrupuloso uso de las encuestas de satisfacción de los alumnos con sus docentes y estudios ([EURYDICE, 2020](#)). El grado (180 ECTS) pertenece a la rama ciencias sociales, y se estructura por semestres, siendo posible comenzar los estudios en dos momentos diferentes del año, lo que se traduce en estos casos en la existencia de dos cohortes en cada año natural.

Para la primera cohorte (125 estudiantes) el entorno fue presencial, mientras que para la segunda (157 estudiantes) fue virtual, reemplazándose las clases presenciales por clases online sincrónicas. Tal y como se observa en la Figura 1, se dispone de las calificaciones finales y de sus componentes de evaluación continua de los 282 estudiantes, mientras que para el análisis valorativo del docente y de la interacción docente-alumno se dispone de los resultados de 114 cuestionarios, dado que rellenar dichos cuestionarios de evaluación era estrictamente anónimo y voluntario, como parte de su acuerdo con la universidad en el consentimiento informado del uso científico de sus valoraciones. La investigación se realizó al comienzo del cierre en Singapur, y los estudiantes se mostraron bastante entusiasmados y solidarios con el nuevo entorno de aprendizaje. Es posible que, si el estudio hubiera sido realizado en una época posterior, después de un mayor tiempo de confinamiento, el ánimo de los estudiantes ante la nueva metodología hubiera cambiado, pudiendo alterar los resultados obtenidos.

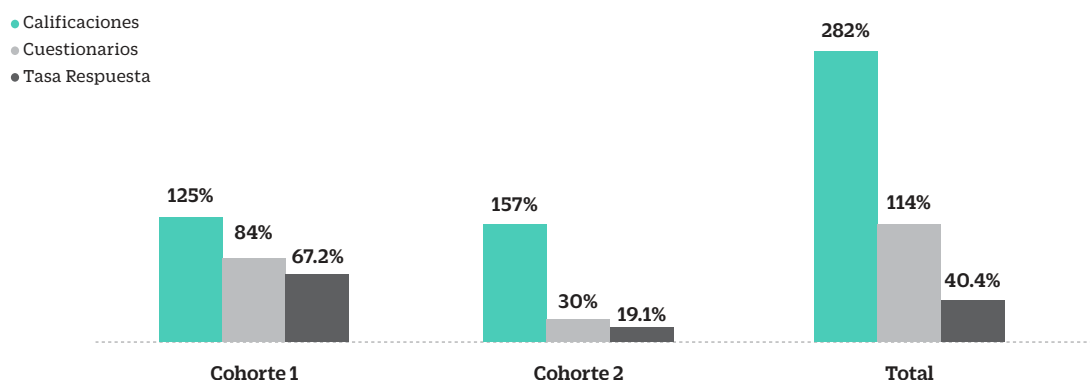
Instrumentos de obtención de datos y variables evaluadas

Las principales variables evaluadas en este estudio se describen a continuación:

- a) Entorno de aprendizaje: Se trata de la variable principal del estudio. Variable nominal con dos niveles: entorno de aprendizaje presencial y entorno de aprendizaje online. Está formado por dos cohortes de una misma asignatura, cursada en el mismo año académico.

Figura 1

Distribución de la muestra según la participación en los distintos componentes evaluativos



- b) Rendimiento académico: Variable cuantitativa continua formada por las calificaciones de los alumnos en cada uno de los componentes evaluativos: Proyecto de la asignatura, Examen final y Calificación global.
- c) Valoración del profesor: Analizada a partir de las encuestas oficiales respondidas por los alumnos a través de nueve ítems en una escala Likert de cinco niveles.
- d) Interacción alumno-profesor: Analizada a partir de las encuestas oficiales respondidas por los alumnos a través de tres ítems en una escala Likert de cinco niveles.

Los instrumentos de evaluación son dos. Por un lado, para el estudio del rendimiento académico se dispone de las calificaciones finales de los estudiantes, así como de las calificaciones de los componentes de dicha calificación final (un proyecto y un examen final escrito online). La confiabilidad de la evaluación del rendimiento académico de los alumnos ha sido medida a través del coeficiente *Alfa de Cronbach*, dando como resultado una confiabilidad adecuada ($\alpha = 0.786$). Por otro lado, la valoración del profesor y la interacción alumno-profesor se analiza en base a los resultados obtenidos en las evaluaciones oficiales efectuadas por los estudiantes en una escala Likert de cinco niveles. Dichas valoraciones de los alumnos fueron gestionadas por una empresa externa, *Kaplan Higher Education* ([Kaplan Higher Education, 2022](#)), en virtud de las exigencias establecidas por el EEES ([EURYDICE, 2020](#)), garantizando su anonimato y

protección de sus datos desde un proceso estrictamente neutro, previamente validado y gestionado externamente, certificando el cumplimiento de todos los requisitos éticos relativos al tratamiento de los datos. Lamentablemente, dicha evaluación externa reporta solamente el resultado final, por lo que no permite cálculo alguno de su confiabilidad de forma externa, aunque cumple con todas las garantías de validez y fiabilidad establecidas por el EEES ([EURYDICE, 2020](#)).

Procedimientos

Los datos se recopilaron del formulario de evaluación estándar diseñado previamente por la universidad, por lo que los estudiantes no sabían que eran parte de ninguna investigación específica, por lo que no se esperaban efectos Hawthorne, ni deseabilidad social relativa a respuestas específicas. En todo momento se siguió el protocolo estandarizado de la universidad, tanto para la evaluación mediante examen como en las encuestas de satisfacción de los estudiantes. Dicho proceso exige una evaluación de la interacción alumno-profesor por parte de los alumnos previa a su evaluación de aprendizaje, para que no sea un factor que influya en dicha evaluación ([EURYDICE, 2020](#)). Este sistema estandarizado cumple con todas las garantías establecidas por el EEES y por la autoridad competente de Singapur en materia de educación universitaria para garantizar objetividad y respetar todos los derechos y anonimato de los alumnos.

Tabla 1
Resultados de la Prueba Kolmogorov-Smirnov

Variable	K-S	Significación
Calificación Proyecto	0.159	0.000**
Calificación Examen	0.159	0.000**
Calificación Final	0.143	0.000**

***Correlación significativa superior al nivel de confianza del 99%*

Análisis de datos

Para valorar la utilización de las pruebas estadísticas inferenciales adecuadas, previamente se realiza la *Prueba de Kolmogorov-Smirnov (K-S)*, destinada a verificar si los datos de análisis siguen una distribución normal. Los resultados de la prueba K-S se muestran en la Tabla 1 a continuación.

Se puede observar en la Tabla 1 como ninguna de las distribuciones de los datos se corresponde a una distribución normal, por lo que es recomendable el uso de pruebas no paramétricas. Por lo tanto, para el estudio de las calificaciones obtenidas por los estudiantes se efectuará mediante un análisis correlacional de su distribución en ambas cohortes, utilizando la *Rho de Spearman* como estadístico de contraste, al ser un estadístico no paramétrico que permite la combinación de variables ordinales inclusive, y se realizará un tratamiento estadístico inferencial mediante la *Prueba t para muestras independientes* y la *Prueba U de Mann-Whitney para muestras independientes* como prueba no paramétrica.

Se establece un nivel de confianza superior al 95% ($\alpha=0.05$) o al 99% ($\alpha=0.01$) para aceptar la significación estadística de los resultados de los contrastes inferenciales o correlacionales que se realicen y se calcula el tamaño del efecto de los resultados inferenciales mediante ETA al cuadrado.

Los datos relativos a ambas cohortes y a todas las calificaciones se analizaron estadísticamente usando el Software estadístico IBM SPSS, versión 24. Igualmente, se compararán los resultados obtenidos en las encuestas de valoración del profesor efectuadas por los estudiantes de las

dos cohortes. Para analizar la interacción alumno-profesor se revisarán las valoraciones de los estudiantes acerca de tres aspectos del nuevo entorno online, recogidos en el cuestionario respondido por los estudiantes de la segunda cohorte.

Resultados

Respecto a la evaluación comparada del rendimiento académico, la Tabla 2 muestra el análisis descriptivo resultante en ambas cohortes según los estadísticos descriptivos: *Media Aritmética* y *Desviación Típica* y mediante los parámetros de distribución muestral asimetría y curtosis. Los estudiantes de la cohorte presencial (N=125) cursaron la asignatura en un entorno presencial durante el primer semestre, y los de la cohorte online (N=157) lo hicieron en un entorno online a lo largo del segundo semestre del año. El sistema de calificación es alfabético, estableciéndose el aprobado mínimo en D- y la calificación más alta en A+. Para su definición numérica se convierten las puntuaciones en una escala 0-12 (de 0= suspenso a 12= A+). La calificación final se obtiene a partir de las calificaciones de un proyecto (con una ponderación del 40%) y de un examen final escrito (con una ponderación del 60%). En la Tabla 2, puede observarse con claridad cómo existen diferencias notables entre ambos grupos en la Calificación Examen y en Calificación Final, siendo las diferencias en la Calificación Proyecto mucho menores. El análisis visual de los resultados, que se realiza a continuación en la Figura 2, puede ayudar a clarificar dichas diferencias.

Tabla 2 Estadísticos descriptivos

	Cohorte	n	M	DE	g1	g2
Calificación Proyecto	Presencial	125	8,50	2,395	-0.951	0.841
	Online	157	8,53	2,171	-0.923	0.834
Calificación Examen	Presencial	125	6,82	2,821	-0.841	-0.024
	Online	157	4,13	3,061	0.181	-1.017
Calificación Final	Presencial	125	7,47	2,131	-0.926	1.038
	Online	157	5,82	2,492	-0.151	-0.580

En la Figura 2, se puede ver que el porcentaje de suspensos en ambos casos es similar (inferior al 2%), pero la distribución de las calificaciones de los estudiantes que superaron la asignatura muestra una mayor concentración de calificaciones bajas (entre D- y C+) en la cohorte 2 online (57%) y de

calificaciones altas (entre B- y A-) en la Cohorte 1 presencial (77%). La distribución de calificaciones de la cohorte 2 es muy similar a la distribución de calificaciones del conjunto de asignaturas del departamento, presentando una distribución más normalizada que la cohorte 1 presencial.

Figura 2 Distribución de las Calificaciones Finales

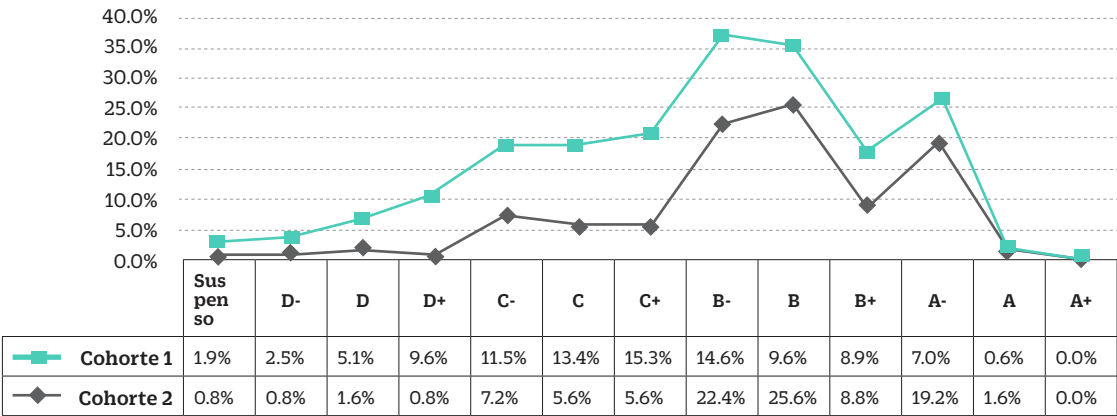


Figura 3 Distribución de las Calificaciones del Proyecto

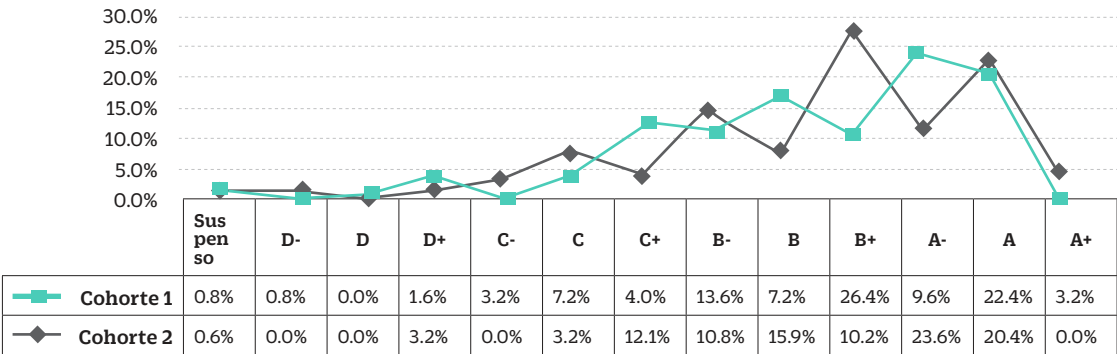
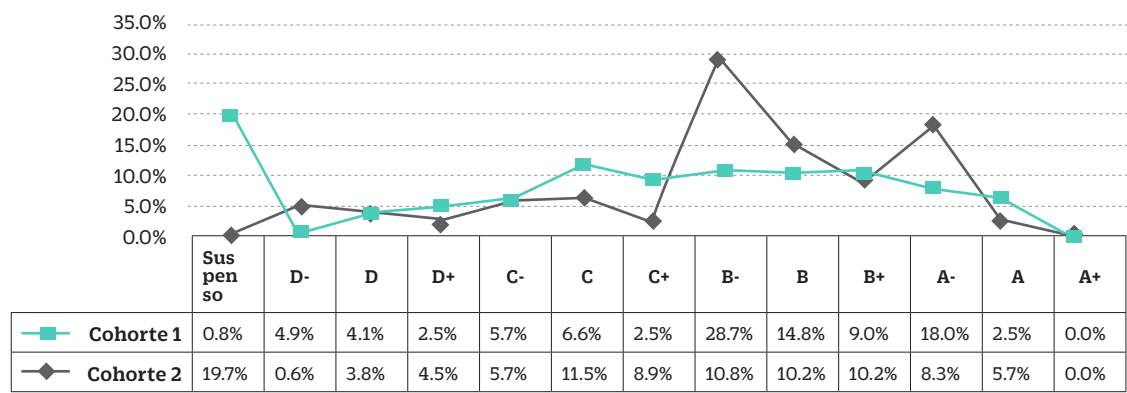


Figura 4
Distribución de las Calificaciones de la Prueba Final



En las Figuras 3 y 4 se reflejan las distribuciones de calificaciones de ambos componentes que conforman la evaluación final.

Como se puede observar gráficamente en la Figura 3, la distribución de las calificaciones de los proyectos es muy similar para ambas cohortes.

En el caso de las distribuciones de las calificaciones del examen final, las diferencias entre ambas cohortes son relevantes. En la cohorte 2 online el porcentaje de suspensos es muy superior, mientras que en la cohorte 1 presencial hay más calificaciones altas (entre B- y A-). El análisis visual de los resultados parece indicar que la diferencia entre ambas cohortes podría deberse a las diferencias en

la distribución de calificaciones del examen Final, tal y como puede verse en la Figura 4, por lo que se torna pertinente un análisis estadístico inferencial de los datos para determinar la significación de las diferencias entre ambos grupos.

En primer lugar, para valorar la significación de las diferencias entre ambas cohortes se utiliza la *Prueba t para muestras independientes*, usando la *t de Student* como estadístico de contraste y la *Prueba U de Mann-Whitney* como prueba no paramétrica, también se calcula *ETA al cuadrado* para mostrar el tamaño del efecto de cada contraste. Los resultados obtenidos pueden observarse en la Tabla 3.

Tabla 3
Resultados de la Prueba t de Student y de la Prueba U de Mann-Whitney para muestras independientes

Variable	t	p ^a	p ^b	η ²
Calificación Proyecto	-,120	0,905	0,902	0.000
Calificación Examen	7,656	0,000**	0,000**	0.170
Calificación Final	5,991	0,000**	0,000**	0.110

Nota. pa: significancia estadística de la prueba t Student; pb: significancia estadística de la U Mann Whitney; η²: Eta al cuadrado

Se puede observar en la Tabla 3 como, tanto la Calificación Examen como la Calificación Final, muestran diferencias estadísticamente significativas según cohortes en los resultados, mientras que la Calificación Proyecto no muestra diferencias estadísticamente significativas. Teniendo en cuenta que la Calificación Final depende de ambas calificaciones, es posible afirmar que es la Calificación Examen la que genera las diferencias entre ambos grupos. Por lo que respecta al tamaño del efecto, el tamaño del efecto de la Calificación Examen, aunque baja, es considerablemente mayor que en las otras dos mediciones, lo que apoya la explicación de que es en el examen donde se observan las diferencias estadísticamente significativas entre ambas cohortes.

Finalmente, se analizan las correlaciones entre las distintas calificaciones para observar el grado de covariación existente entre calificaciones, utilizando la *Rho de Spearman* como estadístico de contraste, ya que permite incluir variables de tipo ordinal y se trata de un estadístico no paramétrico; así puede incluirse en la comparación también la variable cohorte convertida en escala ordinal, tal y como puede verse en la Tabla 4.

El análisis de las correlaciones de la Tabla 4 reflejan que todas son estadísticamente significativas, lo que refleja una alta coherencia intrasujeto (es decir, las calificaciones dependen enormemente del esfuerzo y talento del alumno, lo cual es deseable). No obstante, es posible observar, al igual que en las comparaciones de medias, como la mayor asociación se produce entre la Calificación Final y la Calificación Examen y cómo la única correlación no significativa se da entre la Cohorte y la Calificación Proyecto.

Por lo que respecta a la encuesta de valoración

del profesor, ésta fue respondida por 84 de los 125 estudiantes de la cohorte 1 presencial, y por 30 de los 157 estudiantes de la cohorte 2 online. En la Tabla 5 se muestra el valor medio de las respuestas en una escala Likert de cinco niveles de cada cohorte según los nueve ítems planteados:

En la Tabla 5 se observa cómo existen muy pocas diferencias entre cohortes en la valoración de los profesores, por lo que se puede afirmar que el cambio de entorno de aprendizaje no fue un factor que alterara esta variable y las diferencias en los resultados académicos entre cohortes seguramente no sean debidas a una labor del profesor diferente. A continuación, la Figura 5 muestra visualmente estos resultados para ampliar la información sobre la que establecer conclusiones:

Al igual que ocurre con la Tabla 5, la Figura 5 muestra una alta consistencia en las valoraciones del profesor por parte del alumno, por lo que se refuerza la impresión de que las diferencias entre cohortes no estriban en la labor diferente del profesor. Lamentablemente, al ser una evaluación externa a este estudio, no es posible determinar la significación de las diferencias mediante estadística inferencial, lo cual hubiera servido para determinar la significación estadística de estas impresiones.

En el caso de la interacción alumno-profesor, se utiliza nuevamente la información de las encuestas de valoración del profesor realizadas por los estudiantes con metodología online, en las cuales, en relación con las encuestas de valoración de los estudiantes con metodología presencial, hay tres preguntas adicionales relativas al entorno. La encuesta fue respondida por 30 de los 157 estudiantes de la cohorte 2 online. La Tabla 6 muestra el valor medio de las respuestas según la escala Likert de cinco niveles en estos tres ítems:

Tabla 4
Correlaciones entre las variables de estudio

	1	2	3	4
C. Proyecto	-			
C. Examen	.209**	-		
C. Final	.504**	.930**	-	
Cohorte	-.009	-.424**	-.344**	-

Nota. ** rs Correlación significativa superior al nivel de confianza del 99% (bilateral).

Tabla 5
Resultados de las Encuestas de Valoración

Pregunta	Cohorte 1	Cohorte 2
1.En la guía docente se establecen con claridad los objetivos de la asignatura, los resultados de aprendizaje y los contenidos	4,1	4,1
2. Los requisitos del proyecto y los criterios de calificación se explicaron con claridad	4,0	4,1
3. El proyecto hizo que aumentase mi interés en la asignatura	3,9	4,0
4. Entendí el nivel esperado de mi trabajo en la asignatura	4,0	3,9
5. El profesor estimuló mi interés en la asignatura	3,9	4,0
6. Los métodos de enseñanza y evaluación del profesor hicieron la asignatura interesante	3,8	3,9
7. Las clases del profesor estaban bien organizadas	3,9	4,1
8. El profesor impulsó la participación con debates y otros ejercicios	4,1	4,0
9. El contenido de la asignatura era intelectualmente desafiante	4,0	4,0
Media	4,0	4,0

Figura 5
Valores Medios de las Encuestas de Valoración

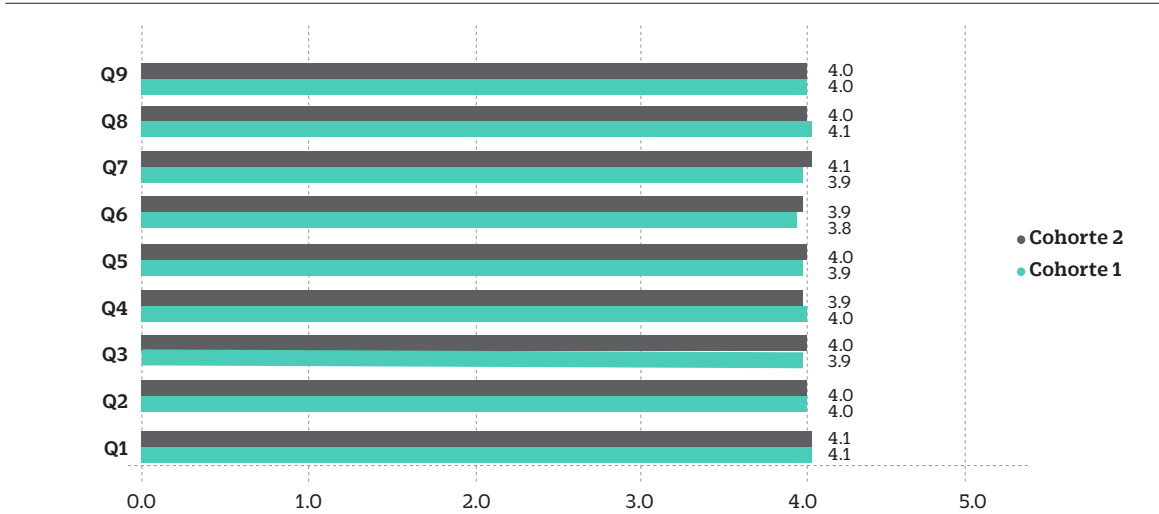
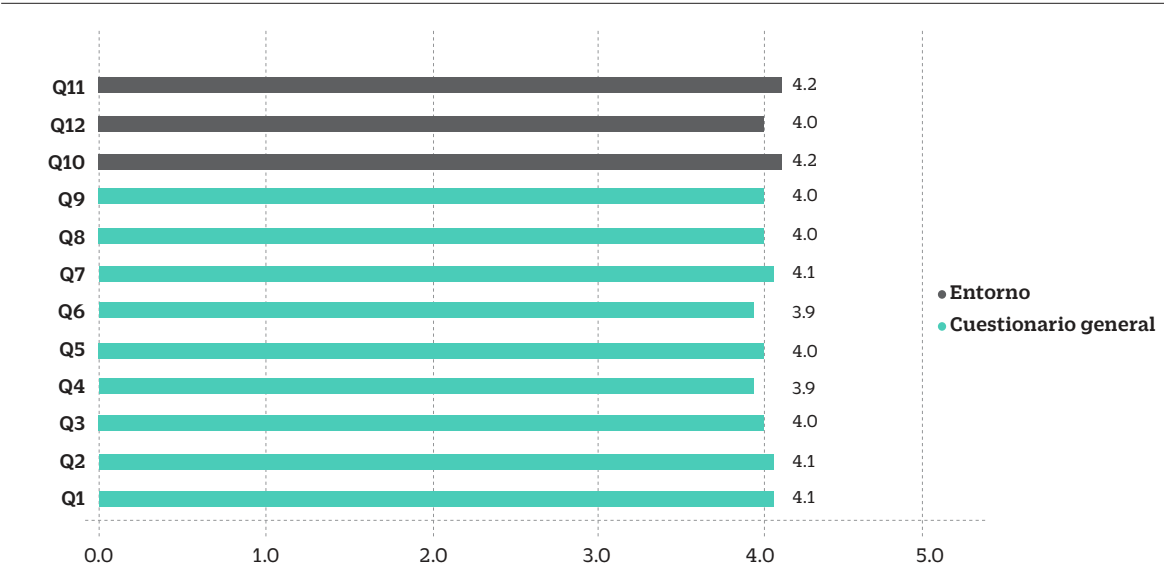


Tabla 6
Resultados de las encuestas de valoración

Pregunta	Cohorte 2
10. La guía docente daba indicaciones claras acerca de cómo acceder a las clases online a través del aula virtual	4,2
11. Las clases online ayudaron a facilitar mi aprendizaje	4,0
12. La navegación a través en el aula virtual fue sencilla	4,2
Media	4,1

Figura 6
Valores Medios de las Cuestiones Generales y del Entorno Online



Es observable en la Tabla 6 cómo las calificaciones obtenidas en las encuestas al alumnado sobre la interacción alumno-profesor en el entorno online reflejan que las diferencias en rendimiento académico entre ambas cohortes no residen tampoco en esta cuestión, puesto que las valoraciones son igualmente positivas a las valoraciones sobre el profesor, e incluso ligeramente mejores, por lo que también se puede afirmar que el cambio de entorno de aprendizaje no fue un factor que alterara esta variable. A continuación, la Figura 6 muestra visualmente estos resultados, integrando los resultados de las tres preguntas relativas al entorno con las nueve preguntas del cuestionario general para ampliar la información sobre la que establecer conclusiones:

La Figura 6 muestra una alta consistencia en las valoraciones sobre la interacción alumno-profesor en el entorno online por parte del alumnado, por lo que se refuerza la impresión de que las diferencias entre cohortes no estriban en una diferente interacción alumno-profesor. Lamentablemente, al ser una evaluación externa a este estudio, no es posible determinar la significación de las diferencias mediante estadística inferencial, lo cual hubiera servido para determinar la significación estadística de estas impresiones.

Discusión

Respecto al primer objetivo de investigación, enfocado en comparar el rendimiento académico de la última cohorte con entorno de aprendizaje totalmente presencial con la primera cohorte con entorno de aprendizaje totalmente online, teniendo en cuenta, no solo la evaluación final global, sino también los distintos componentes evaluativos que la componen, a pesar de que la tasa de éxito se sitúa sobre el 98% para ambas cohortes, se observa una mayor concentración de calificaciones altas en el caso de la primera cohorte, que alejan la distribución de calificaciones de la normalidad esperable, tomando como referencia la distribución de calificaciones de las materias del área. Parece razonable asumir entonces cierta anomalía en la distribución de dichas calificaciones. Tras considerar los diferentes factores que pudieran explicar esta diferencia, en consonancia con las conclusiones aportadas por el estudio de [Sánchez-Cabrero et al. \(2021a\)](#), todo apunta hacia la excepcionalidad de la prueba final de la asignatura, lo que se pudo traducir en una inadecuación de ésta al entorno en que se produjo. Como ya se ha indicado, la docencia de la asignatura para esta cohorte fue presencial, pero una semana an-

tes de la realización de la prueba final se decretó un confinamiento nacional que impidió realizar las pruebas finales de forma presencial.

Dado el protocolo existente en la facultad para la aprobación de las pruebas finales (requieren ser aprobadas por el departamento y por un examinador externo perteneciente a otra universidad) no hubo tiempo material para rediseñar, y revisar todas las pruebas finales de todas las asignaturas del grado para adecuarlas al entorno online. Ante esa situación de emergencia, el centro decidió mantener las pruebas diseñadas para un entorno presencial, a pesar de que iban a ser realizadas de forma remota, poniéndose todos los esfuerzos en ese momento en garantizar que todos los estudiantes pudiesen realizar las pruebas desde sus hogares. En el segundo semestre, tanto la docencia como las actividades de evaluación se diseñaron para un entorno online. El centro comunicó esta decisión con mucho tiempo de antelación, indicando que la decisión era definitiva e independiente del estado de confinamiento, o no, en el momento de realizar las pruebas. En este caso se desarrolló una prueba final de 'libro abierto', asumiendo que todos los estudiantes podían tener acceso a cuantas fuentes de información tuviesen a su alcance. Dadas las características de la prueba, los estudiantes dispusieron de cinco horas para su realización, en lugar de las dos horas de los exámenes presenciales, lo que encaja con las propuestas de [Myvry y Joutsenvirta \(2015\)](#). Para garantizar la integridad, se introdujo un programa antiplagio en el sistema de entrega de la prueba final. Los resultados fueron excepcionales, y la prueba final diseñada fue destacada por el departamento, siendo tomada como referencia del tipo de evaluación deseable en el nuevo entorno.

Para confirmar que la anomalía en los resultados académicos de los estudiantes de la cohorte 1 presencial fue debida a la prueba final, basta con observar las distribuciones de las calificaciones obtenidas por ambas cohortes en los dos elementos que componen la calificación final, el proyecto (Figura 3) y la prueba final (Figura 4). La gran similitud en la distribución de las calificaciones del proyecto pone de manifiesto que el origen de la divergencia radica en la prueba final. La similitud en la distribución de las califi-

nes del proyecto se confirma en el análisis estadístico presentado en la Tabla 3. Con la prueba de Levene se comprobó la existencia homocedasticidad (igualdad de varianzas) en las variables, pues dependiendo de si hay igualdad de varianzas o no, en la prueba *t* se usarán dos estadísticos de contraste diferentes. Teniendo en cuenta que la igualdad de varianzas se confirma cuando el grado de significación *p-valor* (Sig. en la Tabla 3) correspondiente a cada *F* es superior a 0,05, en este caso se confirma para el proyecto (0,512), pero no para el examen (0,03) ni para la calificación final (0,015). En la *Prueba t para igualdad de medias*, se confirma la igualdad para el proyecto (sig. bilat. de 0,905), mientras que se rechaza para el examen y la calificación final (sig. bilat. de 0).

El relativamente elevado porcentaje de estudiantes que no superan la prueba final en la segunda cohorte es consistente con el histórico de resultados en dicha prueba. El motivo es que en el momento de realizar la prueba final los estudiantes ya conocen sus calificaciones del proyecto, y dado que éstas suelen ser altas, algunos tienden a ajustar su esfuerzo en la preparación de la prueba final a la obtención del aprobado en la calificación final (por normativa de la universidad no está permitido establecer calificaciones mínimas en los distintitos componentes de la calificación final). Como se puede observar en la Tabla 4, esto se traduce en una relativamente baja correlación entre las calificaciones de los proyectos y de los exámenes (0,209), muy inferior a la existente entre las calificaciones finales y las de los exámenes (0,930), pero también sensiblemente inferior a la que presentan las calificaciones finales con las de los propios proyectos (0,504).

Teniendo en cuenta la paridad de las tasas de éxito de ambas cohortes, y explicada la anomalía en la prueba final, los resultados obtenidos son coherentes con las conclusiones de [Paul y Jefferson \(2019\)](#) y [Toyne et al. \(2019\)](#), quienes no encontraron diferencias estadísticamente significativas entre los resultados de los estudiantes de los entornos presencial y online. No obstante, sí que hubo diferencias estadísticamente significativas en el examen final, pero en consonancia con el estudio de [Sánchez-Cabrero et al. \(2021a\)](#), las diferencias en rendimiento académico vienen

justificadas por las circunstancias de diseño de la prueba evaluativa y las expectativas de los profesores sobre la evaluación, ya que hubo cambios destacables en este proceso (diferente duración del examen, no adaptación de las cuestiones al medio y medidas compensatorias derivadas de las circunstancias especiales de evaluación) que pueden explicar estas diferencias.

Respecto al segundo objetivo de investigación, consistente en evaluar si el cambio de entorno de aprendizaje ha variado la valoración del profesor y la interacción docente-alumno, y si estos factores han influido en el rendimiento académico. Lo primero que llama la atención es la diferencia en la tasa de respuesta. En la cohorte 1 presencial se situó en un 67% (84 de 125), mientras que en la cohorte 2 online fue de tan solo el 19% (30 de 157). No obstante, la explicación es sencilla. En el entorno presencial las encuestas se realizan durante la penúltima sesión (en ausencia del profesor), de manera que todos los estudiantes que están en clase en ese momento realizan la encuesta. Esta posibilidad no existe en el entorno virtual.

Los datos obtenidos de ambas cohortes son muy similares (Figura 5) y muestran con claridad como la percepción de los estudiantes acerca de la calidad de la asignatura y el docente apenas varió. En dos de las preguntas la valoración media fue la misma, en seis la diferencia de los resultados para ambas cohortes fue de tan solo una décima, y en una fue de dos décimas. La valoración media fue la misma (4 sobre 5). Estos valores, que indican que los estudiantes no percibieron en el nuevo entorno una disminución de la calidad, están en consonancia con las observaciones de [Johnson et al. \(2000\)](#) en su estudio, en la que observaron resultados muy similares en la valoración del docente en el entorno de aprendizaje online y presencial.

Comparando los resultados obtenidos con los recogidos en Johnson et al. (2000) se puede apreciar una sutil diferencia, pues en este trabajo se señala una percepción ligeramente más favorable del docente en el entorno presencial. En el caso del presente estudio, la valoración del docente es la misma en ambos entornos, lo que supone una mejor valoración relativa del docente en el entorno virtual si la comparamos con lo esperable del

estudio anterior. El contexto excepcional (confinamiento motivado por una pandemia) en el que se desarrolló el proceso de enseñanza-aprendizaje de la cohorte 2 podría ser la explicación, pues la predisposición del alumnado hacia el nuevo entorno fue muy positiva desde el principio, mostrando en todo momento una gran empatía, y asumiendo la solución propuesta de cambio de entorno como la mejor de las alternativas posibles, lo que encaja con las conclusiones de [Baczek et al. \(2021\)](#).

La información obtenida de las encuestas de valoración del profesor acerca de la interacción alumno-profesor en el nuevo entorno no es especialmente relevante debido a la reducida muestra, pues se limita a tres cuestiones. A pesar de ello, los resultados obtenidos (valor medio de 4,1) son consistentes con los del resto del cuestionario (media de 4,0). Esto significa que los tres aspectos analizados del nuevo entorno online se valoraron de manera similar al resto de aspectos reflejados en el cuestionario (Figura 6), lo que no altera la valoración conjunta del estudiante de forma estadísticamente significativa, lo cual está en la línea de las conclusiones de [Alonso y Blázquez \(2009\)](#).

Dado el carácter limitado de la información aportada por las preguntas del cuestionario acerca de la interacción alumno-profesor en el nuevo entorno, podría ser útil ampliar esta información con los resultados de un estudio llevado a cabo a nivel institucional. A diferencia de los datos presentados hasta ahora, en este caso la información no proviene directamente de los estudiantes de la materia analizada, sino que lo hace de un estudio efectuado por el centro al finalizar el primer semestre con docencia en formato online con el fin de mejorar la calidad, y en el cual participaron tanto grupos de estudiantes como de profesores. En dicho estudio, entre otros aspectos, se recoge un uso excesivo de los emails para aclarar dudas, en detrimento de las sesiones de tutoría online, las cuales fueron poco utilizadas por los estudiantes. Los estudiantes manifestaron que era importante que el profesor definiese claramente el medio de comunicación en su asignatura, sin mostrar una preferencia clara entre el email y las comunicaciones a través de la plataforma virtual. Igualmente, se recoge la valoración positiva del

uso de la herramienta 'sondeos' de la plataforma virtual durante las clases, por hacer las sesiones más participativas y ayudar a evitar distracciones. Los docentes se manifestaron en el mismo sentido en este aspecto, y destacaron la gran participación en el chat (hasta el punto de hacerse inmanejable por momentos) en contraposición a la baja participación oral. La información presentada en este estudio inicial puede ser útil de cara a diseñar procesos de enseñanza-aprendizaje más efectivos en entornos remotos, así como para desarrollar nuevas líneas de investigación.

El cambio de entorno de aprendizaje propiciado por la pandemia no ha tenido una influencia significativa sobre los resultados del aprendizaje en el caso presentado, si bien la distorsión en los resultados académicos de la prueba final de la cohorte 1 presencial evidencia la importancia de la adecuación de todos los elementos del proceso de enseñanza-aprendizaje al entorno en el que éste se produce, incluidas las estrategias e instrumentos para la evaluación. La calidad, medida a través de las valoraciones del profesor y de la asignatura reflejadas en las encuestas respondidas por los estudiantes, no se ha visto afectada por el cambio de entorno, y la interacción entre alumno y profesor no es percibida de una forma más negativa en el entorno virtual que en el presencial. No obstante, es importante prestar atención a la variedad de formas de comunicación que se presentan en el entorno online (email, mensajería a través del campus virtual, chats, foros, video-tutorías), pues la coexistencia de diferentes alternativas puede generar confusión entre los estudiantes y alejarlos del proceso.

La principal limitación de este estudio parte de que su realización se deriva de un suceso incidental y por ello no se han podido incluir y controlar todas las posibles variaciones metodológicas entre ambos entornos. No se disponen de datos demográficos de edad y sexo, y en los cuestionarios respondidos por los estudiantes los datos medios disponibles no permiten realizar los análisis estadísticos inferenciales oportunos. Como resultado de las limitaciones descritas, el estudio se podría ampliar en un futuro controlando y ampliando las variables de ambos entornos, de forma intencionada, desde el inicio de la investigación.

También podría resultar de interés ampliar la investigación incorporando los resultados de futuras cohortes que cursen la asignatura en su formato presencial original, y analizar en qué medida el cambio de formato inicialmente experimentado (de presencial a virtual) como consecuencia de la pandemia puede influir en la configuración del proceso enseñanza-aprendizaje en el momento en el que se retome el formato presencial. Igualmente, el análisis de los resultados en otras asignaturas de los mismos estudios dotaría de mayor solidez las conclusiones obtenidas.

Finalmente, se debe destacar su aportación a la comparativa entre la docencia y evaluación presencial u online, tanto al ámbito teórico como práctico. A nivel teórico se ha confirmado cómo la docencia online puede garantizar el aprendizaje y desarrollo de competencias al mismo nivel que la docencia presencial, incluyendo la calidad de las interacciones docente-alumno. Por otro lado, a nivel práctico se debe destacar que los procesos evaluativos deben ser adaptados a la modalidad de evaluación para no introducir variaciones en los resultados. El mayor riesgo evaluativo no descansa sobre la modalidad de evaluación, sino sobre las decisiones que se toman al respecto por parte de los docentes, que pueden introducir su sesgo relacionado con expectativas y miedos sobre la evaluación.

