
archivos analíticos de políticas educativas

Revista académica evaluada por pares, independiente,
de acceso abierto y multilingüe



Arizona State University

Volumen 31 Número 11

7 de febrero 2023

ISSN 1068-2341

Metaevaluación del Sistema de Evaluación Docente en una Universidad Pública Mexicana

Alma Delia Hernández-Villafaña



Edna Luna Serrano

Universidad Autónoma de Baja California
México

Citación: Hernández-Villafaña, A. D., & Luna, E. (2023). Metaevaluación del sistema de evaluación docente en una universidad pública mexicana. *Archivos Analíticos de Políticas Educativas*, 31(11). <https://doi.org/10.14507/epaa.31.7501>

Resumen: La metaevaluación es la revisión sistemática de un programa de evaluación, orientada a identificar aciertos y oportunidades de mejora. El objetivo de esta investigación fue realizar una metaevaluación del sistema de evaluación docente de licenciatura una universidad mexicana con base en la opinión de los docentes. El diseño corresponde a un estudio de caso que retoma elementos del modelo de evaluación responsiva; se utilizaron técnicas de análisis univariado y análisis cualitativo del contenido en una muestra de 1,431 docentes que contestaron un cuestionario en línea. Los resultados dan cuenta de la complejidad de la evaluación de la docencia; los hallazgos muestran que los docentes usan los puntajes del Cuestionario de Evaluación de la Competencia Docente por los Estudiantes para identificar fortalezas y debilidades de su práctica docente, proponen revisar su extensión e incluir elementos como respeto y puntualidad, mejorar la comunicación de los resultados a los docentes e incluir otras estrategias de evaluación. Al mismo tiempo, hubo quienes censuraron el uso punitivo de la evaluación por parte de los estudiantes y su falta de conocimiento del proceso. Se concluye que es recomendable la evaluación periódica de los sistemas de evaluación docente con el fin de abonar a su uso y utilidad.

Palabras clave: práctica docente; educación superior; cuestionarios de evaluación docente; retroalimentación

Meta-evaluation of the teaching evaluation system of a Mexican public university

Abstract: Meta-evaluation is a systematic review of an evaluation program, oriented to identify strengths and opportunities for improvement. The objective of this work was to carry out a meta-evaluation of the teaching evaluation system of bachelor's degrees based on the opinion of its faculty. The design corresponds to a case study that resumed elements of the responsive evaluation model. Univariate analysis techniques and content qualitative analysis were used in a sample of 1,431 faculty members who answered an online questionnaire. The results show the complexity of the evaluation of teaching. For example, findings showed faculty use the students' score results of the Teaching Competence Evaluation Questionnaire to identify the strengths and weaknesses of their teaching practice. They also propose extend the scope of the questionnaire and include elements such as respect and punctuality, to improve communication of the results to the faculty, and other evaluation strategies. At the same time, some faculty criticized student's punitive uses of the evaluation and their lack of knowledge about the evaluation process. By way of conclusion, it is recommended that periodic assessments of teaching evaluation systems are conducted with the sole purpose of contributing to its practice and usefulness.

Key words: teaching practice; higher education; teacher-assessment questionnaires; feedback

Meta-avaliação do sistema de avaliação docente em uma universidade pública mexicana

Resumo: A meta-avaliação é a revisão sistemática de um programa de avaliação, orientada a identificar acertos e oportunidades de melhoria. O objetivo desta pesquisa foi realizar uma meta-avaliação do sistema de avaliação docente de licenciatura em uma universidade mexicana com base na opinião dos professores. O desenho corresponde a um estudo de caso que retoma elementos do modelo de avaliação responsivo; foram utilizadas técnicas de análise univariada e análise qualitativa do conteúdo em uma amostra de 1.431 professores que responderam a um questionário online. Os resultados mostram a complexidade da avaliação da docência; os achados mostram que os professores utilizam as pontuações do Questionário de Avaliação de Competência Docente pelos Estudantes para identificar as fortalezas e as fraquezas de sua prática de ensino, propõem rever sua extensão e incluir elementos como o respeito e a pontualidade, melhorar a comunicação dos resultados aos professores e incluir outras estratégias de avaliação. Ao mesmo tempo, houve quem censurasse o uso punitivo da avaliação por parte dos alunos e o seu desconhecimento do processo. Conclui-se que é recomendável a avaliação periódica dos sistemas de avaliação docente com a finalidade de pagar seu uso e utilidade.

Palavras-chave: prática docente; educação superior; questionários de avaliação docente; retroalimentação

Metaevaluación del Sistema de Evaluación Docente en una Universidad Pública Mexicana

El trabajo docente es un elemento clave para el logro de los objetivos educativos, además de ser una de las funciones sustantivas en las Instituciones de Educación Superior (IES). La docencia se asume como una actividad compleja y multidimensional, definida como un sistema no lineal constituido por múltiples características, criterios y procesos que crean relaciones causa y efecto (Wood, 2017). Así pues, la docencia no se limita a presentar información, es adaptativa e interactiva y necesariamente compromete una evaluación (Penuel y Shepard, 2016).

La evaluación educativa alude a una labor profesional realizada de forma objetiva y sistemática (Scriven, 2007), que brinda evidencia de efectividad, confiabilidad, viabilidad e integridad (Stufflebeam y Coryn, 2014). De igual forma, se fundamenta la función de la evaluación como sustento para la toma de decisiones en el ámbito educativo (Cronbach, 2000). Por otra parte, la evaluación de la docencia se reconoce como una práctica social con implicaciones de carácter público y privado, y repercusiones para la sociedad, las instituciones y los docentes (Red Iberoamericana de Investigadores de la Evaluación de la Docencia [RIIED], 2008).

Scriven (2015) propuso a la metaevaluación como una especialidad de la evaluación, cuya finalidad es valorar a la evaluación con el propósito de sustentar sus beneficios. De esta manera, se espera que la metaevaluación provea un procedimiento válido que busca evidenciar los resultados especialmente importantes y dudosos (Stake, 2017).

Por tanto, la metaevaluación resulta relevante para los programas de evaluación docente en función de conocer la calidad de las evaluaciones y profundizar en su desarrollo (Stake, 2017). Cabe mencionar que el concepto de metaevaluación no es nuevo, pero los procesos de metaevaluación de programas de evaluación de la docencia sí los son (Stake, 2017); la publicación de diseños de metaevaluación es escasa debido a la falta de análisis detallados de los problemas reales encontrados en la evaluación de la práctica docente, también, por la carencia de explicaciones organizadas de los procedimientos y herramientas dispuestas para la metaevaluación. En esta línea, se requiere generar y ahondar en estudios que abonen a la consolidación de conocimiento científico sobre procesos de metaevaluación en evaluación de la docencia en educación superior.

La presente investigación se enmarca en la Universidad Autónoma de Baja California (UABC), donde se cuenta con un Sistema de Evaluación de la Docencia (SED), su propósito es administrar la aplicación y entrega de resultados de la evaluación docente. El SED recaba información por medio del Cuestionario de Evaluación de la Competencia Docente (CECD) con base en la opinión de los estudiantes.

Hasta ahora, el SED carece de una metaevaluación, por ello, se planteó la necesidad de llevar a cabo una metaevaluación al sistema para brindar un panorama de la evaluación de la docencia en la universidad, que posibilite establecer y defender conclusiones en beneficio de la comunidad educativa, y favorecer la generación de un cuerpo de conocimiento evaluativo. En este marco, el objetivo general fue realizar una metaevaluación al SED de la UABC con base en la opinión de los docentes. Como objetivo particular, se estableció identificar las fortalezas y cuestiones críticas a mejorar en: el uso de los resultados de la evaluación docente; en el CECD y su administración; y en el sistema gestor del SED.

