

Propuesta de cuestionario para explorar la cultura de la evaluación del aprendizaje de docentes universitarios

Questionnaire proposal to explore the culture of learning evaluation of university professors

Por Sofía CONTRERAS ROLDÁN¹ y Juan Carlos RODRÍGUEZ MACÍAS²

Contreras Roldán, S. y Rodríguez Macías, J. C. (2022). Propuesta de cuestionario para explorar la cultura de la evaluación del aprendizaje de docentes universitarios. *Revista RAES*, XIV(25), 138-159.

Resumen

El estudio tuvo como propósito la exploración de la cultura de la evaluación de docentes universitarios, a través del cálculo de dos índices: uno sobre cultura tradicional y otro, de una cultura centrada en el aprendizaje y la mejora educativa. Los índices se crearon a partir de un cuestionario que explora las creencias, actitudes y valores hacia la evaluación del aprendizaje en el aula. La muestra estuvo conformada por 545 docentes de una universidad pública mexicana. Se llevó a cabo un Análisis Factorial Exploratorio para las escalas y después, se calcularon ocho sub índices derivados del factorial, lo cual permitió la construcción de dos índices: uno que sintetizó todas las variables relacionadas con la evaluación sumativa y tradicional; y otro para la formativa y constructivista. En los resultados, se observó que la cultura evaluativa de los docentes se orienta más hacia la formativa y constructivista.

Palabras Clave Educación superior/ evaluación del aprendizaje/ docentes/ evaluación formativa/ cultura evaluativa.

Abstract

The purpose of the study was to explore the culture of the assessment of university teachers, through the calculation of two indices: one on traditional culture and the other, a culture focused on learning and educational improvement. The indices were created from a questionnaire that explores beliefs, attitudes and values towards evaluating learning in the classroom. The sample consisted of 545 professors from a Mexican public university. An exploratory factor analysis was carried out for the scales and then, eight sub-indices derived from the factorial were calculated, which allowed the construction of two indices: one that synthesized all the variables related to the

¹ Instituto de Investigación y Desarrollo Educativo, Universidad Autónoma de Baja California, México/ sofia.contreras@uabc.edu.mx - <https://orcid.org/0000-0003-1987-3445>

² Instituto de Investigación y Desarrollo Educativo, Universidad Autónoma de Baja California, México/ juancr_mx@uabc.edu.mx - <https://orcid.org/0000-0002-1115-9848>

summative and traditional assessment; and another for the formative and constructivist. In the results, it was observed that the assessment culture of teachers is more oriented towards the formative and constructivist.

Key words Higher Education/ Learning Assessment/ Teachers/ Formative Assessment/ Assessment Culture.

Introducción

De acuerdo con Perozo et al. (2012), en las instituciones de educación superior, la cultura evaluativa comprende el

conjunto de valores universitarios que se constituyen en creencias, principios y normas, impulsoras y reguladoras del comportamiento cotidiano de estos actores en el quehacer universitario, para así lograr los objetivos educativos con calidad, enmarcados en la misión, visión, proyecto institucional y características de cada institución universitaria, lo cual no sería posible lograr de forma individual, sino con valores compartidos entre sus miembros. (Perozo et al., p. 24)

La cultura evaluativa se propaga y promueve a nivel institucional, a través de diversos artefactos como reglamentos, estatutos escolares u otros medios oficiales, cuyo objetivo sea regular la práctica evaluativa (Izci & Caliskan, 2017). Sin embargo, a nivel aula la cultura de la evaluación también se conjuga con los valores, creencias, significados y prácticas que comparten docentes y estudiantes hacia la evaluación del aprendizaje (Reategui, 2015).

Si bien la cultura evaluativa del docente puede estar influenciada y condicionada por los valores, normas y políticas de la institución, hay otros factores propios del docente que configuran sus predisposiciones hacia la evaluación en el aula. Por ejemplo, su propia alfabetización sobre la evaluación de los aprendizajes, derivada de experiencias formativas previas, como cursos y capacitaciones sobre el tema (Rodríguez & Salinas, 2020; Tomasik et al., 2018); así como de la influencia que tiene el área de conocimiento a la que pertenece sobre la tradición de estrategias evaluativas, o de la naturaleza del contenido (práctico o teórico) de la asignatura que imparte (Amaro de Chacín et al., 2008; Henríquez et al., 2020).

En definitiva, tanto las características de la cultura evaluativa institucional como las del docente, tienen un impacto en los estudiantes (Reategui, 2015). Pues en el transitar de su vida escolar, desarrollan determinadas creencias, actitudes y valores hacia la evaluación de los aprendizajes (González, 2000; Quinquer, 2000). El explorar estos aspectos es relevante, dado que múltiples estudios señalan que los métodos y estrategias de evaluación que emplean los docentes, condicionan:

- las prácticas de estudio y el aprendizaje de los estudiantes (Gibbs, 2006; Santos, 1999; Thomson & Falchikov, 1998);
- las actitudes hacia la evaluación, como el miedo o el rechazo (Birenbaum, 2007; Thomson & Falchikov, 1998; Tiwari et al., 2005); y
- los valores, como el considerar que las evaluaciones solo sirven para otorgar una calificación; o por el contrario, que son útiles para proporcionar retroalimentación sobre su aprendizaje (Amaro de Chacín et al., 2008; Henríquez et al., 2020).

En la revisión de la literatura se encontraron pocos estudios que exploren como tal la cultura de la evaluación del aprendizaje en el aula, con excepción de los trabajos de Mendoza (2019) y Reategui (2015). Pero proliferan aquellos que han examinado de manera atomizada, las creencias y prácticas de evaluación de docentes (como Álvarez, 2008; Henríquez et al., 2020; Mendoza, 2019; Moreno, 2009), así como, las actitudes y percepciones de los estudiantes hacia la evaluación en el aula (como Birenbaum, 2007; Gil-Flores, 2012; Thomson & Falchikov, 1998; Tiwari et al., 2005; Rodríguez-Espinosa et al., 2016).

1. Nociones sobre cultura de la evaluación y su aplicación en la educación superior

1.1 ¿Qué es la cultura de la evaluación?

De acuerdo con Murphy (2002), el término de cultura de la evaluación es válido para cualquier tipo de organización que se caracterice por el conocimiento grupal sobre: los objetivos que persiguen las evaluaciones; cuándo se realizan; cuál es el uso de los resultados; y cómo se diseñan. Además, mencionó que las distintas actitudes que los miembros de una organización tengan hacia la evaluación, apuntarán a la formación de diferentes tipos de culturas evaluativas. En la tabla 1, se presentan diferentes definiciones sobre la cultura de la evaluación en el ámbito educativo.

Tabla 1. Conceptualizaciones sobre cultura de evaluación

| Autor | Conceptualización |
|--------------------------|--|
| Barham et al. (2013) | Son aquellos valores, normas, creencias y conductas que reflejan una apreciación compartida hacia la práctica de la evaluación como un valor institucional. |
| Bolseguí y Fuguet (2006) | Es la necesidad de toda institución de evaluarse de manera permanente, con orientación al cambio y a la continua transformación de las prácticas escolares hacia la mejora educativa. Se conforma por reglas, hábitos, rutinas, creencias, valores, estructuras, símbolos, métodos y técnicas que se comparten en una comunidad con relación a los procesos de evaluación. |
| Corti et al. (2011) | Es la capacidad institucional de autoevaluarse para mejorar sus procesos, más allá de la práctica de evaluar como un mecanismo de control. |
| Valenzuela et al. (2011) | Es una construcción social compartida por los miembros de una institución, donde se aprecia a la evaluación como un medio y no un fin. Además, implica el trabajo colaborativo de los actores educativos en todos procesos sistemáticos que estén encaminados a mejorar la calidad académica estudiantil. |
| Weiner (2009) | Son las actitudes y conductas predominantes que caracterizan el funcionamiento del apoyo institucional hacia la evaluación de los resultados educativos de los estudiantes. |

Nota: Elaboración propia.

En las anteriores conceptualizaciones, se aprecia una característica reiterativa en una cultura de la evaluación, todas las actividades y procesos evaluativos van orientados hacia la mejora. Este elemento es importante, porque se aleja de las conceptualizaciones que consideran a la evaluación como solo un proceso de medición y emisión de juicios de valor (Boud, 2019).

Dentro de la cultura de evaluación en una institución, subyace la que corresponde a la cultura de cada docente; esto es, las creencias y prácticas propias de su ejercicio en el aula (Martínez, 2010). De acuerdo con este autor, en el espacio áulico reside un subsistema de la cultura de la evaluación educativa, denominada *cultura de la evaluación del aprendizaje*. Esta comprende las “concepciones, valores, actitudes, interacciones, tradiciones, prácticas, comportamientos, símbolos y significados” (Martínez, 2010, p. 11) sobre la forma de usar a la evaluación en el aula; y la cual puede estar, o no, alineada con la cultura que promueven las instituciones.