Referencias

- Alharthi, M. (2020). Students' attitudes toward the use of technology in online courses. *International Journal of Technology in Education*, 3(1), 14-23. <https://doi.org/10.46328/ijte.v3i1.18>
- Alonso, L., y Blázquez, F. (2009). Are the functions of teachers in e-learning and face-to-face learning environments really different?. *Educational Technology & Society*, 12(4), 331-343. <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.544.2836&rep=rep1&type=pdf>
- Bączek, M., Zagańczyk-Bączek, M., Szpringer, M., Jaroszyński, A. & Woźakowska-Kapłon, B. (2021). Students' perception of online learning during the COVID-19 pandemic. *Medicine*, 100(7), 24821. <https://10.1097/MD.00000000000024821>
- Bengtsson L. (2019). Take-home exams in higher education:

- a systematic review. *Education Science*, 9(4), 267. <https://doi.org/10.3390/educsci9040267>
- Espino-Díaz, L., Fernández-Caminero, G., Hernández-Lloret, C. M., González-González, H., y Álvarez-Castillo, J.L. (2021). Analyzing the impact of Covid-19 on education professionals. Toward a paradigm shift: ICT and Neuroeducation as a binomial of action. *Sustainability*, 12(14), 5646. <https://doi.org/10.3390/su12145646>
- Deeley, S. J. (2018) Using technology to facilitate effective assessment for learning and feedback in higher education. *Assessment and Evaluation in Higher Education*, 43(3), pp. 439-448. [doi:10.1080/02602938.2017.1356906](https://doi.org/10.1080/02602938.2017.1356906)
- de Oliveira, F. J., de Lima, L. S. A., Cidade, P. I. M., Nobre, C. B., & Neto, M. L. R. (2020). Impact of Sars-Cov-2 and its reverberation in global higher education and mental health. *Psychiatry research*, 288, 112977. <https://doi.org/10.1016/j.psychres.2020.112977>
- EURYDICE (2020). Key Features of the Education System in Ireland. *European Commission*. https://eacea.ec.europa.eu/national-policies/eurydice/content/ireland_en
- Fuller, R., Joynes, V., Cooper, J., Boursicot, K., & Roberts, T. (2020). Could COVID-19 be our 'there is no alternative' (TINA) opportunity to enhance assessment?. *Medical Teacher*, 42(7), 781-786. <https://doi.org/10.1080/0142159X.2020.1779206>
- García-Martín, J., y García-Martín, S. (2021). Uso de herramientas digitales para la docencia en España durante la pandemia COVID-19. *Revista Española de Educación Comparada*, 38(38), 151-173. <https://doi.org/10.5944/REEC.38.2021.27816>
- Gaytan, J. (2005). Effective assessment techniques for online instruction. *Information Technology, Learning, and Performance Journal*, 23(1), 25-33. <https://www.proquest.com/openview/445e23def9d81566f83c5ea1108ab95e/1?pq-origsite=gscholar&cbl=9497>
- Gaytan, J., y McEwen, B. C. (2007). Effective online instructional and assessment strategies. *American Journal of Distance Education*, 21(3), 117-132. <https://doi.org/10.1080/08923640701341653>
- Hewson, C. M. (2012). Can online course based assessment methods be fair and equitable? Relationships between students' preferences and performance within online and offline assessments. *Journal of Computer Assisted Learning*, 28(5), 488-498. <https://onlinelibrary.wiley.com/doi/10.1111/j.1365-2729.2011.00473.x>
- Hosseinzadeh, P., Zareipour, M., Baljani, E., & Moradali, M. R. (2022). Social Consequences of the COVID-19 Pandemic. A Systematic Review. *Investigación y Educación en Enfermería*, 40(1). <https://doi.org/10.17533/udea.iee.v40n1e10>
- Johnson, S. D., Aragon, S. R., Shaik, N., & Palma-Rivas, N. (2000). Comparative analysis of learner satisfaction and learning outcomes in online and face-to-face learning environments. *Journal of Interactive Learning Research*, 11(1), 29-49. [http://faculty.weber.edu/eamsel/Research%20Groups/On-line%20Learning/Johnson%20et%20al%20\(2000\).pdf](http://faculty.weber.edu/eamsel/Research%20Groups/On-line%20Learning/Johnson%20et%20al%20(2000).pdf)
- Kaplan Higher Education (2022). Bachelor programme. <https://www.kaplan.com.hk/higher-education>
- López-Pérez, M.V., Pérez-López, M.C., Rodríguez-Ariza, L., & Argente-Linares, E. (2013). The influence of the use of technology on student outcomes in a blended learning context. *Educational Technology Research and Development*, 61, 625-638. <https://doi.org/10.1007/s11423-013-9303-8>
- Myyry, L., & Joutsenvirta, T. (2015). Open-book, open-web online examinations: developing examination practices to support university students' learning and self-efficacy. *Active Learning in Higher Education*, 16(2), 119-132. <https://doi.org/10.1177/1469787415574053>
- Paechtr, M., & Maier, B. (2010). Online or face-to-face? Students' experiences and preferences in e-learning. *Internet and Higher Education*, 13, 292-297. <https://doi.org/10.1016/j.iheduc.2010.09.004>
- Paul, J., & Jefferson, F. (2019). A comparative analysis of student performance in an online vs. face-to-face environmental science course from 2009 to 2016. *Frontiers in Computer Science*, 1(7), 1-9. <https://doi.org/10.3389/fcomp.2019.00007>
- Pérez, F., y Hernández, L. (2020). *Los retos del Covid-19 y los centros educativos: ¿cuáles son los riesgos de aprovechar poco las nuevas tecnologías?*. Instituto Valenciano de Investigaciones Económicas. https://www.ivie.es/wp-content/uploads/2020/07/20.Covid19_IvieExpress_Los-retos-del-COVID-19-y-los-centros-educativos.pdf
- Sánchez-Cabrero, R., Arigita-García, A., Gil-Pareja, D., Sánchez-Rico, A., Martínez-López, F., & Sierra-Macarrón, L. (2022). Measuring the Relation between Academic Performance and Emotional Intelligence at the University Level after the COVID-19 Pandemic Using TMMS-24. *Sustainability*, 14(6), 3142. <https://doi.org/10.3390/su14063142>
- Sánchez-Cabrero, R., Casado-Pérez, J., Arigita-García, A., Zubiaurre-Ibáñez, E., Gil-Pareja, D., & Sánchez-Rico, A. (2021a). E-Assessment in E-Learning Degrees: Comparison vs. Face-to-Face Assessment through Perceived Stress and Academic Performance in a Longitudinal Study. *Applied Sciences*, 11(16), 7664. <https://doi.org/10.3390/app11167664>
- Sánchez-Cabrero, R., Estrada-Chichón, J. L., Abad-Mancheño, A., & Mañoso-Pacheco, L. (2021b). Models on Teaching Effectiveness in Current Scientific Literature. *Education Sciences*, 11(8), 409. <https://doi.org/10.3390/educsci11080409>
- Szczęśniak, D., Gładka, A., Misiak, B., Cyran, A., & Rymaszewska, J. (2021). The SARS-CoV-2 and

mental health: From biological mechanisms to social consequences. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 104, 110046.

<https://doi.org/10.1016/j.pnpbp.2020.110046>

Toyne, M. F., Briley, J. E. & Jalbert, T. (2019). Comparing learning outcomes on face-to-face and online teaching platforms: evidence from Major Field Test Scores. *Business Education & Accreditation*, 11(1), 61-69. <https://ssrn.com/abstract=3460711>

Vlachos, J., Hertegård, E. & Svaleryd, H. B. (2021). The effects of school closures on SARS-CoV-2 among parents and teachers. *Proceedings of the National Academy of Sciences*, 118(9). <https://doi.org/10.1073/pnas.2020834118>

Cultura de evaluación para el aprendizaje según estudiantes y docentes de una universidad pública mexicana

Sofía Contreras Roldán ^{*1}; Patricio Henríquez Ritchie²

^{1,2} Universidad Autónoma de Baja California, Ensenada, México <https://orcid.org/0000-0003-1987-3445> sofia.contreras@uabc.edu.mx

² <https://orcid.org/0000-0002-1026-3379> phenriquez@uabc.edu.mx

Citar como: Contreras Roldán, S., & Henriquez Ritchie, P. (2023). Cultura de evaluación para el aprendizaje según estudiantes y docentes de una universidad pública mexicana. *Revista Digital de Investigación en Docencia Universitaria*, 17(2), e1831. <https://doi.org/10.19083/ridu.2023.1831>

Recibido: 25/02/2023 **Revisado:** 10/03/2023. **Aceptado:** 20/05/2023. **Publicado:** 30/06/2023.

Resumen

Introducción: la cultura de la evaluación para el aprendizaje es aquella donde las estrategias evaluativas empleadas buscan promover y mejorar el aprendizaje del estudiantado. **Objetivo:** conocer la medida en que la cultura de la evaluación en una universidad pública mexicana está orientada hacia estrategias de evaluación para el aprendizaje. **Método:** se aplicaron dos encuestas a 1,849 estudiantes y 545 docentes durante 2022-1, considerando muestras aleatorias no probabilísticas. **Resultados:** los docentes reportaron que la estrategia de evaluación que mejor dominan es la elaboración de exámenes objetivos, y en menor medida la autoevaluación y coevaluación; por su parte, cerca de la mitad de estudiantes valoraron que el uso de éstas es de calidad deficiente a regular, sin embargo afirman estar satisfechos con el empleo de las estrategias de evaluación formativas de sus docentes. **Discusión:** se observaron convergencias y divergencias (docentes/estudiantes) en torno al uso, calidad y dominio de estas estrategias evaluativas.

Palabras claves: evaluación; aprendizaje; evaluación formativa; universidad; docentes

Culture of evaluation for learning according to students and teachers of a Mexican public university

Abstract

Introduction: The culture of assessment for learning is one where the evaluation strategies employed seek to promote and improve student learning. **Objective:** to know the extent to which the culture of evaluation in a Mexican public university is oriented towards evaluation strategies for learning. **Method:** Two surveys were applied to 1,849 students and 545 teachers during 2022-1, considering random non-probability samples. **Results:** The teachers reported that the evaluation strategy that they dominate the most is the development of objective examinations, and to a lesser extent self-evaluation and co-evaluation; for its part, about half of students assessed that the use of these is of poor quality to regular, however, they claim to be satisfied with the use of formative assessment strategies by their teachers. **Discussion:** Convergences and divergences (teachers/students) were observed regarding the use, quality, and mastery of these evaluation strategies.

Keywords: assessment; learning; formative assessment; university; teachers

***Correspondencia:**

Sofía Contreras Roldán

sofia.contreras@uabc.edu.mx



Introducción

En la última década, la mayoría de los modelos educativos en la educación superior tienen una orientación pedagógica hacia el enfoque constructivista ([Moreno-Olivos, 2007](#); [Tünnermann, 2011](#)). Dentro de esta perspectiva, la evaluación educativa ha tomado una diferenciación entre la evaluación *del* y *para* el aprendizaje ([Sánchez, 2022](#)). La evaluación *del* aprendizaje se caracteriza por tener un propósito sumativo, es decir, para determinar la medida en que se han logrado los objetivos de aprendizaje esperados. Por su parte, en la evaluación *para* el aprendizaje se busca que las estrategias evaluativas empleadas, sean por sí mismas un motor para promover y mejorar el aprendizaje del estudiantado, a través de la reflexión de sus fortalezas y áreas de mejora, también conocida como evaluación formativa ([Black & Wiliam, 2004](#); [Sánchez, 2022](#); [Wiliam, 2011](#)). Algunos ejemplos de esta forma de evaluación serían, la autoevaluación y la coevaluación, la evaluación de los aprendizajes previos, la devolución de una retroalimentación efectiva mientras se evalúa, entre otros ([Hattie & Timperley, 2007](#); [Sadler, 1989](#)).

En función de estos dos propósitos que puede cumplir la evaluación (sumativos y formativos), se desprenden dos formas de cultura de la evaluación en una institución educativa ([Berlanga & Juárez-Hernandez, 2020](#)). Como tal, la “cultura de la evaluación” es un término que se ha empleado a nivel macro, es decir, para hablar de las características de la dinámica organizacional de una institución, en torno a la evaluación ([p.e. Fuller, 2013](#); [Valenzuela et al. 2011](#)). Sin embargo, Martínez (2010) afirmó que, en el aula universitaria (nivel micro) se gesta también una cultura de la evaluación propia entre docentes y estudiantes, la cual podría estar o no alineada con la cultura institucional, y que se centra específicamente en la evaluación del aprendizaje y en el uso que se le da. Diferentes autores han conceptualizado esta cultura de la evaluación del aprendizaje y los elementos que la integran. Por ejemplo, para [Bolseguí y Fuguet \(2006\)](#) comprende las reglas (explícitas e implícitas), hábitos, rutinas, creencias, estructuras, símbolos, métodos y técnicas, que comparte

la comunidad educativa respecto a los procesos de evaluación; para [Reategui \(2015\)](#) es una combinación entre lo institucional y los valores, creencias, significados y prácticas que comparten docentes y estudiantes hacia la evaluación.

Para este estudio, se optó por emplear la conceptualización de [Allal \(2016\)](#) y [Martínez \(2010\)](#), dada la limitación de disponibilidad de información recabada en los dos instrumentos utilizados, la cual comprende tres elementos principalmente: las creencias de docentes y estudiantes sobre los propósitos que debe perseguir la evaluación en el aula, las prácticas evaluativas que emplea el docente y en las que el estudiantado se involucra, y las herramientas de evaluación del aprendizaje que el docente utiliza para apoyar estas prácticas. En la educación superior, de acuerdo con [Contre-ras \(2010\)](#), [Corti et al. \(2011\)](#), [Fuller \(2013\)](#), [Moreno \(2009\)](#) y [Price et al. \(2011\)](#), la evaluación de los estudiantes se ha caracterizado mayoritariamente por una cultura de la evaluación *del* aprendizaje. Es decir, las prácticas evaluativas están más centradas en el enfoque tradicional, que se caracteriza por un uso prioritario de exámenes con fines sumativos y una visión de la evaluación como sinónimo de calificación. Cabe mencionar que esta diferenciación entre la evaluación *del* y *para* el aprendizaje, ha sido retomada por otros autores, tanto en la literatura anglosajona como la hispanohablante ([p.e. Gibbs & Simpson, 2006](#); [Moreno, 2016](#); [Rodríguez & Salinas, 2020](#); [Schuwirth & Van der Vleuten, 2011](#); [Padilla & Gil, 2008](#)) y data desde 1989, con el trabajo de [Martínez & Lipson](#).

Son diversos los estudios que han explorado la evaluación de los aprendizajes en el ámbito universitario, desde la opinión de docentes y estudiantes; en los cuales, suele indagarse el uso de estrategias de evaluación de orientación tradicional y de tipo constructivista. En cuanto a las percepciones del estudiantado universitario, el estudio de [Mendoza \(2019\)](#) exploró la satisfacción de los estudiantes de una universidad peruana, en torno a las prácticas evaluativas y el uso de los resultados de la evaluación. Los resultados muestran que el 46.3% de los estudiantes consideró que los docentes tenían capacidades insuficientes para llevar a cabo evaluaciones formativas, y para llevar a cabo una retroalimentación efectiva

para promover el aprendizaje (el 44.4%). Además, solo el 15% del estudiantado consideró que los docentes tomaban en cuenta los saberes previos de manera satisfactoria. En otro estudio, Gil-Flores (2012) exploró las concepciones de los estudiantes sobre las prácticas evaluativas de los docentes en nueve universidades españolas. Los resultados indicaron que, de acuerdo con los estudiantes, la evaluación del aprendizaje: se situaba únicamente al final del proceso de aprendizaje; se enfoca mayormente en el dominio de contenidos teóricos; era poca la retroalimentación que recibían por parte del personal docente; no se tomaban en cuenta las particularidades de los estudiantes a la hora de evaluar; las evaluaciones no se diseñan en función de la aplicabilidad de los conocimientos y habilidades en escenarios de un futuro contexto profesional; y no había participación de los estudiantes en la definición de criterios o selección de distintas estrategias para ser evaluados.

En cuanto a estudios sobre la opinión de docentes universitarios sobre las prácticas y el uso de las estrategias de evaluación que emplean, Halinen et al. (2013) exploraron las creencias y el uso de estrategias de evaluación de docentes de una universidad finlandesa. Se encontró que, los docentes preferían el empleo de exámenes escritos debido a la facilidad de su calificación y a la familiaridad de los estudiantes con este tipo de evaluación. Estos resultados coinciden con los reportados por Moreno-Olivos (2007), quien encontró que la aplicación de exámenes es la herramienta de evaluación más empleada por los docentes de una universidad mexicana, ya que, por reglamento escolar, es un requisito de evaluación que deben utilizar para la rendición de cuentas sobre los aprendizajes de los estudiantes. Al respecto, Panadero et al. (2018), encontraron que, de acuerdo con el análisis de los planes de estudio de nueve universidades españolas, la mayoría de las facultades universitarias priorizan el uso masivo de exámenes de opción múltiple o de respuesta corta, mientras que estrategias como la autoevaluación y la coevaluación son métodos raramente utilizados.

En suma, los estudios revisados, tanto de la percepción de estudiantes como de docentes, muestran que la evaluación del aprendizaje continúa con una mayor presencia en las aulas uni-

versitarias, pues se priorizan más los usos sumativos que formativos.

Conocer la percepción de los estudiantes universitarios sobre los procesos y estrategias de evaluación que realizan sus docentes, es relevante dado que influyen en diversos aspectos de su vida académica, como:

- los hábitos de estudio y de aprendizaje (Gibbs, 2006; Santos, 1999; Thomson & Falchikov, 1998);
- las actitudes de miedo o rechazo hacia las evaluaciones (Birenbaum, 2007; Thomson & Falchikov, 1998; Tiwari et al., 2005); y
- los valores, como la utilidad de la evaluación para retroalimentar su aprendizaje (Amaro de Chacín et al., 2008).

Este trabajo tuvo como propósito general el conocer la medida en que la cultura de la evaluación en una universidad pública mexicana está orientada al empleo de estrategias de evaluación para el aprendizaje. Esto, a través de la exploración de la percepción del estudiantado sobre las estrategias de evaluación para el aprendizaje que emplean sus profesores, en correspondencia con lo que el personal docente opina sobre las estrategias de evaluación para el aprendizaje que afirma dominar y emplear. Esta información permitirá a las autoridades universitarias identificar si las características de la evaluación en el aula, se alinean con el modelo educativo universitario orientado al enfoque constructivista, en el cual se promueve el uso de la evaluación con fines formativos (ver UABC, 2018).

Método

Diseño

El enfoque metodológico del presente estudio fue cuantitativo, con un diseño exploratorio, descriptivo y no experimental. En este sentido, se buscó describir algunas dimensiones en torno a la cultura de la evaluación para el aprendizaje desde la perspectiva de estudiantes y docentes de una universidad pública estatal del noroeste de México (Universidad Autónoma de Baja California, UABC).

Tabla 1
Población y muestra de estudiantes y docentes de UABC, período 2022-1

Participantes	N	n	%
Estudiantes	64,199	1,849	2.9
Docentes	5,572	545	9.8

Fuente: elaboración propia

Participantes

La UABC es una universidad pública del Estado de Baja California (México) fundada hacia mediados del siglo XX (1957) y focalizada en la oferta de servicios de enseñanza superior, investigación y extensión. Actualmente, cuenta con más de 100 programas de licenciatura y más de 50 programas de posgrado (maestrías y doctorados), además de institutos de investigación en múltiples áreas del conocimiento. La matrícula total de estudiantes de licenciatura durante el período 2022-1 fue de 64,199 alumnos, mientras que la población de profesores (de tiempo completo, medio tiempo y de asignatura) fue de 5,572 docentes durante el mismo período¹ (ver Tabla 1).

Para este estudio, a partir de los sujetos de estudio, se seleccionaron dos tipos de muestras con el mismo diseño: una muestra no probabilística de estudiantes de licenciatura y una muestra no probabilística de autoselección de docentes de la UABC. En el caso de los estudiantes, la muestra alcanzó a 1,849 sujetos (2.9% de la población), los cuales fueron invitados a participar contestando un cuestionario de evaluación del desempeño docente (Henríquez, Arámburo & Boroel, 2018; Henríquez & Arámburo, 2021) de manera presencial en el salón de clases. El criterio de inclusión fue solamente su asistencia los días de la aplicación del instrumento y la disponibilidad del docente de turno para interrumpir su clase por un lapso de 20-25 minutos. En el caso de los docentes, la muestra fue de 545 sujetos (9.8% de la población), a quienes se les aplicó el cuestionario sobre la cultura evaluativa de docentes universitarios (Contreras, 2022) de manera virtual y de participación voluntaria, a través de la plataforma abierta de formularios de Google. En este caso, el criterio de

inclusión fue su interés y la voluntad de responder el instrumento, cuyo enlace fue enviado vía correo electrónico institucional.

Instrumentos

- Cuestionario de evaluación del desempeño docente (estudiantes).
Dentro de la Facultad de Ciencias Administrativas y Sociales (FCAyS) de la UABC, cada semestre se ejecuta un proceso de evaluación del desempeño docente desde la opinión de los estudiantes. Para llevar a cabo este proceso, se utiliza un instrumento denominado Cuestionario de evaluación del desempeño docente, el cual está compuesto por cuatro dimensiones (Henríquez, Arámburo & Boroel, 2018; Henríquez & Arámburo, 2021):
 - Planeación de la enseñanza: la cual remite a elementos organizativos del curso, como por ejemplo, la entrega y el cumplimiento del plan de clases; la explicación en el plan de clases de aspectos como el propósito del curso, las competencias a desarrollar, los métodos y estrategias de enseñanza, las evidencias de desempeño y métodos de evaluación del aprendizaje, entre otros. Esta dimensión constó de 2 reactivos nominales dicotómicos y 10 reactivos medidos con escala de Likert de cuatro puntos (No explicó, La explicación fue confusa, Se explicó parcialmente claro y Se explicó claramente). El índice de confiabilidad obtenido, a través del coeficiente de Alfa, fue de .906.
 - Dominio de contenidos: que se refiere al grado de experticia en el manejo de los contenidos por el docente, considerando aspectos como la claridad y la secuencia con que los enseña; la conexión con otras materias; la asociación con situaciones, ex-

¹ https://indicadores.uabc.mx/indicadores/Ind_Publicos/

perencias o problemas de la vida cotidiana; el dominio teórico de los contenidos; entre otros. Esta dimensión se compuso de 16 reactivos medidos con escala de Likert de cuatro puntos (*Nunca, Algunas veces, Casi siempre y Siempre*). El índice de confiabilidad obtenido, a través del coeficiente de Alfa, fue de .947.

3. *Estrategias de enseñanza*: la cual remite a las estrategias didácticas que utiliza el docente en su práctica de enseñanza, con relación a los objetivos de aprendizaje que promueve (memorizar y reforzar la retención de contenidos, asociar conocimientos previos con nuevos contenidos, promover el aprendizaje grupal colaborativo, entre otros) y respecto a las actividades o técnicas de enseñanza (exposición oral de contenidos, uso de apoyos visuales, promoción del aprendizaje situado, uso de debates y diálogos grupales, entre otras). Esta dimensión se compuso de cinco reactivos medidos con escala de Likert de cuatro puntos (*Nunca, Algunas veces, Casi siempre y Siempre*) y 14 reactivos con la misma escala y puntos, pero distintas categorías (*Nula, Mala, Regular y Excelente*). El índice de confiabilidad obtenido, a través del coeficiente de Alfa, fue de .945.
4. *Evaluación del aprendizaje*: referida a las estrategias e instrumentos utilizados por el docente para evaluar el aprendizaje de los estudiantes y a las actitudes de éstos frente a los recursos de evaluación utilizados por el profesor. Esta dimensión estuvo conformada por cinco reactivos medidos con escala de Likert de cinco puntos (*No hace, Insatisfecho, Poco Satisfecho, Satisfecho y Muy satisfecho*), 11 reactivos con escala de Likert de cuatro puntos (*Nula, Mala, Regular y Excelente*) y nueve reactivos medidos con la misma escala y puntos, pero distintas categorías (*Muy en desacuerdo, Desacuerdo, De acuerdo y Muy de acuerdo*). El índice de confiabilidad obtenido, a través del coeficiente de Alfa, fue de .953.

Cabe destacar que este instrumento tiene como objetivo principal recolectar información

acerca del desempeño de los docentes con base en la opinión de los estudiantes, con miras a la retroalimentación, el favorecimiento de la formación continua y la toma de decisiones de mejora del desempeño de los docentes de la facultad. Con base en los objetivos del presente documento y para lograr su cumplimiento cabal, se utilizó solamente la información recolectada en torno a la dimensión de "Evaluación del aprendizaje", la cual considera tres subdimensiones: 1) estrategias de evaluación del aprendizaje, 2) instrumentos de evaluación del aprendizaje y 3) actitudes en torno a las estrategias e instrumentos de evaluación del aprendizaje.

- *Cuestionario sobre la cultura evaluativa de docentes universitarios.*

Por su parte, para recolectar información desde la mirada de los docentes, se utilizó el *Cuestionario sobre cultura evaluativa de docentes universitarios*, desarrollado por [Contreras \(2022\)](#). Este instrumento consta de tres dimensiones:

1. *Creencias sobre la evaluación del aprendizaje*: esta dimensión se compuso de dos escalas: usos principales de la evaluación del aprendizaje que le asigna el docente en el aula, conformada por ocho reactivos medidos en escala nominal dicotómica (*No lo domino, Lo domino poco, Lo domino y Lo domino mucho*) y cuyo índice de confiabilidad (coeficiente de Alfa) fue de .43; y autopercepción que tienen los docentes sobre el nivel de dominio de la evaluación del aprendizaje, conformada por cinco reactivos medidos con la misma escala y puntos, cuyo índice de confiabilidad (coeficiente de Alfa) fue de .81.
2. *Actitudes docentes hacia la evaluación del aprendizaje*: esta dimensión se compuso de tres escalas: usos de diferentes estrategias de evaluación del aprendizaje, conformada por 16 reactivos medidos con escala de Likert de cuatro puntos (*Nunca, Pocas veces, Casi siempre y Siempre*) y cuyo índice de confiabilidad (coeficiente de Alfa) fue de .79; ideas y preconcepciones que tienen acerca de la evaluación del aprendizaje en el aula, de 14 reactivos medidos escala de Likert de cuatro puntos (*Totalmente en desacuerdo, En Desacuerdo,*

De acuerdo y Totalmente de acuerdo), con un índice de confiabilidad (coeficiente de Alfa) de .83; y sentimientos hacia la evaluación del aprendizaje en el aula, compuesta por 11 reactivos medidos con la misma escala y puntos, cuyo índice de confiabilidad (coeficiente de Alfa) fue de .73.

3. *Valores éticos en la evaluación del aprendizaje en el aula*: esta dimensión se compuso de dos escalas: importancia que los docentes le otorgan a los valores ético-técnicos de las evaluaciones en el aula, compuesta de seis reactivos medidos con escala de Likert de cuatro puntos (*Nada importante, Algo importante, Importante y Muy Importante*), con un índice de confiabilidad (coeficiente de Alfa) de .83; y objetividad con la que ejecutan la evaluación del aprendizaje en el aula, compuesta de tres reactivos medidos con la misma escala y puntos, pero distintas categorías (*Nunca, Algunas veces, Casi siempre y Siempre*), cuyo índice de confiabilidad (coeficiente de Alfa) fue de .53.

En este caso, cabe destacar que, en función de los objetivos planteados en el presente estudio, para este reporte se utilizó únicamente la información relacionada con dos dimensiones: creencias sobre la evaluación del aprendizaje (escala de auto-percepción sobre el nivel de dominio de la evaluación del aprendizaje) y actitudes docentes hacia la evaluación del aprendizaje (escala de uso de diferentes estrategias de evaluación del aprendizaje).

Procedimiento

Respecto al *Cuestionario de evaluación del desempeño docente* desde la opinión de los estudiantes, el procedimiento de aplicación fue aprobado por la dirección de la FCAyS de la UABC y supervisado por la coordinación de evaluación docente de la facultad. Como se mencionó anteriormente, este proceso es parte de una estrategia interna de evaluación de la FCAyS que se ejecuta en cada ciclo escolar. Durante el período 2022-1, la aplicación se ejecutó en el mes de abril dentro del horario de clases, en cada salón seleccionado aleatoriamente, solicitando permiso al docente de turno e informando a los estudiantes de manera previa los objetivos, procedimientos y manejo

confidencial de la información a recolectar.

Por su parte, el *Cuestionario sobre cultura evaluativa de docentes universitarios* fue aplicado en línea durante enero de 2022 a través de la plataforma de formularios de Google. Se distribuyó con apoyo de la Coordinación General de Formación Profesional de la UABC, dependencia a través de la cual se solicitó a los directores de facultad y coordinadores de carrera reenviar, vía correo institucional, la liga de acceso del instrumento a su respectiva planta docente. El cuestionario estuvo disponible durante 15 días para ser contestado de manera voluntaria por los docentes de la UABC.

Análisis de datos

Una vez recolectada la información, se procedió a configurar el archivo con las variables de ambos instrumentos, vaciar las respuestas y purificar las bases de datos (suprimir los valores o casos perdidos), para ejecutar los análisis estadísticos por medio del programa *IBM Statistical Package for Social Science* (SPSS), versión 26. En función del cumplimiento de los objetivos de investigación, se obtuvieron los descriptivos básicos (distribuciones de frecuencia, índices de tendencia central y de dispersión) en torno a las variables de interés de ambos instrumentos, con el propósito de describir la cultura de la evaluación para el aprendizaje, tanto desde la óptica de estudiantes como de docentes. En el caso de los estudiantes, las variables claves fueron: *Frecuencia de uso de estrategias de evaluación del aprendizaje, Grado de satisfacción respecto al uso de instrumentos de evaluación del aprendizaje y Actitudes en torno a las estrategias e instrumentos de evaluación del aprendizaje, utilizadas por sus docentes*. Para los docentes, las variables analizadas fueron: *Autopercepción sobre el nivel de dominio de la evaluación del aprendizaje y Uso de estrategias de evaluación del aprendizaje de tipo constructivista y centradas en el estudiante*. A su vez, para dar mayor rigurosidad a los resultados presentados en este documento, se ejecutaron análisis inferenciales no paramétricos (chi cuadrado, X^2) con el propósito de analizar el grado de significatividad estadística entre las diferencias observadas y las teóricamente esperadas, en torno a las variables analizadas para estudiantes y docentes.

Resultados

Resultados de estudiantes

En la Tabla 2 se muestran las distribuciones de frecuencia sobre del grado de satisfacción de los estudiantes en torno a la implementación de diversas estrategias de evaluación por parte de sus docentes. Para esta variable se empleó una escala ordinal de tres puntos (*No hace*: no implementa la estrategia; *Baja*: insatisfecho, poco satisfecho; *Alta*: satisfecho, muy satisfecho). En general, se observan altos niveles de satisfacción en la mayoría de los estudiantes acerca de la implementación de todas estas estrategias por parte de sus docentes: los porcentajes de satisfacción alta se concentran en torno al 80% del estudiantado de la UABC. Se llevaron a cabo análisis estadísticos

para contrastar las frecuencias observadas y las teóricamente esperadas (chi-cuadrado, X^2) en todas las categorías: se observaron valores estadísticamente significativos (sig.=.000; 95% de confianza) para todos los casos. Para complementar los resultados anteriores, se llevaron a cabo comparaciones pareadas de proporciones entre las categorías extremas (*No hace* y *Alta*), por medio de la prueba Z (Kolmogórov-Smirnov), las cuales arrojaron diferencias significativas (sig.=.000) entre las frecuencias observadas en todas las estrategias de evaluación del aprendizaje.

Por su parte, en la Tabla 3 se observa la percepción de los estudiantes acerca del grado de calidad de uso de estrategias de evaluación por parte de los docentes, para lo cual se utilizó una escala ordinal de tres puntos (*Deficiente*, *Regular*,

Tabla 2
Grado de satisfacción de los estudiantes en torno al uso de estrategias de evaluación del docente

Estrategias de evaluación	No hace		Baja		Alta		X^2	Comparaciones pareadas
	n	%	n	%	n	%	Sig.	Sig.
Asesorías y apoyos individuales.	157	8.5	223	12.1	1458	79.4	,000	,000
Retroalimentación a tareas y otras actividades.	64	3.5	232	12.6	1541	83.9	,000	,000
Retroalimentación a evaluaciones parciales.	101	5.5	276	15.0	1459	79.5	,000	,000
Adecuaciones respecto a evaluaciones.	77	4.2	255	13.9	1505	81.9	,000	,000
Diagnóstico de los conocimientos previos.	105	5.7	261	14.2	1473	80.1	,000	,000

Fuente: elaboración propia

Tabla 3
Percepción de los estudiantes respecto al grado de calidad de los docentes en el uso de estrategias de evaluación del aprendizaje

Estrategias de evaluación	Deficiente		Regular		Excelente		X^2	Comparaciones pareadas
	n	%	n	%	n	%	Sig.	Sig.
Exámenes orales	436	23.9	715	39.2	673	36.9	,000	,000
Exámenes de opción múltiple	252	13.8	556	30.3	1024	55.9	,000	,000
Ensayos individuales	285	15.6	670	36.6	876	47.8	,000	,000
Ensayos grupales	328	18.0	626	34.3	873	47.7	,000	,000
Cuestionarios	222	12.1	585	31.9	1024	55.9	,000	,000
Reportes de lectura	242	13.2	657	35.9	933	50.9	,000	,000
Trabajos de investigación	176	9.6	590	32.2	1069	58.2	,000	,000
Exposiciones	253	13.8	606	33.1	973	53.1	,000	,000
Portafolios	239	13.1	603	33.0	988	53.9	,000	,000
Autoevaluación	239	13.1	564	30.8	1026	56.1	,000	,000
Coevaluación	258	14.1	589	32.2	982	53.7	,000	,000

Fuente: elaboración propia

Excelente) y se propusieron 11 tipos de estrategias de evaluación. Se aprecia que, para cada una de estas estrategias de evaluación, alrededor de la mitad de la muestra de estudiantes opinaron que sus docentes los dominan con un grado de excelencia. Sobresalen los trabajos de investigación (el 58.2%), la autoevaluación (el 56.1%), los cuestionarios (el 55.9%) y los exámenes de opción múltiple (el 55.9%), con los porcentajes más altos de calidad de uso. Por el contrario, los instrumentos con niveles de calidad de uso percibidos como regulares/deficientes fueron los exámenes orales (el 63.1%), ensayos individuales (el 52.2%) y ensayos grupales (el 52.3%). Las frecuencias anteriores se contrastaron por medio de la prueba estadística χ^2 , a través de la cual se constataron valores estadísticamente significativos (sig.= .000; 95% de confianza) para todos los casos. Como comple-

mento a estos resultados, se ejecutaron comparaciones pareadas de proporciones entre las categorías extremas (Deficiente y Excelente), por medio de la prueba Z (Kolmogórov-Smirnov), las cuales arrojaron diferencias significativas (sig.=.000) entre las frecuencias observadas en todas las estrategias de evaluación del aprendizaje. Respecto a las actitudes de los estudiantes acerca de las estrategias de evaluación empleadas por sus docentes, se utilizó una escala ordinal de cuatro puntos (*Totalmente en desacuerdo, Desacuerdo, Acuerdo y Totalmente de Acuerdo*). Para fines prácticos, las categorías de respuesta se transformaron a dicotómicas (ver Tabla 4). A nivel general, se observó que gran proporción de los estudiantes poseen actitudes positivas en torno a las estrategias evaluativas que utilizan los docentes. Sin embargo, también la mayoría con-

Tabla 4
Grado de acuerdo de los estudiantes según sus actitudes en torno a las estrategias de evaluación del aprendizaje que utilizan los docentes

Actitudes	Totalmente desacuerdo y Desacuerdo		Totalmente de acuerdo y Acuerdo		X ²	H de Cohen
	n	%	n	%	Sig.	
Me gustan las estrategias de evaluación que utiliza el docente.	222	12.3	1582	,000	,000	,377
Creo que aprendo de mejor manera a través de las evaluaciones que utiliza el docente.	319	17.7	1483	,000	,000	,323
El docente promueve formas de apoyo al aprendizaje paralelas a los exámenes parciales.	281	15.6	1520	,000	,000	,344
El docente se interesa en el mejoramiento del aprendizaje de los estudiantes, más allá de la calificación obtenida.	240	13.3	1565	,000	,000	,367
El docente se preocupa por establecer formas de evaluación relacionadas con problemáticas de la vida real.	217	12.0	1586	,000	,000	,380
El docente se preocupa por diferenciar entre los estudiantes que aprenden con mayor y menor facilidad, adaptando sus estrategias de enseñanza y las formas de evaluación.	392	21.8	1410	,000	,000	,282
Creo que el docente debería incluir otras estrategias de evaluación del aprendizaje en beneficio de todos los estudiantes.	516	28.6	1286	,000	,000	,214
Creo que el docente debería mejorar las estrategias de evaluación del aprendizaje que emplea, en beneficio de todos los estudiantes.	266	14.8	1532	,000	,000	,352

Fuente: elaboración propia

sideró que sus docentes deberían incluir otras estrategias de evaluación a las utilizadas (el 71.4%) y que deberían mejorar estas estrategias en beneficio de todos los estudiantes (el 85.2%). En este caso, también se realizaron análisis estadísticos para contrastar las frecuencias observadas y las teóricamente esperadas (chi-cuadrado, X^2) en todas las categorías: de la misma manera, se observaron valores estadísticamente significativos (sig.=.000; 95% de confianza) para todos los casos. Asimismo, con el propósito de complementar los resultados anteriores, se ejecutaron análisis para comparar las proporciones de cada categoría (estimador de tamaño del efecto, ETE) por medio del indicador H de Cohen. En todos los casos, los valores obtenidos se encuentran entre <0.2 y >0.5, lo que denota un grado de tamaño del efecto intermedio (Khalilzadeh & Tasci, 2017).

Resultados de docentes

En la Tabla 5 se observan las distribuciones de frecuencia en torno al grado de dominio de los docentes, según diferentes estrategias de evaluación en el aula. Se utilizó una escala ordinal de cuatro puntos (*No lo domino, Lo domino poco, Lo domino, Lo domino mucho*). Como se puede observar, la estrategia de evaluación que mayor proporción de docentes afirma *Dominar/ Dominar mucho*, fue la elaboración de exámenes abiertos y de opción

múltiple (el 91.1%). Por su parte, la actividad que menor proporción de docente afirma tener dominio, fue el empleo de estrategias de autoevaluación y coevaluación (el 58.7%). Las frecuencias en torno a esta variable se contrastaron por medio de la prueba estadística X^2 , a través de la cual se constataron valores estadísticamente significativos (sig.=.000; 95% de confianza) para todos los casos.

Por último, respecto a la frecuencia de uso de diversas estrategias de evaluación, se utilizó una escala ordinal de cuatro puntos (*Nunca, Algunas veces, Casi siempre, Siempre*). Como se aprecia en la Tabla 6, la mayoría de los docentes aseveran que utilizan de manera frecuente (*Casi siempre/ Siempre*) estrategias de evaluación de tipo formativa. En particular, evaluaciones auténticas que requieren la aplicación de conocimientos en situaciones reales (el 93.0%) y la retroalimentación grupal (el 89.5%). Llama la atención, los bajos porcentajes de frecuencia de uso (*Algunas veces/Nunca*) de estrategias de autoevaluación (el 47.2%) y, en particular, de la coevaluación entre estudiantes (el 68.0%). Igualmente, se realizaron análisis estadísticos para contrastar las frecuencias observadas y las teóricamente esperadas (chi-cuadrado, X^2) en todas las categorías: de igual forma, se observaron valores estadísticamente significativos (sig.=.000; 95% de confianza) para todos los casos.

Tabla 5
Grado de dominio docente sobre diferentes estrategias de evaluación

Estrategias de evaluación	No lo domino		Lo domino poco		Lo domino		Lo domino mucho		X^2 Sig.
	n	%	n	%	n.	%	n	%	
Elaboración de exámenes abiertos y de opción múltiple.	9	1.7	39	7.2	329	60.4	168	30.7	,000
Elaboración de otros tipos de evaluaciones: rúbricas, escalas, listas de cotejo.	21	3.9	118	21.7	305	56.0	101	18.4	,000
Ejecución de actividades de autoevaluación y coevaluación entre estudiantes.	54	9.9	171	31.4	259	47.5	61	11.2	,000
Retroalimentación detallada sobre fortalezas y áreas de oportunidad a cada estudiante.	23	4.2	111	20.4	301	55.2	110	20.2	,000

Fuente: elaboración propia

Tabla 6
Frecuencia de uso de estrategias de evaluación por parte de los docentes

Estrategias de evaluación	Nunca		Algunas veces		Casi siempre		Siempre		X ² Sig.
	n	%	n	%	n	%	n	%	
Al inicio del curso, evaluó los conocimientos previos	29	5.3	155	28.4	189	34.7	172	31.6	,000
Durante una actividad, evaluó en todo momento dificultades y avances de los estudiantes.	3	0.6	96	17.6	246	45.1	200	36.7	,000
Proporciono retroalimentación grupal sobre las principales debilidades de aprendizaje.	6	1.1	51	9.4	188	34.5	300	55.0	,000
Promuevo que los estudiantes se autoevalúen.	55	10.1	202	37.1	172	31.6	116	21.2	,000
Promuevo que los estudiantes evalúen a otros compañeros (coevaluación).	156	28.6	215	39.4	119	21.8	55	10.2	,000
Las estrategias y métodos de evaluación del aprendizaje buscan que los estudiantes apliquen sus conocimientos en situaciones reales.	5	0.9	33	6.1	189	34.7	318	58.3	,000
Adapto las estrategias o métodos de evaluación en los casos particulares de los estudiantes que así lo requieran.	20	3.7	111	20.4	203	37.2	211	38.7	,000

Fuente: elaboración propia

Discusión

Como se mencionó al inicio, este trabajo tuvo el propósito de explorar la cultura de la evaluación para el aprendizaje desde la mirada de estudiantes y docentes de una universidad pública del noroeste de México. Para ello, se analizó la concordancia entre las percepciones de ambos actores educativos con relación a los siguientes aspectos: el dominio y la frecuencia de uso de diversos tipos de evaluaciones por parte del personal docente, así como la percepción del estudiantado hacia la calidad del uso de diferentes estrategias evaluativas que emplean los docentes en el aula y sus actitudes en torno a estas estrategias utilizadas. Respecto a los resultados del personal docente, se encontró que, en cuanto a su autopercepción sobre el dominio de diversas estrategias de evaluación de tipo constructivista, 6 de cada 10 docentes afirmó dominar la autoevaluación y la coevaluación (58.7%). Sin embargo, casi la mitad de la muestra de estudiantes, calificaron como

Deficiente a Regular la calidad del uso de la autoevaluación (43.9%) y la coevaluación (46.3%) que emplean los docentes. Estos hallazgos son similares a los encontrados por Mendoza (2019), cuyos estudiantes universitarios consideraron que sus docentes poseían capacidades insuficientes para llevar a cabo evaluaciones de tipo formativas. Por su parte, la estrategia de evaluación que la mayor parte de la muestra docente (el 91.1%) afirma dominar/dominar mucho, fue la elaboración de exámenes abiertos y de opción múltiple. Lo anterior coincide con lo encontrado por otros estudios (Moreno-Olivos, 2007; Halinen et al. 2013; Panadero et al. 2018) cuyos resultados señalaron que los exámenes son el mecanismo de evaluación con el que tienen mayor familiaridad los docentes universitarios. Sin embargo, cuatro de cada 10 estudiantes de la UABC consideraron como regular/deficiente la calidad en que sus docentes utilizan los exámenes de opción múltiple (el 44.2%). Lo anterior deja en evidencia que, mientras la mayoría docente percibe que su capacidad para elaborar

estos recursos es alta, casi la mitad de los estudiantes opinan que la calidad de uso por parte de sus profesores es regular/deficiente.