Contexto del Estudio

La UABC es una institución descentralizada en tres campus: Mexicali, Tijuana y Ensenada en el estado de Baja California, México. Los programas educativos se encuentran distribuidos en siete áreas del conocimiento: Ingeniería y Tecnología, Ciencias Agropecuarias, Ciencias de la Salud,

Ciencias Naturales y Exactas, Ciencias de la Educación y Humanidades, Ciencias Sociales y Ciencias Administrativas. Se ofertan 134 programas de licenciatura y 70 de posgrado. En el primer periodo semestral de 2020 la matrícula de estudiantes de la UABC fue de 65,736, y los académicos en ese periodo fueron 5,852 (UABC, 2020).

En la UABC los documentos institucionales reconocen a la evaluación docente como una tarea permanente y formativa (UABC, 2021b), realizada con el fin de brindar a los docentes retroalimentación de su desempeño (UABC, 2021c, sección “Evaluación Docente”). Al mismo tiempo, los puntajes de los estudiantes tienen un peso en la puntuación global del programa de compensación salarial. Si bien, el valor otorgado a los puntajes ha variado a través del tiempo -desde 1994 se institucionalizó su uso-, en 2021 fue del 13.33% (UABC, 2021a). Además, los puntajes sustentan la toma de decisiones en la contratación del profesorado de asignatura (2021c, sección “Evaluación Docente”). De esta forma, los resultados de la evaluación docente son utilizados tanto con fines formativos, de control administrativo y compensación salarial.

La evaluación docente se realiza por medio del Sistema de Evaluación Docente (SED), que gestiona y almacena las respuestas sobre la evaluación del desempeño de los docentes con base en la opinión de los estudiantes (UABC, 2021a). El SED atiende los programas educativos de licenciatura presencial y cursos en línea. En el caso de la modalidad de docencia presencial, el SED dispone de dos cuestionarios: el Cuestionario de Evaluación de la Competencia Docente (CECD) para las asignaturas teóricas de 25 reactivos, y otro para el componente práctico con 20 reactivos. El estudiante evalúa a cada docente según las unidades de aprendizaje que haya cursado en el semestre que se encuentre vigente, la aplicación del CECD se realiza de manera automatizada y en línea (UABC, 2017a), misma que se encuentra programada en el calendario escolar durante el último mes de cada periodo semestral. El sistema gestor del SED cuenta con tres módulos de acceso en función de los usuarios: estudiantes, docentes y directivos de unidades académicas. A partir de la información que recupera el SED se generan distintos tipos de reportes para cada audiencia: docentes, directivos de unidad académica y directivos de la administración central.

De manera similar, a lo que sucede en la mayoría de las universidades estatales en México donde el instrumento privilegiado de evaluación son los cuestionarios de apreciación estudiantil (Rueda et al., 2011), en la UABC el SED utiliza como único instrumento de evaluación al CECD, y bajo un mismo procedimiento se llevan a cabo evaluaciones que son utilizadas en las funciones de control administrativo, compensación salarial -dado que los puntajes tienen un peso en la beca al desempeño- y para la mejora.

Sustento Teórico

En este trabajo, la docencia alude a la actividad profesional de enseñar que se realiza en contextos institucionales. Por tanto, refiere la actuación del profesorado que se desempeña en un ámbito institucional, lo que implica que las acciones individuales del docente no pueden analizarse de forma independiente al contexto donde se desarrolla (Jornet et al., 2020). Desde esta perspectiva la calidad de la enseñanza involucra tanto la interacción entre docente y estudiante, como el conjunto de programas dirigidos al profesorado, a la atención del estudiante, así como, la infraestructura y el equipamiento que la institución pone a disposición de ellos para el desarrollo de la formación profesional (Luna, 2019).

La docencia es uno de los elementos esenciales del proceso educativo, se reconoce como una actividad compleja, multidimensional e incierta en la que intervienen multitud de variables (Coll y Solé, 2014). Al mismo tiempo, la docencia como actividad profesional se define por saber enseñar e implica poseer los conocimientos suficientes sobre la lógica y condiciones que inciden sobre su desarrollo (Zabalza, 2007). El objetivo primordial de la enseñanza es que los estudiantes aprendan (Hativa, 2000). Dada la consideración de que la función docente depende del entorno de trabajo su

actuación puede ser muy variada (Jornet et al., 2020), responde a las características de la realidad institucional, del modelo pedagógico y escenario educativo.

Desde esta aproximación, en la docencia universitaria se han estudiado diversas manifestaciones de la práctica docente en función del modelo pedagógico y contexto educativo, como son: la enseñanza por competencias tanto en la modalidad presencial como en línea (García-Cabrero et al., 2014; García-Cabrero et al., 2018), la enseñanza en escenarios de prácticas clínicas en medicina (Hamui-Sutton et al., 2017), y las tutorías (Torquemada y Jardínez, 2019).

La evaluación docente como práctica social se considera un recurso fundamental para la mejora de los procesos educativos. Al mismo tiempo, se advierte que no posee características intrínsecas que aseguren su bondad; un programa inadecuado de evaluación puede tener efectos nocivos en la comunidad educativa. Por ello, se puntualizan una serie de elementos a considerar en el diseño de programas de evaluación, entre los que destacan: la explicitación de los propósitos y repercusiones de la evaluación, considerar la complejidad de la enseñanza y las particularidades del contexto institucional (como la filosofía institucional y el modelo pedagógico que sustenta la práctica docente), solo así se logrará que los resultados de la evaluación sean útiles para mejorar la docencia (RIIED, 2008).

En este ámbito es fundamental tener en cuenta las dos funciones críticas de la evaluación: la formativa y la sumativa. La función formativa es una evaluación diseñada, realizada y utilizada para el mejoramiento de lo evaluado; y la función sumativa diseñada y usada con el propósito de juzgar el mérito o valor del objeto de evaluación (Scriven, 1991, 1996). Esta aparentemente sencilla clasificación ha tenido un gran impacto en las comunidades académicas y ha suscitado múltiples debates por lo que cabe tener presentes algunas precisiones. Para Scriven (1996) el elemento que distingue a una y otra es el propósito, la intención original de la evaluación. La distinción no es intrínseca al procedimiento, depende del contexto. La evaluación formativa desde su origen se realiza sobre la base del mejoramiento mientras que la sumativa pone el acento de su orientación en la emisión de un juicio de valor. Así, los dos tipos de evaluaciones pueden considerarse como igualmente útiles, y dependerá de las circunstancias que se asuma alguna de ellas.

Además, el concepto de validez en evaluación de la docencia se define como el grado de representatividad y utilidad que permite la interpretación de los resultados de evaluación; involucra la validez interna -refiere a la totalidad de criterios de calidad de consistencia de los diferentes componentes del proceso-, y la validez externa -los resultados del proceso son de utilidad para el uso previsto-, así validar la evaluación es hacerla útil para los involucrados (Jornet et al., 2020).