De acuerdo con autores como Ahmad et al. (2020) y Birenbaum (2016), actualmente se observan dos paradigmas de la cultura de la evaluación en el ámbito educativo: la *evaluación tradicional* y la *evaluación centrada en el aprendizaje y mejora educativa* (de tipo constructivista). La primera cumple con el propósito de proveer información para la toma de decisiones relativas a: la acreditación, la promoción, la aprobación, la selección o el otorgamiento de incentivos (Cizek et al, 2019). También se le refiere como *tradicional*, en aquellas donde se reduce a la evaluación como sinónimo de calificación (Díaz-Barriga & Hernández, 2010; Fernández-Pérez, 1986; López, 2005; Santos, 1999). En ella, la evaluación del aprendizaje que realiza el docente tiene las siguientes características:

- es una herramienta de uso exclusivamente sumativo y están referidas a una norma (Bailey, 1998);
- se orienta a la rendición de cuentas y certificación de los aprendizajes (Ravela, 2006; Shepard, 2008)
- se hace a través de pruebas de lápiz y papel (en particular, de opción múltiple) o computarizadas, que promueven la memorización y repetición de contenidos (Black & William, 1998; Simonson et al, 2000);
- no se proporciona retroalimentación (Bailey, 1999);
- se emplea como un mecanismo de poder y control hacia los sujetos (Perrenoud, 2008; Coll & Onrubia, 2002); y
- tiene un papel pasivo el sustentante.

En contraparte, la *evaluación centrada en el aprendizaje y la mejora educativa*, se fundamenta en las aportaciones de las teorías socioculturales y constructivistas en el campo educativo (Tayyebi, 2022; Van Bergen & Parsell, 2019). En este enfoque la función docente radica en conducir experiencias que guíen al alumnado a construir por sí mismos el conocimiento, para lograr que estos aprendizajes sean significativos y a largo plazo (Coll et al. 1999; Bellocchio, 2010; Jonassen, 1991). Además, enfatiza el proceso activo del estudiante en la construcción de su propio aprendizaje; es decir, la evaluación tiene como propósito conducir al aprendizaje, más allá de solo cuantificarlo (Díaz-Barriga & Hernández, 2010; Shepard, 2008; Stobart, 2010).

Este tipo evaluación comparte comunalidades con las características del uso formativo de la evaluación en el aula de Scriven (1967) y Bloom et al. (1971), pues se enfoca hacia la devolución de información para orientar, tanto al docente como al estudiante, hacia la mejora educativa (Ahmad et al., 2020; Ravela et al., 2017). En la *evaluación para el aprendizaje*, la evaluación:

- es concebida como una actividad horizontal, compartida entre docente-alumnos, que se caracteriza por ser democrática y participativa (Guba & Lincoln, 1989; Moreno, 2016);
- promueve procesos de autoevaluación y coevaluación de estudiantes (Sadler, 1989; Barrientos-Hernán et al., 2020);
- se busca el aprendizaje significativo a través de las experiencias de evaluación (Díaz-Barriga & Hernández, 2010);
- provee retroalimentación sobre las fortalezas y debilidades (Wiggins, 1990);
- hace uso diversificado de estrategias de evaluación (Moreno, 2016); y
- permite que el alumnado participe en la toma de decisiones sobre la evaluación (Alfaro, 2000), entre otros.

1.2 Lo que sabemos de la cultura de evaluación en la Educación Superior.

Diversos autores como Álvarez (2008), Fuller (2013), Harrison (2017), Moreno (2009), Price et al. (2011) han estudiado la cultura de las universidades entorno a la evaluación. Estos, han señalado que la evaluación se centra más en el enfoque tradicional, el cual se caracteriza por el uso prioritario de exámenes con fines sumativos y una visualización de la evaluación como un sinónimo de calificación. Si bien, el enfoque sumativo es inevitable, pues las evaluaciones también deben cumplir con la función técnico-administrativa de certificar las competencias esperadas de los estudiantes universitarios; por su parte Boud & Falchikov (2007), Contreras (2010) y Martínez (2010) han señalado que debería existir un equilibrio con las evaluaciones de tipo formativas.

Al respecto, se han destacado los desafíos que afrontan los docentes universitarios para emplear estrategias de evaluación desde una orientación formativa; por ejemplo, la tasa de matrícula de estudiantes por clase (Amaro de Chacín et al., 2008; Henríquez et al., 2020), así como la falta de formación y alfabetización docente en la didáctica de la evaluación (Rodríguez & Salinas, 2020). Esta falta de formación docente, puede conducir a que los docentes continúen adhiriéndose a los métodos de evaluación tradicionales (Gibbs & Coffey, 2004).

Además, la cultura de la evaluación del docente está influenciada y condicionada por los valores, normas y políticas

del sistema educativo, o de la institución donde labore (Izci & Caliskan, 2017). Dentro de esta cultura institucional, el personal docente tiene un rol clave, pues es el responsable de valorar la medida en que el estudiantado cumple con los objetivos de aprendizaje. Por tanto, la selección de las estrategias de evaluación empleadas en el aula pueden ser resultado de múltiples factores; algunos inherentes al docente, como la formación profesional previa, la experiencia laboral, las creencias y percepciones personales hacia la enseñanza y el aprendizaje -p. e. poseer una postura constructivista- (Izci & Caliskan, 2017); y otros factores de tipo externo, como los propios valores y normas institucionales que exigen ciertas prácticas a las que deben suscribirse los docentes en su quehacer educativo (Courts & McInerney, 1993).

Conocer cuáles son las creencias, actitudes y valores de los docentes hacia la evaluación de los aprendizajes en el aula, resulta relevante dado que diversos autores como Coll y Onrubia (2002), Lambert y Holgado (2001), Padilla y Gil (2008), Rodríguez-Espinosa et al. (2016), Thomson y Falchikov (1998), y Tiwari et al. (2005), afirmaron que la postura del docente frente a las formas de evaluar el aprendizaje, influye de manera importante en distintas áreas del estudiante, como: sus estrategias de estudio, la autorregulación de su aprendizaje, sus valores (p. e. ver a la evaluación como un proceso formativo que los retroalimenta sobre su nivel de aprendizaje), así como en sus actitudes (p. e. tener ansiedad, miedo o resistencia hacia la evaluación).

De acuerdo con Martínez (2010) y Valenzuela et al. (2011), las características de la cultura de la evaluación del aprendizaje que se vive en el aula, pueden o no estar alineadas con la cultura institucional (p. e. con el modelo educativo que se promueve). La posibilidad de que la cultura de la evaluación del aprendizaje no sea congruente con el modelo educativo podría tener serias implicaciones sociales. Lo anterior, en consideración de la función social que tiene la evaluación del aprendizaje en el nivel superior: otorgar a los estudiantes una certificación del dominio de determinadas competencias profesionales que desempeñarán en el ámbito laboral. Esta certificación, se da a través de los resultados de los diversos procesos evaluativos que lleva a cabo la planta docente durante la trayectoria formativa del estudiantado (González, 2000) o por la aplicación de un examen al final del proceso formativo. Por lo tanto, es importante estudiar cuál es la cultura de la evaluación de los aprendizajes que se vive en el aula desde la perspectiva docente. Toda vez, que permite conocer cómo son y se viven las evaluaciones que, al final, se reflejan en una calificación que certifica o no las competencias que los estudiantes universitarios adquirieron en su proceso formativo.

Este trabajo de investigación tuvo dos objetivos. Primero, elaborar una propuesta para la exploración de las características de la cultura de evaluación del aprendizaje, del personal docente de la Universidad Autónoma de Baja California (UABC). Segundo, la elaboración de dos índices; uno para la medición de la *cultura evaluativa tradicional*, y otro, de *cultura evaluativa centrada en el aprendizaje* y la mejora educativa. Además, el contar con estos índices, favorece el conocimiento sobre la importancia que dan los agentes educativos a la cultura de la evaluación (Perozo et al., 2012) y su efecto en el logro académico de los estudiantes (Yali, 2018).

2. Método

2.1 Tipo de estudio

Es una investigación de corte cuantitativo y de tipo exploratorio, porque busca conocer cuál es la cultura que los docentes universitarios tienen hacia la evaluación de los aprendizajes en el aula. Esto, a través de la exploración de sus creencias, actitudes y percepción sobre los valores éticos asociados a la evaluación del aprendizaje.