A su vez, en cuanto al nivel de frecuencia con que los docentes de la UABC afirmaron que emplean diversas estrategias de evaluación de tipo constructivistas y centradas en el estudiante, el 93% de los docentes afirma llevar a cabo siempre/casi siempre, formas de evaluación donde los estudiantes puedan aplicar el conocimiento en situaciones reales. Este hallazgo contrasta con lo encontrado por [Gil-Flores \(2012\)](#), donde los estudiantes afirmaron que las evaluaciones del docente no se diseñaban en función de su aplicabilidad en contextos de la vida profesional.

En cuanto a las estrategias de evaluación para el aprendizaje con menos frecuencia de aplicación por parte de los docentes, fueron la autoevaluación (el 52.9%) y, en particular, la coevaluación (el 31.9%). Sin embargo, llama la atención que, más de la mitad de los estudiantes perciben que estas estrategias de evaluación son aplicadas con un grado de calidad excelente por parte del profesorado: el 56.1% de los estudiantes opina de esta manera respecto a la autoevaluación, y el 53.7% con relación a la coevaluación. Estas inconsistencias podrían deberse a que las muestras de estudiantes y docentes en el presente estudio, al ser no probabilísticas (y en el caso de los docentes, de autoselección), no estuvieron alineadas a los mismos campus y facultades, por lo que podrían interferir otras variables relacionadas con el área de conocimiento a la cual pertenecen ambos tipos de participantes.

Limitaciones

Este trabajo presentó tres principales limitaciones. Primero, los instrumentos utilizados no se diseñaron para comparar las respuestas entre estudiantes y docentes; es decir, no se diseñó un mismo instrumento donde las mismas dimensiones e indicadores se pudieran explorar según la percepción de ambos actores educativos. Esto, porque ambos instrumentos se enmarcaron en diferentes proyectos de la UABC. Sin embargo,

resultó de interés comparar los resultados, dadas las comunilidades en los constructos que ambos instrumentos evaluaron. Segundo, por este mismo motivo, la muestra de estudiantes puede no corresponder a los mismos campus y licenciaturas que la muestra de profesores que respondieron el cuestionario. Aún así, la intención de este ejercicio fue explorar la percepción general que existe entre ambos actores educativos respecto al empleo de estrategias de evaluación para el aprendizaje, en el aula universitaria. Y tercero, dentro del instrumento para docentes, se presentaron dos escalas con un coeficiente de confiabilidad Alfa demasiado bajo (usos principales de la evaluación del aprendizaje que le asigna el docente en el aula, con un valor de .43; y objetividad con la que ejecutan la evaluación del aprendizaje en el aula, con un valor de .53) lo cual limita su finalidad interpretativa. Sin embargo, cabe recordar que dentro del presente documento no se consideraron dichas escalas.

Conclusión

A manera de cierre, este trabajo invita a que futuros estudios puedan integrar la exploración de la cultura de la evaluación para el aprendizaje entre estudiantes y docentes universitarios, bajo un mismo diseño de instrumentos, que permita evaluar las mismas dimensiones, subdimensiones e indicadores, para ambos actores. También, aplicar estos instrumentos a muestras poblacionales que permitan la representatividad de docentes y estudiantes de diferentes áreas de conocimiento; con el fin de analizar si existen diferencias en la apropiación de una cultura de la evaluación para el aprendizaje según el grupo disciplinar al que pertenecen las diferentes carreras.

Finalmente, es importante mencionar que la educación superior requiere tanto de la evaluación del como para el aprendizaje, pues constituyen dos funciones necesarias; ambas son importantes para el propósito que persiguen, pero resulta imprescindible su uso equilibrado en el aula. Para ello, es necesario que los docentes universitarios coadyuven en la generación de una

adecuada cultura de la evaluación en las universidades, donde el peso del enfoque tradicional se equilibre con la incorporación y diversificación de estrategias de evaluación que motiven al estudiantado hacia la reflexión, motivación y regulación de su propio aprendizaje. Solo así, se podrá también configurar una cultura de la evaluación en los estudiantes, donde las evaluaciones no sean percibidas como una actividad burocrática, punitiva y orientada únicamente al control y a la rendición de cuentas.

Referencias

- Allal, L. (2016). The Co-Regulation of Student Learning in an Assessment for Learning Culture. *The Enabling Power of Assessment*, 259–273. doi:10.1007/978-3-319-39211-0_15
- Amaro de Chacín, R., Cadenas, M., y Altuve, J. (2008). Diagnóstico de los factores asociados a la práctica pedagógica desde la perspectiva del docente y los estudiantes. *Revista de Pedagogía*, 29(85), 215-244. <http://www.redalyc.org/articulo.oa?id=65911809002>
- Berlanga, M. L., y Juárez-Hernández, L. G. (2020). Paradigmas de evaluación: del tradicional al socioformativo. *Diálogos sobre educación*, 11(21), 1-14. <https://doi.org/10.32870/dse.v0i21.646>
- Birenbaum, M. (2007). Assessment and instruction preferences and their relationship with test anxiety and learning strategies. *Higher Education*, 53(6), 749-768. <https://doi.org/10.1007/s10734-005-4843-4>
- Black, P. & Wiliam, D. (2004). The Formative Purpose: Assessment Must First Promote Learning. *Yearbook of the National Society for the Study of Education*, 103, 20-50. <http://dx.doi.org/10.1111/j.1744-7984.2004.tb00047.x>
- Bolseguí, M. y Fuguet, A. (2006). Cultura de la evaluación: una aproximación conceptual. *Investigación y postgrado*, 21(1), 77-98. <https://www.redalyc.org/pdf/658/65821104.pdf>
- Contreras, G. (2010). Diagnóstico de dificultades de la evaluación del aprendizaje en la universidad: un caso particular en Chile. *Educación y Educadores*, 13. <https://www.redalyc.org/articulo.oa?id=83416998004>
- Contreras, S. (2022). *Características y variables asociadas a la cultura de la evaluación del aprendizaje, en docentes universitarios: el caso de la Universidad Autónoma de Baja California* [Tesis de doctorado, Universidad Autónoma de Baja California]. http://iide.ens.uabc.mx/documentos/divulgacion/tesis/DCE/2019/Sofia_Contreras_Roldan.pdf
- Corti, A. M., Oliva, L., y Gatica, M. L. (Diciembre 7-9, 2011). Cultura de la calidad, evaluación y acreditación. XI Colóquio Internacional Sobre Gestao Universiária na América do Sul. Florianópolis. <https://repositorio.ufsc.br/handle/123456789/25990>
- Fuller, M. (2013). An Empirical Study of Cultures of Assessment in Higher Education. *Education Leadership Review*, 14(1), 20-27. <https://files.eric.ed.gov/fulltext/EJ1105259.pdf>
- Gibbs, G. (2006). How assessment frames student learning. In Bryan, C. & Clegg, K. (Eds.) *Innovative Assessment in Higher Education: A Handbook for Academic Practitioners* [1st ed.]. Routledge. <https://cetl.ppu.edu/sites/default/files/publications/Innovative%20Assessment%20in%20Higher%20Education.pdf>
- Gil-Flores (2012). La evaluación del aprendizaje en la universidad, según la experiencia de los estudiantes. *Estudios sobre Educación*, 22, 133-153.
- Halinen, K., Ruohoniemi, M., Katajavuori, N., & Virtanen, V. (2014). Life science teachers' discourse on assessment: a valuable insight into the variable conceptions of assessment in higher education. *Journal of Biological Education*, 48(1), 16-22. <https://doi.org/10.1080/00219266.2013.799082>
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- Henríquez, P., Arámburo, V., y Boroel, B. (2018). Análisis de las estrategias de enseñanza según áreas de conocimiento en el nivel educativo superior: Percepciones de estudiantes de la Facultad de Ciencias Administrativas y Sociales (FCAYS), Universidad Autónoma de Baja California (UABC), México. *Compendio Investigativo de Academic Journals Celaya 2018*, tomo 14, 2320-2325. <https://static1.squarespace.com/static/55564587e4b0d1d3fb1eda6b/t/6009d7d8cf9fb56607ab6e41/1611257830213/Compendio+Investigativo+de+Academia+Journals+Celaya+2018+-+Tomo+14.pdf>
- Henríquez, P., Boroel, B., y Arámburo, V. (2020). Percepciones docentes en torno a la evaluación del aprendizaje en el nivel educativo superior: el caso de la UABC (México) y la UCM (España). *Revista Actualidades Investigativas en Educación*, 20(1), 1-23. <https://revistas.ucr>

- ac.cr/index.php/aie/article/view/40122
- Henríquez, P., y Arámburo, V. (2021). Evaluación del desempeño docente por áreas de conocimiento: El caso de la Facultad de Ciencias Administrativas y Sociales de la Universidad Autónoma de Baja California, México. *Actualidades Investigativas en Educación*, 21(3), 631-652.
- Khalilzadeh, J. & Tasci, A.D. (2017). Large sample size, significance level, and the effect size: Solutions to perils of using big data for academic research. *Tourism Management*, 62, 89-96. Disponible en: <https://doi.org/10.1016/j.tourman.2017.03.026>
- Lambert, E., y Holgado, M. A. (2001). La creación de cultura de evaluación institucional. *Revista Anales de Pedagogía*, 19. <http://revistas.um.es/analespedagogia/article/download/284971/206641>
- Martínez, N. (2010). Una aproximación a la cultura de la evaluación. *Diálogos*, 6(4), 7-20. <https://core.ac.uk/download/pdf/47265032.pdf>
- Martínez, M. E., & Lipson, J. I. (1989). Assessment for learning. *Educ Leader*, 47, 73-75.
- Mendoza, M. L. (2019). *Cultura evaluativa y satisfacción percibida en los estudiantes de la Facultad de Educación de la Universidad Nacional Mayor de San Marcos* [Tesis de maestría, Universidad Nacional Mayor de San Marcos]. Cybertesis Repositorio de Tesis Digitales. <https://cybertesis.unmsm.edu.pe/handle/20.500.12672/10984>
- Moreno, T. (2009). La evaluación del aprendizaje en la universidad: tensiones, contradicciones y desafíos. *Revista Mexicana de Investigación Educativa*, 14(41), 563-591. https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-6662009000200010
- Moreno-Olivos, T. (2007). La evaluación del aprendizaje en educación superior. El caso de la carrera de Derecho. *REencuentro. Análisis de Problemas Universitarios*, (48), 61-67. <https://www.redalyc.org/articulo.oa?id=34004809>
- Moreno, T. (2016). *Evaluación del y para el aprendizaje: reinventar la evaluación en el aula*. Universidad Autónoma de Morelos (UAM).
- Padilla, M.T. y Gil, J. (2008). La evaluación orientada al aprendizaje en Educación Superior: condiciones y estrategias para su aplicación en la docencia universitaria. *Revista Española de Pedagogía*, 6(241), 467-486. <https://dialnet.unirioja.es/servlet/articulo?codigo=2709011>
- Panadero, E., Fraile, J., Fernández, J., Castilla, D., & Ruiz, M. A. (2018). Spanish university assessment practices: examination tradition with diversity by faculty, *Assessment & Evaluation in Higher Education*. <https://doi.org/10.1080/02602938.2018.1512553>
- Price, M., Carroll, J., O'Donovan, B., & Rust, C. (2011). If I was going there I wouldn't start from here: a critical commentary on current assessment practice. *Assessment & Evaluation in Higher Education*, 36(4), 479-492. <http://dx.doi.org/10.1080/02602930903512883>
- Ricoy, M. C. y Fernández-Rodríguez, J. (2013). La percepción que tienen los estudiantes universitarios sobre la evaluación: Un estudio de caso. *Educación XX1*, 16(2), 321-342. <http://www.redalyc.org/articulo.oa?id=70626451006>
- Reategui, B. I. R. (2015). *Cultura evaluativa del aprendizaje en una universidad privada de Lima en el 2012* [Tesis de maestría, Universidad San Ignacio de Loyola, Perú]. <https://repositorio.usil.edu.pe/entities/publication/5d53e2a9-b2b6-4be0-874b-9272e77aca61>
- Rodríguez, H. M., y Salinas, M. L. (2020). La Evaluación para el Aprendizaje en la Educación Superior: Retos de la Alfabetización del Profesorado. *Revista Iberoamericana De Evaluación Educativa*, 13(1), 111-137. <https://doi.org/10.15366/rie2020.13.1.005>
- Rodríguez-Espinosa, H., Restrepo-Betancur, L. F. y Luna-Cabrera, G. C. (2016). Percepción del estudiantado sobre evaluación del aprendizaje en la educación superior. *Revista Electrónica Educare*, 20(3), 1-17. <https://www.redalyc.org/pdf/1941/194146862018.pdf>
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144.
- Sánchez, M. (2022). Evaluación del, para y como aprendizaje. En M. Sánchez y A. Martínez (Eds.) *Evaluación y aprendizaje en educación universitaria: estrategias e instrumentos* (Cap. 1, pp. 17-35). UNAM.
- Santos, M. A. (1999). 20 paradojas de la evaluación del alumnado en la universidad española. *Revista electrónica interuniversitaria de formación del profesorado*, 2(1), 369-392. <http://redined.mecd.gob.es/xmlui/handle/11162/147563>
- Rodríguez-Espinosa, H., Restrepo-Betancur, L. F. y Luna-Cabrera, G. C. (2016). Percepción del estudiantado sobre evaluación del aprendizaje en la educación superior. *Revista Electrónica Educare*, 20(3), 1-17. <https://www.redalyc.org/pdf/1941/194146862018.pdf>
- Schuwirth, L. W. T., & Van der Vleuten, C. P. M. (2011). Programmatic assessment: From assessment of learning

- to assessment for learning. *Medical Teacher*, 33(6), 478–485. https://www.researchgate.net/publication/51162010_Programmatic_assessment_From_assessment_of_learning_to_assessment_for_learning
- Thomson, K., & Falchikov, N. (1998). Full on until the sun comes out: the effects of assessment on student approaches to studying. *Assessment and Evaluation in Higher Education*. <https://srhe.tandfonline.com/doi/abs/10.1080/0260293980230405>
- Tiwari, A., Lam, D., Yuen, K. H., Chan, R., Fung, T., & Chan, S. (2005). Student learning in clinical nursing education: Perceptions of the relationship between assessment and learning. *Nurse Education Today*, 25(4), 299–308. <http://dx.doi.org/10.1016/j.nedt.2005.01.013>
- Tünnermann, C. (2011). El constructivismo y el aprendizaje de los estudiantes. *Universidades*, 48, 21–32. <http://www.redalyc.org/articulo.oa?id=37319199005>
- UABC (2018). Modelo educativo de la UABC [Cuadernos de Planeación y Desarrollo Institucional]. <http://web.uabc.mx/formacionbasica/documentos/ModeloEducativodelaUABC2018.pdf>
- Valenzuela, J. R., Ramírez, M. S., y Alfaro, J. A. (2011). Cultura de evaluación en instituciones educativas: Comprensión de indicadores, competencias y valores subyacentes. *Perfiles educativos*, 33(131), 42–63. https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0185-26982011000100004
- Wiliam, D. (2011). What is assessment for learning? *Studies in Education Evaluation*, 37(1), 2–13. <http://dx.doi.org/10.1016/j.stueduc.2011.03.001>

Autoevaluación en educación superior: percepción de los estudiantes sobre su utilidad en el proceso de aprendizaje

María Isabel Ramírez Olivares ^{*1}; Katherina Edith Gallardo ²

^{1,2} Tecnológico de Monterrey, Escuela de Humanidades y Educación, Monterrey, Nuevo León, México ¹<https://orcid.org/0009-0008-3866-9048> misabel.ramirezo@gmail.com, ²<https://orcid.org/0000-0001-8343-9518> katherina.gallardo@tec.mx

Citar como: Ramírez Olivares, M., & Gallardo Córdova, K. (2023). Autoevaluación en educación superior: percepción de los estudiantes sobre su utilidad en el proceso de aprendizaje. *Revista Digital de Investigación en Docencia Universitaria*, 17(2), e1574. <https://doi.org/10.19083/ridu.2023.1574>

Recibido: 27/10/2022 **Revisado:** 03/02/2022. **Aceptado:** 07/06/2023 **Publicado:** 30/06/2023.

Resumen

Introducción: la evaluación del aprendizaje es un proceso donde es importante incorporar a los estudiantes a partir de la autoevaluación, para involucrarlos y hacerlos responsables de su aprendizaje. **Objetivo:** indagar sobre la utilidad de la autoevaluación desde la perspectiva de estudiantes de Educación Superior. **Método:** la metodología es mixta, con un análisis cualitativo realizado en ATLAS.ti y un análisis cuantitativo desarrollado en SPSS y Excel. **Resultados:** la autoevaluación no es usada asiduamente ya que solo el 23% de los participantes la han usado con una frecuencia alta. **Conclusiones:** a) un área de oportunidad es que los estudiantes comparen su aprendizaje con el de compañeros, b) para una autoevaluación eficaz los estudiantes deben fortalecer la confianza en sí mismos, c) es necesario que los docentes incorporen actividades posteriores a la autoevaluación, d) el uso de autoevaluaciones es útil si su implementación tiene un objetivo y cuenta con un proceso estructurado.

Palabras clave: evaluación del aprendizaje; autoevaluación; educación superior; utilidad.

Self-assessment in higher education: students' perception of its usefulness in the learning process

Abstract

Introduction: Learning assessment is a process where it is important to incorporate students through self-assessment, to involve them and make them responsible for their learning. **Objective:** To investigate the usefulness of self-assessment from the perspective of higher education students. **Method:** The methodology is mixed, with a qualitative analysis carried out in ATLAS.ti and a quantitative analysis developed in SPSS and Excel. **Results:** Self-assessment is not frequently used since only 23% of the participants have used it with a high frequency. **Conclusions:** a) An area of opportunity is for students to compare their learning with their peers', b) for effective self-assessment, students must strengthen their confidence in themselves, c) it is necessary for teachers to incorporate activities after self-assessment, d) the use of self-assessment is useful if its implementation has a goal and a structured process.

Keywords: learning assessment; self-assessment; higher education; metacognition; usefulness.

*Correspondencia:

María Isabel Ramírez Olivares
misabel.ramirezo@gmail.com



Introducción

La evaluación en la educación tiene dos funciones que son, la pedagógica o del aprendizaje y la de rendición de cuentas que abarca criterios de acreditación de programas de estudio y certificación de la calidad ([Muskin, 2015](#)). No obstante, en el ámbito de la Educación Superior en México, la evaluación como proceso ha perdido el foco en la valoración de aprendizajes y se ha enfocado principalmente en la acreditación de programas de estudio que, si bien es un aspecto importante, debe ser complementario a los fines didácticos y no excluyente. Esto se observa en las instituciones gubernamentales y no gubernamentales que se encargan de la evaluación en la Educación Superior ([García y Hervás, 2020](#)).

En la mayoría de estas instituciones como la Asociación Nacional de Universidades e Instituciones de Educación Superior ANUIES y el Consejo Nacional de Ciencia y Tecnología CONACyT, que se encargan de la acreditación de la calidad sin enfocarse específicamente en la valoración de los aprendizajes de los estudiantes que egresan de las diferentes instituciones de Educación Superior. En este sentido, solo el Centro Nacional de Evaluación para la Educación Superior CENEVAL, se enfoca en la evaluación de los aprendizajes y se aboca únicamente al uso del examen como instrumento de evaluación ([García y Hervás, 2020](#)).

En este sentido, debemos revalorar la evaluación del aprendizaje como un proceso complejo, en el que cada uno de sus componentes es fundamental para tomar decisiones en el contexto educativo. Debido a esto, no se debe perder de vista la importancia de todos los actores educativos involucrados, como son los estudiantes, docentes, directivos, entre otros. Especialmente se debe integrar la perspectiva de los estudiantes en cada paso a realizar, sin embargo, en los diferentes niveles educativos es constante que se repliquen prácticas e instrumentos de heteroevaluación periodo tras periodo, que se enfocan en su mayoría, en la importancia del juicio de los docentes y no en el logro de aprendizajes de los estudiantes ([Díaz Barriga 2001 y Copado, 2022](#)).

Lo anterior, hace que los estudiantes enfoquen sus esfuerzos en cubrir las expectativas de otros,

dejando de lado la valía de la autoevaluación al no incluirlos en la valoración de su propio aprendizaje y mengua la percepción de responsabilidad que tienen sobre este ([Taras, 2015](#) y [Panadero, Brown y Strijbos, 2016](#)). Esto, sin duda no es un hecho generalizable, pero sí algo comúnmente realizado y que se ve evidenciado en la escasa producción científica sobre el tema de autoevaluación del aprendizaje en América Latina, no solo en educación superior, sino en todos los niveles educativos.

Además, se observa que la autoevaluación ha sido relegada, a partir de la mínima atención del diseño de políticas públicas en México sobre el tema. Por ejemplo, en la Ley General de Educación (2019), a pesar de que existe un capítulo sobre la Educación Superior no se aborda el tema de la evaluación ni de la autoevaluación. También, en la Ley General de Educación Superior (2021) el tema de la evaluación se encauza solo en la acreditación de estudiantes y certificación de planes y programas de estudio.

Con respecto a los antecedentes desarrollados, se establece la importancia de poner atención en integrar la evaluación como un elemento clave dentro de las actividades educativas. Así como de incorporar instrumentos de autoevaluación en diferentes momentos del curso escolar que valoren la consecución del aprendizaje, la construcción de conocimientos y la toma de conciencia sobre elementos metacognitivos que influyen en el aprendizaje.

Asimismo, se coloca en relieve la necesidad de promover la perspectiva pedagógica de la evaluación para ir más allá de la rendición de cuentas institucionales y la transparencia. Por lo tanto, es del interés de la presente investigación indagar sobre el uso de instrumentos y procedimientos de autoevaluación, así como de la utilidad que tienen para estimar el aprendizaje y de hacer conscientes elementos metacognitivos desde la perspectiva del estudiante. En este sentido, es significativo estudiar esta utilidad con referencia a si al autoevaluarse los estudiantes toman decisiones que influyen de forma positiva en sus intereses e identifica el uso que dará a las evaluaciones para contribuir en su aprendizaje (Joint Committee on Standards for Educational Evaluation JCSEE, 2014):

Con base en lo anterior, es necesario desarrollar un marco teórico que permita entender la importancia de la evaluación y la utilidad de la autoevaluación. En este sentido, evaluar en la educación está entendido como un proceso sistemático y fundamental en la formación de estudiantes, que demanda el uso de diversos instrumentos y la intervención de distintos actores que permitan el acopio y análisis de información con la que es posible valorar y promover el aprendizaje, así como tomar decisiones de manera fundamentada e influir en la planeación y la didáctica ([Miller, Linn y Gronlund, 2012](#); [Gallardo, 2013](#) y [Sánchez, 2020](#)).

Con respecto a los elementos a considerar para realizar una adecuada evaluación del aprendizaje, es pertinente que se analice la intención al llevar a cabo este proceso. La intención de la evaluación, a partir de lo establecido por [Paredes \(2016\)](#), puede ser diagnóstica, formativa o sumativa:

- a. Evaluación con intención diagnóstica: es utilizada para definir el punto de partida de algún proceso de aprendizaje o para determinar previamente el nivel de conocimientos, habilidades y actitudes de los estudiantes. Este tipo de evaluación hace posible conocer el estado inicial tanto del grupo como de cada estudiante en individual ([Rosales, 2003](#); [Frade Rubio, 2009](#); [Gallardo, 2013](#) y [Sánchez, 2020](#)).
- b. Evaluación con intención formativa: se usa para realizar un monitoreo continuo y progresivo. Para llevar a cabo esta evaluación se deben planear instrumentos que serán aplicados a lo largo del proceso de aprendizaje, e incluso posterior a terminar. Hacer esto, permite a docentes y estudiantes apreciar el avance, detectar errores y adaptar las estrategias de enseñanza ([Martínez Rizo, 2009, 2013](#); [Gallardo, 2013](#) y [Sánchez, 2020](#)).
- c. Evaluación con intención sumativa: se utiliza para estimar los resultados de aprendizaje al final de un proceso formativo ([Scriven, 1967](#)). Este tipo de valoración se hace al concluir un programa y es útil para identificar qué tanto el estudiante logró los objetivos de aprendizaje, asignar calificaciones o propiciar acreditaciones y certificaciones ([Sánchez, 2020](#)).

Además de considerar la intención se debe tomar en cuenta al actor principal que efectúa la valoración del aprendizaje. En la heteroevaluación, el juicio sobre el desempeño y aprendizajes logrados por los estudiantes es emitido por docentes y educadores ([Fernández y Vanga, 2015](#)). En la coevaluación se considera la voz de los pares y esto hace posible tomar en cuenta la apreciación entre iguales del desempeño y aprendizaje del otro ([Topping, 1998](#); [Torres y Torres, 2005](#); [Jiménez, 2006](#) y [Fernández y Vanga, 2015](#)). En la autoevaluación, el mismo estudiante es el que valora su propio aprendizaje y desempeño, generando a la par un proceso de apoyo para el desarrollo de estrategias metacognitivas y de autorregulación ([Boud y Falchikov, 1989](#); [Boud, 1995](#) y [Kambourova, González-Agudelo y Grisales-Franco, 2021](#)).

En este sentido, la autoevaluación del aprendizaje cuenta con tres condiciones para que el proceso sea adecuado y funcional (Goodrich, 1996). Estas condiciones son; a) la conciencia del valor de la autoevaluación ya que da cabida a abordar el sentido del porqué y el para qué estimar su aprendizaje, b) tener acceso a los criterios en los que se basa la evaluación para que así los estudiantes conozcan las directrices que guían la evaluación y c) la especificidad de la tarea que se va a evaluar, para que los estudiantes tengan claridad sobre la actividad y los pasos a seguir para realizar la autoevaluación.

Con respecto a los tipos de autoevaluación que es posible realizar, existen diferentes tipologías que parten de diferentes criterios de categorización ([Panadero, Brown y Strijbos, 2016](#)). En la Tabla 1, se enlistan cinco de esas tipologías elaboradas con base en la investigación de diversos autores ([Boud y Brew, 1995](#); [Tan, 2001](#); [Brown y Harris, 2013](#); [Panadero y Tapia, 2013](#) y [Taras, 2015](#)).

Sobre el término de metacognición, este ha sido estudiado por diversos autores que han hecho revisiones y propuestas sobre a qué se refiere este concepto que fue introducido en 1976 por Flavell. Por lo tanto, para fines de esta investigación, es necesario especificar que la metacognición se entiende como un proceso cognitivo de orden superior, que abarca el conocimiento y

regulación de la cognición (Flavell, 1979; Brown, 1987; Kim, Zyromski, Mariani, Lee y Carey, 2017 y Valenzuela, 2019).

Asimismo, se toma como base un modelo que abarca lo establecido por Flavell (1979), Brown (1987) y Schraw y Dennison (1994) donde se especifican los componentes del conocimiento y regulación de la cognición. Por un lado, el conocimiento de la cognición se divide en tres tipos que son el declarativo, el procedimental y el condicional. Por otro lado, la regulación de la cognición se divide en tres dimensiones que son la planeación, el monitoreo y la evaluación.

Con respecto a los tipos de conocimiento de la cognición; el conocimiento declarativo comprende el entendimiento que cada estudiante tiene sobre sus habilidades y su desempeño como apren-

diz, el conocimiento procedimental incluye la comprensión de las estrategias y procedimientos para resolver problemas y aprender, y el conocimiento condicional implica el grado de dominio que se tiene sobre cuándo, cómo y por qué usar cada estrategia y/o procedimiento (Brown, 1987 y Schraw y Dennison, 1994).

En relación con las dimensiones de la regulación de la cognición; la planeación se realiza previo a la ejecución de la tarea donde se establece el tiempo a invertir, estrategias y metas, el monitoreo, se lleva a cabo durante la elaboración de la tarea para supervisar el avance y aprendizaje en cuanto a las metas establecidas, y la evaluación se efectúa posterior a la tarea para valorar las estrategias implementadas y el grado de consecución de las metas establecidas (Brown, 1987).

Tabla 1
Tipos de autoevaluación

Autor/es		Categorización
Boud y Brew (1995)	1.	Interés técnico
	2.	Interés comunicativo
	3.	Interés emancipatorio
Tan (2001)	1.	Autoconciencia
	2.	Autoapreciación
	3.	Evaluación autodeterminada
	4.	Prácticas de autoevaluación
	5.	Autoevaluación de tareas
	6.	Autocalificación o autoexamen
Brown y Harris (2013)	1.	Autocalificación
	2.	Autoestimación del desempeño
	3.	Evaluación basada en criterios y rúbricas.
Panadero y Tapia (2013)	1.	Autocalificación
	2.	Rúbrica de evaluación
	3.	Guiones de autoevaluación
Taras (2014)	1.	Autocalificación
	2.	Estándar robusto
	3.	Contrato de aprendizaje
	4.	Autoevaluación estándar
	5.	Modelo de autoevaluación con retroalimentación integrada entre pares y el profesor

Nota: El nombre utilizado en las tipologías de Tan (2001) y Brown y Harris (2013) son traducciones realizadas por la autora, por lo que se podrían encontrar en otros escritos con diferente nombre.

Método

Diseño

La presente investigación se realizó con base en el método mixto, definido por [Plano-Clark e Ivankova \(2016\)](#) como una integración de enfoques cuantitativos y cualitativos de forma intencionada, cuya finalidad es abordar de manera idónea la problemática de investigación. No obstante, no debe mal entenderse el uso del método mixto como la suma de ambas perspectivas, la cuantitativa y la cualitativa. Este método se aborda como una aleación, desarrollada con base a una estrategia que aporta al desarrollo de la investigación ([Ramírez-Montoya y Lugo-Ocando, 2020](#)).

Se eligió hacer este abordaje metodológico, debido a que la aplicación del método mixto en investigaciones del área educativa brinda la posibilidad de comprender y abordar de una forma holística y ajustada, la complejidad y la naturaleza cambiante de los fenómenos de la educación. Lo anterior gracias a la diversidad de diseños, instrumentos y estrategias de análisis que es posible utilizar. Por lo tanto, a pesar de que aún existen retos por superar, es importante aproximarse en mayor medida a temas de innovación educativa y evaluación del aprendizaje desde el enfoque mixto para obtener mayor profundidad y amplitud en los estudios realizados ([Ramírez-Montoya y Lugo-Ocando, 2020](#)).

Participantes

El estudio se llevó a cabo en tres universidades de la zona centro de México, dos de estas instituciones de financiamiento público y una de financiamiento privado. En total participaron 61 estudiantes, 50 personas de escuelas públicas y 11 personas de escuelas privadas. La distribución de estudiantes por carrera es de 26 de la Licenciatura en Pedagogía, 10 de la Licenciatura en Trabajo Social y 25 de la Ingeniería en Biotecnología. Asimismo, de cada carrera hubo participantes de las asignaturas de Bioquímica, Habilidades gerenciales, Gestión de proyectos, Identidad y vinculación profesional, Diseño y evaluación de planes y programas de estudio, Didáctica, Investigación pedagógica y Práctica de especialización. Los participantes fueron elegidos de forma censal, de

los cuales 50 eran mujeres y 11 hombres, el rango de edad va de los 20 a los 29 años.

Instrumentos

Para llevar a cabo este estudio se aplicaron tres cuestionarios que pasaron por un proceso de validez de contenido con base en un juicio de expertos. Cada cuestionario cuenta con la siguiente estructura:

- a. Primer cuestionario: la primera sección contiene ítems de las experiencias previas de autoevaluación, en la segunda sección se hace una proyección sobre si consideran que lograrán los objetivos de aprendizaje de esa asignatura, la tercera sección profundiza en la auto valoración de procesos metacognitivos y afectivos durante su aprendizaje en asignaturas previas y, por último, en la cuarta sección se pregunta de forma abierta sobre la utilidad de ese cuestionario como medio de autoevaluación.
- b. Segundo cuestionario: la primera sección hace una valoración sobre si consideran que están logrando los objetivos de aprendizaje de la asignatura, la segunda sección profundiza en la auto valoración de procesos metacognitivos y afectivos desarrollados en la asignatura y, por último, en la tercera sección se pregunta de forma abierta sobre la utilidad de ese cuestionario como medio de autoevaluación.
- c. Tercer cuestionario: la primera sección hace una valoración sobre si consideran que lograron los objetivos de aprendizaje de la asignatura, la segunda sección profundiza sobre la auto valoración de procesos metacognitivos y afectivos experimentados durante toda la asignatura y, por último, en la tercera sección se pregunta mediante ítems sobre su percepción de la utilidad de la autoevaluación y de forma abierta sobre la utilidad de ese cuestionario en específico como medio de autoevaluación.

Procedimiento

La aplicación de los instrumentos se llevó a cabo con consentimiento previo de los estudiantes y asegurando la confidencialidad de sus datos, dicha aplicación se realizó en los siguientes tres

momentos: a) el primer momento fue durante las primeras semanas de cada curso donde se profundizó sobre sus antecedentes reflexionando sobre sus procesos metacognitivos, experiencias previas de autoevaluación y su percepción ante el primer instrumento, b) el segundo momento se desarrolló aproximadamente a la mitad de cada curso donde se les pidió auto valorar su aprendizaje, procesos metacognitivos y se recabó su percepción ante el segundo instrumento, y c) el tercer momento se realizó en las últimas semanas del curso donde cada estudiante autoevaluó sus procesos metacognitivos y de aprendizaje, comentó su percepción de utilidad de este tercer instrumento y en general la utilidad de la autoevaluación.

Análisis de datos

En este sentido, el análisis de los datos obtenidos se realizó de manera diferenciada. Por un lado, el análisis cuantitativo se realizó con apoyo del software SPSS para presentar la distribución de frecuencias y la correlación de las respuestas de datos no paramétricos, que fueron utilizados así ya que los datos no tienen una distribución normal dado que se obtuvo a los participantes de forma censal. Por otro lado, el análisis cualitativo del discurso se llevó a cabo con el software ATLAS.ti a partir de identificar con códigos elementos esperados con base en la literatura y reconocer aspectos no esperados que hablan sobre la utilidad que perciben los estudiantes de realizar autoevaluaciones durante su proceso de aprendizaje.

Resultados

En primer lugar, se presentan los resultados del análisis cuantitativo, donde es importante resaltar que los 61 participantes que configuran nuestra población contestaron el primer cuestionario donde se estableció un diagnóstico sobre la experiencia previa con la autoevaluación. Con base en estas respuestas, se identifica en la Tabla 2 que la autoevaluación no es una forma de evaluación que sea usada asiduamente, ya que solo el 23% ha usado la autoevaluación en una frecuencia alta. Asimismo, se muestra en la Tabla 3 que la forma más frecuente de autoevaluación es la autovaloración del desempeño, indicando fortalezas y debilidades con apoyo de una rúbrica.

En cuanto a la definición de autoevaluación, la opción más frecuente que se observa en la Tabla 4 con un 67.2%, es la de “reflexionar sobre lo que he logrado y lo que me hace falta lograr para alcanzar los objetivos de aprendizaje”. También, sobre cómo ha sido la experiencia de los participantes con la autoevaluación, se identifica en la Tabla 5 que es poco frecuente que se use para valorar el aprendizaje de un estudiante en comparación con otro compañero. Asimismo, se observa en la Tabla 6 que el momento menos frecuente en el que se realizan autoevaluaciones es al inicio como evaluación diagnóstica y, el más frecuente es al final del curso, con lo cual se puede inferir que ha sido más utilizada como evaluación sumativa.

Tabla 2
Frecuencia con la que los estudiantes han realizado autoevaluaciones

	Frecuencia	Porcentaje
Alta	14	23.0
Regular	33	54.1
Baja	10	16.4
Muy baja	4	6.6
Total	61	100.0

Tabla 3
Formas de autoevaluación que los estudiantes han realizado

	Frecuencia	Porcentaje
Autocalificarme numéricamente, colocando una nota en función de lo que yo pienso aprendí en el curso	26 de 61 personas	42.6
Auto valorar mi desempeño, indicando fortalezas y debilidades con apoyo de una rúbrica	52 de 61 personas	85.2
Narrando mi experiencia de aprendizaje desde mi perspectiva (sin criterios preestablecidos)	14 de 61 personas	22.9

Tabla 4
Mejor definición de autoevaluación

	Frecuencia	Porcentaje
Colocarme una calificación a mi aprendizaje o desempeño	8	13.1
Narrar brevemente lo que he aprendido y lo que me falta por aprender en un curso o módulo	12	19.7
Reflexionar sobre lo que he logrado y lo que me hace falta lograr para alcanzar los objetivos de aprendizaje	41	67.2
Valorar mi aprendizaje o desempeño en comparación con mis pares o compañeros.	0	0
Total	61	100.0

Tabla 5
Experiencia sobre cómo han sido consideradas las autoevaluaciones

	Frecuencia	Porcentaje
Han sido parte de las calificaciones obtenidas en cursos	21	34.4
Han sido parte del proceso de retroalimentación que me han dado los profesores	22	36.1
Han sido reflexiones que me han ayudado a entenderme más como aprendiz	16	26.2
Han sido reflexiones que me han ayudado a compararme con otros pares o compañeros	2	3.3
Total	61	100.0

Por otro lado, al cuestionar a los participantes sobre las expectativas que tenían al autoevaluar-se, la Tabla 7 nos indica que prefieren que el docente comente cuáles fueron sus errores y aciertos. Sobre la utilidad que se atribuye a la autoevaluación, con base en los resultados de la Tabla 8 se infiere que a los participantes les resultaría útil para reflexionar sobre cómo aprenden en términos de estrategias y métodos, dicha respuesta representa la más frecuente con un 52.5%. Asimismo, como se

observa en la Tabla 9, las emociones más presentes durante la autoevaluación fueron el entusiasmo y confianza con un 44.3%, así como la sorpresa e incertidumbre con un 39.3%.
En cuanto a la percepción de comodidad de los participantes durante el proceso de autoevaluación, se observa en la Tabla 10 que prefieren que los criterios establecidos para evaluar surjan de una co-construcción entre el estudiante y el docente.

Autoevaluación en educación superior: percepción de los estudiantes sobre su utilidad en el proceso de aprendizaje

Tabla 6
Momento en el que se realiza la autoevaluación

	Frecuencia	Porcentaje
Al inicio de una materia o curso	11	18.0
Durante el transcurso del proceso de aprendizaje de una materia o curso	20	32.8
Al final de una materia o curso	30	49.2
Total	61	100.0

Tabla 7
Expectativas sobre la autoevaluación

	Frecuencia	Porcentaje
Que el profesor lo lea y comente conmigo para identificar mis errores/lo que puedo mejorar/ identificar por mí mismo mis fortalezas y áreas de oportunidad	29	47.5
Identificar por mí mismo mis fortalezas y áreas de oportunidad	26	42.6
Reflexionar para entender mejor cómo aprendo	5	8.2
Recibir una calificación que me ayude a obtener un mejor promedio	1	1.6
Total	61	100.0

Tabla 8
Utilidad de la autoevaluación

	Frecuencia	Porcentaje
Si te permitiera darte cuenta de los contenidos que has retenido	15	24.6
Si te permitiera darte cuenta de las destrezas que has desarrollado	14	23.0
Si te permitiera reflexionar sobre cómo aprendes en términos de estrategias y métodos	32	52.5
Si te permitiera compararte con lo demás compañeros de tu mismo grupo o clase	0	0
Total	61	100.0

Tabla 9
Emociones durante la autoevaluación

	Frecuencia	Porcentaje
Entusiasmo, confianza	27	44.2
Inseguridad, enfado	7	11.5
Indiferencia, desinterés	3	4.9
Sorpresa, incertidumbre	24	39.3
Total	61	100.0

Tabla 10
Comodidad con la autoevaluación

	Frecuencia	Porcentaje
Proponiendo yo mismo los criterios para el proceso de autoevaluación	12	19.7
Que los criterios los proponga el docente del curso o la institución	5	8.2
Que los criterios sea una co-construcción entre el docente y yo	38	62.3
Que los criterios provengan de una entidad certificadora en la materia	6	9.8
Total	61	100.0

También, en la Tabla 11 un 49.2% de estudiantes refieren que no obtienen indicación alguna posterior a conocer el resultado de la autoevaluación. Además, sobre la utilidad alternativa que brindó la autoevaluación al estudiante, en la Tabla 12 se observa que el 42.6% considera más importante mejorar el diseño instruccional, en cuanto a los tipos de actividades y evaluaciones del curso.