Ante la complejidad que implica evaluar la actividad docente se subraya la necesidad de utilizar diversos instrumentos en su evaluación. En el ámbito universitario los cuestionarios de evaluación de la docencia desde la perspectiva de los estudiantes son la estrategia de evaluación preponderante (Rueda et al., 2021; Smith y Kubacka, 2017). Asimismo, la literatura del campo reitera la necesidad de utilizar otras estrategias de evaluación como son: las rúbricas y cuestionarios de autoevaluación, la evaluación por pares académicos, la observación de clase y los portafolios docentes (de Diego y Rueda 2012; García, 2014). Cabe precisar que ciertos instrumentos de evaluación se encuentran asociados a la evaluación formativa, en particular, las estrategias de autoevaluación y los portafolios docentes.

Los cuestionarios de opinión de los estudiantes se caracterizan por ser los instrumentos más ampliamente estudiados, con una larga historia de investigación que respalda su empleo casi universal en el ámbito universitario (Benton y Cashin, 2014). Si bien, los puntajes derivados de los cuestionarios se consideran válidos, confiables y relativamente libres de sesgo (Linse, 2017), persisten críticas respecto a su utilidad tanto en evaluaciones formativas como sumativas (Spooren et al., 2013). Gran parte de los cuestionamientos críticos al uso de los puntajes se realizan fundamentalmente cuando son utilizados como única fuente de información sobre la docencia

(Spooren et al., 2017). Así, el uso apropiado de los puntajes es uno de los desafíos que enfrentan este tipo de evaluaciones (Linse, 2017).

La metaevaluación es definida por Scriven (2015) como la evaluación de la evaluación, que conlleva un procedimiento enfocado a describir y juzgar una acción de evaluación, sustentado en los referentes bajo los cuales se concibe una buena evaluación (Stufflebeam, 2011). Por consiguiente, una metaevaluación representa el conocimiento evaluativo que incorpora los estándares para el desarrollo y puesta en marcha de un sistema de evaluación (Scriven, 2015).

Scriven (2011) propuso una lista de verificación de seis elementos a tomar en cuenta en el desarrollo de una metaevaluación y advierte que la utilidad de la lista depende del nivel de la consulta y de los asuntos específicos involucrados. Los elementos son: validez, credibilidad, claridad, propiedad, costo-utilidad y generalización. La validez alude a la veracidad de las conclusiones y al establecimiento de las condiciones y restricciones de la metaevaluación; la credibilidad se relaciona con un bajo nivel de sesgo; la claridad representa la precisión y comprensibilidad de los usuarios; la propiedad atiende a la ética, legalidad y pertinencia cultural; el costo-utilidad involucra el estudio de costos y beneficios; y el último elemento, la generalización, conlleva la utilidad del diseño de evaluación, los procedimientos, resultados o conclusiones.

Por otra parte, desde la perspectiva de Stufflebeam (2011) la metaevaluación es una forma de evaluación, y la noción de la metaevaluación responde a ocho premisas: 1. Evaluar el mérito del trabajo evaluativo; 2. Utilidad tanto formativa como sumativa; 3. Valorar la importancia de los objetivos, los diseños y la importancia y calidad de los hallazgos de la evaluación; 4. Describir y juzgar las acciones evaluativas, realizar recomendaciones para mejorar las evaluaciones y el uso conveniente de los resultados; 5. Atender a los evaluadores y a los interesados en su desempeño; 6. Realizar una metaevaluación formativa por evaluadores y obtener opiniones externas respecto al mérito de sus acciones de evaluación; 7. Diseñar preguntas de metaevaluación específicas, obtener, organizar y analizar la información necesaria y emplear la información para la toma de decisiones y rendición de cuentas; y 8. Cumplir los criterios de utilidad, conveniencia y rentabilidad.

La metaevaluación de un sistema de evaluación docente se considera como una actividad a realizarse de forma continua, que integra lo que se sabe sobre la evaluación y la medición de la competencia docente en función de la calidad (Stake, 2017). Por tanto, una metaevaluación aporta un control de calidad para las actividades de la evaluación (Stake, 2011).

En el entorno internacional, la metaevaluación del Programa de Apoyo para la Evaluación de la Actividad Docente del Profesorado Universitario (DOCENTIA), de la Universidad Complutense de Madrid (UCM), tuvo la finalidad de identificar las fortalezas y aspectos susceptibles de mejora del programa, así como, proponer recomendaciones que favorezcan su eficiencia y utilidad (Gallego y Rojo, 2014). Los objetivos específicos atendieron el estudio de los instrumentos, la comparación de las dimensiones destacadas en la literatura sobre la calidad docente, indagación sobre la capacidad de discriminación entre los procedimientos de evaluación de los docentes, y el planteamiento de un sistema adecuado de producción de las puntuaciones para los docentes. El programa DOCENTIA-UCM representa un diseño actualizado, con un plan y desarrollo confiable, viable y sostenible, producto del proceso de reflexión y perfeccionamiento alineado al trabajo docente y ajustado a la normativa institucional; también, se indican procedimientos para asegurar la calidad de la enseñanza y beneficiar su desarrollo y reconocimiento (UCM, 2016).

Un referente de metaevaluación en el escenario mexicano es el trabajo realizado en la Universidad Autónoma del Estado de Hidalgo (UAEH), su objetivo fue sustentar un proceso permanente de evaluación docente apegado al modelo educativo de la universidad, que apoye la calidad de la enseñanza y la formación docente (Elizalde et al., 2008). Algunos de los elementos analizados fueron la planeación, conducción, evaluación, materiales y técnicas didácticas, normatividad, asesorías y satisfacción del aprendizaje. Los docentes (alrededor del 80%) indicaron

que los resultados les posibilitan clarificar sus fortalezas y debilidades; por lo tanto, les proporciona información para orientar la mejora de su desempeño académico.

Otro precedente, en la Universidad Autónoma de Tlaxcala (UAT) se llevó a cabo una metaevaluación al cuestionario de evaluación del desempeño docente por medio de la opinión del estudiante partir de tres dimensiones: personal, disciplinar-pedagógica y profesional (Sánchez, 2019). Algunos de los hallazgos indicaron que docentes y estudiantes priorizan evaluar la atención personalizada hacia el estudiante, planeación didáctica y justicia de la evaluación.

Método

El diseño correspondió a un estudio de caso, desde la concepción de atender el cómo o porqué de una situación (Yin, 2018), con el fin de indagar la complejidad de un contexto particular y la interacción con su entorno (Stake, 1999). También, fueron retomados elementos del modelo de evaluación responsiva, respecto a la atención de las condiciones del contexto y recuperación de las distintas visiones de los actores (Stake, 2011). En esta investigación se realizó una evaluación de la evaluación con el propósito de identificar: los méritos del trabajo de la evaluación docente, su utilidad formativa y sumativa, la calidad del SED, así como realizar recomendaciones para su mejora (Scriven, 2015; Stufflebeam, 2011).

Este trabajo involucró a distintos participantes. Se contó con la colaboración de dos funcionarios de la administración central, seleccionados por su responsabilidad en la definición de las políticas y funcionamiento del SED. Asimismo, 41 directivos de las unidades académicas (equivalente al 60% de los directores y subdirectores de la UABC).

La muestra se conformó por autoselección de 1,431 docentes en el segundo periodo escolar de 2019, representó el 24.6% del total de docentes de la UABC, 612 (42.8%) fueron docentes de tiempo completo y 757 (52.9%) de asignatura, el 4.3% no indicó su categoría de adscripción. Los docentes correspondieron a las siete áreas del conocimiento de la universidad (Ciencias de la Ingeniería y la Tecnología, Ciencias de la Educación y Humanidades, Ciencias de la Salud, Ciencias Sociales, Ciencias Económico Administrativas, Ciencias Naturales y Exactas y Ciencias Agropecuarias); respecto a su antigüedad laboral, el 52.7% refirió contar con ocho o más años; 20.5% de cuatro a siete años y el 21.9% de uno a tres años. El 4.7% de los docentes no informó su tiempo de trabajo. Se utilizó un cuestionario en formato digital diseñado ex profeso para esta investigación.