2.2 Recolección de datos y aplicación

La recolección de datos se llevó a cabo mediante el método de encuesta, a través de la autoadministración de un cuestionario online. Este se cargó en la plataforma *Google Forms* y se generó una liga de acceso. El método de

recolección de datos empleado fue el de investigación por encuesta de autoselección y autoadministrada de tipo online. Si bien, la aplicación de encuestas por Internet contiene un riesgo de cobertura insuficiente o de sesgo por auto-selección de participantes (ver Kerlinger & Lee, 2002, p. 506) es posible estudiar si las características de la muestra obtenida se parecen a las de la población, es decir, es similar a un muestreo intencional de tipo aleatorio. Para lo anterior, se llevó a cabo un estudio para obtener evidencias de la equi-comparabilidad entre la muestra obtenida con la relación a la población. Primero, se estimó el margen de error del tamaño de muestra (con métodos a posteriori), con una tolerancia de error del 5%; y segundo, se examinó la distribución porcentual de distintas variables como edad y sexo, en los parámetros oficiales de la población y la muestra. En el Apéndice A, se presenta el detalle de los análisis que se llevaron a cabo. En los resultados se indica que la muestra obtenida tuvo un margen de error de 0.04, con un nivel de confianza del 95%. Por tanto, los resultados de este trabajo de tesis, pueden extrapolarse a la población docente de la UABC, con un 4% de margen de error.

En cuanto al procedimiento de aplicación del cuestionario, se pidió apoyo al Coordinador General de Formación Profesional de la Universidad Autónoma de Baja California (UABC), para solicitar a los directores de facultad y coordinadores de carrera el reenviar, vía correo institucional, la liga de acceso a su respectiva planta docente. El cuestionario se distribuyó la última semana de enero de 2022 y se mantuvo disponible durante 15 días.

2.3 Participantes

La muestra obtenida estuvo integrada por 545 docentes de nivel licenciatura de la UABC, pertenecientes al ciclo escolar 2022-1. El rango de edad osciló de los 24 a 76 años, con una media de 44.5 años. En la tabla 2, se muestra la distribución de la muestra según sus características de sexo y el área de conocimiento en el que imparten clases. Como se observa, en cuanto a la distribución por sexo, la proporción de hombres y mujeres fue equitativa. En lo que se refiere al área del conocimiento, los porcentajes más altos de respondientes se encontraron en el área de Ciencias de la Salud (27.3%), Ciencias de la Ingeniería (19.3%), Ciencias Sociales (17.4%), y Educación y Humanidades (13.7%). Llama la atención que la encuesta no fue respondida por docentes del área de Pedagogía; si bien se desconocen las razones, posiblemente no fue distribuida a la planta docente por el personal directivo de la coordinación de carrera, dado que la participación en el estudio no fue obligatoria.

Tabla 2. Características de la muestra docente

| Variables | Categorías | n | % |
|----------------------|---------------------------------|-----|------|
| Sexo | Mujer | 270 | 49.5 |
| | Hombre | 275 | 50.5 |
| Área de conocimiento | Arquitectura y Diseño | 13 | 2.4 |
| | Ciencias de la Ingeniería | 105 | 19.3 |
| | Ciencias Sociales | 95 | 17.4 |
| | Pedagogía | 0 | 0.0 |
| | Idiomas | 19 | 3.5 |
| | Ciencias Agropecuarias | 1 | 0.2 |
| | Ciencias Administrativas | 26 | 4.8 |
| | Ciencias Económicas y Políticas | 3 | 0.6 |
| | Educación y Humanidades | 69 | 12.7 |
| | Ciencias Naturales y Exactas | 49 | 9.0 |
| | Ciencias de la Salud | 149 | 27.3 |
| Ciencias Químicas | 16 | 2.9 | |

Fuente: Elaboración propia.

2.4 Instrumentos

El estudio forma parte de una tesis doctoral para la cual se diseñó y desarrolló un cuestionario que exploró la cultura de la evaluación del aprendizaje en el aula de docentes que trabajaban en una universidad pública mexicana. Es importante mencionar, que la construcción de los instrumentos incluidos en el cuestionario, se derivó de una revisión teórica de las dos principales orientaciones de la evaluación del aprendizaje: una tradicional, caracterizada más por el uso de la evaluación con fines sumativos o de rendición de cuentas; y otra constructivista, orientada hacia los usos formativos.

En cuanto al cuestionario sobre cultura de la evaluación del aprendizaje, estuvo integrado por tres secciones: (1) creencias sobre la evaluación del aprendizaje; (2) actitudes hacia la evaluación del aprendizaje; y (3) valores éticos sobre la evaluación del aprendizaje. Cada sección estuvo integrada por diferentes tipos de preguntas y se incluyeron escalas tipo Likert (tabla 3).

Tabla 3. Estructura del cuestionario de cultura de la evaluación del aprendizaje en el Aula

| Sección | Escalas | Subescalas | Categorías de respuesta |
|---|---|--|---|
| Creencias sobre la evaluación del aprendizaje en el aula | Autopercepción sobre el nivel de dominio de la evaluación del aprendizaje | Unidimensional | No lo domino (1) Lo domino poco (2) Lo domino (3) Lo domino mucho (4) |
| | | Uso de estrategias de evaluación constructivistas y centradas en el estudiante | Nunca (1) Algunas veces (2) Casi siempre (3) Siempre (4) |
| Actitudes hacia la evaluación del aprendizaje en el aula | Uso de estrategias de evaluación del aprendizaje | Uso de estrategias de evaluación tradicional y centradas en el docente | Totalmente en desacuerdo (1) Desacuerdo (2) De acuerdo (3) Totalmente de acuerdo (4) |
| | | Ideas y preconcepciones de uso formativo de la evaluación | Totalmente en desacuerdo (1) Desacuerdo (2) De acuerdo (3) Totalmente de acuerdo (4) |
| | Ideas y preconcepciones sobre la evaluación del aprendizaje | Ideas y preconcepciones tradicionales | Totalmente en desacuerdo (1) Desacuerdo (2) De acuerdo (3) Totalmente de acuerdo (4) |
| | | Sentimientos favorables hacia la práctica y usos de las evaluaciones | Totalmente en desacuerdo (1) Desacuerdo (2) De acuerdo (3) Totalmente de acuerdo (4) |
| Valores éticos sobre la evaluación del aprendizaje en el aula | Importancia de valores éticos-técnicos de las evaluaciones | Sentimientos desfavorables hacia la práctica y usos de las evaluaciones | Totalmente en desacuerdo (1) Desacuerdo (2) De acuerdo (3) Totalmente de acuerdo (4) |
| | | Unidimensional | Nada importante (1) Poco importante (2) Importante (3) Muy importante (4) |
| | Objetividad de la evaluación del aprendizaje en el aula | Unidimensional | Nunca (1) Algunas veces (2) Casi siempre (3) Siempre (4) |

Fuente: Elaboración propia.

2.5 Variable de estudio

Como se mencionó en el apartado anterior, para la consideración de los elementos de la cultura de evaluación, a la que se refiere esta propuesta de medición, se hizo una revisión de la literatura para explorar las diferentes definiciones sobre cultura de la evaluación en el ámbito educativo. En la tabla 4, se presenta la definición conceptual que dio origen a la propuesta de una definición operacional, la cual permitió operacionalizar y lograr medir este constructo, a través de distintas dimensiones observables y medibles.

Tabla 4. Definición conceptual y operacional sobre cultura de la evaluación del aprendizaje

| Definición conceptual | Definición operacional |
|---|---|
| La cultura de la evaluación del aprendizaje comprende las “concepciones, valores, actitudes, interacciones, prácticas, comportamientos, símbolos y significados” (Martínez, 2010, p. 11) sobre la forma de usar o no a la evaluación. | La cultura de la evaluación del aprendizaje incluye las creencias docentes sobre el propósito y fines de las prácticas de evaluación en el aula, así como sobre su sentido de autoeficacia para evaluar los aprendizajes de sus estudiantes. Además, integra las prácticas evaluativas, desde los elementos que toman en cuenta para planear las estrategias de evaluación, seguido del qué, cómo y para qué evaluar. Finalmente, se compone de los valores y principios que acompañan los procesos de evaluación del aprendizaje de estudiantes. |

Fuente: Elaboración propia.

2.6 Procedimiento analítico

Para llegar a la propuesta de dos índices para explorar los tipos de cultura de la evaluación de docentes, se llevaron a cabo los siguientes pasos y análisis.

1. Previo al análisis factorial, para cada escala se analizó la adecuación del muestreo y la linealidad de los ítems, a través de la medida Kaiser-Meyer-Olkin (KMO) y la prueba de esfericidad de Bartlett.
2. Se realizó un análisis factorial exploratorio (AFE) para identificar la estructura teórica de los ítems de cada escala y subescalas, que integraron el cuestionario. Para ello, se empleó el método de extracción de factores de Máxima verosimilitud y la rotación Varimax.
3. Se integraron ocho sub índices, a partir de los resultados del AFE, y se calcularon sus medidas de tendencia central y dispersión.
4. Se calcularon dos índices, a partir de la correspondencia de cada sub índice con su tipología de cultura de la evaluación, y se calcularon sus medidas de tendencia central y dispersión.