Por otro lado, en la Tabla 13 se observa que re-

formular y mejorar los procesos de aprendizaje que ejercen, así como clarificar qué y cómo deben seguir trabajando para aprender mejor, son las ayudas más significativas que brinda la autoevaluación. En este sentido, en la Tabla 14 se muestra que el 63.9% de los estudiantes prefieren que los resultados obtenidos de estas autoevaluaciones coincidan con lo que opinan de sí mismos, independientemente de lo que los demás opinen.

Tabla 11
Solicitud del docente posterior a la autoevaluación

	Frecuencia	Porcentaje
Realizar más ejercicios/revisar lecturas para reforzar mis áreas de oportunidad	21	34.4
Poner más atención en clase/participar más	10	16.4
No he recibido indicaciones específicas luego de analizar mis autoevaluaciones	30	49.2
Total	61	100.0

Tabla 12
Utilidad alternativa de la autoevaluación

	Frecuencia	Porcentaje
Mejorar el currículum (temas de las clases)	6	9.8
Mejorar la didáctica del profesor (manera en la que enseña)	16	26.2
Mejorar el diseño instruccional (tipos de actividades y evaluación del curso)	26	42.6
Redefinir los objetivos de aprendizaje	13	21.3
Total	61	100.0

Tabla 13
Ayuda que brinda la autoevaluación

	Frecuencia	Porcentaje
Reafirmar las metas de aprendizaje que tienes	12	19.7
Reformular y mejorar los procesos de aprendizaje que ejerces	22	36.1
Clarificar qué y cómo debes seguir trabajando para aprender mejor	21	34.4
Precisar los mecanismos para aprender mejor	6	9.8
Total	61	100.0

Tabla 14
Resultados con los que debe concordar la autoevaluación

	Frecuencia	Porcentaje
La calificación que mi profesor me otorgó	20	32.8
Lo que mis compañeros opinan de mi desempeño	2	3.3
Lo que yo opino de mí mismo, independientemente de lo que los demás opinen	39	63.9
Total	61	100.0

Con base en los datos anteriores, se han identificado tres correlaciones significativas entre variables:

- a. La primera correlación se representa en la Tabla 15, donde nos muestra que reflexionar sobre lo que se ha logrado y lo que hace falta lograr para alcanzar los objetivos de aprendizaje influye de forma significativa en que se presenten emociones como el entusiasmo, confianza, sorpresa e incertidumbre.
- b. La segunda correlación se observa en la Tabla 16, lo que nos dice que, si los resultados de la autoevaluación concuerdan con lo que los participantes opinan de sí mismos, independientemente de lo que los demás opinen, influye de forma significativa en que se presenten emociones como el entusiasmo, confianza, sorpresa e incertidumbre.
- c. La tercera correlación se vislumbra en la Tabla 17, donde nos muestra que, al no haber solicitudes específicas después de la autoevaluación, influye de forma significativa en que los participantes por

iniciativa propia reformulen y mejoren los procesos de aprendizaje que ejercen y clarifiquen qué y cómo deben seguir trabajando para aprender mejor.

Por otro lado, en el tercer cuestionario donde se obtuvieron respuesta solo de 25 personas de las 61 iniciales, se indagó sobre la utilidad de la autoevaluación. En este sentido, a partir de los datos obtenidos, en la Tabla 18 identificamos que la mayoría de los participantes consideran que es más útil llevar a cabo una autoevaluación aproximadamente a la mitad del curso. Asimismo, en la Tabla 19 se visualiza una alta percepción de utilidad de la autoevaluación para establecer, monitorear y evaluar objetivos de aprendizaje. También se percibe como útil para valorar sus habilidades como aprendiz, identificar cuáles estrategias usar para promover su aprendizaje, así como reconocer cuándo y porqué se deben utilizar esas estrategias. De igual manera, los participantes consideran que es útil para valorar su motivación y monitorear sus emociones con respecto a la asignatura.

Tabla 15
Asociación entre la definición de autoevaluación y las emociones que se presentan al autoevaluarse

	Valor	Significancia
Pearson Chi-Square	14.384a	0.026
Phi	0.486	0.026
Cramer's V	0.343	0.026

Tabla 16
Asociación entre las emociones que se presentan al autoevaluarse y los resultados obtenidos

	Valor	Significancia
Pearson Chi-Square	8.550a	0.073
Phi	0.374	0.073
Cramer's V	0.265	0.073

Tabla 17
Asociación entre la ayuda que brinda la autoevaluación y lo que solicitan los profesores, posterior a realizar estos procesos

	Valor	Significancia
Pearson Chi-Square	13.813a	0.032
Phi	0.476	0.032
Cramer's V	0.336	0.032

Tabla 18
Momento del curso en el que consideran los participantes que es más útil la autoevaluación para su aprendizaje

	Frecuencia
En las primeras semanas del curso	6 de 25 personas
Aproximadamente a la mitad del curso	16 de 25 personas
Al final del curso	10 de 25 personas

Tabla 19
Utilidad de la autoevaluación percibida por los participantes

Afirmaciones sobre la autoevaluación en función de su utilidad	Si	No
La autoevaluación realizada durante su curso es útil para establecer objetivos de aprendizaje	25	0
La autoevaluación realizada durante su curso es útil para elegir cuáles estrategias debían implementar para promover su aprendizaje	25	0
La autoevaluación realizada durante su curso es útil para saber cuándo debían implementar ciertas estrategias para promover tu aprendizaje	25	0
La autoevaluación realizada durante su curso es útil para reflexionar y valorar sus habilidades como aprendiz	24	1
La autoevaluación realizada durante su curso es útil para monitorear y evaluar el logro de los objetivos de aprendizaje	24	1
La autoevaluación realizada durante su curso es útil para valorar su motivación	24	1
La autoevaluación realizada durante su curso es útil para saber la razón por la cual debían implementar ciertas estrategias para promover su aprendizaje	23	2
La autoevaluación realizada durante su curso es útil para monitorear sus emociones con respecto a la asignatura	22	3

En segundo lugar, se presentan los resultados del análisis cualitativo, donde a partir de codificar categorías esperadas y no esperadas, se realiza un análisis del discurso. Por un lado, las categorías no esperadas fueron la identificación de la motivación y su importancia, valoración de la importancia de los contenidos, identificación de la importancia de las emociones, conocimiento sobre elementos de la autoevaluación, valoración

de la importancia de la autoevaluación y la identificación y solución de errores. Por otro lado, las categorías esperadas con base en la literatura sobre la metacognición (Brown, 1987) fueron, identificar habilidades como aprendiz, reconocimiento de estrategias de aprendizaje, identificación de elementos para la resolución de problemas, establecimiento y monitoreo de metas y objetivos y monitoreo del aprendizaje.

Figura 1
Categorías encontradas en el análisis del discurso



Las categorías esperadas que fueron identificadas con mayor frecuencia fueron; a) el monitoreo de aprendizaje, b) la identificación de habilidades como aprendiz y c) el establecimiento y monitoreo de metas y objetivos. En cuanto a las categorías no esperadas encontradas con mayor frecuencia son; a) la identificación de la motivación y su importancia, b) valoración de la importancia de los contenidos y c) la identificación de la importancia de las emociones.

Sobre el monitoreo del aprendizaje, se identificó que los estudiantes a partir de retroalimentaciones que se hacían a ellos mismos o solo reflexionando sobre su desempeño, lograron realizar este seguimiento de los aprendizajes que estaban obteniendo. Un ejemplo de esto son comentarios como: “aprendí que debo llevar un balance y tener retroalimentaciones constantes para monitorear mi avance como estudiante” y “Es importante comparar nuestro grado de satisfacción entre el antes y después de llevar a cabo un proyecto porque de esa forma podemos identificar áreas de oportunidad y dónde reforzar esos conocimientos”.

También, se visualizó que, a partir de las autoevaluaciones los estudiantes identifican y valoran

tanto las habilidades que poseen, como aquellas que necesitan desarrollar o que no tienen como aprendices, un ejemplo sobre esto son comentarios como “Necesito tener más creatividad al momento de estudiar para así sea más fácil” y “creo que es importante hacer una autoevaluación al principio para que los alumnos sepan sus aptitudes y destrezas en general”.

Asimismo, se encontró que la autoevaluación fue útil para hacer a los estudiantes recordar, valorar o replantearse las metas y objetivos que habían determinado que querían lograr en el curso. Lo anterior se demuestra con los siguientes comentarios: “Realmente tengo que establecer mis metas, fijarme en ellas y no dejar que lo negativo factores externos a la materia afecten por completo mi rendimiento y estabilidad emocional” y “Si, la más importante es que pese a las diversas situaciones que se presenten debo tener presente los objetivos planteados para así poder lograrlo.”

En cuanto a las emociones y motivación, ambos resultaron ser elementos significativos sobre los cuales se cuestionaron y analizaron a partir de la autoevaluación. Lo anterior se muestra en comentarios como “me pareció muy bien que nos enfocáramos también en el aspecto emocional ya

Figura 2
Ejemplo de respuestas sobre el monitoreo del aprendizaje

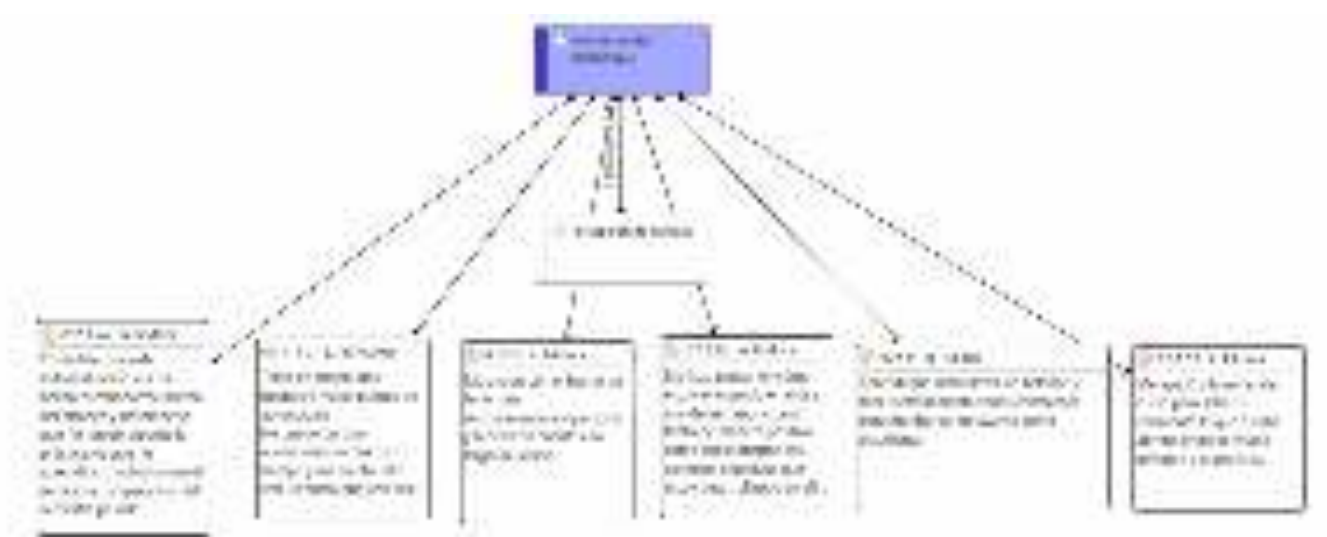
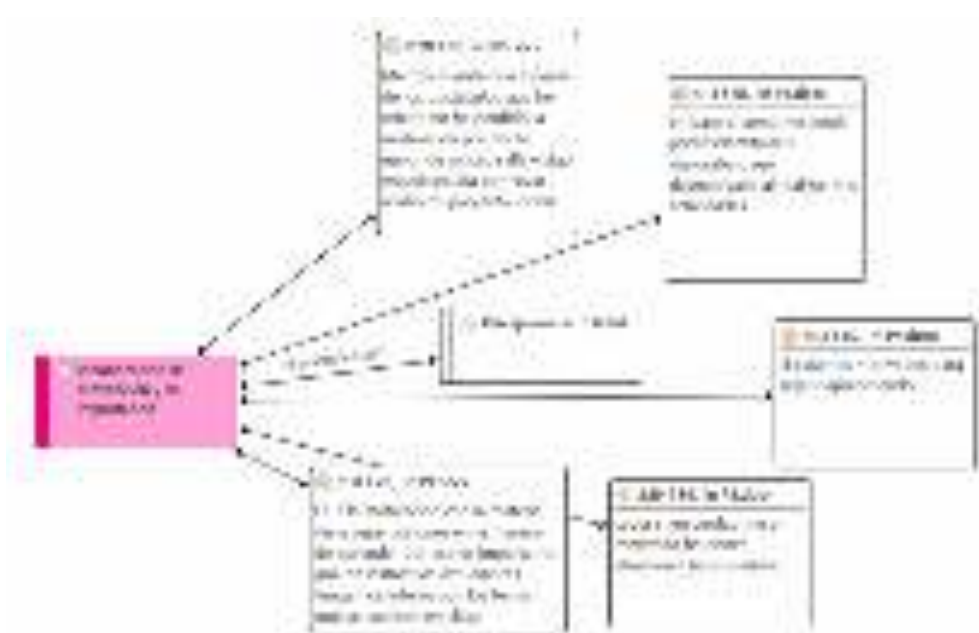


Figura 3
Ejemplo de respuestas sobre la identificación de la motivación y su importancia



que muchas veces ese ámbito es el que nos ayuda o nos retrasa al avanzar” y “que la motivación con la materia tiene repercusiones en la manera de aprender por eso es importante que nosotros los estudiantes tengamos interés con los temas que se explican en clase”.

Además, a partir de las autoevaluaciones los estudiantes reconocieron la importancia de los contenidos que estaban aprendiendo ahora como aspectos importantes que impactarían a futuro en su vida académica y laboral. Es importante subrayar que la importancia de los contenidos no se encontró solo en una, sino en las tres carreras del estudio. Un ejemplo de cómo identifican esta importancia es con comentarios como “esta materia es valiosa tanto para mi desarrollo laboral como para lo que quiero hacer en mi vida” y “que la materia es importante para poder desempeñarme lo mejor posible en mi trabajo”.

Discusión

A partir de los resultados obtenidos de la investigación, las principales conclusiones son las siguientes:

- Un área de oportunidad por explorar en la autoevaluación es que los estudiantes comparen su grado de aprendizaje y logro de objetivos con el de otros compañeros, ya que más allá de competir, unos podrían ser tutores de otros.
- Para llevar una autoevaluación más eficaz los estudiantes deben fortalecer la confianza en sí mismos, ya que una de las expectativas más altas encontradas al realizar un proceso de autoevaluación, es que el docente comente en conjunto con ellos los errores y aciertos que se tuvieron para confirmar con un tercero que en realidad han aprendido.
- Es necesario que los docentes que integren la autoevaluación en sus procesos incorporen también actividades o acciones posteriores a la autoevaluación, ya que sin este plan se reduce la autovaloración a ser un medio para identificar aquello que percibo que aprendí,

pero sin herramientas para lograr o complementar aquellos objetivos o aprendizajes que quedaron pendientes.

- En general, la autoevaluación es útil para; a) establecer, monitorear y evaluar objetivos de aprendizaje, b) valorar las habilidades como aprendizajes, c) elegir y saber cuándo y por qué implementar estrategias de aprendizaje, d) valorar la motivación y monitorear las emociones, e) valorar la importancia de los contenidos para la vida profesional y laboral, y f) para identificar errores que se han cometido y sus posibles soluciones. También, el momento más útil para realizar una autoevaluación es aproximadamente a la mitad del curso o al final.
- Para evitar expectativas erróneas sobre la autoevaluación, es necesario esclarecer desde el inicio del curso cuál es el objetivo, intención e instrumentos a utilizar.

Asimismo, para potenciar la utilidad que los estudiantes perciben de la autoevaluación y aumentar el uso de esta forma de evaluación en la educación superior, se proponen las siguientes intervenciones; a) planear, diseñar y desarrollar actividades que sensibilicen a los estudiantes sobre la importancia de la autoevaluación, b) identificar la cantidad y secuencia idónea para implementar instrumentos de autoevaluación, c) incorporar en cada instrumento de autoevaluación elementos sobre la metacognición, d) dar a conocer a docentes, directivos y administrativos la utilidad que tiene para toda la comunidad escolar llevar a cabo procesos de autoevaluación, y e) realizar protocolos de autoevaluación a partir de grupos colegiados.

En suma, la autoevaluación ha ido ganando terreno durante los últimos años, un factor que influyó fue la pandemia, ya que al ser un hecho mundial que transformó el contexto educativo, hizo necesario que las instituciones se replantearan los mecanismos de evaluación que se han usado por años. También es posible que, como las autoras, otros docentes hayan reflexionado sobre la abundante carga horaria que conlleva una evaluación tradicional y la significatividad real que este tipo de evaluación tiene para los estudiantes.

Referencias

- Boud, D. (1995). *Enhancing Learning through Self-Assessment*. Kogan Page.
- Boud, D. y Brew, A. (1995). Developing a typology for learner self – assessment practices. *Research and development in Higher Education*, 18, 130-135. https://www.researchgate.net/profile/David-Boud/publication/327350493_Developing_a_typology_for_learner_self-assessment_practices/links/5b899f13a6fdcc5f8b75b910/Developing-a-typology-for-learner-self-assessment-practices.pdf
- Boud, D. y Falchikov, N. (1989). Quantitative studies of student self-assessment in higher education: A critical analysis of findings. *Higher Education*, 18 (5), 529-549. <https://doi.org/10.1007/BF00138746>
- Brown, A. (1987). Metacognition, executive control, self-regulation and other more mysterious mechanisms. En W. Franz y K. Rainer (Eds.), *Metacognition, motivation and understanding* (p.p. 65-116). Hillsdale.
- Brown, G., y Harris, L. (2013). Student self-assessment. En J. McMillan (Ed.), *The SAGE handbook of research on classroom assessment* (pp. 367-393). SAGE.
- Copado, A. (2022). Evaluación del aprendizaje. Un siglo de educación a distancia en México. *Revista Innova Educación*, 4(4), 7-19. <https://revistainnovaeducacion.com/index.php/rie/article/view/614/642>
- Díaz Barriga, A. (2001). *El examen: textos para su historia y debate*. Universidad Nacional Autónoma de México, Centro de Estudios sobre la Universidad/Plaza y Valdés.
- Fernández, A. y Vanga, M. (2015). Proceso de autoevaluación, coevaluación y heteroevaluación para caracterizar el comportamiento estudiantil y mejorar su desempeño. *Revista San Gregorio*, 1(9), 6-15. <https://dialnet.unirioja.es/servlet/articulo?codigo=5225628>
- Flavell, J. (1976). Metacognitive aspects of problem solving. En L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231-236). Lawrence Erlbaum.
- Flavell, J. (1979). Metacognition and cognitive monitoring: a new area of cognitive-developmental inquiry. *American Psychologist*, 10(34), 906-911. <https://doi.org/10.1037/0003-066X.34.10.906>
- Frade, L. (2009). *Desarrollo de competencias en educación: desde preescolar hasta bachillerato*. Inteligencia educativa.
- Gallardo, K. (2013). *Evaluación del aprendizaje: retos y mejores prácticas*. Editorial Digital del Tecnológico de Monterrey.
- García, C. y Hervás, M. (2020). Los sistemas de evaluación de la Educación Superior en México y España. Un estudio comparativo. *Revista de la Educación Superior*, 49(194), 115-136. <https://www.scielo.org.mx/pdf/resu/v49n194/O185-2760-resu-49-194-115.pdf>
- Goodrich, H. W. (1996). *Student self-assessment: At the intersection of metacognition and authentic assessment* (Publicación No. 9638747) [Disertación doctoral, Harvard University]. ProQuest Information y Learning.
- Jiménez, G. (2006). Obtención de notas individuales a partir de una nota de grupo mediante una evaluación cooperativa. *Revista Iberoamericana de Educación*, 38(5), 1-14. <https://rieoei.org/RIE/article/view/2624>
- Joint Committee on Standards for Educational Evaluation (JCSEE). (2014). *The Student Evaluation Standards*. <http://www.jcsee.org/ses>
- Kambourova, M., González-Agudelo, E. y Grisales-Franco, L. (2021). La autoevaluación del estudiante universitario: revisión de la literatura. *Teoría de la Educación. Revista Interuniversitaria*, 33(2), 217-264. <https://doi.org/10.14201/teri.23672>
- Kim, B., Zyromski, B., Mariani, M., Lee, S. y Carey, J. (2017). Establishing the Factor Structure of the 18-Item Version of the Junior Metacognitive Awareness Inventory. *Measurement and Evaluation in Counseling and Development*, 50(1-2), 48-57. https://www.researchgate.net/publication/321200427_Establishing_the_Factor_Structure_of_the_18-Item_Version_of_the_Junior_Metacognitive_Awareness_Inventory
- Ley General de Educación [L.G.E.], Diario Oficial de la Federación [D.O.F.], 30 de septiembre de 2019 (México).
- Ley General de Educación Superior [L.G.E.S.], Diario Oficial de la Federación [D.O.F.], 20 de abril de 2021 (México).
- Martínez Rizo, F. (2009). Evaluación formativa en aula y evaluación a gran escala: hacia un sistema más equilibrado. *Revista Electrónica de Investigación Educativa*, 11(2), 1-18. <http://redie.uabc.mx/redie/article/view/231>
- Martínez Rizo, F. (2013). Dificultades para implementar la evaluación formativa: Revisión de literatura. *Perfiles Educativos*, 35(139), 128-150. <http://www.scielo.org.mx/pdf/peredu/v35n139/v35n139a9.pdf>
- Miller, M., Linn, R. y Gronlund, N. (2012). *Measurement and Assessment in Teaching* (11va ed.). Pearson.

- Muskin, J. (2015). *Evaluación del aprendizaje del estudiante y el currículo: Problemas y consecuencias para la política, el diseño y la aplicación*. OIE-UNESCO. <https://www.ibe.unesco.org/es/documento/evaluacion-C3%B3n-del-aprendizaje-del-estudiante-y-el-curriculo-C3%ADculo-problemas-y-consecuencias-para-la>
- Panadero, E., Brown, G., y Strijbos, J. (2016). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review*, 28(4), 803-830. doi:10.1007/s10648-015-9350-2
- Panadero, E. y Tapia, A. (2013). Revisión sobre autoevaluación educativa: evidencia empírica de su implementación a través de la autocalificación sin criterios de evaluación, rúbricas y guiones. *Revista de investigación en educación*, 2(11), 172-197. <https://dialnet.unirioja.es/servlet/articulo?codigo=4734937>
- Paredes, E. (2016, 2 de septiembre - agosto de 2017). *De la evaluación diagnóstica a la evaluación sumativa: logros y fracasos en los aprendizajes* [Presentación de escrito]. Debates en Evaluación y Currículo Congreso Internacional de Educación Evaluación Año 2 No. 2, México. <https://centrodeinvestigacioneducativauatx.org/pdf2016/A208.pdf>
- Plano-Clark, V. e Ivankova, N. (2016). *Mixed Methods Research: A Guide to the Field*. SAGE Publications, Inc. <https://doi.org/10.4135/9781483398341>
- Ramírez-Montoya, M. y Lugo-Ocando, J. (2020). Revisión sistemática de métodos mixtos en el marco de la innovación educativa. *Revista Científica de Educación*, 28(65), 9-20. <https://doi.org/10.3916/C65-2020-01>
- Rosales, C. (2003). *Criterios para una evaluación formativa. Objetivos, contenidos, profesor, aprendizaje y recursos*. Narcea.
- Sánchez, M. (2020). Evaluación del aprendizaje. En M. Sánchez y A. Martínez (Eds.), *Evaluación del y para el aprendizaje: instrumentos y estrategias* (pp. 17-40). Coordinación de Desarrollo Educativo e Innovación Curricular UNAM. https://cuaieed.unam.mx/descargas/investigacion/Evaluacion_del_y_para_el_aprendizaje.pdf
- Schraw, G. y Dennison, R. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 4(19), 460-475.
- Scriven, M. (1967). The methodology of evaluation. En R. W. Tyler, R. M. Gagne, y M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39-83). Rand McNally.
- Tan, K. (2001). *Conceptions of student self-assessment* [Disertación doctoral inédita]. University of Technology.
- Taras, M. (2014). *Student-centred learning and assessment: fact or fiction* [Manuscrito inédito].
- Taras, M. (2015). Autoevaluación del estudiante: ¿Qué hemos aprendido y cuáles son los desafíos?. *RELIEVE*, 21(1), 1-16. DOI: 10.7203/relieve.21.1.6394
- Topping, K. (1998). Peer Assessment between Students in Colleges and Universities. *Review of Educational Research*, 68(3), 249-276. <https://www.jstor.org/stable/1170598>
- Torres, M. y Torres, C. (2005). Formas de Participación en la Evaluación. *Revista Venezolana de Educación*, 9(31), 487-496. <https://www.redalyc.org/pdf/356/35603109.pdf>
- Valenzuela, A. (2019). ¿Qué hay de nuevo en la metacognición? Revisión del concepto, sus componentes y términos afines. *Educ. Pesqui*, 45, <http://dx.doi.org/10.1590/S1678-4634201945187571>

Invarianza factorial del Maslach Burnout Inventory Student Survey (MBI-SS) en estudiantes universitarios chilenos y españoles

María Pía Valdés Castro¹; Kevin Rodrigo Mancilla Mancilla²; Javiera Ignacia Morales Olivares³; Gabriela Belén Acevedo Meza⁴; Ricardo Jorquera Gutiérrez ^{*5}

^{1,2,3,4,5} Universidad de Atacama, Chile; ¹<https://orcid.org/0000-0001-5395-8313> mvaldes097@gmail.com ²<https://orcid.org/0000-0002-1801-019X> kevinmancilla7@hotmail.com ³<https://orcid.org/0000-0001-9009-1943> javieram.olivares@hotmail.com ⁴<https://orcid.org/0000-0002-9347-156X> Gabiacevedo1517@gmail.com ⁵<https://orcid.org/0000-0002-7059-8488> ricardo.jorquera@uda.cl

Citar como: Valdés Castro, M., Mancilla Mancilla, K., Morales Olivares, J., Acevedo Meza, G., & Jorquera Gutiérrez, R. (2023). Invarianza factorial del Maslach Burnout Inventory Student Survey (MBI-SS) en estudiantes universitarios chilenos y españoles. *Revista Digital de Investigación en Docencia Universitaria*, 17(2), e1694. <https://doi.org/10.19083/ridu.2023.1694>

Recibido: 31/08/2022. **Revisado:** 25/09/2022. **Aceptado:** 10/06/2023. **Publicado:** 30/06/2023.

Resumen

Introducción: el Maslach Burnout Inventory-Student Survey (MBI SS) es uno de los instrumentos más reconocidos en todo el mundo para evaluar el burnout en población estudiantil. **Objetivo.** estimar las propiedades psicométricas del Maslach Burnout Inventory Student Survey (MBI-SS) en estudiantes universitarios chilenos y españoles. **Método:** se efectuó un estudio instrumental en una muestra no probabilística por cuotas en 103 estudiantes universitarios chilenos y 104 estudiantes españoles. Se realizó un análisis factorial confirmatorio, se analizó la confiabilidad del instrumento y su invarianza factorial para la muestra española y chilena. **Los resultados** confirman su estructura factorial de tres factores y muestran que el instrumento presenta una adecuada confiabilidad y una invarianza factorial en su versión propuesta de 12 ítems. **Discusión:** se concluye que este instrumento y sus propiedades psicométricas permiten evaluar el burnout académico en estudiantes chilenos y españoles.

Palabras clave: burnout; MBI SS; Maslach Burnout Inventory Student Survey; estudiantes universitarios.

Factorial Invariance of the Maslach Burnout Inventory Student Survey (MBI-SS) in Chilean and Spanish University Students

Abstract

Introduction: The Maslach Burnout Inventory-Student Survey (MBI SS) is one of the most recognized instruments worldwide to assess burnout in the student population. **Objective:** To estimate the psychometric properties of the Maslach Burnout Inventory Student Survey (MBI-SS) in Chilean and Spanish university students. **Method:** An instrumental study was carried out in a non-probabilistic sample by quotas in 103 Chilean university students and 104 Spanish students. A confirmatory factorial analysis was performed, the reliability of the instrument and its factorial invariance for the Spanish and Chilean samples were analyzed. **Results:** The results confirm its three-factor factor structure and show that the instrument has adequate reliability and factor invariance in its proposed 12-item version. **Discussion:** It is concluded that this instrument and its psychometric properties allow the evaluation of academic burnout in Chilean and Spanish students.

*Correspondencia:

Ricardo Jorquera Gutiérrez
ricardo.jorquera@uda.cl



Keywords: burnout; MBI SS; Maslach Burnout Inventory Student Survey; university students.

Introducción

El estrés es considerado el primer paso dentro de un proceso paulatino conocido como *Burnout*, que comienza en el momento que las demandas personales sobrepasan los recursos materiales y humanos que posea un individuo ([Álvarez y Fernández, 1991](#)). [Freudenberger \(1980\)](#) lo describió como un estado de fatiga o frustración dada la devoción a una causa, modo de vida o relación que no produce la recompensa deseada. [Pines, Aaronson y Kafry \(1981\)](#) lo asemejaron a un estado de agotamiento físico, emocional y mental causado por estar implicada la persona durante largos periodos de tiempo en situaciones que le afectan emocionalmente.

[Maslach y Jackson \(1981\)](#) poco después lo definieron como respuesta al estrés laboral crónico que conlleva la vivencia de encontrarse emocionalmente agotado, el desarrollo de actividades y sentimientos negativos hacia las personas con las que se trabaja (actividades de despersonalización), y la aparición de procesos de devaluación del propio rol profesional. Entienden que el *burnout* se configura como un síndrome tridimensional caracterizado por: a) agotamiento emocional, b) despersonalización y c) reducida realización personal.

Una de las etapas de la investigación del *burnout* está caracterizada por la ampliación de la población afectada por el síndrome, hasta considerar a los estudiantes en sus procesos académicos. Es aquí donde es posible hablar propiamente del síndrome del *burnout* académico. Vale la pena distinguir en esta línea dos vertientes: la primera de ellas, que corresponde al estudio del síndrome en alumnos universitarios de carreras de medicina y enfermería y, la segunda, que corresponde a estudiantes de diferentes carreras ([Caballero, et al., 2010](#)). La segunda vertiente mencionada es la generalización del síndrome a todo el ámbito académico, y surge ante la presunción de que los estudiantes universitarios, así como cualquier profesional, se encuentran con presiones y sobre-

cargas propias de la labor académica ([Garcés de los Fayos, 1995](#)).

[Bresó \(2008\)](#) señala que los estudiantes, y cualquier trabajador, mantienen una relación de compensación directa e indirecta con la institución universitaria, evidenciada en apoyos económicos, becas, reconocimientos o premios. Esta presunción permite investigar las respuestas del individuo ante la tensión y sus implicaciones en su sensación de bienestar frente a los estudios ([Bresó, 2008](#); [Caballero, Abello y Palacio, 2007](#)).

En cuanto a los estudios realizados en *burnout* académico, el enfoque tridimensional es el más predominante. En este enfoque, el agotamiento emocional se refiere al sentimiento de haber agotado los recursos emocionales propios, y es considerado como el componente de estrés individual básico del *burnout*. La despersonalización se refiere a las respuestas negativas, cínicas o excesivamente desinteresadas hacia otras personas en el trabajo, este es el componente interpersonal del síndrome. Y finalmente, la baja realización personal que alude a sentimientos de disminución de la competencia de uno mismo, así como la productividad, y representa el componente de autoevaluación del *burnout* ([Moneta, 2011](#)).

Las variables asociadas al síndrome son muy diversas y, en el ámbito académico se han estudiado los factores que pueden restringir o facilitar el desempeño académico del estudiante. [Sala-nova, et al. \(2005\)](#) diferencian entre las variables consideradas “obstáculos” (factores del contexto académico que dificultan el desempeño del estudiante), y los llamados “facilitadores” (factores del contexto académico que posibilitan el buen funcionamiento y ayudan a mitigar los obstáculos). Entre las primeras se destacan las dificultades en el servicio de reprografía, el realizar o dejar demasiadas tareas, los horarios de clase muy cargados y muchos créditos por programa. Entre los facilitadores anotan el adecuado servicio de biblioteca, el compañerismo, el apoyo social de la familia y los amigos, así como recibir becas de estudio.

La operacionalización del *burnout* académico ha sido posible a partir de la estandarización del MBI-GS en estudiantes universitarios, lo cual dio origen al cuestionario de Burnout Inventory-Student Survey (MBI-SS) de [Schaufeli et al. \(2002\)](#).

El Inventario de Burnout de Maslach (MBI) se ha constituido en el *gold standard* para la medición del *burnout*, puesto que es el instrumento más utilizado con estos fines y ha sido validado en múltiples contextos culturales y laborales (Dyrbye et al; 2009; Aguayo et al; 2011; Gil-Monte y Olivares, 2011). Para entender el impacto de este instrumento sobre la investigación en *burnout*, hay que mencionar que la aparición del MBI como técnica e instrumento de evaluación probablemente constituye el hecho diferencial que marca el paso de un período inicial de formulación a otro riguroso y sistemático (Hernández et al., 2005).

El MBI se ha aceptado extensamente en todos los países de América Latina, UE y en los EE.UU. Esto es una ventaja porque permite comparar resultados y desarrollar estrategias de prevención y tratamiento del trastorno, a la vez que impulsa el desarrollo de adaptaciones del cuestionario (Shirom, 2003). Sin embargo, el hecho de que el MBI sea el instrumento utilizado con mayor frecuencia para medir el *burnout* no implica necesariamente que su validez y confiabilidad haya sido consolidada (Grajales, 2001).

El inventario que se examinará es el Maslach Burnout Inventory - Student Survey (MBI-SS), de Schaufeli, et al. (2002), este es un cuestionario para la evaluación del síndrome de *burnout* académico de aplicación colectiva y diligenciamiento individual. Evalúa la sensación de no poder dar más de sí mismo, tanto física como psíquicamente (agotamiento), la presencia de una actitud negativa de desvalorización y pérdida del interés por el estudio (cinismo) y la existencia de dudas acerca de la propia capacidad para realizar el trabajo académico (autoeficacia académica) (Bresó, 2008). Todos los ítems de cada una de estas tres subescalas se puntúan en una escala de frecuencia de siete puntos, que va desde 0 (nunca) a 6 (siempre). Cinco ítems evalúan agotamiento, cuatro evalúan cinismo y seis ítems evalúan autoeficacia académica, para un total de 15 reactivos.

Es importante señalar que a pesar de que este instrumento originalmente se diseñó en inglés ya se ha traducido al español en muchas ocasiones para su aplicación en poblaciones latinoamericanas y españolas (Aguilar-Bustamante et al., 2013; Boada et al., 2015; Merino et al., 2013; Moyano y

Riaño-Hernández, 2013; citado en Hederich-Martínez y Caballero-Domínguez, 2016).

En México, Banda, Robles y Lussier (2021), encontraron alfas de Cronbach de .856 para Agotamiento Emocional, .851 para Cinismo y .717 para Eficacia Académica, realizando además un análisis factorial exploratorio que evidenció sus tres dimensiones. Resultados consistentes con esto obtuvieron Ornelas, Jurado, Blanco, Peinado y Blanco (2020) quienes evidenciaron una estructura de tres factores (agotamiento, cinismo y eficacia) con adecuados indicadores de ajuste de fiabilidad y validez. Sus indicadores de ajustes mostraron resultados aceptables para una submuestra 1 (GFI = .904; RMSEA = .082; CFI = .904) y una submuestra 2 (GFI = .916; RMSEA = .077; CFI = .916), pero que mejora eliminando siete reactivos, con menor carga factorial, encontrando indicadores de ajuste adecuados en la primera submuestra (GFI = .984; RMSEA = .039; CFI = .987) y también en la segunda (GFI = .982; RMSEA = .043; CFI = .985). Los indicadores de confiabilidad omega son de .814 en agotamiento, .741 en cinismo y .776 en eficacia en la primera submuestra y de .802 en agotamiento, .711 en cinismo y .818 en eficacia en la segunda submuestra. En España (Pérez-Fuentes, Molero Jurado, Simón Márquez, Oropesa Ruiz, & Gázquez Linares, 2020), en una muestra de 1,209 estudiantes de la comunidad autónoma de Andalucía mediante una factorial exploratorio y confirmatorio, los autores encontraron que una versión de 12 ítems del instrumento presentaba los mejores niveles de ajuste, de acuerdo con un modelo de tres factores ($\chi^2/df = 3.155$, CFI = .966, TLI = .953, GFI = .961, RMR = .167, RMSEA = .061). En relación con la confiabilidad sus resultados muestran para el factor agotamiento emocional = .83, para el factor cinismo = .82 y para el factor eficacia académica = .79. Un reciente estudio con estudiantes universitarios españoles (Aguayo-Estremera, Cañadas, Ortega-Campos, Ariza & De la Fuente-Solana, 2023) concluye que la estructura interna del MBI-SS queda bien reflejada por el modelo oblicuo congenérico de tres factores, alcanzando buenos valores de fiabilidad y validez convergente y discriminante.

Mientras que en Chile refiriendo al área estudiantil, Pérez et al. (2012) dicen que no existe un

instrumento validado para evaluar el síndrome en universitarios chilenos, es por esto por lo que evaluó la estructura factorial y la confiabilidad de una versión del MBI adaptada a estudiantes. En este estudio, mediante Análisis Factorial Exploratorio, se identificaron tres factores próximos a los propuestos por Maslach, los que mostraron –junto con la escala general– confiabilidades adecuadas. Estos resultados indican que el instrumento tiene una apropiada estructura factorial y consistencia interna, siendo útil para evaluar el desgaste emocional, particularmente interpersonal, en población universitaria. El estudio coincidió con la estructura trifactorial del *burnout* propuesta originalmente por [Maslach y Jackson \(1981\)](#), evidenciando concordancia entre el comportamiento empírico del instrumento y el modelo teórico que lo fundamenta. Sin embargo, los autores confirman que la distribución de los ítems no es exactamente igual a la propuesta originalmente. Donde factores como los ítems 6 (“Estar todo el día con otras personas es un esfuerzo”) y 16 (“Trabajar directamente con personas me produce estrés”) –ambos factores del agotamiento emocional– fueron asignados al factor Cinismo o Despersonalización al presentar mayores cargas factoriales. Por otro lado, los ítems muestran capacidades discriminativas adecuadas de correlación con el resto de los ítems de la escala superiores a $r = .30$, lo que demuestra que presentan una relación suficiente con el factor al que pertenecen.

En esta línea, la presente versión modificada del MBI para estudiantes universitarios, ha mostrado tener una estructura factorial definida y confiabilidades adecuadas que permiten su utilización para estos fines en población universitaria chilena ([Perez et al., 2012](#)).

Es importante mencionar que el MBI como instrumento de evaluación para *burnout* tiene una estructura sistémica que beneficia su aplicación debido a que constituye el paso inicial para formar otros instrumentos rigurosos y sistemáticos. ([Hernández et al., 2005](#)).

La mayoría de las investigaciones coinciden en destacar que el *burnout* correlaciona negativamente la satisfacción frente a los estudios ([Caballero, et al., 2007](#)), la intención de abandonar los estudios y la felicidad frente a los estudios ([Sala-](#)

[nova, et al., 2005](#)) y las menores expectativas de éxito en los mismos ([Martínez y Marques-Pinto, 2005](#)). Estos datos muestran que, si bien es poco clara aún la consecuencia del síndrome en el rendimiento académico, aquel sí tiene un valor predictivo importante en el deterioro de diversas variables relevantes relacionadas con el desempeño académico y con la satisfacción en los estudios, así como con la prevención del abandono de estos por parte del estudiante universitario ([Caballero, Bresó y González, 2015](#)).

[Faye Dumanget \(2017\)](#) analizó este instrumento en una muestra de estudiantes universitarios de diversas carreras. Utilizó un análisis factorial confirmatorio para probar la idoneidad del modelo, evidenciando que la versión francesa del modelo MBI-SS tiene un ajuste aceptable. De igual forma, [Mostert et. al \(2007\)](#) evaluaron el instrumento en estudiantes de habla afrikaans y setswana. Este estudio reveló, mediante un análisis factorial confirmatorio, un buen ajuste para los modelos de referencia de dos factores, apoyando la tesis de que la estructura de dos factores del MBI-SS es equivalente para ambos idiomas.

El objetivo general de este trabajo es describir las propiedades psicométricas del instrumento MBI-SS en estudiantes universitarios chilenos y españoles. En concreto se busca: a) estimar la estructura factorial del instrumento MBI-SS en estudiantes universitarios chilenos y españoles, b) estimar la confiabilidad del instrumento de evaluación aplicado en estudiantes universitarios chilenos y españoles y c) estimar la invarianza factorial del instrumento de evaluación aplicado en estudiantes universitarios chilenos y españoles.