Procedimiento

El procedimiento se organizó en tres etapas: Etapa 1. Diseño del cuestionario, Etapa 2. Aplicación del cuestionario y Etapa 3. Análisis de la información.

Etapa 1. Diseño del Cuestionario

El cuestionario fue elaborado en función de: revisión teórica, identificación de las consideraciones de los responsables del SED y directivos de las unidades académicas. El sustento teórico se derivó de los planteamientos de la RIIED ([2008]; Rueda, 2010), por orientar la reflexión en el diseño e implementación de programas de evaluación de la docencia y su pertinencia al contexto mexicano. Además, se realizaron entrevistas individuales a dos funcionarios de la administración central, con el fin de conocer sus perspectivas y expectativas hacia el SED. Por tal motivo, se les cuestionó sobre: el procedimiento de aplicación del CECD (calendario de aplicación, difusión, aplicación del cuestionario, sobre el modelo docente que se promueve); el contenido y estructura del CECD; la entrega de resultados; el uso de los resultados; los aspectos positivos del SED; los principales efectos del SED; las problemáticas que enfrenta el SED; los cambios o

aspectos a mejorar; y propuestas de otras formas alternativas de evaluación docente. Así, se elaboró una primera versión del cuestionario, mismo que fue aplicado a los directivos con el propósito de conocer su postura sobre el SED, y se incorporaron los aspectos considerados oportunos en la versión final del cuestionario.

De esta forma, el cuestionario quedó conformado por cuatro dimensiones con un total de 30 ítems; de los cuales 26 fueron de respuestas cerradas (17 de escala, 7 de opción múltiple y 2 dicotómicos) y 4 de respuesta abierta. La escala establecida fue del 1 con la valoración más baja al 10 la más alta, la opción múltiple se determinó conforme a la naturaleza del ítem y la respuesta dicotómica fue sí y no. Los ítems de respuesta abierta fueron opcionales y de libre extensión, se incluyó uno por cada dimensión, a saber: sobre el uso de los resultados de la evaluación docente, la pertinencia del CECD, el proceso de aplicación y operación, y utilidad del SED (véase la Tabla 1).

Tabla 1

Estructura del Cuestionario

Dimensiones	Ítems				Total
	Respuesta cerrada			Respuesta abierta	
	Escala del 1 al 10	Opción múltiple	Dicotómicos		
Información general		4			4
1. Uso de los resultados de evaluación docente	1	2		1	4
2. Cuestionario de evaluación docente para los alumnos	8	1	2	1	12
3. Procedimiento de aplicación	4			1	5
4. Valoración del SED	4			1	5
N.º de ítems respuesta cerrada	17	7	2		
Total de ítems de respuesta cerrada y abierta		26		4	
Total de ítems en el cuestionario					30

Nota. N.º= número.

Etapa 2. Aplicación del Cuestionario

El cuestionario se administró en el *software Limesurvey* versión 1.92. Los docentes fueron convocados a participar por medio de un mensaje a su correo electrónico institucional (se invitó a toda la planta docente); el tiempo estimado para contestar el cuestionario fue de 10 minutos; se definió un margen de 15 días de recuperación de la información; y la tasa de respuestas se monitoreó por medio del *software* (a los siete días fue enviado un recordatorio).

Etapa 3. Análisis de la Información

Los datos correspondientes a los ítems cerrados se integraron en una base de datos del *software SPSS* versión 26, donde se realizó un análisis univariado con el propósito de visualizar las variables de forma individual, para identificar valores atípicos, frecuencias bajas y otras particularidades de los datos (Afifi et al., 2020). Por tanto, se determinaron las frecuencias y porcentajes de cada ítem de manera separada para conocer la estructura de los datos.

La información derivada de los ítems abiertos se concentró en un archivo del *software Atlas.ti* versión 7.5.4. Para el estudio de los datos se aplicó la técnica de análisis cualitativo de contenido (ACC), y a partir de un enfoque inductivo-deductivo se construyeron categorías de análisis (Mayring, 2014). Por consiguiente, con el propósito de esclarecer y organizar la información, se elaboraron cuatro agendas de códigos, una por cada dimensión del cuestionario, orientadas a la pregunta del tema central de cada dimensión. En la Dimensión 1 Uso de los resultados de evaluación docente el número de datos analizados fue 443, en la Dimensión 2 Cuestionario de evaluación docente para los alumnos 220, en la Dimensión 3. Procedimiento de aplicación 199 y en la Dimensión 4. Valoración del SED correspondieron 254; en suma, los datos examinados fueron 1,116.

La organización de las agendas de códigos consistió en definir cada categoría, que a su vez se compuso por códigos, reglas de codificación y ejemplos, a fin de describir con precisión bajo qué condiciones se codificaron los datos (Mayring, 2014). Por ejemplo, una de las categorías de las agendas fue Uso de los resultados, definida como la delimitación de los usos de la evaluación (RIIED, 2008; Rueda, 2010); dentro de los códigos que la integraban se encontraba: Utilidad de la evaluación para mejorar la enseñanza (UTIL); la regla de codificación indicaba: codificar cuando se declare, de manera explícita o implícita, la percepción del docente sobre la utilidad, importancia o valor del cuestionario o de la evaluación para mejorar su práctica; y el ejemplo consistió en la cita textual con el número de línea del dato en la unidad hermenéutica. Así, las categorías emergieron del estudio continuo de los datos y de la identificación de referentes teóricos (RIIED, 2008; Rueda, 2010) considerados alineados a éstas.

El ACC condujo al establecimiento de cinco categorías en las agendas de códigos: Metodológica-procedimental, Uso de los resultados, Política de la evaluación, Ética y Competencia docente. La categoría Metodológica-procedimental integró datos vinculados a la estructura, contenido y procedimiento de aplicación del CECD, y a la comunicación de los resultados (Rueda, 2010). La categoría Uso de los resultados concentró el empleo dado a los resultados de la evaluación docente, referente a temas relacionados a la toma de decisiones administrativas, asignación de recursos, promoción y estímulos económicos (RIIED, 2008; Rueda, 2010). La categoría Ética refirió aspectos éticos en el proceso de evaluación, relacionados a las “buenas prácticas, los valores, las actitudes, los posicionamientos teórico-ideológicos, entre otras cuestiones” (Rueda, 2010, p. 347); la categoría Política de la evaluación reunió datos sobre los propósitos y las repercusiones de la evaluación de la docencia como condición indispensable de un sistema de evaluación (RIIED, 2008). En el caso de la categoría Competencia docente, emergió a partir de datos asociados con el dominio del docente relativo al conocimiento disciplinar, pedagógico y compromiso profesional.

Además, se calculó la confiabilidad entre codificadores, el tipo obtenido fue la reproducibilidad (también nombrada concordancia intersubjetiva y confiabilidad de formas paralelas), entendida como un proceso que puede ser replicado por diferentes analistas y por considerarse una medida sólida de confiabilidad (Krippendorff, 2004). La confiabilidad entre codificadores supone una comparación de dos analistas que codifican un mismo contenido con el fin de generar una medida de objetividad (Mayring, 2014).