Todos los análisis se llevaron a cabo en el programa *Statistical Package for Social Sciences* (SPSS), en la versión 25.

3. Resultados

3.1 Análisis de dimensionalidad de las escalas

En cuanto a la escala *Autopercepción sobre el dominio de la evaluación del aprendizaje*, fue diseñada para medir un solo constructo. La prueba KMO y Bartlett arrojaron un coeficiente de 0.81, con una Chi cuadrada significativa. De acuerdo con Kaiser y Rice (1974) los valores por encima de 0.70 se consideran como aceptables para proceder al AFE. El análisis factorial realizado, probó la unidimensionalidad de la escala, cuyo único componente extraído explicó el 57.32% de la varianza de la autopercepción que tienen los docentes sobre su nivel de dominio de la

evaluación. Respecto a los resultados de la matriz factorial, se encontró que los coeficientes, o pesos factoriales, oscilaron entre los valores de 0.600 al 0.723 (ver tabla 5).

Tabla 5. Características factoriales de la escala Autopercepción sobre el dominio de la evaluación del aprendizaje en el aula

| Factor 1 | Peso factorial |
|--|----------------|
| 24a. Dominio para elaborar exámenes abiertos y de opción múltiple. | 0.60 |
| 24b. Dominio para elaborar otro tipo de evaluaciones, además de los exámenes (p. e. rúbricas de calificación para calificar proyectos, escalas, listas de cotejo). | 0.68 |
| 24c. Dominio para llevar a cabo actividades de coevaluación y autoevaluación en los estudiantes. | 0.70 |
| 24d. Dominio para elaborar evaluaciones que sean técnicamente válidas y confiables. | 0.72 |
| 24e. Dominio para elaborar una retroalimentación detallada sobre las fortalezas y áreas de oportunidad a cada estudiante. | 0.72 |

Nota. N = 454. Método de extracción de Máxima verosimilitud con rotación Varimax y normalización Kaiser.

En lo referente a la escala sobre *Uso de estrategias de evaluación del aprendizaje*, se exploraron dos tipos de estrategias de evaluación: las constructivistas y centradas en el estudiante; así como las tradicionales y centradas en el docente. De acuerdo con la medida KMO y la prueba de Bartlett, la escala obtuvo un valor de 0.83 y una significancia estadística, cuyo resultado indicaron que los datos eran idóneos para llevar a cabo el análisis factorial.

Como producto del AFE, emergieron tres factores los cuales explicaron el 51% de la varianza total. En la tabla 6, se observa que las agrupaciones de los ítems del primer y segundo factor, corresponden a las subdimensiones planeadas. Sin embargo, dos ítems que pertenecen teóricamente al Factor 1 (sobre uso de estrategias de evaluación constructivista y centrada en el alumno), se agruparon aparte.

Con relación a la escala sobre *Ideas y preconcepciones de uso de la evaluación del aprendizaje*, los ítems se diseñaron para explorar dos tipos: las de uso formativo y uso sumativo. La totalidad de la escala obtuvo un coeficiente de 0.87 con una significancia estadística en la prueba KMO y Bartlett. Se procedió al análisis factorial, donde se encontraron dos factores subyacentes que explicaron el 52.9% de la varianza total.

Tabla 6. Características factoriales de la escala Uso de estrategias de evaluación del aprendizaje

| Ítems | Peso factorial | | |
|--|----------------|------|------|
| | 1 | 2 | 3 |
| Factor 1: Estrategias de evaluación constructivista y centradas en el estudiante | | | |
| 26h. Proporciono información sobre cómo mejorar los aprendizajes que no se lograron, derivada de los resultados de las evaluaciones. | 0.68* | 0.12 | 0.25 |
| 26j. Adapto las estrategias o método de evaluación en los casos particulares de estudiantes que así lo requieran. | 0.65* | 0.10 | 0.11 |
| 26k. Empleo estrategias de evaluación de diversos tipos para recabar información sobre el grado de aprendizaje de los estudiantes. | 0.63* | 0.17 | 0.17 |
| 26i. La selección de las estrategias y métodos de evaluación de mi clase, buscan que los estudiantes puedan aplicar sus conocimientos en situaciones reales. | 0.62* | 0.19 | 0.04 |
| 26d. Proporciono retroalimentación grupal, sobre las principales debilidades de aprendizaje identificadas en los resultados de las evaluaciones. | 0.61* | 0.08 | 0.09 |

| | | | |
|--|-------|-------|-------|
| 26b. Durante una actividad, evalúo en todo momento las dificultades y avances de los estudiantes. | 0.60* | 0.04 | 0.08 |
| 26e. Proporciono a cada estudiante información detallada sobre los aprendizajes no logrados en cada actividad que se les evalúa. | 0.59* | 0.04 | 0.20 |
| 26a. Al inicio del curso o de un nuevo tema, evalúo los conocimientos previos del estudiantado | 0.47* | -0.03 | 0.16 |
| 26c. Acuerdo con mis estudiantes, los criterios que se utilizarán para la evaluación de los aprendizajes. | 0.38* | -0.02 | 0.03 |
| Factor 2: Estrategias de evaluación tradicional y centradas en el docente | | | |
| 26n. La selección de estrategias y métodos de evaluación de los aprendizajes solo los determino yo. | -0.13 | 0.88* | 0.07 |
| 26ñ. La selección de criterios y porcentajes de los métodos de evaluación solo los determino yo. | -0.14 | 0.85* | 0.07 |
| 26m. Para aprobar mi asignatura, la calificación de cada estudiante se calcula a partir de la evaluación de su desempeño individual en el curso. | 0.24 | 0.42* | -0.09 |
| 26l. Independiente de que la asignatura sea teórica o práctica, prefiero evaluar a mis estudiantes por medio de exámenes o pruebas objetivas. | 0.17 | 0.38* | -0.14 |
| 26o. En mis evaluaciones busco que mis estudiantes reflejen los contenidos tal y como los revisamos en clase o en las lecturas. | 0.09 | 0.30* | 0.03 |
| Factor 3: | | | |
| 26g. Promuevo que los estudiantes evalúen a otros compañeros (coevaluación). | 0.27 | -0.06 | 0.79* |
| 26f. Promuevo que los estudiantes se autoevalúen. | 0.36 | 0.01 | 0.64* |

Nota. N = 454. Método de extracción de Máxima verosimilitud con rotación Varimax y normalización Kaiser. Las cargas factoriales por encima de 0.30 se señalan en asterisco.

Como se aprecia en la tabla 7, los ítems se agruparon de la forma en que se diseñaron previamente; sin embargo, se puede observar que el ítem 27i contribuye muy poco al factor que pertenece. Al respecto, algunos autores sugieren la eliminación de ítems con cargas factoriales menores de 0.30 (Field, 2013; Tabachnick & Fidell, 2007). Lo anterior sugiere, que la afirmación “Todas las carreras deberían elaborar exámenes departamentales” no correspondió empíricamente a ninguno de los constructos subyacentes de la escala y, por tanto, debió eliminarse.

Tabla 7. Características factoriales de la escala Ideas y preconcepciones de uso de la evaluación del aprendizaje

| Ítems | Peso factorial | |
|---|----------------|------|
| | 1 | 2 |
| Factor 1: Ideas y preconcepciones de uso formativo | | |
| 27a. En el marco de esta asignatura, yo puedo elaborar evaluaciones auténticas (enfocadas en el desempeño) de las competencias que deben desarrollar mis estudiantes en el curso. | 0.68* | 0.01 |
| 27b. Elaboro evaluaciones para promover la continuidad del aprendizaje de los estudiantes (la evaluación como actividad de aprendizaje). | 0.71* | 0.08 |
| 27c. Es fundamental que los docentes cuenten con conocimientos técnicos para elaborar evaluaciones válidas y confiables. | 0.69* | 0.02 |
| 27d. Las evaluaciones las uso para conocer a detalle cada uno de mis estudiantes. | 0.59* | 0.29 |
| 27e. Las evaluaciones deberían utilizarse para identificar a los estudiantes que requieren de mayor apoyo del docente. | 0.67* | 0.15 |
| 27f. Las evaluaciones contribuyen para que los estudiantes autorregulen su aprendizaje. | 0.66* | 0.11 |
| 27g. Las evaluaciones deben ir siempre acompañadas de recomendaciones individuales sobre cómo mejorar las áreas de oportunidad de los estudiantes. | 0.72* | 0.07 |

| | | |
|--|-------|-------|
| 27h. Dada la complejidad del aprendizaje, se deben emplear diferentes formas de evaluación de acuerdo a la naturaleza de lo que se quiere evaluar. | 0.70* | -0.05 |
| 27i. Todas las carreras deberían elaborar exámenes departamentales. | 0.29 | 0.27 |
| Factor 2: Ideas y preconcepciones tradicionales | | |
| 27j. Las evaluaciones son una forma de controlar a los estudiantes. | 0.03 | 0.68* |
| 27k. Las evaluaciones solo sirven para evaluar conocimientos declarativos o conceptuales de los estudiantes. | 0.03 | 0.77* |
| 27l. Las evaluaciones sirven para ordenar a los estudiantes respecto a su rendimiento académico según los puntajes obtenidos. | 0.12 | 0.73* |
| 27m. Las evaluaciones se usan solo para cuantificar el nivel de aprendizaje de los estudiantes (p. e. asignar una calificación). | 0.04 | 0.68* |
| 27n. Con un examen de opción múltiple solo se pueden evaluar habilidades de pensamiento de bajo orden cognitivo (recordar, identificar y reconocer). | 0.10 | 0.57* |