Respecto de lo anterior, se busca proporcionar información del concepto *burnout* frente a futuras investigaciones sobre el instrumento Maslach Burnout Inventory Student Survey (MBI-SS) a través de comprobar la estructura factorial del instrumento y la equivalencia psicométrica del mismo en grupos de estudiantes españoles y chilenos; ya que en el contexto de la comparación intergrupar es indispensable plantearse la necesidad de llevar a cabo la adaptación de un instrumento de medida psicológica que cumpla con todos los criterios de equivalencia, pero so-

bre todo plantearse si la misma estructura factorial es aplicable a distintos grupos de sujetos o, de modo más genérico, a distintas poblaciones ([Abalo et al., 2006](#)).

Método

Diseño

La metodología empleada en esta investigación es de carácter cuantitativo con un diseño transversal. Además, este estudio se define como instrumental ya que “en esta categoría se incluyen todos aquellos trabajos que analizan las propiedades psicométricas de instrumentos de medida psicológicos” ([Ato, López-García & Benavente, 2013, p. 5](#)). También el objetivo de las investigaciones instrumentales es proponer la creación de un nuevo instrumento o la adaptación de una herramienta ya existente a un nuevo contexto de aplicación pertinente al momento actual ([Argumedo et al., 2016](#)).

Participantes

El muestreo es de tipo no probabilístico y por cuotas. La muestra estuvo conformada por 207 estudiantes, de los cuales 103 pertenecían a una universidad pública del norte de Chile (49.8%) y 104 a una universidad pública española ubicada en Cataluña (50.2%). Según sexo 127 (61.4%) eran mujeres y 80 (38.6%) hombres. Sus edades fluctuaban entre los 18 y los 45 años ($M = 22.37$; $DS = 4.13$).

Los criterios de inclusión que se consideraron para los participantes fueron el cumplimiento de la mayoría de edad establecida por su país de residencia, el contar con la matrícula vigente en diferentes semestres o grados académicos, contemplados desde primer año hasta sexto año, dentro de una carrera de nivel universitario en las universidades consideradas en este estudio. Mientras que en los criterios de exclusión fueron ser estudiantes de intercambio o encontrarse circunstancialmente en la universidad.

Instrumentos

El instrumento que se utilizó en este estudio fue el Maslach Burnout Inventory Student Survey (MBI-SS) el cual es una versión modificada a estudian-

tes del Maslach Burnout Inventory General Survey (MBI-GS) ([Schaufeli et al., 1996](#)). Para esto se modificó y adaptó la actividad académica en el rol del estudiante universitario como el equivalente al rol de un trabajador formal. Inicialmente el instrumento tenía 16 preguntas, pero se adaptó a 15 preguntas luego del proceso de validación de [Schaufeli et al., \(2002\)](#).

El cuestionario que mide el Síndrome de Burnout Académico en estudiantes evalúa tres dimensiones: Agotamiento, Cinismo y Eficacia.

En este instrumento el estudiante debe responder a través de la frecuencia en que ha experimentado las descripciones de los ítems y dimensiones mencionados, para esto debe usar una escala en formato Likert de seis alternativas (0 = Nunca; 1 = Pocas veces al año o menos; 2 = Una vez al mes o menos; 3 = Unas pocas veces al mes; 4 = Una vez a la semana; 5 = Pocas veces a la semana; 6 = Todos los días). Los puntajes altos en Agotamiento y Cinismo y las puntuaciones bajas en Eficacia son indicativos de *burnout*.

[Schaufeli](#) y colaboradores publicaron en el 2002 las características psicométricas del MBI-SS en 1,661 estudiantes universitarios de tres países: España, Portugal y Holanda. Los resultados mostraron, por análisis factorial, que tres factores se ajustaban perfectamente a los dominios de estudio (Agotamiento, Cinismo y Rendimiento académico), con buenos niveles de confiabilidad según el alfa de Cronbach (Agotamiento entre .74 y .80 dependiendo del país; Cinismo entre .79 a .86 y Rendimiento académico entre .67 y .76) ([Loayza-Castro et al., 2016](#)).

Procedimientos

Para acceder a los 207 participantes estipulados en este estudio, se realizó como primer paso un contacto con ambas universidades, con la finalidad de tener un acceso formal a la realización del cuestionario en los estudiantes. Posterior a este permiso, a los estudiantes que accedieron participar de la investigación se les otorgó un enlace con un cuestionario en la plataforma Google Form, el cual le mostró inicialmente un consentimiento informado, donde se dio conocimiento y se detallaron los criterios éticos de la investigación, los objetivos de esta y la metodología utilizada. Una vez que el participante aceptó su participa-

ción mediante el consentimiento informado, rellenando sus datos y garantizando que lo estaba realizando bajo su voluntad. Posterior a esto se le presentó el cuestionario Maslach Burnout Inventory-Student con 15 ítems.

El estudio consideró como guía los estándares éticos de la ANID (Agencia Nacional de Investigación y Desarrollo), el cual cuenta con leyes de protección a la integridad de un individuo dentro de Chile. En España, este estudio se apegó a los estándares éticos de CSIC (Consejo Superior de Investigaciones Científicas), que consta del cuidado de datos personales de personas que participan en investigaciones científicas. También se apegó a un formato universal con los Principios éticos de los psicólogos y código de conducta American Psychological Association (APA, 2017).

Análisis de datos

En primer lugar, se analizaron los ítems del MBI-SS a través de estadísticos descriptivos: media, desviación estándar, asimetría y curtosis. Se verificó su normalidad considerando valores de asimetría y curtosis entre -1 y 1 (Darlington y Hayes, 2017). La normalidad multivariante se corroboró mediante el coeficiente de Mardia (Esnaola et al., 2011).

El método de cálculo de la fiabilidad del instrumento se realizó mediante el coeficiente Omega de McDonald, dado que este indicador logra superar diversas falencias que tiene el Alpha de Cronbach (Ventura-León & Caycho-Rodríguez, 2017).

Para la evaluación de la validez del instrumento se usó el análisis factorial confirmatorio que por su marco estadístico permite evaluar y corroborar empíricamente la validez y la estructura conceptual del constructo (Herrero, 2010). Se usó un procedimiento basado en el método de Mínimos Cuadrados Ponderados Diagonalizados (DWLS Diagonal Weighted Least Squares), recomendado cuando el número de opciones de respuesta decrece, el tamaño de la muestra es pequeño, o cuando los datos observados no poseen una distribución normal (Rigo y Donolo, 2019).

En esta investigación se utiliza la invarianza factorial, que se conceptualiza como el proceso de verificación en que las propiedades de medida de los instrumentos o sus ítems son independientes de las características de los grupos evaluados,

pero no del constructo que el instrumento está midiendo (Byrne, 2008).

Los índices que se consideraron en el análisis factorial confirmatorio, como en la estimación de la invarianza factorial fueron: Chi2/gl, el índice de Tucker-Levis (TLI), el índice de bondad de ajuste comparativo (CFI). Se tomaron en cuenta las recomendaciones de Hu y Bentler (1999), quienes recomiendan valores superiores a 0,95 en CFI y TLI para considerar un ajuste adecuado. Así mismo, valores inferiores a 0,06 en RMSEA se estima como adecuado. Al comparar los modelos, se estima como mejor aquel que posee valores más altos en CFI y TLI y valores más bajos en RMSEA en comparación a los otros.

Se efectuó un análisis factorial confirmatorio multigrupo, la cual se entiende como una herramienta analítica que se puede utilizar para examinar la validez y la equivalencia/invarianza de las mediciones de las puntuaciones en diversos grupos. Es necesario señalar que este es un proceso progresivo en donde, en el caso de la presente investigación, se exploró en primer lugar la equivalencia de la estructura del MBISS en las muestras chilena y española (invarianza configural), luego se constató la igualdad estadística de las cargas factoriales (invarianza métrica), posteriormente de los thresholds o umbrales (invarianza fuerte), y de los residuales (invarianza estricta) (Pendergast, von der Embse, Kilgus & Eklund, 2017). Para lo anterior, se estimó las diferencias entre los modelos (configural, métrico, escalar y estricto), empleando las variaciones de los índices CFI (ΔCFI) y RMSEA ($\Delta RMSEA$), considerando valores $\Delta CFI \leq .01$ y $\Delta RMSEA \leq .015$ (Cheung & Rensvold, 2002).

Resultados

En primera instancia se efectuó el análisis descriptivo de los ítems. Al respecto, es posible señalar que las medias más altas se encontraron en los ítems "Con qué frecuencia te sientes entusiasmado cuando alcanzas tus metas estudiantiles" ($M = 4.21$; $DS = 1.22$), perteneciente a la dimensión Agotamiento, y los ítems "Con qué frecuencia has aprendido cosas interesantes a lo largo de tus estu-

dios" ($M = 4.79$; $DS = 1.11$) y "Con qué frecuencia te sientes agotado al final de un día en la universidad" ($M = 4.42$; $DS = 1.34$), pertenecientes a la dimensión Eficacia académica. En cuanto a las medias más bajas destacamos "Con qué frecuencia haces comentarios burlones sobre la utilidad de lo que estudias" ($M = 2.49$; $DS = 2.20$) y "Con qué frecuencia dudas sobre la importancia de tus estudios" ($M = 1.92$; $DS = 1.78$) ambas de la subescala de Cinismo.

Se analizaron los ítems de la escala a fin de establecer el comportamiento normativo de estos. Se consideró que valores de asimetría y curtosis entre -1 y 1 indican normalidad de la variable (Darlington y Hayes, 2017). Los reactivos presentaron valores menores a 1, a excepción del ítem "Con qué frecuencia te sientes entusiasmado cuando alcanzas tus metas estudiantiles", lo que concluiría en una variable anormal.

La normalidad multivariada fue evaluada mediante un análisis Mardia para la asimetría y cur-

tosis multivariada. Se evidenció un coeficiente de asimetría de 19.76 ($X^2 = 681,54$, $gl = 364$, $p < ,001$) y un coeficiente de curtosis de 185.98 ($z = 7.06$; $p < .001$), con lo que se constató la ausencia de una distribución normal multivariada de los datos. Dado lo anterior, y considerando también el sistema de respuesta tipo Likert del instrumento, se decidió desarrollar el análisis factorial confirmatorio a través del método de mínimos cuadrados ponderados diagonalizados (DWLS).

La confirmación de la estructura del instrumento se verificó por medio de análisis factorial confirmatorio (AFC). En este caso, se efectuó la estimación de bondad de ajuste del modelo mediante el método de mínimos cuadrados ponderados diagonalizados (DWLS). Los índices que se consideraron en el AFC fueron: χ^2 , χ^2/gl , el Índice de Bondad de Ajuste (GFI), el Índice de Tucker-Lewis (TLI), y el Índice de Bondad de Ajuste Comparativo (CFI).

Tabla 1
Estadísticos Descriptivos Ítems

Ítem	Media	Desviación estándar	Asimetría	Curtosis
Ítem 1	4.29	1.36	-.91	.26
Ítem 2	4.42	1.34	-.73	.13
Ítem 3	3.93	1.60	-.37	-.82
Ítem 4	3.24	1.55	-.19	-.42
Ítem 5	3.95	1.40	-.54	-.07
Ítem 6	2.88	1.62	.06	-.73
Ítem 7	3.56	1.58	-.17	-.81
Ítem 8	2.49	2.20	.26	-1.47
Ítem 9	1.92	1.78	.53	-.93
Ítem 10	4.35	1.10	-.36	-.14
Ítem 11	3.56	1.40	-.40	-.22
Ítem 12	3.77	1.47	-.42	-.33
Ítem 13	4.82	1.22	-1.18	1.29
Ítem 14	4.79	1.11	-.85	.70
Ítem 15	3.87	1.38	-.55	.09

El modelo de tres factores evidenció indicadores de ajustes adecuados para la muestra total ($X^2 = 159.30$, $gl = 87$, $p < .001$; $X^2/gl = 1.83$; $TLI = .959$; $CFI = .962$; $RMSEA = .064$) y para la muestra chilena y española por separado (Ver tabla 2). Sin embargo, se evidenciaron bajas cargas factoriales (inferiores a ,4) en algunos ítems de la muestra chilena y española (Ver tabla 3). En este caso se encontró el ítem 8 (“Con qué frecuencia haces comentarios burlones sobre la utilidad de lo que estudias”) en la muestra chilena y los ítems 11 y 13 (“Con qué frecuencia crees que puedes hacer una contribución efectiva en las clases que asistes”, “Con qué frecuencia te sientes entusiasmado cuando alcanzas tus metas estudiantiles”) en la muestra española.

Considerando lo planteado en el párrafo anterior, se realizó la verificación de un nuevo modelo con base en un instrumento de 12 ítems, evidenciándose una mejora en los indicadores de bondad de ajuste en la muestra total ($X^2 = 82.04$, $gl = 51$, $p < .01$; $X^2/gl = 1.609$; $TLI = .972$; $CFI = .979$; $RMSEA = .054$). Las cargas factoriales estandarizadas y confiabilidad de este modelo se puede apreciar en la Tabla 4.

También se consideró la confiabilidad de las dimensiones del instrumento en la muestra chilena y española a través del ω de McDonald, considerando valores de .70 o mayores para una fiabilidad aceptable (George y Mallery, 2003), apreciándose valores superiores a este rango en todas las dimensiones en la muestra chilena y española (ver Tabla 4).

Tabla 2
Modelos Estructurales Muestra Total, Muestra Chilena y Muestra Española

	X2	DF	p	CFI	TLI	RMSEA	SRMR
Tres Factores 15 ítems	159.298	87	.000	.962	.954	.064 [.048-.079]	.084
Tres Factores 12 ítems	82.038	51	.004	.979	.972	.054 [.031-.075]	.077
Tres Factores 15 ítems (Chile)	102.519	87	.122	.985	.982	.042 [.000-.071]	.103
Tres Factores 12 ítems (Chile)	59.411	51	.196	.989	.986	.040 [.000-.078]	.077
Tres Factores 15 ítems (España)	82.930	87	.604	1.000	1.005	.000 [.000-.048]	.082
Tres Factores 12 ítems (España)	51.792	51	.443	.999	.999	.012 [.000-.065]	.082

Tabla 3
Cargas Factoriales Estandarizadas y Confiabilidad del Instrumento Original en la Muestra Chilena y Española

	Chile			España		
	Agotamiento	Cinismo	Eficacia	Agotamiento	Cinismo	Eficacia
Ítem 1	.828			.808		
Ítem 2	.601			.580		
Ítem 3	.674			.760		
Ítem 4	.811			.774		
Ítem 5	.873			.830		
Ítem 6		.863			.744	
Ítem 7		.875			.864	
Ítem 8		.358			.642	
Ítem 9		.672			.562	
Ítem 10			.603			.511
Ítem 11			.724			.379
Ítem 12			.775			.801
Ítem 13			.540			.349
Ítem 14			.406			.522
Ítem 15			.757			.626
McDonald's ω	.869	.792	.831	.865	.783	.739
IC del 95%	[.828-.909]	[.727-.856]	[.781-.881]	[.823-.906]	[.715-.852]	[.664-.813]

Tabla 4
Cargas Factoriales Estandarizadas y Confiabilidad del Instrumento de 12 Ítems en la Muestra Chilena y Española

	Chile			España		
	Agotamiento	Cinismo	Eficacia	Agotamiento	Cinismo	Eficacia
Ítem 1	.839			.799		
Ítem 2	.605			.600		
Ítem 3	.671			.764		
Ítem 4	.799			.781		
Ítem 5	.878			.811		
Ítem 6		.837			.738	
Ítem 7		.884			.853	
Ítem 9		.624			.506	
Ítem 10			.571			.531
Ítem 12			.672			.718
Ítem 14			.449			.447
Ítem 15			.739			.622
McDonald's ω	.869	.838	.772	.865	.750	.726
IC del 95%	[.828-.909]	[.785-.891]	[.702-.841]	[.823-.906]	[.674-.826]	[.641-.811]

Invarianza factorial según país

Al efectuar el procedimiento de invarianza factorial en el instrumento de 15 ítems (Ver tabla 5), comparando los modelos respectivos, los resultados de la invarianza métrica y configural muestran diferencias significativas ($\Delta CFI = .018$ y $\Delta RMSEA = -.022$), por lo cual, no es posible constatar una equivalencia de las cargas factoriales de ambos instrumentos. Igual cosa ocurre al comparar el modelo métrico y escalar ($\Delta CFI = .016$ y $\Delta RMSEA = -.011$), lo que permite señalar que los interceptos sí son invariantes. Con lo anterior, no existiría evidencia para afirmar que el instrumento de 15 ítems sea invariante en ambas muestras.

Como se aprecia en la Tabla 6, los índices de ajuste orientados a determinar la invarianza factorial del instrumento de 12 ítems en la muestra de la universidad chilena y española evidencian valores que permiten reconocer su equivalencia en ambos grupos. Los ajustes del modelo con el cual se determinó la invarianza configural mostró valores adecuados ($TLI = .992$; $CFI = .994$; $RM-$

$SEA = .030$). Igual panorama se aprecia en los valores que representan el ajuste de la invarianza métrica ($TLI = .993$; $CFI = .994$; $RMSEA = .028$). Al comparar los resultados de la invarianza métrica y configural no se aprecian diferencias significativas ($\Delta CFI = .000$ y $\Delta RMSEA = .002$). Los resultados permiten establecer que las cargas factoriales son invariantes entre ambos grupos. El modelo con el cual se determinó la invarianza escalar también evidenció valores favorables ($TLI = .983$; $CFI = .985$; $RMSEA = .043$). Al comparar los resultados de la invarianza métrica y escalar no se constatan cambios importantes ($\Delta CFI = .009$ y $\Delta RMSEA = -.015$), lo que permite señalar que los interceptos son invariantes. Finalmente, el modelo con el cual se determinó la invarianza estricta también evidenció valores adecuados ($TLI = .987$; $CFI = .987$; $RMSEA = .037$). Las diferencias entre los valores de ajuste del modelo escalar y estricta fueron adecuadas ($\Delta CFI = -.002$ y $\Delta RMSEA = .006$), permitiendo entregar un soporte empírico a la invarianza estricta.

Tabla 5
Invarianza Factorial del Instrumento de 15 Ítems entre la Muestra Chilena y Española

	X ²	gl	p	CFI	TLI	RMSEA	ΔCFI	ΔRMSEA
Configural	185.449	171	.213	.993	.991	.029		
Métrico	232.148	183	.008	.975	.971	.051	.018	-.022
Escalar	277.102	198	.000	.959	.957	.062	.016	-.011
Estricto	296.005	213	.000	.957	.958	.062	.002	.000

Tabla 6
Invarianza Factorial del Instrumento de 12 Ítems entre la Muestra Chilena y Española

	X ²	gl	p	CFI	TLI	RMSEA	ΔCFI	ΔRMSEA
Configural	111.203	102	.251	.994	.992	.030		
Métrico	119.969	111	.264	.994	.993	.028	.000	.002
Escalar	142.734	120	.077	.985	.983	.043	.009	-.015
Estricto	151.026	132	.123	.987	.987	.037	-.002	.006

Discusión

El presente estudio tenía como objetivo constatar las propiedades psicométricas del Maslach Burnout Inventory Student Survey (MBI-SS) en estudiantes de España y Chile. Si bien el instrumento original evidenció adecuadas propiedades psicométricas, en cuanto a su confiabilidad y su estructura factorial, fue posible apreciar un mejoramiento de sus indicadores, especialmente estructurales, al acortar su extensión a solo 12 ítems. En otras investigaciones se han encontrado excepciones a esto, demostrando que el comportamiento de los elementos de la escala es inestable, sugiriendo corregir o eliminar ítems del instrumento (Grajales, 2001). Asimismo, otros autores como Pérez-Fuentes, Molero Jurado, Simón Márquez, Oropesa Ruiz, & Gázquez Linares (2020) también lograron visualizar un mejor ajuste del instrumento en una versión de 12 ítems.

Con respecto al instrumento, todas las dimensiones que lo componen presentan una adecuada confiabilidad (Agotamiento, Cinismo y Eficacia académica). En esta línea, la versión tanto de 15 como la de 12 ítems del MBI-SS para estudiantes universitarios, han mostrado tener una estructura factorial definida y confiabilidades adecuadas que permiten su utilización para estos fines en población universitaria tanto chilena como española, evidenciando valores omegas de McDonald superiores a ,70 en todos los factores del instrumento. Asimismo, se aprecia una adecuada bondad de ajuste del modelo de tres factores del instrumento.

Conforme a la potencial fuente de sesgo evidenciada en el instrumento original de 15 ítems, este fenómeno se podría asociar a la procedencia cultural de los evaluados y al sistema universitario en el cual están inmersos. Aunque esto se aplicaría a las muestras que consideraron estudiantes europeos y latinoamericanos (Campos y Marôco, 2012), no se replica en Charry y colaboradores (2018) cuyos grupos tenían esas características (Colombia y España).

Si bien no se logró constatar la invarianza factorial de la versión del instrumento original de 15 ítems, sí se logró verificar la invarianza estructural del instrumento para la muestra de estudiantes de Chile y España considerando la versión

propuesta de 12 ítems. Los resultados de otras investigaciones no han mostrado lo mismo (Campos y Marôco, 2012; Schaufeli et al., 2002). También ha habido cuestionamientos al MBI, creando una duda razonable sobre la equivalencia cultural del constructo (Kulakova et al., 2017). Es por esto la importancia de los estudios transculturales, en diferentes continentes, con distintas costumbres y culturas, como son Europa y Latinoamérica (Kulakova et al., 2017).

Para concluir, este estudio permitió conocer que el instrumento MBI-SS de 12 ítems muestra un mejor ajuste estructural que el original de 15 ítems. A su vez, que el instrumento de 12 ítems permite la equivalencia de su aplicación tanto en Chile como en España. Además, se comprobó que todas las dimensiones, Agotamiento, Cinismo y Eficacia, tienen una buena y/o adecuada confiabilidad. En cuanto al modelo, este es invariante tanto para la muestra española como chilena en estricto rigor, con los 12 ítems, es posible comparar los resultados.

En cuanto a las limitaciones que se presentan, son que este estudio consta de una muestra no probabilística, y es acotada en cada una de las poblaciones estudiadas. Sin embargo, permite visualizar al instrumento comparativamente en ambas realidades, por lo tanto, en futuras investigaciones se sugiere el uso de muestras más grandes, de carácter representativas. Es importante contar con instrumentos válidos para el diagnóstico y la investigación en salud mental en estudiantes universitarios en países como Chile y España, para, de esta forma contribuir a su bienestar.

Referencias

- Abalo, J., Lévy, J., Rial, A., & Varela, J. (2006). Invarianza factorial con muestras múltiples. En J. Lévy (Ed.), *Modelización con Estructuras de Covarianzas en Ciencias Sociales*, (pp. 259-278). Netbiblio.
- Aguayo, R., Vargas, C., de la Fuente, E. & Lozano, L. (2011). A meta-analytic reliability generalization study of the Maslach Burnout Inventory. *International Journal Of Clinical and Health Psychology*, 11 (2), 343-361.
- Aguayo-Estremera, R., Cañadas, G., Ortega-Campos, E., Ari-

- za, T. & De la Fuente-Solana, E. (2023). Validity Evidence for the Internal Structure of the Maslach Burnout Inventory-Student Survey: A Comparison between Classical CFA Model and the ESEM and the Bifactor Models. *Mathematics*, 11(6), 1-21. <https://doi.org/10.3390/math11061515>
- Aguilar-Bustamante, M., & Riaño-Hernández, D. (2013). Propiedades psicométricas del "School Burnout Inventory" SBI en población colombiana adolescente. Paper presented at the 34th Interamerican Congress of Psychology, 15-19 July, Brasilia (Brazil).
- Álvarez Gallego, E., & Fernández Ríos, L. (1991). El Síndrome de "Burnout" o el desgaste profesional. *Revista de la Asociación Española de Neuropsiquiatría*, 11(39), 257-265.
- APA (2017). Ethical Principles of Psychologists and Code of Conduct. <https://www.apa.org/ethics/code>
- Argumedo, D., Nóbrega, M., Bárrig, P. & Otiniano, F. (2016). Criterios Homologados de Investigación en Psicología (CHIP). Investigaciones instrumentales. Versión 1.0. Pontificia Universidad Católica del Perú.
- Ato, M., López-García, J. J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología/Annals of Psychology*, 29(3), 1038-1059.
- Banda, J., Robles, V., & Lussier, R. (2021). Validación del Maslach Burnout Inventory en estudiantes universitarios de El Bajío mexicano. *RIDE. Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, 12(23), e052. <https://doi.org/10.23913/ride.v12i23.1092>
- Boada, J., Merino, E., Sánchez, J. C., Prizmic, A. J., & Vigil, A. (2015). Adaptation and psychometric properties of the SBI-U scale for Academic Burnout in university students. *Anales de Psicología*, 31(1), 290-297.
- Bresó, E (2008). Well-being and performance in academic settings: The predicting role of self-efficacy. Disertación doctoral no publicada, Universidad Jaume I de Castellón, Castellón, España.
- Byrne, B. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicothema*, 20(4), 872-882. <https://reunido.uniovi.es/index.php/PST/article/view/8744>
- Caballero, C., Abello, R., & Palacio, J. (2007). Relación del burnout y rendimiento académico con la satisfacción frente a los estudios. *Avances en Psicología Latinoamericana*, 25(2), 98-111.
- Caballero, C., Bresó, E, & González. (2015) Burnout en estudiantes universitarios. *Psicología desde el Caribe*. 32 (3), 424-441.
- Caballero, C., Hederich, C. & Palacio, J. (2010). El burnout académico: delimitación del síndrome y factores asociados con su aparición. *Revista Latinoamericana de Psicología*, 42(1), 131-146
- Charry, C., Garzón, A., Pozo, C., & Bretones, B. (2018). Invarianza factorial del burnout académico para estudiantes hispanohablantes. *Revista mexicana de investigación educativa*, 23(79), 1195-1215. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-66662018000401195&lng=es&tlng=es
- Campos, J. A. D. B., & Marôco, J. (2012). Maslach Burnout Inventory – Student Survey: Portugal – Brazil cross – cultural adaptation. *Revista de Saúde Pública*, 46(5), 816-824.
- Cheung, G.W., & Rensvold, RB. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233-255. https://doi.org/10.1207/S15328007SEM0902_5
- Darlington, R. B., & Hayes, A. F. (2017). *Regression Analysis and Linear Models: Concepts, Applications, and Implementation*. The Guilford Press
- Dyrbye, L., Thomas, M., Harper, W., Massie, F.S., Power, D., Eacker, A., Durning, S., Moutier, C., Szydlo, D., Novotny, P., Solan, J. y Shanafelt, T. (2009). The learning environment and medical student burnout: a multicentre study. *Medical Education*, 43, 274-282
- Esnaola, I., Infante, G. & Zulaika, L. (2011). The multidimensional structure of physical self-concept. *The Spanish Journal of Psychology*, 14(1), 304-312.
- Faye Dumanget, C., Carré, J., Le Borgne, M., & Boudoukha, P. A. H. (2017). French validation of the Maslach Burnout Inventory Student Survey (MBI SS). *Journal of Evaluation in Clinical Practice*, 23(6), 1247-1251.
- Freudenberger, H. J. (1980). Burnout: The high cost of high achievement. Doubleday.
- Garcés de los Fayos, E. (1995). Burnout en niños y adolescentes: Un nuevo síndrome en psicopatología infantil. *Psicothema*, 7 (1), 33-40.
- George, D., y Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference*. 11.0 update (4th ed.). Allyn y Bacon.
- Gil-Monte, P. y Olivares, V. (2011) Psychometric properties of the "Spanish Burnout Inventory" in chilean professionals working to physically disabled people. *The Spanish Journal of Psychology*, 14 (1), 441-451
- Grajales, T. (2001). Estudio de la validez factorial del Maslach Burnout Inventory versión española en una pobla-

- ción de profesionales mexicanos. *Memorias del CIE*, 2, 63-82. <http://tgrajales.net/mbivalidez.pdf>
- Hernández, E. G., Jiménez, B. M., Muñoz, A. R., & Benadero, M. E. M. (2005). Breve historia del Burnout a través de sus instrumentos de evaluación. En *Quemarse en el trabajo: 11 perspectivas del burnout* (pp. 161-182). Egado Editorial.
- Hederich, M. C., & Caballero, D.C. (2016). Validación del cuestionario Maslach Burnout Inventory-Student Survey (MBI-SS) en contexto académico colombiano. *Revista CES Psicología*, 9(1), 1-15.
- Herrero, J. (2010). El Análisis Factorial Confirmatorio en el estudio de la Estructura y Estabilidad de los Instrumentos de Evaluación: Un ejemplo con el Cuestionario de Autoestima CA-14. *Psychosocial Intervention*, 19(3), 289-300. http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1132-05592010000300009&lng=es&tlng=es
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, 6(1), 1-55.
- Kulakova, O., Moreno Jiménez, B., Garrosa, E., Sánchez Hernández, M. O., & Aragón, A. (2017). Universalidad del constructo del Maslach Burnout Inventory en un contexto latinoamericano. *Acta de investigación psicológica*, 7(2), 2681-2688. <https://doi.org/10.1016/j.ai-pprr.2017.05.001>
- Loayza-Castro, J. A., Correa-López, L. E., Cabello-Vela, C. S., Huamán-García, M. O., Cedillo-Ramírez, L., Vela-Ruiz, J. M., Pérez-Acuña, L. M., Gonzales-Menéndez, M. J. M., & De La Cruz-Vargas, J. A. (2016). Síndrome de Burnout en estudiantes universitarios: tendencias actuales. *Revista De La Facultad De Medicina Humana*, 16(1). <https://revistas.urp.edu.pe/index.php/RFMH/article/view/333>
- Maslach, C. & Jackson, C. (1981). The measurement of experienced burnout. *Journal of Occupational Behavior*, 2, 99-113.
- Martínez, I., & Marques-Pinto, A. (2005). Burnout en estudiantes universitarios de España y Portugal y su relación con variables académicas. *Aletheia* (21), 21-30. http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1413-03942005000100003&lng=pt&tlng=es
- Merino, C., Delgadillo, A., & Caballero, R. (2013). ¿Burnout en adolescentes?: Validez estructural del inventario de burnout escolar (SBI) (Burnout in adolescents: Structural validity School Burnout Inventory (SBI)). Paper presented at the 34th Interamerican Congress of Psychology, 15-19 July, Brasilia (Brazil).
- Moneta G. B. (2011) Need for Achievement, Burnout, and Intention to Leave: Testing an Occupational Model in Educational Settings. *Personality and Individual Differences*, 50, 274-278. <https://doi.org/10.1016/j.paid.2010.10.002>
- Mostert, K., Pienaar, J., Gauche, C., & Jackson, L. (2007). Burnout and engagement in university students: A psychometric analysis of the MBI-SS and UWES-S. *South African Journal of Higher Education*, 21(1), 147-162. <https://doi.org/10.4314/sajhe.v21i1.25608>
- Moyano, N., & Riaño-Hernández, D. (2013). Burnout escolar en adolescentes españoles: Adaptación y validación del School Burnout Inventory. *Ansiedad y Estrés*, 19, 95-103.
- Ornelas, M., Jurado, P. J., Blanco, H., Peinado, J. E., & Blanco, J. R. (2020). Composición factorial del Inventario de Burnout de Maslach para Estudiantes en universitarios mexicanos. *Acta Universitaria* 30, e2516. <http://doi.org/10.15174.au.2020.2516>
- Pendergast, L. L., von der Embse, N., Kilgus, S. P., & Eklund, K. R. (2017). Measurement equivalence: A non-technical primer on categorical multi-group confirmatory factor analysis in school psychology. *Journal of school psychology*, 60, 65-82. <https://doi.org/10.1016/j.jsp.2016.11.002>
- Pérez, V.C., Parra, P.P., Fasce, H.E., Ortiz, M., Bastías, V.N., & Bustamante, D.C. (2012). Estructura Factorial y Confiabilidad del Inventario de Burnout De Maslach En Universitarios Chilenos. *Revista Argentina De Clínica Psicológica*, 21, 255-263.
- Pérez-Fuentes, M. C., Molero Jurado, M. M., Simón Márquez, M. M., Oropesa Ruiz, N. F., & Gázquez Linares, J. J. (2020). Validation of the Maslach Burnout Inventory-Student Survey in Spanish adolescents. *Psicothema*, 32(3), 444-451. <https://doi.org/10.7334/psicothema2019.373>
- Pines-Malakh, A., Aronson, E., y Kafry, D. (1981). Burnout: from tedium to personal growth. Free Press. http://books.google.com/books?id=9_NGAAAMAAJ
- Rigo, D. Y., & Donolo, D. (2019). Modelos de ecuaciones estructurales usos en investigación psicológica y educativa. *Revista Interamericana De Psicología/Interamerican Journal of Psychology*, 52(3), 345-357. <https://doi.org/10.30849/rip.ijp.v52i3.388>
- Salanova, M., Martínez, I., Bresó, E., Llorens, S. & Grau, R. (2005). Bienestar psicológico en estudiantes universitarios: Facilitadores y obstaculizadores del desempeño aca-

- démico. *Anales de Psicología*. 21(1), 170-180.
- Schaufeli WB, Leiter MP, Maslach, C, Jackson SE. (1996). The Maslach Burnout Inventory: General Survey (MBI-GS). En: Maslach C, Jackson SE, Leiter MP, ed. *Maslach Burnout Inventory Test Manual*, Consulting Psychologists Press, 19-26.
- Schaufeli, W., Martínez, I., Marques Pinto, A., Salanova, M., & Bakker, A. (2002) Burnout and engagement in university students: A cross-national study. *Journal of Cross - Cultural Psychology*. 33(5), 464-481.
- Schaufeli, W., Salanova, M., González-Romá, V. & Bakker, A. (2002). The measurement of burnout and engagement: A confirmatory factor analytic approach. *Journal of Happiness Studies*, 3, 71-92.
- Shirom, A. (2003). Job-related burnout: A review. En Quick J. C. y Tetrick L. E. (Eds), *Handbook of occupational health psychology* (pp. 245-264). American Psychological Association. <https://doi.org/10.1037/10474-012>
- Ventura-León, J. & Caycho-Rodríguez, T. (2017). El coeficiente Omega: un método alternativo para la estimación de la confiabilidad. *Revista latinoamericana de ciencias sociales, niñez y juventud*, 15 (1), 625-627.

Representaciones semánticas de la bioética por jóvenes universitarias desde la ética del cuidado

Pedro César Cantú-Martínez

Universidad Autónoma de Nuevo León, Nuevo León, México <https://orcid.org/0000-0001-8924-5343> pedro.cantum@uanl.mx

Citar como: Cantú-Martínez, P.C. (2023). Representaciones semánticas de la bioética por jóvenes universitarias desde la ética del cuidado. *Revista Digital de Investigación en Docencia Universitaria*, 17(2), e1592. <https://doi.org/10.19083/ridu.2023.1592>

Recibido: 06/01/2022. **Revisado:** 07/02/2022. **Aceptado:** 10/06/2023. **Publicado:** 30/06/2023.

Resumen

Introducción: en la actualidad los estudios de género cuentan con mucha importancia; por lo cual, esta investigación nos permitirá conocer el discernimiento de las mujeres acerca de la bioética. **Objetivo:** reconocer las representaciones semánticas que jóvenes universitarias tienen en relación con la bioética. **Metodología:** este estudio se realizó en estudiantes universitarias del área de las ciencias biológicas. Se empleó la técnica de redes semánticas naturales y la postura de ética del cuidado para su examinación. **Resultados:** se evaluaron 191 palabras que representan una riqueza de 81 enunciados. Las principales expresiones fueron la "confianza", "diálogo", "debate", "deber" y "principios". **Discusión:** las estudiantes universitarias ejercen su razón práctica eligiendo lo mejor para los otros, confirmando los supuestos de la ética del cuidado.

Palabras clave: ética; sistema de valores; creencias; toma de decisiones; razonamiento

Semantic Representations of Bioethics by Young University Students From the Ethics of Care

Abstract

Introduction: At present, gender studies are very important, so research is carried out that will allow us to know the discernment of women about bioethics. **Objective:** To recognize the semantic representations that young university students have in relation to bioethics. **Methodology:** This study was carried out in university students in the area of biological sciences. The technique of natural semantic networks and the posture of ethics of care were used for its examination. **Results:** 191 words were evaluated, representing a richness of 81 statements. The main expressions were "trust", "dialogue", "debate", "duty" and "principles". **Discussion:** College students exercise their practical reason by choosing what is best for others, confirming the assumptions of the ethics of care.

Keywords: ethics; value systems; belief; decision making; reasoning.

Introducción

Los estudios de género han ido adquiriendo cada vez más relevancia en el marco de la investigación. En este trabajo de investigación se examinarán las representaciones semánticas que un grupo de jóvenes universitarias tienen en relación con la bioética, las cuales serán analizadas mediante la ética del cui-

*Correspondencia:

Pedro César Cantú-Martínez
pedro.cantum@uanl.mx



dado, propuesta por Carol Gilligan ([García, 2015](#)), para de esta forma comprender cómo esta postura teórica se manifiesta en un grupo concreto de mujeres universitarias. Esta propuesta del desarrollo moral surge como respuesta a la ética de la justicia planteada por Lawrence Kohlberg ([Alonso & Fombuena, 2006](#)), que se caracteriza por el interés en la reciprocidad y justicia. En tanto, en la ética del cuidado, se interioriza en las apreciaciones emanadas de la experiencia femenina que hace hincapié en los asuntos relacionados con el apego y auxilio de las personas ([Medina-Vicent, 2016](#)).

Por otra parte, las representaciones –en acuerdo con [Jodelet \(2020\)](#)– enmarcan la vinculación que tiene una persona o un grupo de ellas frente a su entorno, creando así un imaginario social, en el cual discurre una corriente de ideas, creencias y posturas somatizadas. Mientras las representaciones semánticas emergen de la experiencia y organización del conocimiento –de un individuo o un sector social– que provee información referente a la estructura semántica y ordenamiento interno de los significados, que configuran la interpretación de un concepto ([Maldonado, et al., 2020](#)).

En una visión más amplia, se debe comprender que la ética del cuidado incide en la acción de asistir a las personas con el propósito de brindar el mayor bienestar biopsicosocial posible. Donde prevalecen actitudes maternalistas por parte de las mujeres, donde se privilegia la empatía, acompañamiento y confidencialidad, elementos que sustentan una favorable relación ([García, 2015](#); [Domingo-Moratalla, 2019](#)). Estas particularidades son la que permiten llevar a cabo un tránsito de cualidades objetivas y subjetivas, que hacen evidente la estructuración histórica de las mujeres, particularmente en su contexto de identidad moral.

Por otro lado, esta postura de la ética del cuidado, que se ha desarrollado ampliamente en las últimas décadas, pretende reconocer la forma distinta de evaluar y enfrentar las tomas de decisiones desde el conocimiento de las mujeres; es decir, desde la posición que le ha sido asignada por la sociedad, en cuyo caso prevalecen actividades que preferentemente se inclinan a la compasión y afinidad ([Mesa, 2005](#); [Tamanini & Da Costa, 2022](#)). Por consiguiente, la ética del cuidado tiene que ver con ese valor que emana de

cada una de las mujeres y que se vincula al contenido de deliberar y calificar, así como también a la voluntad de proceder. Desde la perspectiva de [Camps \(1998\)](#):

No se trata de postular una naturaleza femenina específicamente cuidadora o cuidadora, sino de constatar la existencia de una cultura que no han hecho suya los hombres y, por lo mismo, no ha sido parte de la vida pública, sino que ha sido considerada más bien un estorbo para los comportamientos públicos (p. 74-75).

A juicio de [León-Correa \(2008\)](#), se requiere una ética del cuidado que proceda del escrutinio de todas las perspectivas, trascendiendo por encima de las diferencias y buscando con sumo cuidado que no retorne a las tesis paternalistas ya existentes, enfatizando así el principio de complementariedad. Por lo cual, al llevar a cabo esta investigación permitirá conocer la percepción de las mujeres sobre lo que es la bioética. Es así, que se torna sumamente valioso ya que la erudición femenina ha alimentado también la historia y construcción de la sociedad, aportando nuevas maneras de solucionar los conflictos entre las personas y distintos modos de satisfacer las necesidades de las mismas, como bien lo ha señalado [Comins \(2003\)](#). Por lo tanto, a continuación, se hará una aproximación a la representación semántica de la bioética y su relación con el contenido de la ética del cuidado que simultáneamente subyace por la perspectiva de género. La postura y abordaje de esta investigación se lleva a cabo desde la fenomenología, cuya intervención permite estudiar un suceso, contemplando el punto de vista de quienes participan en ella ([Lambert, 2006](#)).

Marco Conceptual

Con la concreción de la ética del cuidado llevada a cabo por Carol Gilligan –en 1982– se da un vuelco a los conceptos prevalecientes y propuestos por Lawrence Kohlberg, de tal manera que se bosqueja un distinto paradigma que permite ampliar la contextualización de la ética. De hecho, comenta [Camps \(2013\)](#) al respecto:

A través del estudio y el análisis directo del sentir y el razonar de las niñas, Gilligan descubrió el valor del cuidado, un valor [...] que debiera ser tan importante como la justicia, pero no lo era porque se desarrollaba sólo en la vida privada y doméstica protagonizada por las mujeres (p. 7).