El análisis se realizó sustentado en la medida de confiabilidad de Holsti (1969, como se cita en Mayring, 2014), con la participación de dos analistas: codificador 1 (c1, analista titular) y codificador 2 (c2, analista invitado). El coeficiente fue determinado en función del número de acuerdos en la codificación de c1 y c2 y el número de codificaciones realizadas por c1. El ejercicio se llevó a cabo en dos momentos con la finalidad de retroalimentar y asegurar que el ACC de los datos se generó bajo condiciones exentas de alteraciones o sesgos —accidentales o intencionales— y que el análisis representa lo mismo para todos los que lo emplean (Krippendorff, 2004).

Resultados

Los resultados obtenidos se estudiaron desde una postura de complementariedad analítica para favorecer el objetivo de la investigación. A continuación, se presentan los resultados del análisis univariado y después el ACC, así como los análisis de confiabilidad.

Análisis Univariado

Los ítems de respuesta cerrada fueron contestados por el 100% de la muestra. A partir de la estructura del cuestionario, en la Dimensión 1. Uso de los resultados de evaluación docente, referente a la pregunta ¿Qué uso da a los resultados de evaluación docente?, la mayor frecuencia de respuestas se obtuvo en: identificar fortalezas y debilidades (71.7%), seguida de buscar alternativas de mejora de la práctica docente (60.9%), y seleccionar cursos de formación (23.6%). Por otro lado, destacó el porcentaje que respondió no hacer uso de los puntajes (3.2%).

Respecto a la Dimensión 2. Cuestionario de evaluación docente para los alumnos, al preguntar el nivel de acuerdo para incorporar otros elementos en el CECD, los puntajes más altos se obtuvieron en: respeto (79.0), puntualidad (63.2), y uso de las Tecnologías de la Información y la Comunicación (TIC) en el aula (48.6); en lo general, consideraron que el CECD incluye los elementos importantes de la práctica docente (71 puntos en adelante suman los valores 8, 9 y 10) (véase la Tabla 2). En este sentido, se valoró positivamente el contenido del CECD y se puntualizaron elementos a incorporar en el CECD.

Tabla 2

Promedios de Valoración sobre Elementos del CECD

Ítem	Promedios de valoración respecto a la escala									
	1	2	3	4	5	6	7	8	9	10
¿Qué tan de acuerdo estaría en incluir en el CECD valorar el respeto?	.7	.2	.3	.4	2.0	.4	1.3	5.6	10.0	79.0
¿Qué tan de acuerdo estaría en incluir en el CECD valorar la puntualidad?	2.5	.7	.9	.9	3.4	1.3	3.3	9.4	14.3	63.2
¿Qué tan de acuerdo estaría en incluir en el CECD valorar el uso de TIC en el aula?	3.0	.4	1.4	.7	6.6	2.4	5.2	16.3	15.6	48.6
¿Considera que el CECD incluye los elementos importantes de la práctica docente?	1.4	.7	1.2	1.8	8.0	4.3	10.9	30.9	22.0	18.8

Además, ante el cuestionamiento de considerar necesario el desarrollo de otros CECD para los estudiantes dirigidos a programas educativos particulares, la opinión de los docentes estuvo dividida, por un lado, el 57.3% de los docentes señalaron que no era indispensable, y por otra parte el 42.6% de los docentes estuvieron de acuerdo en realizar cuestionarios específicos. Al respecto, la propuesta de CECD con mayor porcentaje fue para Ciencias de la Salud (26.7), seguida de Artes Plásticas y Danza con el 15.8%, el 12.8% de los docentes indicaron que se requiere un cuestionario para cada programa educativo, el 8.9% refirió indispensable un cuestionario para la modalidad semiescolarizada, así como un 5.9% por modalidad educativa, y con el mismo porcentaje Ciencias Sociales y Humanas (5.9). Asimismo, con menor presencia se sugirieron cuestionarios para el Posgrado (2.9%) y por áreas del conocimiento (2.9%).

También, el 75.0% de los docentes consideraron pertinente utilizar otras estrategias de evaluación docente junto con el CECD. De esta forma, se indicó la pertinencia de incorporar otras estrategias de evaluación de la docencia, tales como evaluación de pares académicos (67.6%) y autoevaluación (59.6%).

Referente a la Dimensión 3. Procedimiento de aplicación, con los puntajes más altos, los docentes estimaron adecuado responder el CECD de forma obligatoria (61.9). Asimismo, los docentes apuntaron conveniente la difusión para promover que los estudiantes contesten el CECD (45), además de aprobar la manera en que la unidad académica realiza la aplicación del CECD (39.3) y estimaron oportunas las fechas de aplicación del cuestionario (58.7 puntos al sumar los tres valores más altos de la escala) (véase la Tabla 3). De esta manera, se valoró positivamente el procedimiento de aplicación del CECD.

Tabla 3

Promedios de Valoración sobre el Procedimiento de Aplicación del CECD

Ítem	Promedios de valoración respecto a la escala									
	1	2	3	4	5	6	7	8	9	10
¿Le parece adecuado que sea obligatorio para los estudiantes contestar el CECD?	4.4	1.0	1.2	.5	4.8	1.8	3.0	8.4	13.0	61.9
¿Es pertinente la difusión para promover que los estudiantes contesten el CECD?	2.2	1.3	2.0	1.5	6.6	3.4	6.1	14.4	17.5	45.0
¿Considera adecuada la manera en que en su unidad académica se realiza la aplicación del CECD?	3.3	1.6	1.3	1.3	6.7	2.9	6.7	16.9	20.0	39.3
¿Son oportunas las fechas del CECD?	7.4	1.8	4.1	2.7	8.0	5.2	9.1	18.4	15.5	24.8

Acerca de la Dimensión 4. Valoración del SED, en alusión a los valores más altos de la escala, con 40.6 puntos los docentes manifestaron que el SED es útil, y con 30.3 puntos los docentes refirieron que el reporte de resultados de la evaluación es claro (los puntajes ascienden a 72.5 y 67.3 respectivamente con la sumatoria de los tres valores más altos de la escala). Y con puntajes menores, los docentes indicaron disponer de los resultados de la evaluación cuando los requieren con 22.1 puntos (y un puntaje de 54 con los tres valores más altos), y en esta línea, la devolución oportuna de resultados de la evaluación concentró el valor más alto en 8 con un puntaje de 19. De esta forma, si bien de manera general se valoró como útil el SED, se identificaron aspectos centrales a mejorar en favor de la usabilidad (véase Tabla 4).