Nota. N = 454. Método de extracción de Máxima verosimilitud con rotación Varimax y normalización Kaiser. Las cargas factoriales por encima de 0.30 se señalan en asterisco.

En lo tocante a la escala *Sentimientos hacia la práctica y usos de las evaluaciones*, se sondearon dos tipos: sentimientos favorables y sentimientos desfavorables. La medida de prueba de KMO señaló un valor de 0.79, el cual se considera dentro del rango de una adecuación de la muestra, y cuya prueba de esfericidad de Bartlett fue significativa.

En tabla 8 se observa que la agrupación de los ítems no corresponde a la estructura bidimensional de la escala, sino que subyacen tres factores. La agrupación de ítems en el primer factor corresponde a los designados como actitudes favorables hacia la evaluación; mientras que la agrupación de ítems en el segundo, a los sentimientos desfavorables. Por su parte, los ítems 28f y 28g, que se agruparon en el tercer factor (pese a pertenecer teóricamente dentro del segundo), parecen evaluar un constructo distinto, cuya redacción pareciera reflejar más un uso que una actitud. Por tanto, se tomó la decisión de eliminar ambos ítems de la escala, y así conservar únicamente los dos primeros factores que corresponden teóricamente al diseño planeado.

Tabla 8. Características factoriales de la escala *Sentimientos hacia la práctica y usos de las evaluaciones*

| Ítems | Peso factorial | | |
|--|----------------|-------|-------|
| | 1 | 2 | 3 |
| Factor 1: Sentimientos favorables hacia la práctica y usos de las evaluaciones | | | |
| 28a. Me gusta diseñar evaluaciones que entusiasmen a mis estudiantes a aplicar lo aprendido. | 0.76* | -0.05 | 0.03 |
| 28b. Estimulo el interés de mis estudiantes para mejorar los resultados de sus evaluaciones. | 0.81* | -0.04 | 0.10 |
| 28c. Me gusta aprender sobre nuevas técnicas para evaluar mejor a mis estudiantes. | 0.76* | -0.03 | 0.02 |
| 28d. Me preocupa que mis estudiantes se sientan nerviosos o estresados con las evaluaciones del curso. | 0.49* | 0.09 | 0.06 |
| 28e. Me gusta que a mis estudiantes les vaya bien en los trabajos o tareas que se evalúan. | 0.68* | 0.00 | -0.05 |
| Factor 2: Sentimientos desfavorables hacia la práctica y usos de las evaluaciones | | | |
| 28h. Siento que se sobrevalora el uso que se le puede dar a las evaluaciones de los estudiantes. | 0.10 | 0.57* | 0.16 |
| 28i. Siento que las evaluaciones son poco útiles para el proceso de enseñanza y aprendizaje. | -0.10 | 0.78* | 0.10 |
| 28j. Me aseguro que a partir de mis evaluaciones, solo aprueben los mejores estudiantes. | -0.04 | 0.48* | 0.42 |
| 28k. Las evaluaciones solo etiquetan a los estudiantes. | 0.01 | 0.65* | 0.19 |
| Factor 3: | | | |
| 28f. Las evaluaciones son una manera de imponer respeto sobre los estudiantes. | 0.05 | 0.28 | 0.75* |
| 28g. Las evaluaciones son un medio para controlar el aprendizaje de los estudiantes. | 0.08 | 0.19 | 0.80* |

Nota. N = 454. Método de extracción de Máxima verosimilitud con rotación Varimax y normalización Kaiser. Las cargas factoriales por encima de 0.30 se señalan en asterisco.

En cuanto a la escala *Importancia de valores éticos-técnicos de las evaluaciones en el aula*, la prueba de KMO resultó en un valor de 0.81, y la prueba de esfericidad de Bartlett fue significativa. Lo anterior corroboró, que la muestra es suficiente para realizar el análisis factorial. De acuerdo con el AFE, del conjunto de ítems se extrajo un solo factor (la cual corresponde con la estructura unidimensional a confirmar) que explica el 58% de la varianza total. En la tabla 9, se puede observar que el Ítem 29a posee una carga factorial menor a 0.30; por tanto, no aporta al constructo subyacente y se debe eliminar.

Tabla 9. Características factoriales de la escala *Importancia de valores éticos-técnicos de las evaluaciones en el aula*

| Factor 1 | Peso factorial |
|--|----------------|
| 29e. Honestidad en la calificación y emisión de los resultados. | 0.95* |
| 29f. Transparencia en el proceso de evaluación y entrega de resultados. | 0.93* |
| 29d. Justicia en los criterios de evaluación que se establecen. | 0.77* |
| 29c. Pertinencia de las evaluaciones respecto a los aprendizajes a evaluar en el aula. | 0.59* |
| 29b. Responsabilidad para elaborar evaluaciones que atiendan la diversidad de mis estudiantes. | 0.46* |
| 29a. Rigurosidad en la calidad técnica del diseño de las evaluaciones que realiza. | 0.29 |

Nota. N = 454. Método de extracción de Máxima verosimilitud con rotación Varimax y normalización Kaiser. Las cargas factoriales por encima de 0.30 se señalan en asterisco.

Respecto a la escala de *Objetividad de la evaluación del aprendizaje en el aula*, la prueba KMO y Bartlett arrojó un coeficiente de 0.63, que indicó la imposibilidad de llevar a cabo el análisis factorial, dada la falta de adecuación de los datos para dicho análisis. Es decir, que los ítems que pertenecen a esta escala no están correlacionados en la muestra obtenida, o que las interrelaciones entre los ítems son igual a cero.

3.2 Creación del índice y subíndices

Con base en los resultados de los diferentes AFE realizados, donde se probaron los supuestos de normalidad y linealidad de las escalas, se procedió a una reducción de variables. Para ello, se crearon diferentes subíndices que permitieron sintetizar el número de ítems que corresponden a cada constructo.

En la tabla 10, se presentan los diferentes subíndices creados, así como sus respectivos estadísticos descriptivos. Es importante mencionar, que estos se elaboraron a partir de las puntuaciones brutas (suma aritmética de los ítems correspondientes a cada factor encontrado) y fueron convertidas en proporción de porcentajes.

La agrupación de ítems que integran estos sub índices permitieron, a su vez, la creación de dos índices: uno que sintetizó todas las variables relacionadas con la evaluación sumativa y tradicional (Índice de Cultura Evaluativa Tradicional); y otro para la formativa y constructivista (Índice de Cultura Evaluativa Formativa).

Tabla 10. Estadísticos descriptivos de los sub índices creados

| Sub índices | Estadísticos descriptivos | | | |
|--|---------------------------|-----------|----------|-----------|
| | <i>M</i> | <i>DE</i> | Curtosis | Asimetría |
| Sub índice de autopercepción sobre el dominio de la evaluación del aprendizaje | 72.20 | 14.04 | 0.38 | -0.19 |
| Sub índice de uso de estrategias de evaluación constructivistas y centradas en el estudiante | 80.74 | 12.74 | -0.17 | -0.45 |
| Sub índice de uso de estrategias de evaluación tradicional y centrada en el docente | 53.15 | 18.22 | -0.29 | 0.43 |
| Sub índice de ideas y preconcepciones de uso formativo de la evaluación | 83.83 | 12.85 | 3.03 | -1.20 |
| Sub índice de ideas y preconcepciones de uso sumativo de la evaluación | 53.14 | 18.22 | -0.29 | 0.43 |
| Sub índice de sentimientos favorables hacia la práctica y usos de la evaluación | 87.54 | 12.35 | 3.69 | -1.36 |
| Sub índice de sentimientos desfavorables hacia la práctica y usos de la evaluación | 37.56 | 13.17 | 1.04 | 0.97 |
| Sub índice de importancia de valores ético-técnicos de las evaluaciones | 90.50 | 10.59 | 2.70 | -1.31 |

Nota. El rango para todos los índices fue de 75 puntos (valor mínimo de 25 y máximo de 100).