Gilligan logró contrastar los supuestos universales que delineaban las diferencias existentes entre los roles masculinos como femeninos, donde hace un recuento de las distintas etapas del ciclo vital de las personas, observando que el desarrollo de las personalidades –entre varones y mujeres– se percibe de forma distinta; mientras que en las mujeres la construcción es mediante el apego con la madre; en cambio, los varones lo llevan a cabo mediante el desapego con esta; por lo tanto, las experiencias de la construcción como personas es totalmente distinto ([Santacruz, 2006](#)).

En este sentido, argumenta [Faerman \(2015\)](#) que Carol Gilligan refuta la postura en la que Kohlberg manifiesta en su teoría de la moral por solo considerar el vértice masculino como pauta y contemplar que la forma de razonar del género femenino fuese inferior solo por la condición de ser mujer. Es así como surge la ética del cuidado mediante el trabajo que desarrolla Gilligan, en el que encuentra como hallazgos –bastante relevantes– que las mujeres cuentan con un razonamiento moral diferente al de los varones. Así se construye la ética del cuidado como una disyuntiva a la propuesta por Kohlberg denominada ética de la justicia. Uno de los sustentos que soporta la teoría de Gilligan es que aduce que:

atribuye estas diferencias no a cuestiones de anatomía sino al hecho de que las mujeres, universalmente, son responsables en gran parte del cuidado de los recién nacidos. Como resultado de ello, en cualquier sociedad con estas características, la personalidad femenina llega a definirse en relación y conexión con otras personas más de lo que suele hacerlo la personalidad masculina ([Faerman, 2015, p. 125](#)).

Este proceso de carácter subjetivo evidenciado por Gilligan no era considerado por la realidad existente de orden patriarcal –donde tanto mujeres como varones eran erigidos mayormente por los estereotipos esgrimidos por el sexo–, que con-

llevó que las mujeres no ostentaran un pronunciamiento propio.

En cambio, la postura de Gilligan da apertura a una nueva forma de parlamentar acerca de las eventualidades morales, esencialmente de forma razonada y promover que ninguna postura de orden moral –masculina– se le puede considerar como universal. En todo caso, se constituye junto con las otras en la completitud de la valoración de un contexto indicado, poniendo de manera palpable que el derrotero y desarrollo de la moral de las mujeres es distinto al de los varones, pero igual de apreciable.

A juicio de [Alvarado \(2004\)](#) la ética del cuidado considera que entre dos personas sucede la interacción y el contacto moral, en el que subsiste una atención y que, además, enlaza a los individuos como parte de la vinculación humana reinante. Atendiendo a este último argumento, Gilligan permite acercarse a valorar los dilemas con un talante puesto en observar las relaciones humanas de los individuos involucrados, ocupándose por el bien de los otros; esto es comprender el vínculo entre la persona y la otredad, mediante la consideración del compromiso, de tal manera que lo bueno lo constituye la atención y el interés hacia los demás.

Todo lo anterior, conduce a llevar a cabo la presente indagación con la finalidad de acercarse a la configuración semántica del concepto de bioética por parte de un grupo de mujeres estudiantes universitarias y su correspondencia con la ética del cuidado que postuló Carol Gilligan.

Método

Diseño

La presente investigación de carácter transversal y descriptiva, abordando las representaciones semánticas que surgen del concepto bioética por un grupo de mujeres universitarias. Esta investigación se lleva a cabo en el marco de la línea de investigación y aplicación del conocimiento “Desarrollo Humano Sustentable” que se encuentra registrada en el Cuerpo Académico Ciencias Exactas y Desarrollo Humano UANL-CA-181 con “Nivel Consolidado”,

adscrito a la Facultad de Ciencias Biológicas de la Universidad Autónoma de Nuevo León.

Participantes

Estudiantes universitarias del área de las ciencias biológicas, que se encontraban estudiando la carrera de Químico Bacteriólogo Parasitólogo –que cursaban el tercer semestre– en la Universidad Autónoma de Nuevo León –San Nicolás de los Garza, N.L., México–. El número de participantes fue de 38 estudiantes, contando con edades entre 19 y 20 años. El muestreo fue no probabilístico e intencional.

Instrumentos

Para su implementación se utilizó la técnica de redes semánticas naturales formulada por [Vera-Neiriga et al. \(2005\)](#).

Procedimientos y análisis de los datos

Inicialmente se aplicó un cuestionario con consentimiento verbal, a las participantes, recabándose información correspondiente a la edad Para proceder a los siguientes pasos: 1) A partir de la palabra detonante bioética las participantes enuncian cinco palabras que consideren y posteriormente se les solicita que las enumeren de la más sobresaliente a la menos preponderante. 2) Una vez llevada la acción anterior se revisa la información y se procede a hacer una normalización de las palabras anotadas por las estudiantes universitarias. 3) Seguidamente se llevó a cabo

un examen de la información emanada para obtener el valor J que conforma el capital semántico, el valor M que enuncia el peso semántico de cada una de las expresiones inscritas por las participantes como palabras definidoras, el Grupo SAM que compone el conjunto de enunciados con pesos específicos más elevados, para esto se consideró conveniente recurrir a el valor del tercer cuartil, y determinar las locuciones más sobresalientes, y seguidamente identificar las primeros cinco campos semánticos naturales para inmediatamente agenciar el indicador FMG, que representa la distancia referida en porcentaje entre las palabras ([Cantú-Martínez, 2019](#)).

Resultados

Las estudiantes universitarias tras la palabra detonante de bioética expresaron sus representaciones semánticas, donde esgrimieron 190 palabras. Se encontró después de la normalización de los enunciados un valor de J de 81 pronunciamientos coligados al campo semántico detonante de bioética. Además, se encontró que las primeras cinco posiciones de este inventario las conformaron –con mayor peso semántico– la “confianza” y “diálogo” con un peso de 19 cada una, siguiéndole el “debate” con un peso de 18, continuando el “deber” con peso de 17 y finalmente el pronunciamiento de “principios” con un valor de 16 (véase Tabla 1).

Tabla 1
Grupo SAM, Valor J, Menciones, Valor M e Indicador FMG del Constructo “Bioética”, Señalado por las Participantes.

Grupo SAM	Menciones	Valor M	Indicador
			FMG (%)
Confianza	5	19	100
Diálogo	5	19	100
Debate	6	18	94.7
Deber	4	17	89.5
Principios	5	16	84.2

Nota. Grupo SAM (conjunto de enunciados): Valor J (capital semántico): 81; Valor M (peso semántico); Indicador FMG (distancia entre los enunciados).

El Grupo SAM, que integra las nociones que las estudiantes universitarias han pronunciado para concretar las representaciones sociales de lo que significa “bioética”, se explica en la Tabla 1. Los más relevantes por su posición –que se puede considerar como núcleo– son la “confianza” y “diálogo”, prosiguiendo la palabra “debate” a una distancia de 94.7%, continuando el “deber” con un promedio de 89.5% y posteriormente el vocablo de “principios” con un valor de 84.2%

También se observa que en el Grupo SAM los campos semánticos mayormente concurrentes –FMG– fueron la “confianza” y el “diálogo”, ambos con un peso de 19, y fueron enunciados por el 26.3% de los participantes; en tanto que los cinco términos más relevantes constituyeron el 15.1% de todas las menciones hechas por las estudiantes. En lo correspondiente a la correlación entre el número de las menciones y el valor M –que representa– el peso semántico se encontró que los enunciados en su mayoría se ordenan de manera muy similar, lo cual se comprobó mediante un valor de correlación de Spearman de 0.864 con valor de $p < 0.01$.

Discusión

Como se ha podido observar en la sección de los resultados, las estudiantes universitarias participantes han expresado sus imaginarios semánticos en relación con el término detonante que fue el vocablo “bioética”; con la finalidad de articular la visión femenina del contexto de la realidad que estas estudiantes recrearon mediante sus pronunciamentos, y que dan cuenta –sin lugar a dudas– de sus propias experiencias, costumbres, valores y comprensión de lo que les rodea.

La ética del cuidado, que fue propuesta por Carol Gilligan, sostiene entre sus principales argumentaciones que la acción moral cuenta como nodo central las relaciones interpersonales y la generosidad, donde es palpable la respuesta de una persona ante la necesidad de otra (Kittay, 2011). De tal manera, que Botes (2000), argumenta también que la ética del cuidado se caracteriza por las relaciones armónicas y requerimientos de la otredad,

elementos que constituyen sumamente relevantes y que finalmente conducen a la acción.

De esta forma, la construcción de bioética por las estudiantes participantes, a partir de las representaciones semánticas –de forma nuclear– se estableció primordialmente en cinco campos semánticos: “confianza”, “diálogo”, “debate”, “deber” y por último los “principios”. Por lo cual, se llevó a cabo un análisis de estos conceptos expresados con el propósito de caracterizarlos en el marco de la ética del cuidado, es decir con esta perspectiva de género, propuesta por Carol Gilligan.

Inicialmente, el término de “confianza” desde la perspectiva de Lozano (2003) establece que la locución se vincula con el saber, que es primordial en toda relación donde interactúan personas, cuyo escenario se da en circunstancias con cualidades de valencia positiva para el actor social y bajo condiciones de incertidumbre y no de certeza. De esta manera, la confianza es una forma que expresa interés por el otro y se constituye en el vínculo de reciprocidad que media de una manera cordial y razonable, entre la forma de proporcionar ayuda y hacer eco de la solicitud sobre una necesidad, como lo han hecho saber las participantes (Santacruz, 2006). Este hallazgo además revela que las estudiantes consideran a la otra persona digna para otorgar esta confianza, que representa –en otros términos– una configuración simbólica de una correspondencia entre las partes involucradas, con apoyo incondicional en situaciones comprometidas.

Mientras que el “diálogo” –indicado por las estudiantes– se constituye en una correspondencia comunicativa con una visión no autoritaria, que permite comprender de mejor manera la relación acordada y existente de aquellas personas que participan en él, o bien pone en evidencia la realidad moral del acto en el que se inmiscuyen (Cabra-Torres, 2010). Tal como lo expone Zapata y Mesa (2009), al señalar, que un diálogo se da entre dos o más personas que conversan, que, de manera alternativa, haciendo uso de sus intervenciones, dan cuenta de sus propósitos, opiniones y deseos, mientras se suscita un proceso de negociación entre los interlocutores. Este hecho resalta ya que concuerda con el marco de la teoría de Gilligan (Domingo-Moratalla, 2019), donde

sobresale una razón práctica de orden raciovitalista, que se pone al servicio de la vida desde la perspectiva de observar en la otredad la vulnerabilidad, la fragilidad y el vínculo emocional, como lo hacen saber las mujeres participantes en el estudio.

En tanto el “debate” – aludido por las alumnas universitarias– se erige tangiblemente en la forma estructurada de aducir diferentes razonamientos entre dos o más personas en las que se examinan las opiniones. Esto es, favorece la toma de decisiones al exponerse los argumentos, beneficiando particularmente, el concebir un pensamiento crítico o defensor, y que genera juicios morales informados en un ambiente de comunicación afable y efectiva. De hecho, el debate en acuerdo con [Rangel \(2007\)](#), es un proceso dialéctico, con la finalidad de conseguir un acuerdo que se motive en el empleo de la razón. Y que, en los últimos años, mediante los debates sociales también las mujeres han podido superar los papeles –estereotipados– que la sociedad les ha impuesto sobre sus personas, en tanto, que ahora han demostrado una evidente y marcada participación social ([Alcaraz & Vázquez, 2020](#)).

Por otra parte, el “deber” es la conducta de los seres humanos que se sostiene en las reglas morales que rigen a una sociedad, y que se hacen tangibles en motivos y acciones morales de orden subjetivo que se expresa en distintos niveles de compromisos y responsabilidades hacia los otros miembros de la sociedad, develando la conciencia moral que se posee. Por consiguiente, aunque el deber se ha constituido en un ordenamiento jurídico, aquí se plasma el altruismo presente en las mujeres universitarias que participaron, en términos de una acción solidaria y de subsidiaridad con los demás ([Sant’Ana, 2014](#)).

En cambio, la locución de “principios”, última de este núcleo de términos que simboliza la red semántica natural de las principales representaciones que manifiestan las participantes universitarias, connota las disposiciones y las decisiones que toma un ser humano de manera general y que rigen su conducta moral. Además, demarca un camino de actuación –del ser humano y en este caso de las mujeres– que le permite hacer hincapié que lo relevante es mantener y sostener las

buenas relaciones con las demás personas. Todo esto en el marco del respeto a los contenidos que engloban la autonomía, la beneficencia, no maleficencia y justicia ([Siurana, 2010](#); [Beauchamp, 2020](#)). En este contexto todo ser humano –en uso de la razón– cuenta con el atributo de conducirse y obrar en acuerdo con la representación que de carácter moral ostenta. En este caso, las mujeres ejercen su razón práctica en la elección de lo que contemplan mejor para los otros –ética del cuidado– sin considerar sus propias aspiraciones, generando entonces un comprensivo compromiso de orden íntimo, que se yergue como imperativo de lo que considera que es bueno para los demás, lo será también para ella.

En suma, en el presente estudio se ha indagado sobre las representaciones semánticas de la bioética en jóvenes universitarias desde la perspectiva de la ética del cuidado, advirtiendo que la metodología empleada posibilitó conocer la adjudicación semántica que las mujeres participantes otorgaban al vocablo bioética. No obstante, el nivel de reflexividad de las participantes, estos resultados son ratificados solo para el grupo intervenido, y aproximan fenomenológicamente al pensar de las mujeres; sobre lo cual se pudo comprobar que las mujeres tienden más al contexto de la responsabilidad por los demás, como lo aseveró Carol Gilligan.

Así mismo, se recomienda insistir en investigaciones con abordajes metodológicos que se concentren en un análisis desde la perspectiva y miramiento del propio actor, es decir, descubriendo la postura de la otredad. Por último, se puede proyectar que los resultados obtenidos de las participantes revelan la experiencia vivida y evocan los argumentos racionales de estas, como también las sensibilidades, que al conjugarse dan una puntualización descriptiva como también vivencial de su existencia.

Referencias

- Alcaraz, A. & Vázquez, J.C. (2020). Las barreras del desarrollo laboral de las mujeres. Una aproximación latinoamericana. *América Crítica*, 4(1), 59-65.

- DOI: [10.13125/americanacritica/3867](https://doi.org/10.13125/americanacritica/3867)
- Alonso, R. & Fombuena, J. (2006). La ética de la justicia y la ética de los cuidados. *Portularia*, 6(1), 95-107. <https://www.redalyc.org/pdf/1610/161016087008.pdf>
- Alvarado, A. (2004). La ética del cuidado. *Aquichan*, 4(4), 30-39. <https://aquichan.unisabana.edu.co/index.php/aquichan/article/view/47>
- Beauchamp, T.L. (2020). Principialismo bioético y biojurídico: ¿necesitan la bioética y el bioderecho europeos un marco diferente de principios? *Revista Principia Juris*, 17(36), 11-33. <http://revistas.ustatunja.edu.co/index.php/piuris/article/view/2060/1800>
- Botes, A. (2000) A comparison between the ethics of justice and the ethics of care. *Journal of Advanced Nursing*, 32(5), 1071-1075. <https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-2648.2000.01576.x>
- Cabra-Torres, F. (2010). El diálogo como fundamento de comunicación ética en la evaluación. *Educación y Educadores*, 13(2), 239-252. <https://www.redalyc.org/pdf/834/83416998005.pdf>
- Camps, V. (1998). *El siglo de las mujeres*. Cátedra.
- Cantú-Martínez, P.C. (2019). La apreciación semántica de la noción desafíos en bioética por alumnos universitarios del campo de las Ciencias Biológicas. *Revista Iberoamericana de Bioética*, 10, 1-11. <https://doi.org/10.14422/rib.i10.v2019.002>
- Comins, I. (2003). La ética del cuidado: contribuciones a una transformación pacífica de los conflictos. *Feminismo/s*, 9, 93-105. <http://rua.ua.es/dspace/handle/10045/3643>
- Domingo-Moratalla, A. (2019). Cuidado y responsabilidad: de Hans Jonas a Carol Gilligan. *Pensamiento*, 75(283), 357-373. <https://doi.org/10.14422/pen.v75.i283.y2019.019>
- Faerman, R. (2015). Ética del Cuidado: una mirada diferente en el debate moral. *Revista de Teoría del Derecho de la Universidad de Palermo*, 2(1), 123-146. https://www.palermo.edu/derecho/pdf/teoria-del-derecho/n3/TeoriaDerecho_06.pdf
- García, L. (2015). La ética del cuidado y su aplicación en la profesión de enfermería. *Acta Bioethica*, 21(2), 311-217. <https://scielo.conicyt.cl/pdf/abioeth/v21n2/art17.pdf>
- Jodelet, D. (2020). Sobre el espíritu del tiempo y las representaciones sociales, *Revista Cultura y Representaciones Sociales*, 15(29), 19-36. www.culturayrs.unam.mx
- Kittay, E.F. (2011). The Ethics of Care, Dependence, and Disability. *Ratio Juris*, 24(1), 49-58. <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9337.2010.00473.x>
- Lambert, C. (2006). Edmund Husserl: la idea de la fenomenología, *Teología y Vida*, 47, 517-529. <https://scielo.conicyt.cl/pdf/tv/v47n4/art08.pdf>
- León-Correa, F.J. (2008). Ética del cuidado feminista y bioética personalista. *Persona y Bioética*, 12(30), 53-61. <https://www.redalyc.org/pdf/832/83203006.pdf>
- Lozano, J. (2003). En torno a la confianza, *CIC Cuadernos de Información y Comunicación*, 8, 61-70. <https://revistas.ucm.es/index.php/CIYC/article/view/CIY-CO303110061A>
- Maldonado, A.C. Tapia, M.I. & Arancibia, B.M. (2020). *Perfiles Educativos*, 42(167), 138-157. <https://doi.org/10.22201/issue.24486167e.2019.167.59208>
- Medina-Vicent, M. (2016). La ética del cuidado y Carol Gilligan: una crítica a la teoría del desarrollo moral de Kohlberg para la definición de un nivel moral postconvencional contextualista. *Daimon Revista Internacional de Filosofía*, 67, 83-98. <https://revistas.um.es/daimon/article/view/199701>
- Mesa, J.A. (2005). La ética del cuidado y sus implicaciones en la formación de la escuela. En: J.A. Mesa, A. Restrepo, J. Barrera, F. Vázquez, A. Sanz, P. Escobar, C.J. Cuartas, G. Iago, J. Martín, E. Chaux, B.C. Daza, L.M. Vega, E. Velázquez, H. Gutiérrez, J. Parra & F. De Roux. (Autores). *La educación desde las éticas del cuidado y la compasión*. (pp. 21-33). Editorial Pontificia Universidad Javeriana. https://books.google.com.mx/books/about/La_educaci%C3%B3n_desde_las_%C3%A9ticas_del_cuid.html?hl=es&id=VrtZKvyBKcIC&redir_esc=y
- Rangel, M. (2007). *El debate y la argumentación. Teoría, técnicas y estrategias*. Trillas.
- Sant'Ana, A. (2014) Los deberes de las personas y la realización de los derechos fundamentales. *Estudios Constitucionales*, 12(2), 13-27. <https://www.redalyc.org/pdf/820/82032680002.pdf>
- Santacruz, M.C. (2006). Ética del cuidado. *Revista de la Facultad de Ciencias de la Salud de la Universidad del Cauca*, 8(2), 45-51. <https://revistas.unicauca.edu.co/index.php/rfcs/article/view/927>
- Siurana, J.C. (2010). Los principios de la bioética y el surgimiento de una bioética intercultural. *Veritas*, 22, 121-157. https://www.scielo.cl/scielo.php?script=sci_art-text&pid=S0718-92732010000100006
- Tamanini, M. & Da Costa, H. (2022). Uma voz diferente e as diferenças em meio ao vozerio: Gilligan revisitada à luz da teoria feminista. Schème: *Revista Eletrônica de Psicologia e Epistemologia Genéticas*, 14(Número especial), 104-144. <https://doi.org/10.36311/1984-1655.2022.v14.esp.p104-144>
- Vera-Noriega, J.A., Pimentel, C.E., & Batista de Albuquerque,

- F.J. (2005). Redes semánticas: aspectos teóricos, técnicos, metodológicos y analíticos. *Ra Ximhai*, 1(3), 439-451. <https://doi.org/10.35197/rx.01.03.2005.01.jv>
- Zapata, C.M., & Mesa, J.E. (2009). Los modelos de diálogo y sus aplicaciones en sistemas de diálogo hombre-máquina: revisión de la literatura. *Dyna*, 76(160), 305-315. <https://www.redalyc.org/pdf/496/49612068021.pdf>

Principles and assumptions of psychometric measurement

Gavin T. L. Brown

The University of Auckland, New Zealand <http://orcid.org/0000-0002-8352-2351> gt.brown@auckland.ac.nz

Citar como: Brown, G. (2023). Principles and assumptions of psychometric measurement. *Revista Digital de Investigación en Docencia Universitaria*, 17(2), e2022. <https://doi.org/10.19083/ridu.2023.1834>

Recibido: 13/12/21. **Revisado:** 24/02/22. **Aceptado:** 21/06/22. **Publicado:** 30/06/2023.

Abstract

Introduction: the validity of claims about learner self-reports depends on the credibility of the measures used. Inventories developed within the psychometric tradition are expected to provide theoretical and empirical evidence for the validity and reliability of the measures to support subsequent interpretations and decisions. These practices depend on a set of assumptions based on latent trait theory that are essential to understanding what psychometric measures can do. This tutorial outlines essential characteristics of design and reporting of psychometric self-report data.

Keywords: self-report; psychometric methods; validity; reliability; evaluation

Principios y supuestos de la medición psicométrica

Resumen

Introducción: la validez de las afirmaciones sobre los autoinformes depende de la credibilidad de las medidas utilizadas. Se espera que los inventarios desarrollados dentro de la tradición psicométrica proporcionen evidencia teórica y empírica de la validez y confiabilidad de las medidas para respaldar las interpretaciones y decisiones posteriores. Estas prácticas dependen de un conjunto de suposiciones basadas en la teoría del rasgo latente que son esenciales para comprender lo que pueden hacer las medidas psicométricas. Este documento describe las características esenciales del diseño y el informe de datos de autoinforme psicométrico.

Palabras clave: autoinforme, métodos psicométricos, validez, fiabilidad, evaluación

Introduction

Decision-making about instruction, intervention, or treatment requires high-quality information about current status (i.e., strengths and weaknesses), trends in change (if any), and impact of precursor and introduced causal factors. Furthermore, understanding of the merit, worth, or value of the obtained values in the information needs to be exercised so as to lead to appropriate decisions, actions, and consequences (Scriven, 1967). Educational and clinical interventions require robust information from those who are supposed to benefit from those interventions (i.e., students, clients). If we seek

***Correspondencia:**

Gavin T. L. Brown
gt.brown@auckland.ac.nz



to change knowledge, skill, attitude, belief, and behaviours, we need accurate measures of where people are and where they have got to post-intervention ([Messick, 1989](#)). As such then, all methods of systematically sampling individuals' skill, attitude, ability, or any other characteristic constitute 'tests' ([Cizek, 2020](#)).

Psychometrics, then, is the statistical science behind testing and measuring latent human emotions, thoughts, attitudes, beliefs, opinions, ideas, and knowledge that shape manifest behaviours. Because these phenomena exist behind the eyes, between the ears, and in the viscera of human existence, it is essential that the individual provide this information to the tester/assessor. While this information may be biased through optimistic illusions about oneself ([Dunning et al., 2004](#)), there really is no option but to obtain and exploit the information given by the individual about the state of their own feelings, cognitions, and intentions. Unlike research with physical or animal phenomena, the human is capable of knowing and communicating about internal states and traits. While intimate partners (e.g., parents, siblings, children, partners) may have privileged insights into the mind and behaviours of an individual, the person with the potentially most complete and current information is the individual herself. Observable behaviours may give useful insights into the mind and will of an individual; nevertheless, quite contrary beliefs, intentions, or motivations can lead to very similar observed actions. Hence, making sense of what is in the human mind depends on obtaining information from the owner of that mind. Other than exceptional animals that have learned to communicate with sign language, only the human individual seems able to have something to say that others cannot know unless the individual reveals that information. Thus, the reliance of psychometric research on self-reported information is both necessary and inevitable. Nonetheless, the reliability and validity of tools used to test or measure these psychological phenomena has to be demonstrated for the results to be deemed credible.

In this paper, I shall provide a high-level overview of or guide to psychometric theory and principles used to evaluate measurements of human

knowledge, beliefs and attitudes, or intentions. Readers interested in detailed understanding of psychometric principles and practices are encouraged to read authoritative handbooks on psychometrics ([Rao & Sinharay, 2007](#)), educational measurement ([Brennan, 2006](#)), or testing ([Geisinger, 2013](#)).

Psychometric Theory & Practice

Latent Theory Assumptions

Psychometric theory rests on the assumption that manifest behaviours are explained or caused by variation in latent psychological (i.e., cognitive, emotional, volitional, intentional, etc.) attributes within the individual ([Borsboom, 2005](#)). These can be called thoughts, beliefs, attitudes, or any number of invisible, non-material, but causal factors that consistently generate observable behaviours. This assumption arises out of our inability to read each other's psychological phenomena without the ability to communicate with each other verbally ([Corballis, 2002](#)) about these things. That we can talk about psychological phenomena (e.g., feelings, thoughts, beliefs, etc.) and believe that we understand what others are talking about suggests strongly that we think these things exist. Hence, latent trait theory about psychological phenomena appears to be a logical conclusion from the nature of our shared existence.

The origins of statistical or mathematical characteristics of latent trait theory date back to Spearman's application of correlation techniques (e.g., rank-order correlations, correction for attenuation, and estimation of test reliability) to patterns in test scores to infer the idea of general mental abilities ('g') that account for shared variation among scores ([Clauser, 2022](#)). General mental ability or intelligence is presumed to explain why students who do well in one subject tend to do well on other subjects. Unsurprisingly, research into the nature of intelligence has advanced to multi-dimensional models and described the causal influence of both genetic and environmental factors ([Deary, 2001](#)).

The mathematical model of latent trait theory assumes that the underlying ability behind test-taker performance requires just one assumption. [Lord \(1953\)](#) puts that assumption as: “*the trait or ability under discussion can be thought of as an ordered variable represented numerically in a single dimension*” (italics in original, p. 518). By extension, researchers who measure an individual’s underlying orientation toward a specific mental phenomenon assume that the latent cause has ordered unidimensional characteristics so that individuals can be located as to their frequency, intensity, importance, likelihood, or valence for that construct ([Allport, 1935](#)). The widespread use of [Likert’s \(1932\)](#) summated rating scale in social psychology manifests this assumption that there is latent cause that explains the strength and covariance of responses to prompts, questions, or items.

It has been argued that latent factors have not been proven to be quantitative ([Michell, 1999](#)). Because evidence for latent constructs can only be obtained from manifest operations (e.g., answers to questions), research into how individuals perceive themselves or phenomena must always rest on an assumption that the underlying trait is something that can be counted or quantified. That our experiences of the world and our abilities to respond to that world differ is not an assumption but a given. Hence, it is relatively easy to accept that all individuals vary within populations around normative values for such diverse psychological phenomena as the nature of one’s love for their own mother, their ability to solve problems, their ability to use language(s), and so on. Consequently, as [Michell \(2008\)](#) notes, latent trait theory research still has to be honest to admit that “At present we do not know whether this hypothesis [quantitative models of psychological entities] is true, but we will assume it recognizing that at some point in the future someone needs to investigate it” (p. 12).

Consequently, measurements of a latent construct are a proxy for cognitive processes which are always psychologically multidimensional ([Rupp, 2008](#)). Thus, as [Rupp \(2008\)](#) put it:

it **may still be truly meaningful**, and not just statistically convenient, to summarize data via

a unidimensional latent variable model. The resulting latent variable **might** serve as a more effective and efficient pragmatic communication tool that represents a decidedly conscious compromise between cognitive fidelity, empirical feasibility, and utilitarian practicability. (p. 122; emphasis added)

This shows that our measurement, regardless of mathematical approach, of a psychological construct always presumes that we have used a defensibly valid or theoretically robust set of proxies, which, in accordance with scientific fallibility ([Popper, 1945](#)), may turn out to be wrong.

An advantage of the quantitative assumption is that the numeric values, especially when obtained from large populations, will behave like a real-valued random variable so that mathematical manipulations (e.g., correlations, mean, standard deviation, etc.) provide insight into variation within populations as to the latent trait. Choices have to be made as to how the numbers created in measurement are statistically analysed, ranging from classical test theory to item response theory ([Embretson & Reise, 2000](#)). Nonetheless, an individual’s position relative to a social norm is an important factor, within the theory of planned behaviour, in determining intentionality and behaviour ([Ajzen, 1991](#)).

Related to the possibility that latent constructs are not quantitative, we must accept that our measurements do not have true zero values or inherent scale increments. The arbitrary nature of the scales we use to measure psychological phenomena does not invalidate the existence of the trait. For example, intelligence tests tend to set population mean to 100 and standard deviation to 15. This is a convenient way to describe and locate individuals relative to others. The mean value of 100 could just as easily have been 500 or 1000, with a concomitant SD of 50 or 100. The meaning of any score relative to those norm values would allow us pragmatically to plan useful educational or clinical interventions. Hence, the mathematics of latent trait theory when combined with a theoretical framework for the nature and function of that latent trait allow us to understand and respond to needs or strengths.

External Validation

Just because latent theory for psychological phenomena is pragmatic, does not mean that a latent construct matters. Observation of systematic relationships between variation in the latent construct and observable behaviours and outcomes in the real world is needed to establish whether a hypothesised latent trait matters. Evidence from observable behaviours or independently generated outcomes scores overcomes the bias of self-report as the sole source of data. Gold-standard validation of self-reported scores requires comparison with other measures that are theoretically similar or different (i.e., convergence and divergence; [Campbell & Fiske, 1959](#)). A self-report score that is highly correlated with a previously developed measure of a related construct provides convergent evidence. In contrast, low correlations with a measure of a completely different construct provides divergent evidence for the meaningfulness of the measure. Likewise, divergence across time or informant may also call into question the validity of a self-report. Ideally, the convergent and divergent measures will avoid the methodological weakness of self-report data, which is the point of the multi-trait, multi-method approach. Data from biometric evidence (e.g., fMRI), online behaviour, test scores, attendance, and other physical measurements will show if variation in scores is meaningfully related to (i.e., causally or explanatorily) to constructs that should be sensitive to the construct ([Borsboom et al., 2004](#); [Zumbo, 2009](#)).

Modern neuropsychology attempts to find specific organs of the brain that map to psychological processes (e.g., firing of mirror neurons when physically grasping is related to understanding others; [Kaplan & Jacoboni, 2006](#)). An fMRI study revealed that, when participants were given bogus feedback about performance, brain regions associated with negative affect (i.e., posterior cingulate cortex, the medial frontal gyrus, and the inferior parietal lobule) were activated when norm-referenced feedback was given to low-competence participants and also when criterion-referenced feedback was given to high-competence participants ([Kim, Lee, Chung, & Bong, 2010](#)). Further, performance-approach

goal scores correlated positively with activation in the negative emotion brain areas during norm-referenced feedback.

A review of EEG studies identified strong emphasis in the negative emotion brain areas around feedback-related negativity ([Meyer et al., 2021](#)), but suggested that because multiple brain regions are involved in these processes and the limitations of the EEG method, modifications are needed to understand how feedback relates to the brain. The complexity of the brain can also be seen in how visual representations of familiar objects and people are located in multiple brain locations ([Quiroga et al., 2005](#)). While laboratory studies using fMRI or EEG may reveal how the brain interacts with the mind, the challenge is how those studies relate to the complexities of functioning in the real-world environments. There are differences in results when organisms are studied in laboratory glass tubes (i.e., in vitro) and when they are released into living populations (i.e., in vivo) ([Autoimmunity Research Foundation, 2012](#)); let alone how they might behave in a cyber or simulated environment (i.e., in silica). How well human self-reported perceptions about feedback map to brain activity when anticipating or receiving feedback is still not evaluated.

A well-established approach to establishing validity of measures of a psychological construct is analysis of how individual self-reports relate to outcome measures. This is well established in educational testing systems, such as OECD's Program for International Student Assessment (PISA) surveys. [Marsh et al. \(2006\)](#) showed that self-reported psychological constructs (i.e., self-reported interest, self-concept, and self-efficacy) collected in PISA from >100,000 students in 2000 had statistically significant, but modest ($.25 < r < .35$), relationships with performance, with invariance across 25 different jurisdictions. Two things need to be said about this result. First, the relationships are consistent with theoretical expectations of how these psychological constructs function. Second, the effect is modest, in part because in vivo contexts are so complex and because individuals may have variability in the constructs they endorse. Indeed, cluster analysis of motivational variables within assessment

of mathematics performance showed that in almost all of the 12 nations analysed across three waves of data, there were individuals who did not have consistent self-reported scores across the three motivational scores and performance depended more or less on the mix of motivations (Michaelides et al., 2019).

Nonetheless, self-reported scores about one's own psychological phenomena is a fraught domain. Not only do humans suffer from memory problems about their experiences (Schacter, 1999), but they also suffer from ignorance about themselves and their competence (Dunning, Heath & Suls, 2004), in part because being honest about inadequacies or failure may threaten their ego (Boekaerts & Corno, 2005); or among adolescents it may be 'fun' to subvert surveys (Fan et al., 2006). To overcome these threats psychometrics proposes a number of methods outlined below.

Theoretical Grounding

The field of psychometrics has argued extensively about how to establish validity evidence for any measure of psychological phenomena (Kane & Bridgeman, 2022). Prior to Messick (1989), validity tended to be thought of in terms of multiple types (i.e., face, content, concurrent, construct, and predictive). However, contemporary understanding is that validity is a unitary concept best captured as 'construct validity' (Cizek, 2020). Evidence for a degree of validity judgment (e.g., 'preponderance of evidence', 'clear and convincing evidence', 'substantial evidence'; Cizek, 2020, p. 26) is achieved by consideration of the various empirical and theoretical arguments for the proposed interpretation of a measure (Cizek, 2020; Kane, 2006; Messick, 1989).

Validation evidence includes the theoretical and explanatory qualities of the measurement tool or instrument; the stimulus items, prompts, or tasks presented to elicit responses need to be theoretically aligned to expert theoretically informed definitions of a domain that include hypotheses of how the construct will influence responding (American Educational Research Association et al., 2014; Schmeiser & Welch, 2006). This means that considerable effort should have gone into specifying the domain and then developing and testing

the coherence of the various stimuli or prompts used to elicit responses from individuals to that theoretical specification. Pilot studies (International Test Commission, 2018), expert judgement panels (McCoach et al., 2013), participant think aloud studies (van Someren et al., 1994), and cognitive interviews (Karabenick et al., 2007), and so on, are used to demonstrate that there is evidence that the instrument has prima facie alignment with what it is intended to measure.

Reports of how that test or battery is administered and the kinds of data collected are essential to give confidence that the protocols are replicable and theoretically in accord with the domain. Evaluations that test the theoretical preferred model against alternative explanations also provide evidence that the scales are sound (Cronbach, 1988). Hence, evaluation of the internal structure of an inventory should include multiple competing alternatives.

Internal Structure

The psychometric industry takes a scientific approach in which data collection should generate consistent patterns of responses amongst individuals in accordance with their varying responses to a phenomenon. A key constraint on measuring psychological constructs is that they are in and of themselves not directly observable; they are latent (Borsboom, 2005). As such multiple indicators for multiple causes (MIMIC) are used to reduce error in estimation of the strength and direction of attitude, belief, or value and to better represent the phenomenon of interest (Jöreskog & Goldberger, 1975). With sufficient samples and theoretically designed measurements, mathematical modelling of MIMIC response patterns (e.g., estimate of internal reliability, factor analysis) is used to create evidence for the structure and dependability of the proposed scales or factors (Haertel, 2006).

Approaches that provide evidence about the pattern of responses include scale reliability estimation which can be estimated in multiple ways. Although most researchers are familiar with Cronbach's (1951) alpha, extensive research indicates McDonald's (1999) omega and Hancock and Mueller's (2001) coefficient H are superior meth-

ods for establishing reliability of a set of item. Within the Rasch modeling framework, [Wright and Stone's \(1999\)](#) item separation G can be used to claim that items elicit responses coherent with that version of item response theory. Researchers should be aware that very high reliability results (e.g., $\alpha > .90$) can be obtained by writing items that are almost identical in wording (i.e., have high homogeneity) producing a 'bloated specific' ([Cattell & Tsujioka, 1964](#)).

While principal component analysis can identify underlying vectors in a data matrix, psychometric theory, with its emphasis on error, relies on the common factor model to identify shared and unique factors underlying the same data ([Bryant & Yarnold, 1995](#)). Conventional criteria exist to guide interpretation of the statistical results ([Bandalos & Finney, 2010](#)) so that the quality of evidence for the internal structure of a research or measurement tool can be evaluated (i.e., scales with little or poor evidence can be ignored, while those with robust evidence can be used). Once robust measurement of latent constructs is established, scores for each factor can be derived ([DiStefano, Zhu, & Mîndrilă, 2009](#)).

With scale scores, individuals and groups can be distributed by their scores, which allows comparison of one construct to another and the comparison of score differences between groups and over time. These analytic techniques are robustly associated with the field of psychometrics as they address issues of demonstrating statistically that the theoretical expectations have been met. These practices help create an argument for the trustworthiness of the information obtained from humans about themselves and for interpretations and uses of that data.

Replicability

In the spirit of scientific research, replication studies can examine the stability of psychometric properties across samples ([Makel et al., 2012](#)). Statistical techniques such as multigroup confirmatory factor analysis allow researchers to establish whether statistical models of how participants respond to an instrument vary according to whether the new sample is drawn from

the same population or not. Ideally, the psychometric characteristics of the measurement tool should be within chance when applied in a new sample drawn from the same population. Clearly, non-invariance should be expected when samples are from divergent populations. This does not mean the measurement is broken; rather, it suggests that the measurement works differently, or the construct being measured is different across language, age, culture, prosperity, or educational boundaries. Consider the non-invariance found in *Teacher Conceptions of Feedback* inventory between New Zealand and Louisiana which have very different policy frameworks ([Brown, Harris, O'Quin & Lane, 2017](#)). Measurements that are deployed with new samples have greater opportunity to generate validation evidence by overcoming chance artefacts associated with the development of the measurement. Consider the similarity of the *Teacher Conceptions of Assessment* inventory within New Zealand and its lack of invariance across jurisdictions and languages ([Brown, Gebril, & Michaelides, 2019](#)).

An unfortunate side effect of emphasis on and the complexity of statistical and mathematical modelling of the internal structure of a measurement is that validating evidence from external measures tends to be overlooked. Clearly, it is harder to collect evidence from independent samples, to ask participants to complete parallel or divergent measures at the same time as a new one, and even more difficult to collect independent behavioural evidence so as to make a strong case that a new measure is not only psychometrically robust but also has theoretical and empirical evidence for the relationship between what the *mind* reported and what can be seen from the outside. Nonetheless, without the statistical and mathematical evidence of how a new measure actually works it will not be possible to test theoretical claims about how humans think, feel, and behave. Furthermore, given the complexity of factors impinging upon performance, the actual effect of any specific perception may be quite small. When effects are small and responsive to environmental conditions, they are inherently hard to replicate ([Lindsay, 2015](#)).

Challenges

A fundamental problem within psychology is that everything in the life experiences, environments, and physiology of individuals influences everything they think, feel, believe, say, or do. So, it should not surprise us that the impact of any single psychological constructs should be relatively small because it interact with all other things that also matter. Efforts to isolate and understand important psychological factors in human life has unfortunately led to widespread jingle-jangle (i.e., same words with different meanings or same meanings with different words) in the field (Flake & Fried, 2020). Nonetheless, latent trait theory provides us a way into the mind, heart, and mental representations of individuals.

Good psychometric evidence for measurements of any psychological construct should be able to:

1. Demonstrate fidelity to a theoretically robust description of what the construct is, how it functions, and what it should do;
2. Provide evidence that the proposed operationalisation has prima facie credibility against that theory;
3. Demonstrate robust statistical evidence for the coherence of items against the construct design of a measurement model;
4. Provide evidence that the measurements are reproducible from additional samples;
5. Provide evidence that the measurements produce effects on other measures (including self-reports), behaviours, or outcomes that align with theoretical expectations; and
6. Provide evidence that the construct can be manipulated such that measurement scores change and have the theoretically proposed effects.

Readers may be dismayed at the thought that any single report should necessarily achieve all of these things. However, an excellent example of how these concerns can be addressed is visible in Thielsch and Hirschfeld (2019) which in seven studies provided: a theoretical framework, item set development, statistical demonstration of scale or factor properties, demonstrated test-retest reliability, validated the scales against other conver-

gent and divergent measures, experimentally manipulated scores, and created large-sample norms. Indeed, users of any psychometric measure should expect evidence of this kind before settling on the use of a new tool or inventory.