Tabla 4

Promedios de Valoración sobre la Apreciación del SED

Ítem	Promedios de valoración respecto a la escala									
	1	2	3	4	5	6	7	8	9	10
¿Considera que el SED es útil?	3.6	1.9	1.9	2.3	6.6	4.0	7.4	15.2	16.7	40.6
¿Es claro el informe de resultados de su evaluación docente?	4.4	1.8	2.5	2.8	7.8	4.9	8.5	17.6	19.4	30.3
¿Dispongo de los resultados cuando los necesito?	11.2	2.7	3.3	3.1	9.6	7.0	8.8	17.0	15.1	22.1
¿La devolución de los resultados de la evaluación docente es oportuna?	10.4	3.1	4.6	3.5	11.5	6.4	11.5	19.0	14.6	15.4

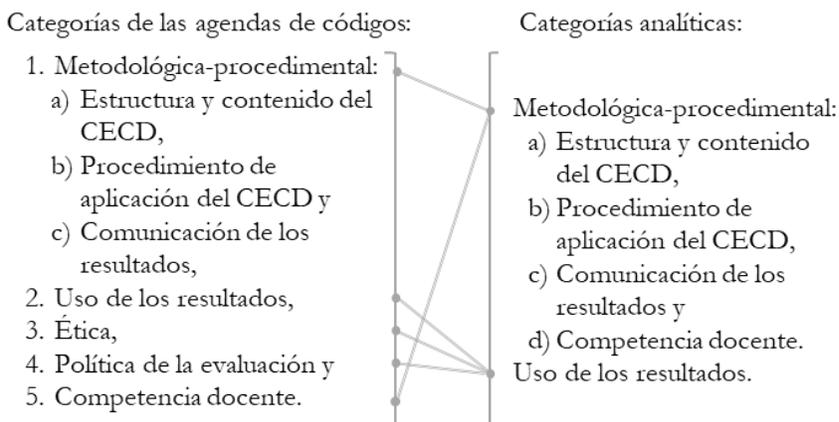
Análisis Cualitativo del Contenido

En inicio, cabe señalar que el ACC fue realizado con los datos obtenidos del segmento de participantes que contestaron los ítems de respuesta abierta. Así, el porcentaje de respuesta de la Dimensión 1. Uso de los resultados de evaluación docente fue de 30.9, a la Dimensión 2. Cuestionario de evaluación docente para los alumnos correspondió el 13.9%, en la Dimensión 3. Procedimiento de aplicación contestaron el 15.3% y en la Dimensión 4. Valoración del SED se contó con una participación del 17.7%. De forma global, la tasa promedio de respuesta fue de 19.4% de la muestra.

A partir de las cinco categorías de las agendas de códigos del ACC se realizó una concentración de la información para su estudio que derivó en la delimitación de dos categorías analíticas: metodológica-procedimental y Uso de los resultados. La categoría analítica metodológica-procedimental se organizó en cuatro subcategorías: Estructura y contenido del cuestionario, procedimiento de aplicación, comunicación de los resultados y competencia docente. La categoría analítica uso de los resultados abordó: el empleo de los resultados, cuestiones éticas y elementos referentes a la política de la evaluación en alusión a los fines y efectos de la evaluación de la docencia (véase Figura 1). Enseguida, los resultados del ACC se presentan de acuerdo a las categorías analíticas señaladas.

Figura 1

Relación entre Categorías de las Agendas de Códigos y Categorías Analíticas



Categoría Analítica Metodológica-Procedimental

Subcategoría Estructura y Contenido del Cuestionario. Una de las cuestiones mayormente mencionada por los docentes fue la extensión del cuestionario, los docentes expresaron que el número de preguntas no favorece el propósito del CECD: “Hay un exceso de reactivos en la evaluación docente, por eso los alumnos lo hacen sin fijarse” (2019_D2-87¹). Asimismo, se mencionó que el cuestionario requería mejorarse: “Hacer la redacción de la evaluación más amigable, menos técnica” (2019_D2-105).

Aparte de la extensión del CECD, los docentes citaron desconocer las preguntas que componen el cuestionario: “Desconozco el cuestionario, nunca he leído los reactivos que les hacen llegar a los estudiantes mediante estas encuestas de evaluación” (2019_D2-17).

Los docentes solicitaron establecer cuestionarios específicos de evaluación según las características particulares del área de estudio: “Por ejemplo en las profesiones del área médica” (2019_D2-2.9[29]), “Artes Plásticas y Danza” (2019_D2-2.9[52]), “Posgrado” (2019_D2-2.9[334]).

Subcategoría Procedimiento de Aplicación. Sobre la obligatoriedad de la evaluación docente, existen opiniones diversas, por un lado, los docentes apuntaron que tal obligación conduce a los estudiantes a realizar la evaluación solo como requisito de inscripción: “La mayoría de los alumnos realizan la evaluación como un simple requisito para poder inscribirse en el siguiente semestre” (2019_D2-2). También, hubo docentes que consideraron conveniente contestar el CECD de forma obligatoria: “El cuestionario debe de ser obligatorio para permitir la inscripción” (2019_D3-156), sin embargo, se realizaron comentarios sobre los efectos que podría generar la imposición de responder el CECD: “Tengo mis dudas sobre la obligatoriedad, creo que debemos hacer más conciencia y que la participación sea de manera voluntaria y consciente sobre el impacto del instrumento en la práctica docente” (2019_D3-86).

Asimismo, los participantes consideraron que la cantidad de docentes que el estudiante tiene que evaluar influye para realizar una evaluación correcta: “Los alumnos comentan que lo tedioso de la evaluación docente es la cantidad de profesores a evaluar. Inician con entusiasmo y terminan marcando cualquier valor para poder concluir” (2019_D2-25), “el objetivo se diluye al saber la cantidad de maestros a evaluar” (2019_D3-22). Además, se indicó que hay una baja participación de estudiantes que contesta el CECD: “en ocasiones solo se tiene evaluación de uno o dos alumnos” (2019_D3-38).

Por otra parte, los docentes realizaron comentarios sobre fallas en el funcionamiento del sistema gestor del SED: “en los últimos semestres ha fallado mucho la página de evaluación” (2019_D3-13), “Continuamente el sistema no funciona adecuadamente y la evaluación no se concluye” (2019_D3-103). También, señalaron conveniente realizar una mayor difusión para contestar el CECD: “Considero que se debe hacer mayor difusión para que el alumno realice la evaluación” (2019_D3-51).

Subcategoría Comunicación de los Resultados. La explicitación de los objetivos e implicaciones de la evaluación de la docencia son una condición necesaria, esto favorece que todos los involucrados aporten al cumplimiento de las metas establecidas (RIIED, 2008). En este sentido, una parte importante de los docentes señalaron tener desconocimiento total o parcial de los resultados de la evaluación: “Nunca he recibido mis resultados de la evaluación docente (...)” (2019_D1-789), “Desconozco los resultados de la evaluación docente hasta el momento” (2019, D1-1014); seguido de la falta de información para acceder a estos: “Desconozco la forma en que puedo

¹ Se empleó un código de identificación respecto al año, dimensión y número de línea de la cita establecido en la base de datos.

acceder a los resultados de mi evaluación docente” (2019, D1-472), “No sé en dónde puedo consultar estos resultados” (2019_D1-634).

En otro asunto, los docentes subrayaron la entrega de los resultados por el sistema gestor del SED fuera de tiempo: “Los resultados del semestre anterior no están listos para cuando se hacen las plantas y para cuando los alumnos se inscriben” (2019_D1-1045), “Los resultados deben estar más a la mano para los interesados "docentes". Todavía hace un mes, no estaban disponibles en la plataforma los resultados de los cursos que impartí” (2019_D1-2575).

Subcategoría Competencia Docente. Un tema emergente incorporado a esta categoría aludió a la competencia docente, el cual representó un hallazgo ajeno a las dimensiones cuestionadas que, aunque tuvo poca incidencia, brindó información sobre la opinión de docentes referente al dominio pedagógico del contenido disciplinar: “(...) hay profesores de asignatura mal evaluados dadas sus limitaciones didácticas (no cumplen el programa, tienen faltas recurrentes) (...)” (2019_D1-580).

Categoría Analítica Uso de los Resultados

Esta categoría analítica refirió a la delimitación de los usos de la evaluación (RIIED, 2008; Rueda, 2010), además, fueron incluidas las categorías Ética y Política de la evaluación de las agendas de códigos por estimarse afines a esta categoría analítica.