El *Índice de Cultura Evaluativa Tradicional (ICET)*, se calculó a partir de las respuestas de los docentes en las siguientes variables:

- Sub índice de uso de estrategias de evaluación tradicional y centrada en el docente.
- Sub índice de ideas y preconcepciones de uso sumativo de la evaluación.

Por su parte, el *Índice de Cultura Evaluativa para el Aprendizaje (ICEA)*, sintetiza las puntuaciones obtenidas en las siguientes:

- Sub índice de uso de estrategias de evaluación constructivistas, y centradas en el estudiante.
- Sub índice de ideas y preconcepciones de uso formativo de la evaluación.

En la tabla 11 se especifican los resultados de las medidas de tendencia central y dispersión de ambos índices. Como se aprecia, el índice sobre cultura formativa fue el que obtuvo una media de puntuación más alta. Es decir, en promedio, la cultura de la evaluación de los docentes de la UABC está más orientada a la formativa que a la tradicional.

Tabla 11. Resultados estadísticos de los índices sobre tipos de cultura evaluativa

| Índices | Estadísticos descriptivos | | | | | |
|--|---------------------------|-----------|-------|------|----------|-----------|
| | <i>M</i> | <i>DE</i> | Mín. | Máx. | Curtosis | Asimetría |
| Índice de cultura evaluativa tradicional | 53.15 | 18.22 | 25.00 | 100 | -0.29 | 0.43 |
| Índice de cultura evaluativa para el aprendizaje | 82.29 | 10.65 | 43.06 | 100 | -0.07 | -0.41 |

Nota. El rango para todos los índices fue de 75 puntos (valor mínimo de 25 y máximo de 100).

4. Discusión

Los resultados y aportaciones de este trabajo se describen en dos sentidos: por un lado, se contribuyó con una definición operativa de *cultura de la evaluación del aprendizaje* de docentes, que detalla los elementos y las características a través de las cuales es posible conocer la cultura evaluativa de docentes universitarios; y por el otro, con la propuesta de un cuestionario para explorar el tipo de cultura evaluativa que poseen.

En cuanto a lo primero, si bien la literatura ofrece diversas definiciones conceptuales sobre cultura de la evaluación (como Barham et al., 2013; Bolseguí & Fuguet, 2006; Corti et al., 2011; Valenzuela et al., 2011; Weiner, 2009), están orientadas en términos institucionales, es decir, a nivel macro. Por tanto, era necesario desarrollar una conceptualización que integrara elementos que pudieran describir la cultura de la evaluación del propio docente (más allá del marco institucional); pues la cultura, de acuerdo con Geertz (1992) se configura también con los significados de cada individuo.

Respecto al segundo aporte, la elaboración de esta definición, condujo al diseño y desarrollo de un cuestionario compuesto por nueve escalas que exploran las creencias, actitudes y valores de los docentes hacia la evaluación del aprendizaje. Este instrumento se diseñó para explorar estos elementos en función de dos tipologías de cultura: una orientada a la evaluación tradicional, y otra a la evaluación de tipo constructivista. Estos tipos, corresponden a los dos actuales paradigmas sobre la evaluación del aprendizaje (Al-Sabbah et al. 2022; Ahmad et al., 2020). Posteriormente, a partir de los análisis de factores de las escalas, se construyeron dos índices, uno para cada tipo de cultura de la evaluación.

Los índices propuestos son valiosos en dos sentidos. Primero, porque representa la medición de dos constructos poco explorados empíricamente en el ámbito educativo a nivel docente, pues se ha estudiado prioritariamente la cultura de la evaluación institucional (Farkas, 2013; Fuller et al., 2016, Perozo et al., 2012). Y segundo, ofrece a la comunidad académica dentro de las universidades, la oportunidad de monitorear en qué medida el cuerpo docente avanza en la adopción de creencias y actitudes orientadas a la evaluación como una estrategia formativa centrada en el aprendizaje y con fines de mejora educativa, o si el colectivo docente continúa concibiendo a la evaluación del aprendizaje como un mecanismo de control y poder sobre los estudiantes. Además, otro propósito de crear estos índices, fue contar con dos variables dependientes que pudieran emplearse en futuros estudios para la aplicación de otras técnicas de análisis (como análisis de regresión múltiple) que permitan, por ejemplo, conocer los factores que se asocian a cada tipo de cultura de la evaluación, o explorar si existen asociaciones entre el tipo de cultura evaluativa y el área de conocimiento del docente, entre otros.

Finalmente, en cuanto al tipo de cultura de la evaluación de la muestra de docentes estudiada, se encontró que, en general, los docentes de la UABC presentan más características asociadas a una cultura de la evaluación orientada a la evaluación constructivista y centrada en el estudiante, lo cual empata con lo que promueve el modelo educativo universitario. Por un lado, este hallazgo contrasta con otros estudios (como Fuller, 2013; Moreno, 2009; Price et

al., 2011) cuyos resultados apuntaron a que la evaluación, en el ámbito universitario, se centra más en el enfoque tradicional. Y que, incluso, eso es uno de los desafíos de la docencia en la educación superior (Hidalgo, 2021; Ibarra-Sáiz y Rodríguez-Gómez, 2020). Pero por el otro, coincide con los hallazgos de Henríquez et al. (2020) quienes encontraron que los docentes universitarios emplean estrategias asociadas a funciones de la evaluación no solo sumativas (tradicional), sino también de tipo formativas (orientadas hacia la retroalimentación del aprendizaje del estudiantado) de manera equilibrada. Cabe señalar que este último estudio, fue llevado a cabo también en una muestra de docentes que pertenecen a la UABC y otra a docentes de una universidad española.

Referencias bibliográficas

Ahmad, S., Sultana, N., & Jamil, S. (2020). Behaviorism vs Constructivism: A Paradigm Shift from Traditional to Alternative Assessment Techniques. *Journal of Applied Linguistics and Language Research*, 7(2), pp. 19-33. <http://www.jallr.com/~jallrir/index.php/JALLR/article/view/1092/1261>

Alfaro, M. (2000). *Evaluación del aprendizaje*. Universidad Pedagógica Experimental Libertador.

Álvarez, I. (2008). Evaluación del aprendizaje en la universidad: una mirada retrospectiva y prospectiva desde la divulgación científica. *Revista Electrónica de Investigación Psicoeducativa*, 14(6), 235-272. https://pdfs.semanticscholar.org/33d1/1c6e8a0767db8dc07bf468c8bdf6fd2d967a.pdf?_ga=2.265169352.630625556.1578593477-755149206.1575757542

Al-Sabbah, S., Almomani, J., Amani, D., & Najwan, F. (2022). Traditional versus Authentic Assessment in Higher Education. *Pegem Journal of Education and Instruction*, 12(1), 283-291. <https://doi.org/10.47750/pegegog.12.01.29>

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>

Bailey, K. M. (1998). *Learning about language assessment: dilemmas, decisions, and directions*. Heinle & Heinle.

Bailey, K. M. (1999). *Washback in Language Testing* [TOEFL Monograph Series MS- 15]. Educational Testing Service.

Barham, J.D., Tschepikow, W.K., & Seagraves, B. (2013). Creating a culture of assessment. In D. Timm, J.D. Barham, A. Knerr, & K. McKinney (Eds.), *Best Practices in Student Affairs Assessment*. SAGE Publications.

Barrientos-Hernán, E. J., López-Pastor, V. M., y Pérez-Brunicardi, D. (2020). Evaluación Auténtica y Evaluación Orientada al Aprendizaje en Educación Superior. Una Revisión en Bases de Datos Internacionales. *Revista Iberoamericana de Evaluación Educativa*, 13(2), 67-83. <https://doi.org/10.15366/riee2020.13.2.004>

Bellochio, M. (2010). *Educación basada en competencias y constructivismo: Un enfoque y un modelo para la formación pedagógica del Siglo XXI* (2da edición). Asociación Nacional de Universidades e Instituciones de Educación Superior/ Universidad de Colima/Universidad Autónoma de Ciudad Juárez. https://issuu.com/huguibetoben/docs/educacion_basada_en_competencias_y

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>

Birenbaum, M. (2016). *Assessment Culture Versus Testing Culture: The Impact on Assessment for Learning. The Enabling Power of Assessment*, 275–292. doi:10.1007/978-3-319-39211-0_16

Black, P. J., & William, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, 5(1), 7–74.

- Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. McGraw-Hill.
- 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>
- Boud, D. (2019, June 26-27). Developing evaluative judgement within courses. *7th International Assessment in Higher Education Conference*. Manchester, UK.
- Boud, D. & Falchikov, N. (2007). Introduction: assessment for the longer term. In Boud, D. & Falchikov, N. (Eds.) *Rethinking Assessment for Higher Education: Learning for the Longer Term*. <https://doi.org/10.4324/9780203964309>
- Cizek, G. J., Andrade, H. L., & Bennett, R. E. (2019). Formative assessment: history, definition, and progress. In H. L. Andrade, R. E. Bennett, and G. J. Cizek (Eds.) *Handbook of Formative Assessment in the Disciplines* (pp. 3–19). Routledge.
- Coll, C. y Onrubia, J. (2002). Evaluar en una escuela para todos. *Cuadernos de Pedagogía*, 318, 50-54. http://tecnoayudas.com/MEMORIAS/Memorias_esperanza/CC_JO_EvaluarEscuelaTodos_02.pdf
- Coll, C., Mrtin, E., Mauri, T., Miras, M., Onrubia, J. Solé, I., y Zabala, A. (1999). *El constructivismo en el aula* (7ma. ed.). Grao. <http://www.terras.edu.ar/biblioteca/3/3Disponibilidad-del-aprendizaje.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>
- 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 Universária na América do Sul. Florianópolis. <https://repositorio.ufsc.br/handle/123456789/25990>
- Courts, P. L., and McInerney, K. H. (1993). *Assessment in higher education: Politics, pedagogy, and portfolios*. Conn.
- Díaz-Barriga, F. y Hernández, G. (2010). *Estrategias docentes para un aprendizaje significativo. Una interpretación constructivista* (3ª ed.). Mc Graw Hill.
- Farkas, M. G. (2013). Building and Sustaining a Culture of Assessment: Best Practices for Change Leadership, *Reference Services Review*, 41(1), 13-31. https://pdxscholar.library.pdx.edu/cgi/viewcontent.cgi?article=1077&context=ulib_fac
- Fernández-Pérez, M. (1986). *Evaluación y cambio educativo: análisis cualitativo del fracaso escolar*. Morata.
- Field, A. (2013). *Discovering Statistics using SPSS* (4th ed.). SAGE.
- 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>
- Fuller, M.; Skidmore, S.; and Bustamante, R. (2016). Empirically Exploring higher Education Cultures of Assessment. *The Review of Higher Education*, 39(3), 395–429. [doi:10.1353/rhe.2016.0022](https://doi.org/10.1353/rhe.2016.0022)
- Geertz, C. (1992). *La interpretación de las culturas*. Gedisa.
- 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 edition). New York. <https://cetl.ppu.edu/sites/default/files/publications/Innovative%20Assessment%20in%20Higher%20Education.pdf>

- Gibbs, G., & M. Coffey. (2004). The impact of training of university teachers on their teaching skills, their approach to teaching and the approach to learning of their students. *Active Learning in Higher Education*, 5, 87–100. <https://doi.org/10.1177/1469787404040463>
- Gil-Flores, J. (2012). La evaluación del aprendizaje en la universidad según la experiencia de los estudiantes. *Estudios sobre Educación*, 22, 133-153. <https://core.ac.uk/download/pdf/161255273.pdf>
- González, M. (2000). Evaluación del aprendizaje en la enseñanza universitaria. *Revista de Pedagogía Universitaria*, 5(2). http://www.sld.cu/galerias/doc/sitios/prevemi/evaluacion_aprendizaje2.doc
- Guba, E. G. & Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. Sage.
- Harrison, C.J., Könings, K.D., Schuwirth, L.W.T., Wass, V., & Van der Vleuten, C. P. M. (2017). Changing the culture of assessment: the dominance of the summative assessment paradigm. *BMC Medical Education*, 17(73). <https://doi.org/10.1186/s12909-017-0912-5>
- 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>
- Hidalgo, M. E. (2021). Reflexiones acerca de la evaluación formativa en el contexto universitario. *Revista Internacional De Pedagogía E Innovación Educativa*, 1(1), 189–210. <https://doi.org/10.51660/ripie.v1i1.32>
- Ibarra-Sáiz, M. S., y Rodríguez-Gómez, G. (2020). Aprendiendo a evaluar para aprender en la educación superior. *Revista Iberoamericana de Evaluación Educativa*, 13(1). <https://dialnet.unirioja.es/servlet/articulo?codigo=7408491>
- Izci, K., & Caliskan, G. (2017). Development of prospective teachers conceptions of assessment and choices of assessment tasks. *International Journal of Research in Education and Science (IJRES)*, 3(2), 464-474. <https://files.eric.ed.gov/fulltext/EJ1148463.pdf>
- Jonassen, D. H. (1991). Evaluating Constructivistic Learning. *Educational Technology*, 31(9), 28-33. <https://www.learnlib.org/p/170798/>
- Kaiser, H. F., & Rice, J. (1974). Little Jiffy, Mark IV. *Educational and Psychological Measurement*, 34(1), 111–117. doi:10.1177/001316447403400115
- Kerlinger, F. y Lee, H. (2002). *Investigación del Comportamiento. Métodos de Investigación en Ciencias Sociales*. McGraw Hill.
- 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>
- López, V. M. (2005). La evaluación como sinónimo de calificación. Implicaciones y efectos en la Educación y en la Formación del Profesorado. *REIFOP*, 8(4). <https://www.redalyc.org/pdf/2170/217017876007.pdf>
- Martínez, N. (2010). Una aproximación a la cultura de la evaluación. *Diálogos*, 6(4), 7-20. <http://www.redicces.org.sv/jspui/bitstream/10972/2042/1/1.%20Una%20aproximacion%20a%20la%20cultura%20de%20evaluacion.pdf>
- Mendoza, M. (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>
- Morales, P. (2012). *Tamaño necesario de la muestra: ¿Cuántos sujetos necesitamos?* Universidad Pontificia Comillas. <https://web.upcomillas.es/personal/peter/investigacion/Tama%20Muestra.pdf>

Moreno, T. (2009). La evaluación del aprendizaje en la universidad: tensiones, contradicciones y desafíos. *Revista Mexicana de Investigación Educativa*, 14(41), 563-591. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-66662009000200010

Moreno, T. (2016). *Evaluación del aprendizaje y para el aprendizaje: reinventar la evaluación en el aula*. Universidad Autónoma Metropolitana-Cuajimalpa.

Murphy, D. (2002). El desarrollo de una cultura de la evaluación. *Revkala, Revista de Lenguaje y Escritura*, 7(13). <http://aprendeenlinea.udea.edu.co/revistas/index.php/ikala/article/viewFile/3211/2981>

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>

Perozo, L. González, M. y Jiménez, C. (2012). Cultura evaluativa en las instituciones universitarias. Valor fundamental para la calidad educativa. *Revista Educación en Valores*, 2(18), 22-35.

Perrenoud, P. (2008). *La evaluación de los alumnos. De la producción de la excelencia a la regulación de los aprendizajes. Entre dos lógicas*. (1ra. ed.) Colihue.

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>

Quinquer, D. (2000). Modelos y enfoques sobre la evaluación: el modelo comunicativo. En *Evaluación como ayuda al aprendizaje. Claves para la innovación educativa*, (pp. 13-20). Graó.

Ravela, P. (2006). *Fichas didácticas. Para comprender las evaluaciones educativas*. PREAL.

Ravela, P., Picaroni, B., y Loureiro, G. (2017). *¿Cómo mejorar la evaluación en el aula? Reflexiones y propuestas de trabajo para docentes* (Colección Aprendizajes Claves para la Educación Integral). INEE.

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]. http://repositorio.usil.edu.pe/bitstream/123456789/1778/1/2015_Reategui_Cultura%20evaluativa%20del%20aprendizaje%20en%20una%20universidad%20privada%20de%20Lima%20en%20el%202012.pdf

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/riee2020.13.1.005>

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(2), 1409-4258. <https://dialnet.unirioja.es/servlet/articulo?codigo=5618890>

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144.

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>

Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagné & M. Scriven (Eds.). *Perspectives of curriculum evaluation* (Vol. 1, pp. 39-83). Rand McNally.

Shepard, L. A. (2008). *La evaluación en el aula*. INEE. www.inee.edu.mx/images/stories/Publicaciones/Otros_textos/Aula/Completo/evaluacion_aula_completo.pdf

Simonson, M., S. Smaldino, M. Albright, & S. Zvacek. (2000). Assessment for distance education. In *Teaching and Learning at a Distance: Foundations of Distance Education* (Chap. 11). Prentice-Hall.