References

- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179-211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Allport, G. W. (1935). Attitudes. In C. Murchison (Ed.), *A Handbook of Social Psychology* (pp. 798-844). Clark University Press.
- American Educational Research Association, American Psychological Association, National Council for Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing (U.S.). (2014). *Standards for Educational and Psychological Testing*. American Educational Research Association.
- Autoimmunity Research Foundation. (2012). *Differences between in vitro, in vivo, and in silico studies*. The Marshall Protocol Knowledge Base. Retrieved 12 November from <http://mpkb.org/home/patients/assessing-literature/in-vitro-studies>
- Bandalos, D. L., & Finney, S. J. (2010). Factor analysis: Exploratory and confirmatory. In G. R. Hancock & R. O. Mueller (Eds.), *The Reviewer's Guide to Quantitative Methods in the Social Sciences* (pp. 93-114). Routledge.
- Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology: An international review*, 54(2), 199-231. <https://doi.org/10.1111/j.1464-0597.2005.00205.x>
- Borsboom, D. (2005). *Measuring the mind: Conceptual issues in contemporary psychometrics*. Cambridge University Press.
- Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, 111(4), 1061-1071. <https://doi.org/10.1037/0033-295X.111.4.1061>
- Brennan, R. L. (Ed.). (2006). *Educational Measurement* (4th ed.). Praeger.
- Brown, G. T. L., Gebriel, A., & Michaelides, M. P. (2019). Teachers' Conceptions of Assessment: A Global Phenomenon or a Global Localism. *Frontiers in Education*, 4(16). <https://doi.org/10.3389/educ.2019.00016>

- Brown, G. T. L., Harris, L. R., O'Quin, C., & Lane, K. E. (2017). Using multi-group confirmatory factor analysis to evaluate cross-cultural research: identifying and understanding non-invariance. *International Journal of Research & Method in Education*, 40(1), 66-90. <https://doi.org/10.1080/1743727X.2015.1070823>
- Bryant, F. B., & Yarnold, P. R. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 99-136). APA.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multi-trait multi-method matrix. *Psychological Bulletin*, 56(2), 81-105.
- Cattell, R. B., & Tsujioka, B. (1964). The importance of factor-trueness and validity, versus homogeneity and orthogonality, in test scales. *Educational and Psychological Measurement*, 24(1), 3-30. <https://doi.org/10.1177/001316446402400101>
- Cizek, G. J. (2020). *Validity: An integrated approach to test score meaning and use*. Routledge.
- Clauser, B. E. (2022). A history of classical test theory. In B. E. Clauser & M. B. Bunch (Eds.), *The history of educational measurement: Key advancements in theory, policy, and practice* (pp. 157-180). Routledge.
- Corballis, M. C. (2002). *From Hand to Mouth: The Origins of Language*. Princeton University Press.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334. <https://doi.org/10.1007/BF02310555>
- Cronbach, L. J. (1988). Five perspectives on the validity argument. In H. Wainer & H. Braun (Eds.), *Test validity* (pp. 3-17). Lawrence Erlbaum.
- Deary, I. J. (2001). *Intelligence: A Very Short Introduction*. Oxford University Press.
- DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research & Evaluation*, 14(20). <https://doi.org/10.7275/da8t-4g52>
- Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, 5(3), 69-106. <https://doi.org/10.1111/j.1529-1006.2004.00018.x>
- Embretson, S. E., & Reise, S. P. (2000). *Item Response Theory for Psychologists*. LEA.
- Fan, X., Miller, B. C., Park, K.-E., Winward, B. W., Christensen, M., Grotevant, H. D., & Tai, R. H. (2006). An exploratory study about inaccuracy and invalidity in adolescent self-report surveys. *Field Methods*, 18, 223-244. <https://doi.org/10.1177/152822X06289161>
- Flake, J. K., & Fried, E. I. (2020). Measurement schmeasurement: Questionable measurement practices and how to avoid them. *Advances in Methods and Practices in Psychological Science*, 3(4), 456-465. <https://doi.org/10.1177/2515245920952393>
- Geisinger, K. F. (Ed.). (2013). *APA Handbook of Testing and Assessment in Psychology*. American Psychological Association.
- Haertel, E. H. (2006). Reliability. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 65-110). Praeger.
- Hancock, G. R., & Mueller, R. O. (2001). Rethinking Construct Reliability Within Latent Variable Systems. In R. Cudeck, S. d. Toit, & D. Sorbom (Eds.), *Structural Equation Modeling: Present and Future - A Festschrift in Honor of Karl Joreskog* (pp. 195-216). Scientific Software International Inc.
- International Test Commission. (2018). ITC Guidelines for Translating and Adapting Tests (Second Edition). *International Journal of Testing*, 18(2), 101-134. <https://doi.org/10.1080/15305058.2017.1398166>
- Jöreskog, K. G., & Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association*, 70(351a), 631-639. <https://doi.org/10.1080/01621459.1975.10482485>
- Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 17-64). Praeger.
- Kane, M., & Bridgeman, B. (2022). The evolution of the concept of validity. In B. E. Clauser & M. B. Bunch (Eds.), *The history of educational measurement: Key advancements in theory, policy, and practice* (pp. 181-203). Routledge.
- Kaplan, J. T., & Iacoboni, M. (2006). Getting a grip on other minds: mirror neurons, intention understanding, and cognitive empathy. *Social neuroscience*, 1, 175-183. <https://doi.org/10.1080/17470910600985605>
- Karabenick, S. A., Woolley, M. E., Friedel, J. M., Ammon, B. V., Blazeovski, J., Bonney, C. R., De Groot, E., Gilbert, M. C., Musu, L., Kempler, T. M., & Kelly, K. L. (2007). Cognitive processing of self-report items in educational research: Do they think what we mean? *Educational Psychologist*, 42(3), 139-151. <https://doi.org/10.1080/00461520701416231>
- Kim, S.-I., Lee, M.-J., Chung, Y., & Bong, M. (2010). Comparison of brain activation during norm-referenced versus

- criterion-referenced feedback: The role of perceived competence and performance-approach goals. *Contemporary Educational Psychology*, 35(2), 141-152. <https://doi.org/10.1016/j.cedpsych.2010.04.002>
- Likert, R. (1932). A Technique for the Measurement of Attitudes. *Archives of Psychology*, 22, 5-55.
- Lindsay, D. S. (2015). Replication in Psychological Science. *Psychological Science*, 26(12), 1827-1832. <https://doi.org/10.1177/0956797615616374>
- Lord, F. M. (1953). The relation of test score to the trait underlying the test. *Educational and Psychological Measurement*, 13(4), 517-549. <https://doi.org/10.1177/001316445301300401>
- Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in Psychology Research. *Perspectives on Psychological Science*, 7(6), 537-542. <https://doi.org/10.1177/1745691612460688>
- Marsh, H. W., Hau, K.-T., Artelt, C., Baumert, J., & Peschar, J. L. (2006). OECD's brief self-report measure of educational psychology's most useful affective constructs: Cross-cultural, psychometric comparisons across 25 countries. *International Journal of Testing*, 6(4), 311-360. https://doi.org/10.1207/s15327574ijt0604_1
- McCoach, D. B., Gable, R. K., & Madura, J. P. (2013). *Instrument Development in the Affective Domain: School and Corporate Applications*. Springer.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed., pp. 13-103). MacMillan.
- Meyer, G. M., Marco-Pallarés, J., Boulinguez, P., & Sescousse, G. (2021). Electrophysiological underpinnings of reward processing: Are we exploiting the full potential of EEG? *NeuroImage*, 242, 118478. <https://doi.org/https://doi.org/10.1016/j.neuroimage.2021.118478>
- Michaelides, M., Brown, G. T. L., Eklöf, H., & Papanastasiou, E. (2019). *Motivational Profiles in TIMSS Mathematics: Exploring Student Clusters across Countries and Time* (Vol. 7). Springer Open & IEA.
- Michell, J. (1999). *Measurement in psychology: A critical history of a methodological concept*. Cambridge University Press <https://doi.org/10.1017/CBO9780511490040>
- Michell, J. (2008). Is Psychometrics Pathological Science? *Measurement: Interdisciplinary Research and Perspectives*, 6(1-2), 7-24. <https://doi.org/10.1080/15366360802035489>
- Popper, K. (1945). *The open society and its enemies: The high tide of prophecy: Hegel, Marx, and the aftermath* (Vol. II). George Routledge & Sons.
- Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature*, 435, 1102-1107. <https://doi.org/10.1038/nature03687>
- Rao, C. R., & Sinharay, S. (Eds.). (2007). *Psychometrics* (1st ed.). Elsevier North-Holland.
- Rupp, A. A. (2008). Lost in Translation? Meaning and Decision Making in Actual and Possible Worlds. *Measurement: Interdisciplinary Research and Perspectives*, 6(1-2), 117-123. <https://doi.org/10.1080/15366360802035612>
- Schacter, D. L. (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist*, 54(3), 182-203. <https://doi.org/10.1037/0003-066X.54.3.182>
- Schmeiser, C. B., & Welch, C. J. (2006). Test development. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 307-353). Greenwood.
- Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagne, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (Vol. 1, pp. 39-83). Rand McNally.
- Thielsch, M. T., & Hirschfeld, G. (2019). Facets of Website Content. *Human-Computer Interaction*, 34(4), 279-327. <https://doi.org/10.1080/07370024.2017.1421954>
- van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. C. (1994). *The Think Aloud Method: A practical guide to modeling cognitive processes*. Academic Press.
- Wright, B. D., & Stone, M. H. (1979). *Best Test Design*. MESA press.
- Zumbo, B. D. (2009). Validity as contextualized and pragmatic explanation, and its implications for validation practice. In R. Lissitz (Ed.), *The concept of validity* (pp. 65-82). Information Age Publishers.

Aprendizagem baseada em projetos, portfólio coletivo e nutrição: integrando saberes no Sistema Único de Saúde (SUS)

Renato Pereira da Silva¹; Ângela Maria Campos Santana²; Ana Íris Mendes Coelho³; Ana Vlândia Bandeira Moreira⁴; Luiza Carla Vidigal de Castro⁵; Tiago Rocha Pinto⁶; Mara Rejane Barroso Barcelos⁷

¹Departamento de Nutrição e Saúde, Universidade Federal de Viçosa, Minas Gerais, Brasil <https://orcid.org/0000-0003-3893-969X>

²Departamento de Nutrição e Saúde, Universidade Federal de Viçosa, Minas Gerais, Brasil <https://orcid.org/0000-0002-5361-5741>

³Departamento de Nutrição e Saúde, Universidade Federal de Viçosa, Minas Gerais, Brasil <https://orcid.org/0000-0001-9098-5584>

⁴Universidade Federal de Juiz de Fora, Instituto de Ciências Biológicas, Juiz de Fora, Minas Gerais, Brasil <https://orcid.org/0000-0001-6899-7378>

⁵Universidade Federal de Viçosa, Departamento de Nutrição e Saúde, Viçosa, Minas Gerais, Brasil <https://orcid.org/0000-0002-7613-1416>

⁶Universidade Estadual Paulista, Faculdade de Medicina de Botucatu, Departamento de Saúde Pública, Botucatu, São Paulo, Brasil <https://orcid.org/0000-0003-4834-2897>

⁷ Universidade Federal do Espírito Santo, Centro de Ciências da Saúde, Vitória, Espírito Santo, Brasil <https://orcid.org/0000-0001-7288-9468>

Citar como: Pereira, R., Campos, A., Mendes, A., Bandeira, M., Vidigal, L., Rocha, T. y Borroso, M. (2023). Aprendizagem baseada em projetos, portfólio coletivo e nutrição: integrando saberes no Sistema Único de Saúde (SUS). *Revista Digital de Investigación en Docencia Universitaria*, 17(2), e1193. <https://doi.org/10.19083/ridu.2023.1193>

Recibido: 17/05/2020. **Revisado:** 04/06/2020. **Aceptado:** 12/05/2023. **Publicado:** 30/06/2023

Resumo.

Introdução: a formação na área de saúde deve contemplar equanamente os domínios psicomotor, cognitivo e afetivo a partir da problematização de situações cotidianas do Sistema Único de Saúde (SUS). **Objetivo:** avaliar a formação do nutricionista a partir da integração de saberes da tríade ensino-serviço-comunidade, a partir da Aprendizagem Baseada em Projetos (ABP) e Portfólio Coletivo (PC) na micropolítica do SUS. **Método:** um “chá de bebê” planejado e executado por estudantes, orientado por um grupo de docentes, destinado a gestantes da rede SUS, oportunizando o trabalho vivo dos conceitos-ferramentas ‘alteridade’, ‘vínculo’, ‘acolhimento’, e ‘gestão participativa’ foi empreendido. **Resultados:** registros no PC evidenciaram articulação entre diferentes saberes e domínios de aprendizagem na formação do nutricionista por competências. **Conclusão:** ABP+PC é uma ferramenta de ensino-aprendizagem, potente e viável para a formação do profissional capaz de atuar no SUS integrando ensino-serviço-comunidade.

Palavras-chave: educação baseada em competências; educação superior; nutrição em saúde pública; sistema único de saúde.

Project Based Learning, Collective Portfolio and Nutrition: integrating knowledge in Unified Health System

Abstract.

Introduction: Graduation in the health area should equally contemplate the psychomotor, cognitive and affective domains from the problematization of everyday situations of the Unified Health System (SUS). **Objective:** To evaluate the training of nutritionists based on the integration of knowledge from the teaching-service-community triad, based on Project-Based Learning (PBL) and Collective Portfolio (PC) in the micropolitics of the SUS. **Method:** A “baby shower” planned and executed by students, guided

***Correspondência:**

Renato Pereira da Silva

renatop.silva@ufv.br



by a group of teachers, aimed at pregnant women in the SUS network, providing the live work of the concept-tools “alterity”, “bond”, “welcoming”, and “participatory management” was undertaken. Results: records in the CP showed articulation between different knowledge and learning domains in the training of nutritionists by skills. **Conclusion:** PBL+CP is a powerful and viable teaching-learning tool for the training of professionals capable of working in SUS integrating teaching-service-community.

Keywords: competency-based education; higher education; public health nutrition; unified health system.

Introdução

Por pouco mais de um século, o “Relatório Flexner” desencadeou uma série de reformulações, pautadas no paradigma biomédico, que se fizeram notar tanto na prática quanto no ensino nos cursos de formação profissional na área de saúde. Contudo, a discussão atual no ensino superior na área de saúde aborda a insuficiência deste modelo na formação de um profissional capaz de traduzir todo o seu conhecimento técnico-científico, por meio de relações sociais com a população que permeiem o processo saúde-doença, em indicadores epidemiológicos favoráveis nas comunidades onde ele está inserido ([Mitre et al., 2008](#); [Pagliosa & Da Ros, 2008](#); [Bravo et al., 2014](#)).

No Brasil, o acesso cada vez mais democrático ao ensino superior, tem evidenciado esta necessidade de extrapolação da educação superior para além do seu domínio estritamente técnico-científico ([Ristoff, 2014](#)). Neste contexto, elementos de relevância social que incidam sobre o processo saúde-doença da população devem ser agregados visando uma aprendizagem significativa, realizada em cenários reais de ensino-aprendizagem, que favoreçam o desenvolvimento equânime e concomitante dos domínios psicomotor, cognitivo e afetivo no estudante ([Ceccim & Feuerwerker, 2004](#)). Contribuem ainda para esta mudança de paradigma no ensino superior em saúde, a adoção, pelos cursos de graduação, de matrizes curriculares e projetos políticos pedagógicos que articulem as tensões no campo do saber (formação como movimento de produzir-se; formação como

território do trabalho vivo; formação como experimentação; e formação como criação), em consonância com as diretrizes curriculares nacionais para formação na área da saúde e também com os princípios doutrinários e organizativos do Sistema Único de Saúde (SUS) ([Brasil, 2001](#); [Bravo et al., 2014](#); [De Negri et al., 2017](#); [Grillo et al., 2019](#)).

A confluência de tais fatores situam a micropolítica do SUS como um espaço privilegiado, onde o singular encontro trabalhador-usuário ocorre ([Abrahão & Merhy, 2014](#); [Moura & Silva, 2015](#)), onde afetos e subjetivações na dimensão profissional da gestão do cuidado são produzidos. Ressalta-se que tais afetos e subjetivações constituem a base para a aprendizagem significativa do processo saúde-doença e uma formação crítica, reflexiva e ética do profissional de saúde ([Mitre et al., 2008](#); [Oliveira et al., 2015](#)).

Nesse contexto, este estudo objetivou avaliar a integração dos saberes da tríade ensino-serviço-comunidade, utilizando a Aprendizagem Baseada em Projetos (ABP) associada ao Portfólio Coletivo na formação do nutricionista dentro da micropolítica do Sistema Único de Saúde (SUS). Relata-se um recorte ensaístico de uma experiência acadêmica de integração de conhecimentos das áreas de Nutrição e Alimentação, Nutrição Clínica, Nutrição Social e Saúde Pública, protagonizada por estudantes na micropolítica do SUS, experienciando a tríade ensino-serviço-comunidade, o desenvolvimento de competências (conhecimentos, habilidades e atitudes) e o currículo oculto na construção do cuidado em saúde e nutrição ([Simões, 2005](#); [Junqueira & Cotta, 2014](#); [Azimpour & Khalilzade, 2015](#)).

Método

Trata-se de um estudo transversal, qualitativo, descritivo e exploratório de dados empíricos gerados por estudantes de uma Instituição Federal de Ensino Superior (IFES), a partir de uma experiência de Aprendizagem Baseada em Projetos (ABP) associada ao Portfólio Coletivo (PC). A coleta de dados foi realizada a partir de dados secundários registrados e descritos no PC. As etapas metodológicas do processo ensino-aprendizagem são apresentadas a seguir.

Figura 1

Etapas da ABP (SIMÕES, 2005)



Aprendizagem Baseada em Projeto e Portfólio Coletivo

A metodologia da ABP foi adotada neste estudo concebendo-se a educação como um ato político, espontâneo, libertador, coletivo, solidário e comprometido, capaz de articular teoria e prática de maneira crítico-reflexiva, na qual liderança, gestão, relação e colaboração interprofissionais são trabalhadas em um contexto social real (Figura 1) (Simões, 2005).

Nesta abordagem a aprendizagem se dá essencialmente a partir de um projeto com aplicação prática, no qual competências e valores são estruturados utilizando métodos sistemáticos de ensino, com acompanhamento do processo e avaliação formativa do estudante (Márquez & Jiménez-Rodrigo, 2014; Queiroz et al., 2016; Zancul et al., 2017).

Visando potencializar a ABP, foi introduzido o portfólio, como um instrumento de aprendizagem, de desenvolvimento de habilidades e de avaliação de competências e da própria aprendizagem, a fim de favorecer a integração de saberes no desenvolvimento de competências para atuação do nutricionista no SUS (Costa et al., 2018).

Desenho e Cenário do Estudo (Etapa de Problematização da ABP)

A ABP foi desenvolvida no contexto da disciplina Gestão de Unidades de Alimentação e Nutrição II, do curso de graduação em Nutrição da Universidade Federal de Viçosa, campus Viçosa, Minas Gerais, Brasil, tendo como prática de ensino o de-

envolvimento de um projeto para a promoção de alimentação saudável de maneira a atender uma determinada coletividade. Essa prática de ensino, vigente desde 2008, consiste do planejamento e execução de um evento, por estudantes da disciplina, tendo como facilitadores o professor coordenador e um técnico de nível superior. Para a edição realizada no 2º semestre/2014, alvo deste estudo, algumas adaptações metodológicas foram necessárias:

- promoção da interação academia-SUS, a partir da inclusão de público-alvo externo (gestantes usuárias e nutricionista do SUS) à academia;
- ampliação da equipe do projeto pela inclusão de 02 docentes facilitadores;
- promoção da interação facilitadores-estudantes de nutrição-estudantes participantes do Programa Nacional de Reorientação da Formação Profissional em Saúde (PRÓ-Saúde)/Programa de Educação pelo Trabalho para a Saúde (PET-Saúde);
- realização de oficinas mensais problematizadoras (03 oficinas no semestre letivo), a fim de desenvolver as competências (conhecimentos, habilidades e atitudes) dos estudantes em situações reais de sua atuação profissional, tendo como base princípios do "Peer Instruction", propostos por Eric Mazur (Zancul et al., 2017).

A questão norteadora do projeto, "Como integrar os conhecimentos das áreas de Técnica Dietética, Gestão de Unidades de Alimentação e

Nutrição, e Saúde Pública no planejamento e execução de um evento especial para contemplar as necessidades nutricionais e de saúde de uma coletividade atendida no SUS?”, foi elaborada a partir destas oficinas. Como resposta foi concebido o evento “chá de bebê” para gestantes do SUS.

Implementação de Ações (Etapa de Desenvolvimento da ABP)

Focando a interação ensino-serviço-comunidade, o evento foi planejado para um público aproximado de 35 pessoas, com duração prevista de 03 horas. Assim, esse público foi constituído por 10 gestantes e 01 nutricionista do SUS, 14 estudantes da disciplina, 02 estudantes de Educação Física/UFV, 04 facilitadores e 04 professores, externos ao projeto, participantes do processo de avaliação. A disponibilidade de recursos foi o fator limitador do número de pessoas.

Durante o processo de planejamento os estudantes da disciplina sentiram necessidade de interagir com outras áreas de saber, visando oferecer um evento diferenciado, proporcionando uma experiência única a seu público. Desta forma, os participantes do PRÓ-Saúde/PET-Saúde auxiliaram os estudantes da disciplina na seleção de unidades de saúde pertencentes às Redes de Atenção à Saúde (RAS) do município, para subsequente seleção das gestantes usuárias do SUS. As gestantes, foram selecionadas, a partir de sua manifestação voluntária, em consultas de pré-natal e atividades educativas coletivas. Suas preferências e restrições alimentares, visando a composição do cardápio do evento e a elaboração das preparações, foram levantadas nestes momentos.

Desse modo, os estudantes de Nutrição atuaram no planejamento do cardápio, construção das fichas técnicas das preparações, elaboração das preparações, considerando as particularidades do período gestacional no que tange alimentação e nutrição, em consonância com o tipo e horário do evento. Os estudantes de Educação Física contribuíram com a realização de uma oficina de exercícios de respiração e relaxamento para gestantes, planejada para ocorrer durante a execução do evento.

Aos facilitadores reservou-se o papel de orientar, intervir pontualmente em intercorrências

não dirimidas pelos próprios estudantes e participar da avaliação formativa por meio do PC solicitado a eles. Ressalta-se que todos os movimentos, sentimentos e impressões experimentados por todos os participantes durante o percurso de execução do projeto foram registrados no portfólio, reafirmando seu potencial como instrumento de ensino-aprendizagem ([Costa et al., 2018](#)).

Realização do Evento e Avaliação do Projeto (Etapa de Síntese da ABP)

A síntese do projeto se deu a partir da execução do evento e avaliação dos dados obtidos com os instrumentos de avaliação, em vigência na disciplina desde o ano de 2008: PC; ‘Ficha de Avaliação do Planejamento e Execução do Evento’ (contendo informações sobre o convite, cardápio e serviço, custo do evento e disponibilidade do portfólio para consulta durante o evento); e ‘Ficha de Avaliação do Estudante por seus Pares e de Autoavaliação’.

O PC foi utilizado como instrumento de avaliação formativa, compartilhada entre estudantes e facilitadores, a cada etapa do evento proposto a fim de analisar a evolução do trajeto percorrido e de registrar os sentimentos e impressões de estudantes, gestantes e demais participantes do evento, ao final da sua execução. A ‘Ficha de Avaliação do Planejamento e Execução do Evento’ foi preenchida pelos facilitadores membros do projeto e 04 membros externos convidados para exercer esta função, conforme realizado anualmente dentro da prática pedagógica utilizada na disciplina. Cada membro da equipe dos estudantes foi avaliado individualmente pelos pares, além de autoavaliar-se, utilizando a ‘Ficha de Avaliação do Estudante por seus Pares e de Autoavaliação’.

Visando assegurar anonimato e sigilo de informações, foi atribuído um número aleatório para cada participante do evento. Esta numeração evidencia que as informações prestadas se originaram de indivíduos diferentes.

Aspectos Éticos

Como o estudo foi elaborado a partir de atividades de uma prática de ensino regular da referida disciplina, utilizando dados secundários obtidos de registros no PC, foi dispensada a apreciação pelo

Comitê de Ética em Pesquisa/Conselho Nacional de Ética em Pesquisa (CEP/CONEP), visto que o estudo configura-se como relato de uma *“atividade realizada com o intuito exclusivamente de educação, ensino ou treinamento sem finalidade de pesquisa científica, de alunos de graduação...”* (Brasil, 2016, p. 2).

Resultados

A ABP possibilitou a resposta à questão norteadora do projeto, a partir da interação ensino-serviço-comunidade.

A realização do evento consistiu em um momento de confraternização entre todos os atores sociais envolvidos, em um ambiente harmonioso, diferenciado e decorado pelos estudantes de Nutrição de acordo com a temática proposta. Toda a dinâmica do evento foi executada com o intuito de acolher as gestantes do SUS, assegurando visibilidade social às sete que compareceram.

Protagonismo e profissionalismo dos estudantes foram evidenciados ao: a) planejarem um cardápio consistente com as necessidades nutricionais e particularidades de saúde e do período gestacional destas mulheres e ao mesmo tempo atenderem aos demais participantes; b) construírem fichas técnicas apropriadas de todas as preparações servidas; c) elaborarem preparações consonantes com a qualidade nutricional, sensorial e higiênico-sanitária e com o tipo e horário do evento; d) criarem e elaborarem manualmente os convites personalizados com a temática do evento.

A oficina para as gestantes, realizada pelos estudantes de educação física durante o evento, proporcionou um momento acolhedor e de integração entre todos os participantes, num contexto diferente do cotidiano, e especialmente dedicado às gestantes.

A articulação com o SUS foi um processo inovador na elaboração do evento realizado. Os desafios e fortalezas identificados na execução do projeto são listados no Quadro 1.

Quadro 1.
Desafios e fortalezas da interação ensino-serviço-comunidade

Desafios do projeto	Resultados alcançados
Aproximar o SUS dos estudantes da disciplina Gestão de Unidades de Alimentação e Nutrição II	• Inovação na prática pedagógica da disciplina, com a inclusão do SUS na ABP; • Inserção dos estudantes em cenários reais de ensino-aprendizagem.
Integrar docentes de diferentes áreas de atuação num projeto único de intervenção com o SUS	• Integração das áreas de Gestão de Unidades de Alimentação e Nutrição, Higiene e Segurança de Alimentos, e Saúde Pública, com ampliação da equipe e desenvolvimento de um projeto multidisciplinar, ampliando as perspectivas de suas ações e de sua abrangência a partir da pluralidade de ideias e integração de conhecimentos; • Inovação educacional na prática docente na área da saúde; • O cenário do SUS como parte da vivência propiciou uma experiência diferenciada para todos os envolvidos integrando ensino-serviço-comunidade.
Multiplicidade de atividades a serem desenvolvidas num tempo determinado e com recursos limitados para a execução do projeto	• O apoio institucional do Departamento de Nutrição e Saúde (DNS) da UFV, em parceria com o SUS, e outras parcerias estabelecidas propiciaram o cumprimento de todas as atividades dentro do planejado; • A utilização da ABP e o suporte dos facilitadores na tomada de decisões e soluções contribuíram para o cumprimento das múltiplas atividades no tempo e com recursos disponíveis.

Necessidade de apoio de atores externos à instituição de ensino (servidores e usuários do SUS), estudantes do curso de graduação em Educação Física e participantes do PET-Saúde e de sensibilizá-los e mobilizá-los para a viabilização do projeto	• Estabelecimento de novas parcerias e vivência do protagonismo pelas gestantes da rede SUS que constituíram o sujeito principal do evento realizado; • Apoio e a participação harmoniosa dos atores sociais contribuíram para o êxito o projeto; • Ampliação da rede de contato dos participantes do projeto e integração de facilitadores, estudantes, profissionais e usuários do SUS contribuíram para proporcionar uma formação universitária capaz de responder às necessidades de saúde do SUS, em consonância com as DCNs.
Limitação de recurso financeiro, material e humano impactando na limitação do número de participantes do evento	• Utilização de habilidades dos membros da equipe e desenvolvimento de novas habilidades obtendo os resultados desejados, com qualidade; • Empreendimento de ações bem sucedidas na busca por patrocinadores para apoiarem a execução do projeto.
Construção do PC pelos estudantes da disciplina	• As orientações dos facilitadores, a disponibilidade, envolvimento e criatividade do grupo propiciaram a construção do PC, evidenciando o trajeto até então percorrido e contribuindo para avaliação formativa.

Fonte: elaborado pelos autores, 2020.

A conjunção entre ABP e PC oportunizou aos estudantes um trabalho harmônico entre os domínios fundamentais da aprendizagem ([Ferraz & Belhot, 2010](#)) conforme observado no Quadro 2:

Quadro 2.

Contribuição da ABP e PC na formação do nutricionista

Domínio Cognitivo	Domínio Psicomotor	Domínio Afetivo
• Imbricação entre os conhecimentos dos ciclos básico e profissionalizante: percebida pelos docentes e autorrelatada pelos estudantes; • Resgate de conhecimentos básicos como, por exemplo anatomia, fisiologia, psicologia, sociologia, técnica dietética e gestão de alimentação coletiva ao selecionar o público 'gestante'; • Avanços nas categorias 'conhecimento', 'compreensão', 'aplicação', 'análise', 'síntese' e 'avaliação' deste domínio a partir do planejamento e execução do evento.	• Execução das preparações e realização de tarefas correlatas (exemplo: confecção de convites, decoração do ambiente) ao evento com criatividade, excelência e zelo; • Aquisição de reflexos, percepções e habilidades físicas e de comunicação não verbal, fora do currículo tradicional, contemplando as categorias deste domínio: 'imitação', 'manipulação', 'articulação' e 'naturalização'.	• Introeção de conceitos-ferramentas como 'alteridade', 'acolhimento', 'vínculo' e 'gestão participativa do SUS' pelos estudantes; • Produção mútua de afetos.

Figura 2
Relatos dos participantes do evento.

Puxa! Nunca imaginei participar de um evento tão bonito assim na vida!
(Gestante 1).
O sorriso de todos e a satisfação em nos servir alegrou o meu dia e o do meu bebê que se alegrou em meu ventre. Nota 1000!
(Gestante 4).
Decoração linda, recepção calorosa e ótimas preparações!
(Servidor do SUS 1).
Mais do que um evento, este deu visibilidade aos invisíveis. Constituiu-se como uma oportunidade aos estudantes de vivenciar cenários e dramas reais de nossa sociedade. Pudemos sentir mais de perto a importância e responsabilidade do curso de Nutrição como protagonista no processo de transformação social necessária ao Brasil
(Docente 3).
A ideia assustadora de realizar um evento para 30 pessoas, com uma equipe de 14 indivíduos tão diferentes e ao mesmo tempo tão iguais foi se tornando realidade. A convivência não foi fácil: pontos de vistas fortes, visões diferentes, pouco dinheiro para cada detalhe planejado, pouco tempo, mas preferimos focar no que cada um tem de bom e fazer acontecer”
(Discente 6).
O dia da realização do evento é um ápice de intensas emoções e, por isso, agradecemos ao Departamento de Nutrição e Saúde por ter a preocupação em colocar disciplinas como essa, que propiciam muito mais do que conhecimento técnico, mas o conhecimento pessoal na grade curricular
(Discente 11).



Por meio do uso de fala livre, registrada no portfólio, foi possível perceber os sentimentos e a produção de afetos dos participantes do evento. A Figura 2 traz alguns relatos expressivos dos participantes.

Os relatos evidenciam como as categorias do ‘domínio afetivo’ (receptividade, resposta, valorização, organização e caracterização) foram plenamente trabalhadas com os estudantes a partir da execução deste projeto (Ferraz & Belhot, 2020). Ademais, conforme registrado no Quadro 2, os domínios ‘cognitivo’ e ‘psicomotor’ também se fizeram presentes na aprendizagem.

Discussão

A inclusão do SUS como cenário real de prática para o desenvolvimento do projeto na disciplina de Gestão de Unidades de Alimentação e Nutrição II mostrou-se viável tendo em vista a excelência com a qual o evento foi realizado. A aproximação

dos usuários do SUS às atividades de ensino proporcionou novos desafios os quais, para serem superados, requereram o desenvolvimento de novos conhecimentos, habilidades e atitudes dos estudantes, viabilizando a integração dos saberes da tríade ensino-serviço-comunidade (Quadro 1). Ressalta-se que essa aproximação atende às Diretrizes Curriculares Nacionais do curso de Nutrição (Brasil, 2001; Bravo et al., 2014), contribuindo para formação de profissionais comprometidos com as necessidades de saúde da população e com o desenvolvimento do SUS.

A conjugação da ABP ao PC, nessa prática pedagógica, mostrou-se adequada para o trabalho dos conceitos-ferramentas utilizados para o desenvolvimento do projeto com os estudantes. A intencionalidade da escolha destas ferramentas visou o trabalho para além da apreensão dos conhecimentos técnico-científicos, buscando a aquisição de outras competências importantes para a formação do nutricionista (Quadro 2). Nes-

te contexto, sentimentos vivos como 'alteridade', 'acolhimento' e 'vínculo', foram possibilitados e externalizados a partir do corpo desta experiência, sob a lógica de humanização no SUS ([Oliveira et al., 2015](#); [Passos & Carvalho, 2015](#); [Bark et al., 2018](#)). Assim, o evento, além da confraternização e promoção da alimentação saudável inseridas na atividade acadêmica, alcançou um status de um chá de bebês real, inclusive cumprindo sua função social, com iniciativa dos estudantes de arrecadação e distribuição de fraldas às gestantes participantes.

Toda a importância e potência da micropolítica do trabalho nas organizações de saúde corroborando com o ensino superior, enfatizando o uso de tecnologias leves na gestão do cuidado, e por conseguinte fomentando a humanização da atenção à saúde no *locus* onde esta é efetivamente produzida ([Castiel et al., 2014](#)) foi revelada, aprendida e apreendida pelos estudantes com a execução deste projeto, conforme observado nos registros realizados no PC. Ainda por meio desses registros foi possível extrair informações acerca da satisfação das gestantes e de outros participantes do projeto com o evento.

A análise dos resultados evidencia o alcance dos objetivos de maneira positiva, com o enfrentamento favorável dos desafios identificados no contexto do projeto (Quadro 1). Possivelmente, contribuíram para os resultados encontrados as fortalezas a seguir: a) boa aceitação da proposta pelo SUS de integrar o serviço com a universidade, facilitando a aproximação com a comunidade; b) desenvolvimento de novas habilidades nos domínios cognitivo, psicomotor e afetivo (Quadro 2), a partir do envolvimento de pessoas de diferentes áreas de atuação, trabalhando em sincronia nas múltiplas atividades do projeto.

Os agenciamentos vivenciados pelos estudantes, a percepção e a compreensão do papel de cada ator do SUS (o profissional de saúde, na dimensão profissional; o gerente da equipe, na dimensão organizacional e; o gestor do sistema, na dimensão sistêmica) foram fundamentais para o trabalho vivo da participação popular e do controle social (um dos princípios organizativos do SUS). Além de fortalecer, deste modo, também a gestão participativa do SUS, junto às gestantes deste projeto, tendo como premissa a alimentação saudável,

para além da nutrição dos indivíduos, promoveu a aproximação de pessoas e culturas, a socialização de conhecimentos e a convivência entre seres humanos, embasando a formação do "cuidado de si" e a corresponsabilização em saúde na produção do comum, sem o risco de se produzir mais do mesmo, dentro do modelo de produção social da saúde ([Ceccim & Feuerwerker, 2004](#); [Moura & Silva, 2015](#); [Castiel et al., 2014](#); [Ely & Toassi, 2018](#)).

Ações de empreendedorismo também foram trabalhadas neste projeto com os estudantes. Gestão de recursos financeiros e busca ativa por patrocinadores externos e parcerias/integração com estudantes de outros cursos de graduação da IFES foram realizadas. A conexão com estudantes de outros cursos propiciou o desenvolvimento de uma educação interprofissional ([Ely & Toassi, 2018](#); [Lawlis et al., 2019](#)), no momento em que o SUS, como cenário de prática, proporciona um espaço de compartilhamento de percepções e experiências, necessário para a sua formação. Ressalta-se que tais afetamentos impactam, para além da formação acadêmica, na educação continuada dos profissionais da rede SUS e na qualidade de vida de seus usuários.

O estudo evidenciou que a integração ensino-serviço-comunidade foi bem sucedida e revelou a importância de superar os novos desafios, por todos os envolvidos, que foram desencadeados com a aproximação das atividades acadêmicas com o SUS.

A metodologia empregada constituiu-se como uma possibilidade de realização de um processo ensino-aprendizagem significativo na formação do nutricionista, integrando demandas de saúde da comunidade e da rede SUS e contemplando a tríade ensino-serviço-comunidade. O impacto positivo do uso conjunto da ABP e PC foi observado nos estudantes e experimentado pelos docentes, uma vez que esta integração se apoia na relação dialógica entre educando e educador.

Conhecimento e aprendizado foram construídos e mediados a partir dos agenciamentos 'estudante-estudante', 'estudante-facilitador', 'estudante-servidor do SUS', 'estudante-usuário do SUS', 'facilitador-facilitador', 'facilitador-servidor do SUS' e 'facilitador-usuário do SUS', integrando-se os quatro pilares da educação, propostos por

Jacques Delors ([Brasil, 2001](#); [Silva, 2008](#); [Abrahão & Merhy, 2014](#); [De Negri et al. 2017](#)). Paralelamente, os registros no PC, evidenciaram a presença e relevância do 'currículo oculto' (Azimpour & Khalilzade, 2015), evidenciando aos docentes que formas de melhor explorá-lo na formação generalista, humanista, crítica e reflexiva do nutricionista, devem ser buscadas. O estudo possibilitou o fornecimento de informações preciosas para a discussão departamental acerca da revisão da matriz curricular e do projeto político-pedagógico do curso de Nutrição em questão, centrada na perspectiva da inserção do estudante nos cenários reais de ensino-aprendizagem e possibilitou também, de maneira secundária, a compreensão viva do conceito de 'integralidade em saúde' e do desenho e *modus operandi* das RAS no SUS.