En alusión a las características éticas en el proceso de evaluación, los docentes sostuvieron de forma recurrente que la realización del CECD no es tomada con seriedad, porque generalmente los estudiantes no atienden al propósito de la evaluación y no realizan una evaluación crítica: “Algunos alumnos lo toman como juego y sólo rellenan el cuestionario sin razonarlo” (2019_D4-195), “Así que considero que la evaluación docente no es tomada en serio por los alumnos, no ven la importancia que tiene para el profesor. Si eres un maestro laxo, los evalúan de maravilla” (2019_D1-148).

Además, los docentes indicaron que la evaluación carece de objetividad: “La evaluación docente resulta ser totalmente subjetiva por parte del alumno” (2019_D1-397), “(...) no existe manera de comprobar que tan objetivos son los alumnos al contestar el cuestionario (...)” (2019_D2-147).

Posteriormente, enunciaron la falta o desconocimiento de información de los estudiantes sobre el propósito de la evaluación: “Los alumnos no tienen idea de que es lo que hacen con este cuestionario” (2019_D1-96).

En otro tema, en su mayoría los docentes expresaron que los estudiantes usan la evaluación como una forma de venganza, generalmente asociada a la baja calificación del estudiante debido a su mal desempeño académico: “La encuesta es instrumento de venganza para algunos alumnos hacia el maestro” (2019_D1-119), “Creo que la evaluación es utilizada por algunos alumnos como venganza por la calificación obtenida” (2019_D1-444). Por tanto, los docentes sugieren sensibilizar a los estudiantes sobre los propósitos e importancia de la evaluación de la docencia: “Creo que hace falta sensibilizar e informar a los estudiantes acerca de la importancia de este tipo de evaluaciones” (2019_D1-741), así como la falta de seguimiento de los resultados o carencia de información al respecto: “Conviene informar qué seguimiento se le da a las evaluaciones docentes por parte de las coordinaciones” (2019_D1-2856), “No hay seguimiento ni estrategias para que la evaluación sea más integral y coadyuve al mejor desempeño docente” (2019_D1-2943).

La UABC (2017a) establece que la evaluación de la docencia se realiza de forma sistemática y permanente, a fin de mejorar el desempeño, la promoción de docentes y la adecuada asignación de recursos. Asimismo, uno de los cinco principios orientadores sustentados en la misión, visión y filosofía de la universidad, marca la evaluación permanente como un proceso de retroalimentación

de los resultados para renovar y conducir las acciones institucionales al cumplimiento de los objetivos de la universidad (UABC, 2017b). En este aspecto, gran parte de los docentes resaltan el uso de los resultados para retroalimentar y fortalecer su práctica, apoyados principalmente en los comentarios realizados por los estudiantes en el CECD: “(...) para mí los comentarios me ayudan a identificar debilidades” (2019_D1-49) “Lo único que reviso son los comentarios, esto debido a que la calificación sale relativamente alta siempre” (2019_D1-1374).

Confiabilidad

Con el fin de retroalimentar y mejorar las agendas de códigos, se realizó el análisis de confiabilidad en dos momentos. En el primer análisis se obtuvo un coeficiente de .59, posteriormente, a partir de la revisión de los dos analistas, se esclareció y precisó la descripción del ejercicio, la concordancia de opiniones, dudas y comentarios. Las principales sugerencias señaladas por el c2 fueron examinar los códigos de la categoría Uso de los resultados, en relación con el uso que los docentes indicaban dar a la evaluación, además de explorar la opinión de los docentes sobre el propósito de la evaluación y su finalidad para mejorar.

En el segundo análisis de confiabilidad se obtuvo un coeficiente de .84, esto brindó una medida de objetividad relacionada a la independencia de los hallazgos de los analistas (Mayring, 2014). Aquí, sobresalió examinar las sugerencias de mejora en la categoría Metodológica-procedimental en función de clarificar su definición y regla de codificación, asimismo, revisar la categoría Ética respecto a la descripción de la evaluación irreflexiva de los estudiantes indicada por los docentes.

Se destaca que, en la investigación cualitativa resulta complejo lograr un completo acuerdo entre analistas, porque las particularidades interpretativas siempre cuentan con un elemento subjetivo, incluso bajo la guía de los lineamientos del ACC (Mayring, 2014).

En la búsqueda de particularidades entre las opiniones de los docentes, por categoría de adscripción -tiempo completo y asignatura-, se compararon sus respuestas, las diferencias se encontraron en la pregunta abierta de la Dimensión 4. Valoración del SED. Los docentes de asignatura, ponderaron la importancia de la evaluación y del CECD como una herramienta para mejorar la práctica docente, y la ausencia o falta de información sobre los resultados de la evaluación. A su vez, los docentes de tiempo completo enfatizaron la falta de información con la que cuenta el estudiante sobre la finalidad e importancia de la evaluación docente, así como, el poco o nulo seguimiento de los resultados por parte de las autoridades educativas y de evidencias de medidas remediales.

Discusión y Conclusiones

La metaevaluación comprende la revisión y comprobación de un proceso de evaluación que permite identificar los aciertos y las cuestiones problemáticas relacionadas a dificultades administrativas, usos inadecuados o fallas técnicas (Stufflebeam, 2011). En consecuencia, proporciona información sistemática para la regulación de las acciones a fin de lograr las metas propuestas. Al mismo tiempo, se reconoce a la validez como elemento fundamental de la evaluación de la docencia, en el entendido de que validar el proceso de evaluación es hacerla útil a los participantes e interesados, y de esta forma buscar la equidad y justicia de la evaluación (Jornet et al., 2020).

Desde esta perspectiva, este trabajo se propuso realizar una metaevaluación al SED de una universidad pública en México con base en la opinión del profesorado implicado, que permitiera dar cuenta de las fortalezas y cuestiones críticas a mejorar en: el uso de los resultados de la evaluación docente, el instrumento de evaluación y su administración, así como el sistema gestor del SED. El

conjunto de resultados mostró la complejidad inherente a los procesos de la evaluación de la docencia y la diversidad de posturas sobre un mismo sistema de evaluación.

En relación con el uso de los resultados, destacan las respuestas que aluden al empleo de los resultados para mejorar la práctica docente (identificar fortalezas y debilidades 71.7 %, buscar alternativas de mejora 60.9% o seleccionar cursos de formación 23.6%), por el contrario, solo un 3.2% del total de la muestra refirió no utilizar los puntajes. Al respecto, cabe comentar que la evaluación docente posibilita la reflexión individual sobre la enseñanza y el aprendizaje, igualmente, brinda las bases para la comprensión de las complejas relaciones sociales que emergen en diferentes escenarios de formación (RIIED, 2008).

Por otra parte, se registraron consideraciones éticas que en conjunto refieren los posibles sesgos en los puntajes, como la falta de seriedad de los estudiantes al contestar los cuestionarios o el uso de la evaluación con fines de venganza. Si bien, los puntajes derivados de los cuestionarios se consideran válidos, confiables y relativamente libres de sesgo (Linse, 2017), persisten críticas respecto a su utilidad tanto en evaluaciones formativas como sumativas por consideraciones como las aquí enunciadas. Además, los docentes expresaron falta de seguimiento a los resultados o repercusiones que incidan en la mejora. En este escenario cabe destacar la necesidad de vincular los puntajes de los estudiantes con estrategias institucionales de mejoramiento profesional, ya que, solamente entregar los reportes de evaluación al profesorado resulta insuficiente para incidir en el mejoramiento de la práctica docente (Luna y Torquemada, 2008).