Stobart, G. (2010). *Tiempos de pruebas: Los usos y abusos de la evaluación*. Morata.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Allyn & Bacon.

Tayyebi, M., Abbasabady, M.M. & Abbassian, GR. (2022). Examining classroom writing assessment literacy: a focus on in-service EFL teachers in Iran. *Lang Test Asia*, 12(12). <https://doi.org/10.1186/s40468-022-00161-w>

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>

Tomasik, M. J., Berger, S., & Moser, U. (2018). On the development of a computer-based tool for formative student assessment: epistemological, methodological, and practical issues. *Frontiers in Psychology*, 9. <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02245/full>

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. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0185-26982011000100004&lng=es&tlng=es

Van Bergen, P., & Parsell, M. (2019). Comparing radical, social and psychological constructivism in Australian higher education: a psycho-philosophical Gubperspective. *Australian Education Research*, 46, 41–58. <https://doi.org/10.1007/s13384-018-0285-8>

Weiner, W. F. (2009). Establishing a Culture of Assessment. *Academe*, 95(4), 28-32. <https://www.aaup.org/article/establishing-culture-assessment#.Xe73BuhKjIV>

Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment, Research, and Evaluation*, 2(2). <https://doi.org/10.7275/ffb1-mm19>

Yali, J. M. (2018). *Incidencia de la cultura evaluativa del docente en el rendimiento académico de los estudiantes del nivel primaria en la Red N° 01 de la provincia de Junín* [Tesis de maestría, Universidad Nacional de Educación].

Apéndice A. Estudio de equi-comparabilidad entre la muestra obtenida versus la población

Este estudio tuvo por objeto conocer en qué medida la muestra (no probabilística) obtenida comparte las mismas características de representatividad, que una muestra poblacional seleccionada al azar. En este apartado, se describen los análisis y cálculos que se llevaron a cabo para comprobar *a posteriori*, si la muestra resultante es adecuada para representar a la población docente de la Universidad Autónoma de Baja California (UABC).

En un primer momento, se verificó la adecuación del tamaño de la muestra obtenida. Morales (2012), mencionó que es posible estimar *a posteriori* si una muestra obtenida sin aleatoriedad es adecuada, con relación al tamaño de la población conocida. Con el fin de verificar esto, se aplicó la siguiente fórmula y, con ello, estimar el margen de error en poblaciones finitas:

$$e = \sqrt{\frac{(pqz^2)(N-n)}{n(N-1)}}$$
$$\sqrt{\frac{(.25)(1.96^2)(6,090 - 545)}{545 (6,090 - 1)}}$$

Donde:

pq: probabilidad de éxito y de fracaso ($p = 0.5$ y $q = 0.5$)

z = valor estandarizado de z

N = tamaño de población

n = tamaño de la muestra

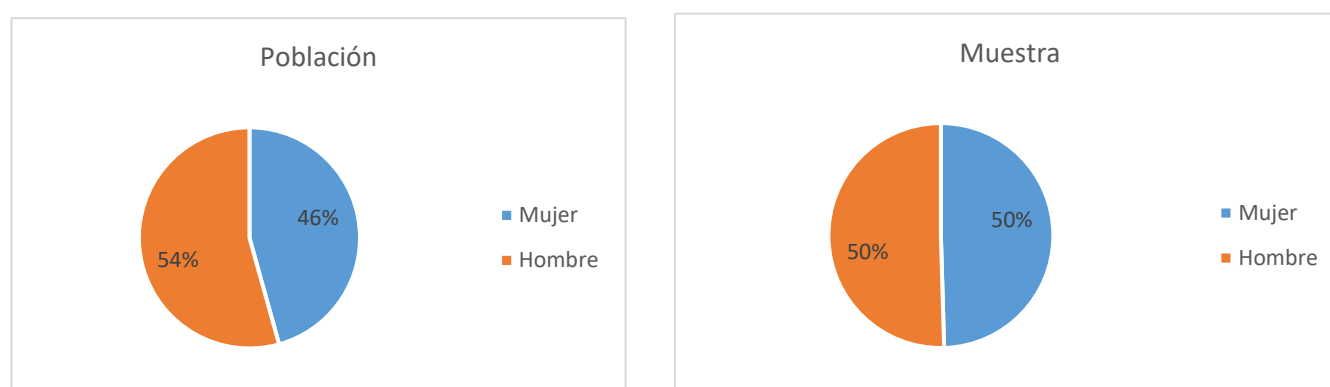
Al sustituir los valores de nuestro estudio en la formula anterior, se procedió a realizar el cálculo. Se obtuvo un margen de error de 0.04, con un nivel de confianza del 95%. Por tanto, esto indica que los resultados de este trabajo de tesis, pueden extrapolarse a la población docente de la UABC, con un 4% de margen de error.

En un segundo momento, se buscaron más evidencias de la equi-comparabilidad de la muestra, con la población. Para ello, se buscaron estadísticas oficiales sobre algunas variables que ayudaran a caracterizar la población docente de la UABC. A través del portal de Transparencia, del sitio web oficial universitario (www.uabc.edu.mx), se identificaron las variables que pudieran contrastarse con los datos de contexto que se exploran en este trabajo de investigación. Se obtuvieron datos poblacionales de la distribución docente por sexo, grupo de edad y años de antigüedad, y se compararon con la muestra.

Es importante mencionar, que no fue posible calcular si estas diferencias eran estadísticamente significativas. Esto, se debe a que solo se contó con los valores por variable de la población en términos generales, y no con los datos puntuales que permitieran realizar un análisis de diferencia de medias. Sin embargo, este ejercicio se llevó a cabo para valorar el tamaño de las diferencias.

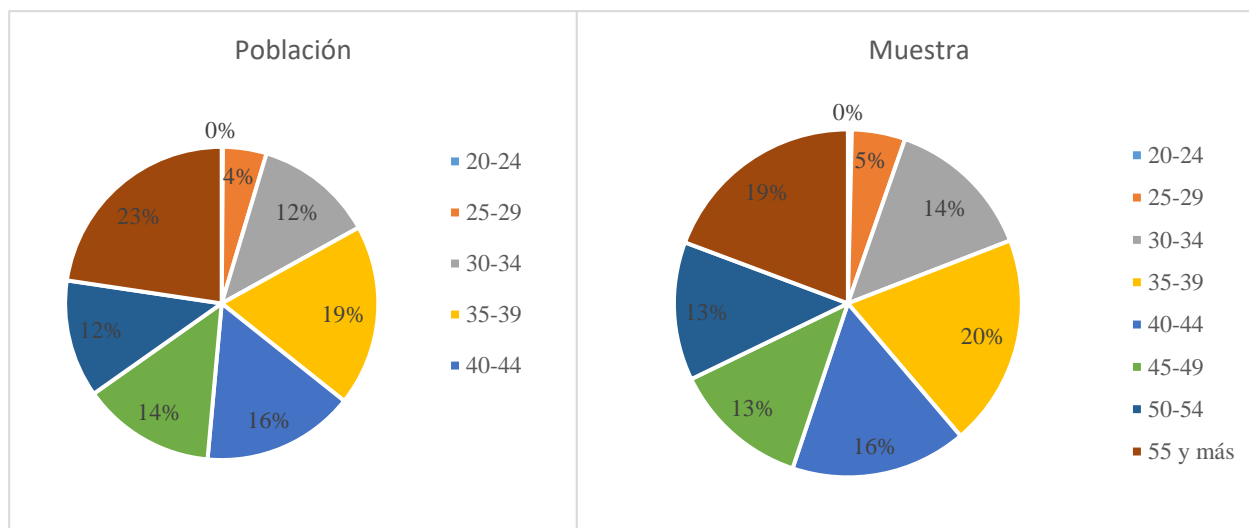
En cuanto a la variable de Sexo, la población docente está integrada por 3,308 hombres y 2,782 mujeres; mientras que la muestra obtenida se distribuye en 275 hombres y 270 mujeres. En la Figura A1, se ilustra el comparativo de diferencias porcentuales en la muestra y población.

Figura A1. Comparativo de porcentaje de docentes por sexo: muestra versus población



Respecto a la variable de Edad, las cifras institucionales despliegan los resultados por grupo de edad, mismos que se crearon para la muestra obtenida. En la Figura A2, se presenta, las distribuciones porcentuales por cada categoría, donde se observa que éstas son similares entre la muestra docente y la población.

Figura A2
Comparativo de porcentaje de docentes por grupo de edad: muestra versus población



Finalmente, se concluyó que las características estudiadas de la muestra de docentes obtenida, se distribuyeron de manera similar a la población docente de la UABC; y que es posible que los resultados de esta investigación puedan extrapolarse a la población.

Fecha de recepción: 31/5/2022

Fecha de aprobación: 1/5/2022