Referências

- Abrahão, A.L., Merhy, E. E. (2014). Formação em saúde e micropolítica: sobre conceitos-ferramentas na prática de ensinar. *Interface*, Botucatu, v. 18, n. 49, p. 313-324. <https://doi.org/10.1590/1807-57622013.0166>.
- Azimpour, E., Khalilzade, A. (2015). Hidden Curriculum. *World Essays Journal*, Faisalabad, 3(1), 18-21.
- Barck, M. M., Posanski, M., Oliveira, K.V., Brancher, J.A., Kriger, L., Gabardo, M.C.L. (2018). Alterity and empathy: essential virtues for the training of the dentist. *Revista da ABENO*, Londrina, 18(2), 104-101. <https://doi.org/10.30979/rev.abeno.v18i2.579>.
- BRASIL. Conselho Nacional de Educação. Câmara de Educação Superior (2001). *Resolução CNE/CES 5/2001*. Diário Oficial da União, Brasília, 9 de novembro de 2001. Seção 1, p. 39.
- BRASIL. Ministério da Saúde. Conselho Nacional de Saúde (2016). *Resolução nº 510, de 07 de abril de 2016*. Diário Oficial da União, Brasília, 24 de maio de 2016. Seção 1, p. 44-6.
- Bravo, V.A.A., Cyrino, E.G., Azevedo, M.A.R. (2014). Ensino na Atenção Primária à Saúde e as Diretrizes Curriculares Nacionais: o papel do projeto político-pedagógico. In: Cyrino, A.P., Godoy, D., Cyrino, E.G. *Saúde, ensino e comunidade: reflexões sobre práticas de ensino na atenção primária à saúde*. (pp. 25-48). São Paulo: Cultura Acadêmica.
- Castiel, L.D., Ferreira, M.S., Moraes, D. R. (2014). Os riscos e a promoção do autocontrole na saúde alimentar: moralismo, biopolítica e crítica parresíasta. *Ciência & Saúde Coletiva*, Rio de Janeiro, 19(5), 1523-1532, 2014. <http://dx.doi.org/10.1590/1413-81232014195.06212013>.
- Ceccim, R.B., Feuerwerker L.C.M. (2004). O quadrilátero da formação para a área da saúde: ensino, gestão, atenção e controle social. *Physis*, Rio de Janeiro, 14(1), 41-65.
- Costa, G.D., Driessen, E., Silva, L.S., Campos, A.A.O., Costa, T. M. T., Donateli, C.P., Cotta, R. M. M. (2018). Collective portfolio: assessment of teaching and learning in health undergraduate courses. *Ciência & Saúde Coletiva*, Rio de Janeiro, 23(11), 3779-3787. <https://doi.org/10.1590/1413-812320182311.27072015>.
- De Negri, S.T., Amestoy, S.C. Heck, R.M. (2017). Reflexões sobre a história da nutrição: do florescimento da profissão ao contexto atual da formação. *Revista Contexto & Saúde*, Ijuí, 17(32), 75-84. <https://doi.org/10.21527/2176-7114.2017.32.75-84>.
- Ely, L.I., Toassi, R.F.C. (2018). Integration among curricula in Health professionals' education: the power of interprofessional education in undergraduate courses. *Interface*, Botucatu, 22(2), 1563-1575 <https://doi.org/10.1590/1807-57622017.0658>.
- Ferraz, A.P.C.M., Belhot, R.V. (2010). Taxonomia de Bloom: revisão teórica e apresentação das adequações do instrumento para definição de objetivos instrucionais. *Gestão & Produção*, São Carlos, 17(2), 421-431. <http://dx.doi.org/10.1590/S0104-530X2010000200015>.
- Grillo, L.P., Ibarra, L.S., Mezadri, T. (2019). Conceito ampliado de saúde na formação dos profissionais de educação física. *Saúde e Pesquisa*, Maringá, 12(2), 265-273. <https://doi.org/10.17765/2176-9206.2019v12n2p265-273>.
- Junqueira, T.S., Cotta, R.M.M. (2014). Matriz de ações de alimentação e nutrição na Atenção Básica de Saúde: referencial para a formação do nutricionista no contexto da educação por competências. *Ciência & Saúde Coletiva*, Rio de Janeiro, 19(5), 1459-1474. <http://dx.doi.org/10.1590/1413-81232014195.11932013>.
- Lawlis, T., Torres, S.J., Coates, A.M., Clark, K. Charlton, K.E., Sinclair, A.J., Wood, L.G., Devine, A. (2019). Development of nutrition science competencies for undergraduate degrees in Australia. *Asia Pacific Journal of Clinical Nutrition*, Beijing, v. 28, n. 1, p. 166-176, 2019. [http://dx.doi.org/10.6133/apjcn.201903_28\(1\).0022](http://dx.doi.org/10.6133/apjcn.201903_28(1).0022).
- Márquez, E. Jiménez-Rodrigo, M. L. (2014). Project-based le-

- arning in virtual environments: a case study of a university teaching experience. *Revista de Universidad y Sociedad del Conocimiento*, Barcelona, 11(1), 76-90. <http://dx.doi.org/10.7238/rusc.v11i1.1762>.
- Mitre, S.M., Siqueira-Batista, R., Giradi-de-Mendonça, J.M., Morais-Pinto, N.M., Meirelles, C.A.B., Pinto-Porto, C., Moreira, T., Hoffmann, L.M.A. (2008). Metodologias ativas de ensino-aprendizagem na formação profissional em saúde: debates atuais. *Ciência & Saúde Coletiva*, Rio de Janeiro, 12(supl. 2), 2133-2144. <https://doi.org/10.1590/S1413-81232008000900018>.
- Moura, R.F.S. de, SILVA, C.R.C. e. (2015) Afetividade e seus sentidos no trabalho do agente comunitário de saúde. *Physis*, Rio de Janeiro, 25(3), 993-1010. <https://doi.org/10.1590/S0103-73312015000300016>.
- Oliveira, J.R., Albuquerque, M.C.S., Brêda, M.Z., Barros, L.A., Lisboa, G.L.P. (2015). Concepções e práticas de acolhimento apresentadas pela enfermagem no contexto da atenção básica à saúde. *Revista de Enfermagem da UFPE on line*, Recife, 9(supl. 10), 1545-1555. DOI: 10.5205/reuol.8463-73861-2-SM.0910sup201522.
- Pagliosa, F.L., Da Ros, M.A. (2008). O Relatório Flexner: para o bem e para o mal. *Revista Brasileira de Educação Médica*, Brasília, 32(4), 492-499. <https://doi.org/10.1590/S0100-55022008000400012>.
- Passos, E., Carvalho, Y.M. (2015). A formação para o SUS abrindo caminhos para a produção do comum. *Saúde e Sociedade*, São Paulo, 24(supl. 1), 92-101. <https://doi.org/10.1590/s0104-12902015s01008>.
- Queiroz, J.R.C., Oliveira, M.F.J., Souza, D.M.C., Araújo, A.L.L.S., Fuscilla, M.A.P. (2016). Aprendizagem por projeto e inovação tecnológica: união por competências. *Revista da ABENO*, Londrina, 16(2), 2-6. <https://doi.org/10.30979/rev.abeno.v16i2.243>.
- Ristoff, D. (2014). O novo perfil do campus brasileiro: uma análise do perfil socioeconômico do estudante de graduação. *Avaliação*, Campinas, 19(3), 723-747. <https://doi.org/10.1590/S1414-40772014000300010>.
- Silva, L.R. (2008). Unesco: Os quatro pilares da "educação pós-moderna". *Inter-Ação*, Goiânia, 33(2), 359-378.
- Simões, J.D. (2005). Pedagogia de projetos. Disponível em: <http://smeduquedecaxias.rj.gov.br/nead/Biblioteca/Formação%20Continuada/Artigos%20Diversos/simoes-pedagogia.pdf>
- Zancul, E.S., Souza-Zomer, T.T., Cauchick-Miguel, P.A. (2017). Project-based learning approach: improvements of an undergraduate course in new product development. *Production*, São Paulo, 27(spe:e20162252). <http://dx.doi.org/10.1590/0103-6513.225216>

El COVID-19 y la valoración del proceso enseñanza - aprendizaje en estudiantes de derecho

Constanza Astudillo Meza*

¹Facultad de Derecho, Universidad Santo Tomas, Antofagasta, Chile <https://orcid.org/0000-0001-9796-3492> castudillom@santotomas.cl

Citar como: Astudillo Meza, C. (2023). El COVID-19 y la valoración del proceso enseñanza - aprendizaje en estudiantes de derecho. *Revista Digital de Investigación en Docencia Universitaria*, 17(2), e1279. <https://doi.org/10.19083/ridu.2023.1279>

Recibido: 04/09/2020. **Revisado:** 06/12/2020. **Aceptado:** 30/05/2023. **Publicado:** 30/06/2023.

Resumen

Introducción: la pandemia COVID-19 modificó todos los aspectos de la vida, incluso en el ámbito educativo, donde instituciones de educación superior implementaron clases por medio de plataformas digitales en todas las áreas del conocimiento, incluyendo el Derecho. **Objetivo:** escribir la valoración de estudiantes de la carrera de Derecho respecto del proceso de enseñanza - aprendizaje virtual e innovación de la docencia en el marco de la pandemia de COVID-19 en una institución de educación superior en el norte de Chile. **Método:** se aplicó un cuestionario exprofeso orientado a evaluar la valoración de los participantes respecto del proceso de enseñanza - aprendizaje virtual e innovación de la docencia en el marco de la pandemia de COVID-19. Se realizaron análisis descriptivos y diferencias de proporciones. **Resultados:** la valoración respecto del proceso de enseñanza-aprendizaje fue mayoritariamente regular. Por su parte dentro de las ventajas destacaron la organización del tiempo y la posibilidad de visualizar las grabaciones de las clases. **Conclusiones:** existe la necesidad de abordar factores que influyen en la calidad del aprendizaje, cuyas causas son ajenas a lo académico.

Palabras clave. pandemia; educación a distancia; aprendizaje a distancia; universidad; Chile.

The COVID-19 and the Assessment of the Teaching - Learning Process in Law Students

Abstract

Introduction: The COVID-19 pandemic modified all aspects of life including the educational field, where higher education institutions implemented classes through digital platforms, in all areas of knowledge, including Law. **Objective:** To describe the assessment of law students regarding the virtual teaching-learning process and teaching innovation in the context of the COVID-19 pandemic in a higher education institution in northern Chile. **Method:** A questionnaire was applied to evaluate the participants' evaluation of the e-learning and teaching innovation process in the context of the COVID-19 pandemic. Descriptive analyses and differences in proportions were performed. **Results:** The evaluation of the teaching-learning process was mostly regular. On the other hand, among the advantages, the organization of time and the possibility of visualizing the class recordings were highlighted. **Conclusions:** There is a need to address factors that influence the quality of learning, the causes of which are not academic.

Keywords: pandemic; distance education; distance learning; university; Chile.

*Correspondencia:

Constanza Astudillo Meza
castudillom@santotomas.cl



Introducción

La enseñanza del Derecho se ha desarrollado desde un enfoque tradicional, es decir, unidireccional, donde es el docente quien recita los contenidos. En este sentido, es el docente el centro del proceso de enseñanza - aprendizaje. Lo anterior responde no sólo a la disciplina del Derecho, sino al modelo conductivista presente por mucho tiempo en la educación en todos sus niveles. Este enfoque educativo, abordado desde la psicología, tiene como presupuesto estudiantes acríticos, pasivos y disciplinados, con un rol secundario dentro del proceso de enseñanza - aprendizaje ([Mazorco Irueta, 2010](#)). Bajo este paradigma, los estudiantes no desarrollan el pensamiento crítico y se limitan a reproducir contenidos, siendo el conocimiento algo unilateral desde el profesor.

Particularmente en Derecho, la enseñanza se caracterizaba en transmitir o reproducir contenidos, del modo más exacto posible ([Valle, 2006](#)). Sin perjuicio de lo anterior, la enseñanza del Derecho ha sufrido profundos cambios en el último tiempo, mudando desde este modelo tradicional o bancario, en palabras de [Paulo Freire \(2005\)](#), y unidireccional de enseñanza basado en la clase magistral ([Elgueta & Palma, 2014](#)), hacia un modelo basado en el desarrollo de competencias, dialogante entre el profesor y el estudiante, orientado a mejorar el perfil de egreso de los estudiantes y dotarlos de competencias genéricas y específicas para el ejercicio de la profesión de abogado.

Paralelamente se ha avanzado, de manera sustancial, en la incorporación de metodologías y estrategias de enseñanza y aprendizaje orientadas a potenciar el saber conceptual y el saber hacer del estudiante ([Salas, 2016](#)). Estos saberes, requieren de técnicas que son llamadas activas y que están centradas en la construcción de conocimientos significativos y profundos para los estudiantes, quienes participan en los procesos como agentes activos en la construcción y reconstrucción del conocimiento y no como simples receptores ([Universidad Santo Tomás, 2019](#)).

A pesar de que se aprecia un avance con la incorporación del enfoque constructivista, de enseñanza y la inclusión de metodologías activas, un número considerable de docentes es renuente a

modificar sus prácticas pedagógicas, y replican el modelo con el que fueron educados, que en la mayoría de los casos es el tradicional. En virtud de ello, las instituciones de educación superior ponen a disposición de sus profesores capacitaciones en el ámbito de la pedagogía universitaria, con el fin de mejorar el proceso pedagógico y poder aplicar técnicas y metodologías activas - participativas. En este contexto, el centro del proceso de enseñanza-aprendizaje está orientado a los estudiantes. Adicionalmente se promueve la incorporación y el uso de las tecnologías de la información y conocimiento (en adelante TIC) que en el contexto del COVID-19 han resultado de suma utilidad, pues la sociedad ha debido adaptarse de manera obligatoria a los cambios producto de la pandemia que ha incidido en todos los ámbitos de la vida cotidiana.

Aspectos como el trabajo, familia y educación se han visto trastocados ante la implementación de medidas sanitarias frente al estado de emergencia, particularmente el mantener la distancia social, la recomendación más consensuada y reiterada. Este escenario condujo al uso de medios de comunicación remotos y telepresenciales, utilizados transversalmente en todos los niveles de educación, incluida la educación superior, que comenzó a impartirse desde plataformas digitales. Este formato posibilita la grabación de las clases, permitiendo que el estudiante repita sus actividades, visualice los errores sin ser cuestionado por esto, contribuyendo de ese modo a una mejoría en su rendimiento académico ([Guerrero & Flores, 2009](#)). Resulta importante destacar que en la actualidad las TIC tienen variadas ventajas y cada vez es más frecuente que los perfiles de egreso declaren como competencia genérica el uso de estas. De hecho, se ha reportado que el uso de las TIC favorece la creatividad, el trabajo en grupo y otras habilidades cada vez más valoradas en el mercado laboral ([Marcello-Ruiz et al., 2019](#)).

Lo expuesto podría resultar un obstáculo en la enseñanza del Derecho que, como se ha afirmado, se imparte aún regularmente de modo tradicional, existiendo una clara verticalidad en la relación alumno - profesor, donde el énfasis está en la enseñanza y no en el aprendizaje ([Veas & González, 2015](#)). La metodología tradicional en la que se imparten las clases presenciales en Derecho po-

dría resultar incompatible con las clases virtuales, pues en estas se requiere un rol proactivo del alumno, orientado al desarrollo de habilidades de autogestión, autorregulación y motivación intrínseca ([Moreno Candil et al., 2021](#)). En consecuencia, el docente debe ser capaz de propiciar un ambiente de aprendizaje que estimule y fomente el pensamiento crítico y la capacidad de análisis de los alumnos. En este escenario, el rol del académico es vital, por ser el encargado de desafiar el pensamiento del estudiante y apoyar el proceso de construcción de conocimiento, a diferencia del papel que ejercía anteriormente centrado en ser un transmisor de conocimientos ([Campillay Briones & Meléndez Araya, 2015, p.3](#)).

La importancia de este estudio radica en que permitirá indagar las percepciones de los estudiantes de Derecho acerca de su proceso de enseñanza y aprendizaje, considerando principalmente que, si bien las universidades lograron implementar en Chile la modalidad virtual de un modo rápido y eficaz, éste no ha estado exento de críticas por parte del alumnado. Con respecto a lo anterior, el Informe del Instituto Internacional de la Unesco para la Educación Superior en América Latina y el Caribe (2020) señala que las dificultades experimentadas en esta modalidad van desde la baja conectividad y la falta de contenido en línea adecuado con los planes de estudio nacionales, hasta un profesorado no preparado para esta nueva modalidad.

Sumado a lo anterior, en el área del Derecho aún existen docentes que no utilizan recursos tecnológicos para sus clases; pues, pese a las capacitaciones en el ámbito de la pedagogía universitaria, continúan con el modelo clásico de enseñanza. Ello, en el contexto de enseñanza a distancia y remota, podría dificultar el proceso de enseñanza - aprendizaje para los estudiantes, debido al escaso uso de tecnologías y aplicativos digitales, porque se utilizaban de manera complementaria y parcial, es decir, como apoyo instrumental a la tarea presencial ([Guzmán Droguett et al, 2021](#)). Cobra especial relevancia lo señalado ya que la educación a distancia requiere de un docente autónomo y autosuficiente.

Resulta necesario recalcar que la evidencia generada con respecto a la percepción de estudiantes de Derecho acerca de su proceso de enseñanza - aprendizaje en escenarios de virtualidad, a pro-

pósito de la pandemia del COVID-19, es inexistente. Lo que existe se ha concentrado en indagar y medir el impacto en la educación superior en general. Por ello, el objetivo de este trabajo es describir la valoración de los estudiantes de la carrera de Derecho respecto del proceso de enseñanza - aprendizaje virtual e innovación de la docencia en el marco de la pandemia de COVID-19 en una institución de educación superior en el norte de Chile.

Método

Diseño

Estudio de tipo no experimental, transversal y descriptivo.

Participantes

El total de la muestra estuvo compuesta por 53 estudiantes de la carrera de Derecho de una Universidad de la zona norte de Chile, de los cuales 30 fueron mujeres (56,6%) y 23 fueron hombres (43,4%).

Instrumento

Se utilizó un cuestionario exprofeso orientado a evaluar la valoración de los participantes respecto del proceso de enseñanza - aprendizaje virtual e innovación de la docencia en el marco de la pandemia de COVID-19. El instrumento estuvo compuesto de cinco preguntas. La primera pregunta, orientada a la valoración del proceso de enseñanza - aprendizaje, tuvo un rango de respuesta de 1 a 4 donde 1 indicaba malo y 4 excelente. Las preguntas 2 y 3, orientadas a la innovación y mantención de la virtualidad, tuvo un rango de respuesta dicotómico. Las preguntas 4 y 5 estuvieron orientadas a determinar ventajas y desventajas de la virtualización de las asignaturas. El contenido del instrumento fue sometido a validación de contenido por parte de un comité de expertos compuesto por seis académicos, tres de ellos expertos en Derecho y tres expertos en educación superior. Luego de considerar las observaciones realizadas por los expertos se depuró el instrumento final.

Procedimientos y Aspectos éticos

La presente investigación fue revisada y aprobada por el Comité Ético Científico de la Universidad

patrocinante, bajo el rol o código de aceptación 01-20. Esta investigación forma parte de un proyecto de mayor dimensión que incluye estudiantes de Derecho que cursan la asignatura de Clínica Jurídica impartida en 11 escuelas a lo largo de Chile. La selección de los 53 participantes fue mediante la técnica de bola de nieve combinada con un muestreo intencionado por conveniencia. Cabe destacar, que cada uno de los participantes al aceptar su participación firmaron dos copias de consentimiento informado, una copia para respaldo del equipo de investigación y la otra para los participantes. La recolección de los datos se llevó a cabo durante el segundo semestre del año 2020.

Análisis estadísticos

Los datos fueron ingresados a una planilla en el programa estadístico IMB SPSS V24. Posteriormente se realizaron los análisis de distribución

de frecuencias para evaluar los resultados de los participantes.

Resultados

En la tabla 1 se describe la valoración de los estudiantes respecto de su proceso de enseñanza-aprendizaje en el contexto de COVID-19. En este sentido, se destaca que solo el 5,7% de los participantes valoró el proceso como excelente.

En la Tabla 2 se describe la percepción de estudiantes respecto de la innovación de profesores en sus prácticas habituales de enseñanza aprendizaje. En este sentido, se reporta que mayormente los estudiantes no consideraron que sus profesores innovaran en sus prácticas pedagógicas.

La Tabla 3 describe la consideración por parte de los estudiantes respecto de la posibilidad de continuar con la modalidad virtual durante los próximos meses. Una amplia mayoría considera que es posible.

Tabla 1
Valoración del Proceso de Enseñanza- Aprendizaje en el Contexto COVID-19

Respuestas	n	%
Excelente	3	5.7
Bueno	21	39.6
Regular	23	43.4
Malo	6	11.3

Tabla 2
Percepción de Innovación de Prácticas de Enseñanza – Aprendizaje Durante la Pandemia COVID-19

Respuestas	n	%
Sí	21	39.6
No	32	60.4

Tabla 3
Continuidad con la Modalidad Virtual en los Próximos Semestres

Respuestas	n	%
Sí	45	84.9
No	8	15.1

La Tabla 4 indica las ventajas percibidas por los estudiantes respecto de su proceso de aprendizaje durante la pandemia de COVID-19. En este sentido, las dos ventajas que reportaron mayor frecuencia fueron una mayor organización del tiempo libre y la reiteración de contenido por la grabación de clases.

La tabla 5 indica las desventajas percibidas por los estudiantes respecto de su proceso de aprendizaje durante la pandemia de COVID-19. En esta línea, la desventaja que reportó mayor frecuencia fue la de falta de concentración y atención.

Tabla 4
Ventajas en su Aprendizaje en el Contexto de COVID-19

Respuestas	n	%
Mayor organización del tiempo disponible	9	16.9
Flexibilidad	1	1.9
Mayor interacción que en el sistema presencial	1	1.9
Disponibilidad inmediata del material de estudio	1	1.9
No salir de casa	1	1.9
Ninguna	5	9.4
Pasar tiempo en familia	1	1.9
Reiteración de contenidos mediante clases grabadas	9	16.9
Se evitan los contagios	4	7.5
Disponer de mayor tiempo para estudiar	8	15.1
Mayor comodidad	4	7.5
Mayor aplicación de la materia porque se realizan más casos prácticos	3	5.7
Mayor dedicación	1	1.9
Asistencia a clases	1	1.9
No responde	1	1.9
Se eliminan los tiempos de traslado hacia la universidad	3	5.7

Tabla 5
Desventajas Percibidas en el Aprendizaje Durante COVID-19

Respuestas	n	%
Falta de espacio en los hogares para prestar atención a las clases.	5	9.4
Falta de concentración y atención	10	18.8
Ansiedad, estrés y falta de motivación	6	11.3
Problemas de conectividad a internet	11	20.7
Falta o inexistencia de retroalimentación e interacción entre docente y estudiantes.	6	11.3
Clases muy largas y exceso de carga académica	3	5.7
No poder aplicar la materia a casos prácticos	4	7.5
No aprendo y se dificulta entender	3	5.7
Exposición a la computadora	1	1.9
No existen desventajas	1	1.9
Extensión de cuarentena ocasiona problemas al dormir	1	1.9
Cambio de método de estudio	1	1.9
No responde	1	1.9

Discusión

El objetivo de este estudio fue describir la valoración de los estudiantes de la carrera de Derecho respecto del proceso de enseñanza-aprendizaje virtual e innovación de la docencia en el marco de la pandemia de COVID-19 en una institución de educación superior en el norte de Chile.

Es necesario remarcar que se ha reportado en la literatura que la pandemia generó ciertos cambios que exigieron una adaptación de forma abrupta, pero que desarrollaron experiencia en el camino para generar verdaderos procesos de cambio en la innovación de la enseñanza del Derecho ([Puebla Fortunato, 2021](#)). Siguiendo la idea anterior, si bien se genera la oportunidad para repensar la enseñanza del Derecho, no se puede desconocer que la virtualidad para esta ciencia y para la educación superior en general, impacta en la manera en que los educandos logran adquirir los conocimientos; pese a que la evidencia bibliográfica es escasa debido a lo reciente del tema investigado. En este sentido, dentro de la literatura se ha evidenciado que el uso de las plataformas virtuales y/o las aplicaciones digitales no contribuyeron a la consolidación de los aprendizajes debido a la falta de conocimiento del funcionamiento de las mismas ([Aguilar Gordón, 2020](#); [Elgueta Rosas, 2020](#); [Ruiz, 2020](#)).

Los estudiantes indicaron que su evaluación del proceso de enseñanza - aprendizaje en el contexto de COVID-19 en términos generales es regular. Ello se condice con los hallazgos de quienes evidenciaron una apreciación esencialmente regular ([Ramírez Jaliri et al., 2020](#)). Esta situación puede explicarse por la inexperiencia de los docentes respecto de metodologías para el desarrollo de clases virtuales y a las dificultades de implementar un formato diametralmente distinto al tradicional en un contexto de incertidumbre e improvisación.

En cuanto a la percepción de los estudiantes acerca de la innovación educativa de sus docentes en el periodo de pandemia, la mayoría considera que no incluyeron actividades innovadoras dentro de las clases telepresenciales. La apreciación de los alumnos puede explicarse en razón de que los docentes no realizaron adaptaciones

en su forma de hacer clases presenciales en función a un nuevo contexto dado por la virtualidad. Por ende, lo que esencialmente se produjo fue la reproducción del escenario presencial en modalidad virtual. Esto ha sido descrito previamente respecto de que los estilos de praxis educativa del docente mediado por las TIC en todos los niveles educativos, aún permanecen tácitos y cada profesor trabaja en forma intuitiva sin una formación adecuada ([Ojeda-Beltrán et al., 2020](#)). Lo expresado no obsta a que los profesores deben intentar adecuarse al contexto de cambio, ya que la innovación docente cobra un papel muy relevante y permite que todos los docentes se preocupen por la enseñanza activa que fomenta el diálogo con los estudiantes ([Castellanos-Claramunt, 2020](#)).

Dentro de las ventajas descritas con mayor frecuencia por los estudiantes, destacan la organización del tiempo y la reiteración de contenidos con clases grabadas. Respecto de la organización del tiempo se torna una oportunidad de mejora ya que se ha reportado que los alumnos de educación superior presentan dificultades en la gestión del tiempo ([Jiménez Carrasco, 2015](#)) porque lo distribuyen de manera arbitraria sin priorizar lo urgente y cediendo a otras distracciones. Si bien el tiempo y su organización es un tema transversal, se considera que se vuelve especialmente relevante en los alumnos de educación superior, pues en esta etapa deben desarrollar sus actividades de modo independiente y autónomo, sin que exista un adulto que les haga un seguimiento o los reporte. Por lo tanto, es de importancia que los ambientes de enseñanza - aprendizaje virtualizados logren ser visualizados como una ventaja en este escenario de pandemia y también, posteriormente, pues toda hora empleada en planear eficazmente ahorra de tres a cuatro horas de ejecución y produce mejores resultados ([Jiménez Carrasco, 2015](#)).

En cuanto a la repetición de contenidos como ventaja identificada por los estudiantes en este trabajo, es posible sostener que es una referencia directa al modelo tradicional de enseñanza del Derecho. Enuncian esta posibilidad como positiva, ya que les permitiría memorizar de mejor manera el discurso del profesor. Sin perjuicio de lo expresado ha de tenerse presente que el proce-

so de enseñanza y aprendizaje incluye al docente, pero también al estudiante, quien debe cumplir un rol que le permita ser protagonista y constructor de su conocimiento, siendo el profesor un mero facilitador de contenidos. Este proceso es la instancia donde se involucran ambas partes, sin evolucionar hacia una enseñanza del Derecho con técnicas y metodologías más activas.

En esta línea, si no se considera un rol distinto en los estudiantes, fuera del rol memorístico, pasivo y de preguntas cerradas bajo el enfoque clásico de la enseñanza del Derecho, resulta complicado incluir técnicas o herramientas innovadoras en el contexto de la enseñanza del Derecho ([Bayuelo Schoonewolff, 2015](#)).

Cabe mencionar que, según la apreciación de algunos estudiantes, no se ha generado ninguna ventaja en su proceso de enseñanza y aprendizaje en el contexto de COVID-19. Las razones para estimar esto son variadas, dado que, reportan que no aprenden, que existe conexión inestable a internet y que esta modalidad no les produce ninguna motivación, lo que se traduce en que no identifiquen ventaja alguna. Estas razones causan inequidad y exclusión, ya que los aprendizajes de los estudiantes se producen en contextos socioeconómicos que están en desventaja ([Lovón Cueva y Cisneros Terrones, 2020](#)). De acuerdo con el razonamiento anterior, y ante la desigualdad, es lógico que no se identifiquen ventajas en la enseñanza virtual.

En relación con las desventajas identificadas es posible indicar que se mencionan la falta de concentración, falta de conectividad, ansiedad y estrés y falta de espacio. Es necesario reconocer la importancia de la concentración en el proceso de enseñanza - aprendizaje, considerada como básica en el acercamiento a nuevos contenidos y procesamiento de información, así como en el desarrollo del aprendizaje, la memoria, el lenguaje y la orientación ([Granados Ramos et al, 2016](#)). La ausencia de esta es perjudicial, porque en la actualidad la educación superior está orientada hacia un modelo por competencias. En este sentido, se busca que el estudiante adquiera una habilidad y que manifieste una cierta capacidad o destreza al desarrollar una tarea ([González Ferreras & Wagenaar, 2003](#)). Por lo tanto, esta demostración

solo se puede lograr si el estudiante comprende y para comprender necesariamente tiene que desarrollar su concentración y atención.

Por otro lado, en cuanto a la falta de conectividad a internet, es importante mencionar que el año 2005 se afirmaba que en Chile existía una barrera económica para acceder a un computador y a internet, lo que genera una brecha digital ([Fernández-Medina, 2005, p.84](#)). A 15 años de este diagnóstico, los factores económicos continúan siendo una determinante en la brecha digital; además, se debe considerar que un buen número de usuarios no sabe utilizar las herramientas tecnológicas y programas que proporcionan los computadores. Lo anterior acrece con la irrupción masiva de los *smartphones*, que presentan versiones adaptadas de los programas de una computadora. En consecuencia, la brecha digital no solo abarca la desigualdad en la adquisición y/o acceso a las TIC, sino también se refiere a las habilidades requeridas para utilizar dichas tecnologías en aras de un fin determinado ([Moreno Candil et al., 2021](#)).

Cabe tener en cuenta que la geografía de Chile impide tener una cobertura completa de internet, pues el país tiene más de 4,000 kilómetros de largo con localidades urbanas y rurales. Las comunas urbanas cuentan con una conexión del 67,8% y las comunidades rurales del 16,7% ([Tapia, 2020](#)). Si la realidad en cuanto a la conectividad ya era compleja antes de la pandemia, en la actualidad se dificulta aún más, pues es un hecho público y notorio que ha aumentado el tráfico de datos de internet, enlenteciendo la conexión. Por todo lo expuesto, resulta razonable que los estudiantes identifiquen este hecho como una desventaja, pues muchos de ellos tienen hermanos que también deben ingresar a clases virtuales o padres que están con modalidad de trabajo a distancia desde sus casas, con velocidad insuficiente de internet para el peso de los videos y audio, lo que dificulta el acceso.

También se identifica como desventaja la ansiedad, estrés y falta de motivación. El que los estudiantes universitarios consideren estas emociones como una desventaja, da cuenta de lo problemático de la salud mental en Chile. En la actualidad la autoridad de salud pertinente, implementó un pro-

grama de acompañamiento llamado “Saludable-Mente” en este escenario de COVID-19.

Se ha reportado una serie de reacciones y estados psicológicos observados en la población, incluyendo efectos en la salud mental, las primeras respuestas emocionales de las personas incluyen miedo e incertidumbre extremos (Urzúa et al., 2020; Asociación Latinoamericana de Psicología de la Salud, 2020). Si bien no se puede afirmar que existe un incremento de la ansiedad, pero es posible sostener que resulta inusual que existan proporciones tan altas de sentimientos negativos dentro del contexto de enseñanza a distancia (Castillo Riquelme et al., 2021).

En este marco, los resultados obtenidos están en armonía con los arrojados por la encuesta #EstamosConectados de la organización Educación 2020 donde el 63% de los estudiantes indicó sentir aburrimiento, un 41% manifestó sentir ansiedad o estrés, seguido por frustración y molestia (35%); mientras que un 21% dijo sentir tranquilidad y solo un 3% indicó sentirse “feliz” de estar en casa.

Por último, en lo que respecta a la falta de espacio en los hogares para prestar atención a las clases, se sostiene que esta desventaja indicada por los estudiantes apunta a la noción de ambiente educativo que remite al escenario donde existen y se desarrollan condiciones favorables de aprendizaje (Duarte Duarte, 2003). Si bien la idea de ambiente de aprendizaje no se limita al aula y puede incluir el uso de otros espacios por los alumnos, el hogar no se encuentra acondicionado del mismo modo que una institución educativa, pues está pensado como un lugar de encuentro familiar y descanso. Por lo anterior, parece coherente que asistir a las clases desde dicho lugar sea considerado una desventaja.

Se debe tener en cuenta que la educación virtual no se agota en impartir las clases por medios remotos y replicar la forma en que los profesores realizan la docencia en el aula, pues existen diversas variables para tener en cuenta (nivel del curso, datos de internet, contenidos teóricos o prácticos, etc.). La creencia de recrear lo que se realiza en el aula sin el contexto de pandemia, no hace otra cosa que generar en estudiantes la sensación de falta de innovación por parte de los docentes que, de acuerdo con los datos obtenidos en este

trabajo, se estima en un 60.4%. Los académicos de educación superior deben poner atención a este nuevo contexto de pandemia, y redoblar los esfuerzos para interactuar con los alumnos, pues hay muchos factores ocultos (como conectividad, ansiedad o estrés) que pueden influir en su rendimiento. Lo anterior ha sido detectado como una desventaja en este trabajo.

A modo de conclusión, se debe señalar que la percepción de los estudiantes acerca del proceso de enseñanza y aprendizaje es variada. De esta forma los alumnos consideran el manejo en la organización de sus tiempos, disponer de más tiempo para estudiar y la posibilidad de reiterar la materia mediante las grabaciones de las clases como las mayores ventajas de este sistema virtual. Por otro lado, consideran la mala conexión, falta de concentración, ansiedad y poca retroalimentación del profesor como las principales desventajas. Se puede observar que estas obedecen a factores externos, ajenos al proceso de enseñanza y aprendizaje desde la universidad, y que están presentes en la comunidad estudiantil en general, sin que sean exclusivamente apreciaciones de los estudiantes de Derecho.

Finalmente, resulta relevante indicar que el estudio presentado tuvo como limitaciones el muestreo intencionado por conveniencia, el número de alumnos, y localidades comprendidas, pues la Facultad está presente en otras ciudades de Chile y no solo en el norte, por lo que las apreciaciones podrían ser diversas a las recogidas en este trabajo. Asimismo, en un futuro podrían incluirse más variables, relevantes para el contexto de la educación virtual o en línea. Sin perjuicio de lo anterior, la investigación propuesta conduce a repensar la enseñanza del Derecho, y la relación inexorable de las tecnologías de la información que será un insumo permanente para los docentes de hoy y en el futuro.

Referencias

Aguilar Gordón, F. (2020). Del aprendizaje en escenarios presenciales al aprendizaje virtual en tiempos de pandemia. *Estudios pedagógicos*, 46(3), 213-223. <https://>

- doi.org/10.4067/S0718-07052020000300213
- Asociación Latinoamericana de Psicología de la Salud (2020). La psicología de la salud en el enfrentamiento a la COVID-19 en América Latina. Editorial CEDEM.
- Bayuelo Schoonewolff, P. (2015). La educación y el Derecho en torno a un nuevo paradigma transformador. *Justicia*, 27, 167-184. <https://doi.org/10.17081/just.3.27.326>
- Camarillo Hinojoza, H., & Barboza Regalado, C. (2020). La enseñanza-aprendizaje del derecho a través de una plataforma virtual institucional. *Revista Pedagogía Universitaria y Didáctica del Derecho*, 7(2), 143-166. <https://doi.org/10.5354/0719-5885.2020.57035>
- Campillay Briones, S. & Meléndez Araya, N. (2015). Análisis de impacto de metodología activa y aprendizaje heurístico en asignaturas de ingeniería. *Actualidades Investigativas en Educación*, 15(2), 330-346. <http://doi.org/10.15517/aie.v15i2.18950>
- Castellanos-Claramunt, J. (2020). Innovación docente y TIC desde la perspectiva de la docencia en Derecho. *Revista Pedagogía Universitaria y Didáctica del Derecho*, 7(2), 167-184. <https://doi.org/10.5354/0719-5885.2020.57150>
- Castillo Riquelme, V., Cabezas, Maureira N., Vera Navarro, C., Toledo Puente, C. (2021). Ansiedad al aprendizaje en línea: relación con actitud, género, entorno y salud mental en universitarios. *Revista Digital de Investigación en Docencia Universitaria*, 15(1), e1284. http://www.scielo.org.pe/scielo.php?pid=S2223-25162021000100004&script=sci_abstract
- Duarte Duarte, J. (2003). Ambientes de aprendizaje: una aproximación conceptual. *Estudios pedagógicos*, 29, 97-113. <https://doi.org/10.4067/S0718-7052003000100007>
- Elgueta, M., & Palma, E. (2014). Una propuesta de clasificación de la clase magistral impartida en la facultad de derecho. *Revista Chilena de Derecho*, 41(3), 907-924.
- Elgueta Rosas, M. (2020). ¿Hay alguien ahí? Interacciones pedagógicas con cámaras apagadas en tiempos de pandemia. *Revista Pedagogía Universitaria y Didáctica del Derecho*, 7(2), 1-8. <https://doi.org/10.5354/0719-5885.2020.60556>
- Fernández-Medina, F. (2005). Brecha e inclusión digital en Chile: los desafíos de una nueva alfabetización. *Comunicar*, 12(24), 77-84.
- Freire, P. (2005). *Pedagogía del oprimido*. (2ª ed.). Siglo XXI editores.
- Guerrero, T. & Flores, H. (2009). Teorías del aprendizaje y la instrucción en el diseño de materiales didácticos informáticos. *Educere*, 13(45), 317-329.
- Guzmán Droguett, M., Cuéllar Becerra, C., Faúndez Bastías, M., & Lizama Zamora, C. (2021). Universidad y educación no presencial en contexto de pandemia. *Revista Pedagogía Universitaria y Didáctica del Derecho*, 8(1), 325-346. <https://doi.org/10.5354/0719-5885.2021.58153>
- González Ferreras, J. & Wagenaar, R. (Eds.). (2003). *Tuning Educational Structures in Europe*. Informe final. Fase 1. Universidad de Deusto. <http://bit.ly/2NcO8fZ>
- Granados Ramos, D., Figueroa Rodríguez, S. & Velásquez Durán, A. (2016). Dificultades de atención y competencias de investigación en estudiantes universitarios de psicología. *Enseñanza e Investigación en Psicología*, 21(2), 131-140.
- Instituto Internacional de la UNESCO para la Educación Superior en América Latina y el Caribe. (2020). COVID-19 y Educación superior: De los efectos inmediatos al día después. Análisis de impactos, respuestas políticas y recomendaciones. <http://www.iesalc.unesco.org/wp-content/uploads/2020/05/COVID-19-ES-130520.pdf>
- Jiménez Carrasco, L. (2015). ¿Cómo organizan su tiempo los universitarios? *Revista de Investigación Psicológica*, (14), 97-108.
- Lovón Cueva, M., & Cisneros Terrones, S. (2020). Repercusiones de las clases virtuales en los estudiantes universitarios en el contexto de la cuarentena por covid-19: El caso de la PUCP. *Propósitos y Representaciones*, 8, 1-15. <https://doi.org/10.20511/pyr2020.v8nSPE3.588>
- Marcello-Ruiz, J., Cabrera Almeida, F., Rodríguez-Esparragón, D. & Eugenio, F. (2019). c. En D. De la Cruz, A. Ravelo, C. Travieso, J. Alonso, J. Canino & S. Pérez (Eds.), *Libro de Actas de las VI Jornadas Iberoamericanas de Innovación Educativa en el Ámbito de las TIC y las TAC* (pp. 177-183). Grupo de Innovación Docente: Aplicaciones Tecnológicas para la Enseñanza de las TIC (ATETIC) de la Universidad de Las Palmas de Gran Canaria.
- Mazorco Irureta, G. (2010). Educación en crisis. *Polis Revista Latinoamericana*, 9(25), 269-288. <https://polis.ulagos.cl/index.php/polis/article/view/684>
- Moreno-Candil, D., Bravo-Verdugo, E. R., Burgos-Dávila, C. J., & Grijalva Verdugo, A. A. (2021). Educación superior y pandemia: estudio piloto del cuestionario de experiencia educativa durante pandemia. *Revista Digital de Investigación en Docencia Universitaria*, 15(1), e1347. <https://doi.org/10.19083/ridu.2021.1347>

- Ojeda-Beltrán, A., Ortega-Álvarez, D. D., Boom-Marcano, E. A. (2020). Análisis de la percepción de estudiantes presenciales acerca de clases virtuales como respuesta a la crisis del covid-19. *Espacios: Ciencia, Tecnología y Desarrollo*, 41(42), 81-92. <https://doi.org/10.48082/espacios-a20v41n42p07>
- Puebla Fortunato, C. (2021). El rol docente en la nueva virtualidad: Un debate sobre la formación de destrezas jurídicas en el contexto actual. *Revista Justicia & Derecho*, 4(2), 1-11. <https://doi.org/10.32457/rjyd.v4i2.548>
- Ruiz, E. (2020). La práctica docente universitaria en ambientes de educación a distancia. Tensiones y experiencias de cambio en Girón, J., (Ed.), *Educación y pandemia: Una visión académica* (pp.10-313). Universidad Nacional Autónoma de México.
- Salas, C. (2016). Experiencia de aprendizaje situado y colaborativo en la formación actitudinal de los estudiantes de derecho. En E. Sandoval (Coord.), *Nuevas experiencias en el ámbito de las Ciencias Jurídicas. Actas del III Seminario de enseñanza del Derecho* (pp. 128-147). Ediciones Jurídicas de Santiago.
- Tapia, G. (24 de junio de 2020). Ministra Hutt y zonas rurales: "Uno de los grandes desafíos que tiene Chile es la conectividad digital y de internet". *Radio Agricultura*. <https://www.infraestructurapublica.cl/ministra-hutt-zonas-rurales-uno-los-grandes-desafios-chile-la-conectividad-digital-internet/#:~:text=En%20la%20instancia%2C%20la%20ministra,el%20estado%20tiene%20que%20actuar>
- Universidad Santo Tomás. (2019). Guía para la reflexión sobre la praxis docente: La planificación didáctica, 1-61.
- Urzúa, A., Vera-Villarroel, P., Caqueo-Úrizar, A., & Polanco-Carrasco, R. (2020). La Psicología en la prevención y manejo del COVID-19. Aportes desde la evidencia inicial. *Terapia psicológica*, 38(1), 103-118. <https://doi.org/10.4067/S0718-8082020000100103>
- Varguillas, C., & Bravo, P., (2020). Virtualidad como herramienta de apoyo a la presencialidad: Análisis desde la mirada estudiantil. *Revista de Ciencias Sociales (Ve)*, 36 (1), 219-232.
- Valle, A. (2006). Cultura jurídica y enseñanza del derecho ¿creencias o competencias? *Revista Escuela de Derecho*, (7), 85-98.
- Veas, M. & González, I. (2015). La coevaluación y la evaluación compartida en la enseñanza del Derecho: Análisis de una experiencia en el aula. En K. Becerra (Coord.), *Nuevos paradigmas e innovaciones en la enseñanza del Derecho. Actas II Seminario de Enseñanza del Derecho* (pp. 49-61). Ediciones Jurídicas de Santiago.



UPC
Universidad Peruana
de Ciencias Aplicadas