Acercas del CECD, los docentes propusieron revisar la extensión del cuestionario e incluir elementos como respeto (79%) y uso de las TIC (48.6%). También, apuntaron la pertinencia de diseñar cuestionarios específicos, en particular para evaluar a los docentes de las áreas de la salud y artes, en atención a las particularidades de la práctica docente en estas áreas del conocimiento. Por consiguiente, cabe tener presente la discusión en la literatura sobre adecuar el contenido de los cuestionarios a las necesidades del contexto educativo y pedagógico (Benton y Cashin, 2012), así como los desarrollos de cuestionarios de evaluación de la docencia para programas específicos como medicina (Flores-Hernández et al., 2017).

Además, el 75% de la muestra señaló la pertinencia de incorporar otras estrategias de evaluación como autoevaluación y evaluación por pares académicos. De esta forma se propuso diversificar las fuentes de información y fortalecer el SED, lo cual alude a la discusión sobre las limitaciones de los cuestionarios de apreciación estudiantil (Hornstein, 2017), y la pertinencia de utilizar diversos instrumentos de evaluación para posibilitar una mejor comprensión de la actividad docente (Nasser-Abu, 2017; Rueda, 2021).

Respecto a la administración del CECD, priorizaron brindar información a los estudiantes sobre el propósito de la evaluación con el fin de sensibilizarlos sobre su importancia, así como mejorar la difusión sobre el periodo para contestar el cuestionario; y sobre la comunicación de los resultados, que los reportes de resultados sean comprensibles, detallados, de tal forma que el profesorado pueda identificar sus fortalezas y aspectos a modificar. En este sentido, se realizaron señalamientos puntuales que abonan a la credibilidad y utilidad de los resultados de la evaluación (Jornet et al., 2020). En consecuencia, los resultados de este trabajo de metaevaluación derivaron en recomendaciones puntuales orientadas hacia la mejora del SED.

Este estudio de metaevaluación brindó un panorama de las expectativas, cuestionamientos y propuestas de los docentes respecto al SED; permitió delimitar los aspectos críticos que requieren ser revisados y modificados que van desde esclarecer los propósitos de la evaluación, enriquecer la conducción del proceso evaluativo y los procesos administrativos que involucra. De esta forma, la función de la metaevaluación no es alcanzar una calificación resumida descriptiva, sino brindar información exhaustiva sobre las observaciones del objeto evaluado (Stake, 2011). Adicionalmente,

esta metaevaluación aportó información valiosa para orientar el diseño de mecanismos que ayuden a una mayor usabilidad de los resultados de evaluación.

En México, la evaluación se ha planteado como un recurso fundamental para mejorar el sistema educativo, simultáneamente, las políticas de evaluación de la docencia en educación superior han favorecido la implementación de programas asociados a estímulos económicos. Al mismo tiempo, la experiencia documentada a lo largo de 30 años ha hecho patente los efectos nocivos en los procesos de evaluación que ocurren cuando los resultados tienen un alto impacto o consecuencias graves en los evaluados (Rueda, 2021). En este escenario, las IES enfrentan el gran reto de conocer y comprender mejor las condiciones y particularidades de la práctica docente, a fin de implementar sistemas de evaluación que orienten la mejora de la actividad de enseñar y aprender. Por tanto, una asignatura pendiente en las políticas de evaluación de la docencia en las IES mexicanas es alentar cursos de acción que favorezcan reconocer la complejidad de la docencia y su evaluación formativa.

En consecuencia, se requiere el desarrollo de sistemas de evaluación docente que expresen claramente su propósito principal y clarifiquen las características esperadas de la docencia. A fin de, conocer y comprender mejor la función docente e identificar las condiciones necesarias para su desarrollo. Al mismo tiempo, al ponderar la validez como elemento fundamental de la evaluación en la perspectiva de hacerla útil para la mejora de la práctica docente, es factible consolidar acciones para mejorar la calidad de la formación profesional.

Referencias

- Afifi, A., May, A., Donatello, R. A., y Clark, V. A. (2020). Data visualization. En *Practical Multivariate Analysis* (6a. ed., pp. 37-58). Taylor & Francis Group.
- Benton, S. L., y Cashin, W. E. (2012). Student ratings of teaching: A summary of research and literature. *IDEA Papers Series*, 50.
https://ideacontent.blob.core.windows.net/content/sites/2/2020/01/PaperIDEA_50.pdf
- Benton, S. L., y Cashin, W. E. (2014). Student ratings of instruction in college and university courses. En M. B. Paulsen (Ed.), *Higher education: Handbook of theory and research*, (Vol. 29, pp. 279-326). Springer. https://doi.org/10.1007/978-94-017-8005-6_7
- Coll, C., y Solé, I. (2014). Enseñar y aprender en el contexto del aula. En C. Coll, J., Palacios y Á. Marchesi (Comps.), *Desarrollo psicológico y educación: 2. Psicología de la educación escolar* (2a ed., pp. 357-386). Editorial Alianza.
- Cronbach L. J. (2000) Course improvement through evaluation. En D. L. Stufflebeam, G. F. Madaus y T. Kellaghan (Eds.), *Evaluation models. Evaluation in education and human services* (Vol. 49, pp. 235-247). Springer. https://doi.org/10.1007/0-306-47559-6_14
- de Diego, M., y Rueda, M. (2012). La evaluación docente en educación superior: Uso de instrumentos de autoevaluación, planeación y evaluación por pares. *Voces y Silencios. Revista Latinoamericana de Educación*, 3(2), 59-76.
<https://revistas.uniandes.edu.co/doi/pdf/10.18175/vys3.2.2012.04>
- Elizalde, L., Pérez, C. G., y Olvera, B. I. (2008). Metaevaluación del proceso de evaluación docente: Universidad Autónoma del Estado de Hidalgo. *Reencuentro*, (53), 113-124.
<https://www.redalyc.org/pdf/340/34005310.pdf>
- Flores-Hernández, F., Gatica, F., Sánchez-Mendiola, M., y Martínez, A. (2017). Evolución de la evaluación del desempeño docente en la Facultad de Medicina; evidencia de validez y confiabilidad. *Investigación en educación médica*, 6(22), 96-103.
<https://doi.org/10.1016/j.riem.2016.06.004>

- Gallego, C., y Rojo, N. (2014). Meta-evaluación del programa Docencia de la UCM. Universidad Complutense de Madrid.
https://eprints.ucm.es/34313/3/Memoria%20del%20proyecto%20Meta-evaluaci%C3%B3n%20del%20programa%20DOCENTIA%20UCM_2014.pdf
- García, (2014). *Modelos de evaluación docente en el mundo e instrumentos utilizados para evaluar: Ventajas, desventajas y consideraciones para su elección*. [Informes Técnicos MIDE UC]. Pontificia Universidad Católica de Chile.
- García-Cabrero, B., Loredó, J., Luna, E., y Rueda, M. (2014). *Competencias docentes en educación media y superior: Desarrollo y validación de un modelo de evaluación*. Universidad Autónoma de Baja California.
- García-Cabrero, B., Luna-Serrano, E., Ponce-Ceballos, S., Cisneros-Cohenour, E., Cordero-Arroyo, G., Espinoza-Díaz, Y., y García-Vigil, M. H. (2018). Las competencias docentes en entornos virtuales: Un modelo para su evaluación. *Revista Iberoamericana de Educación a Distancia*, 21(1), 343-365. <http://dx.doi.org/10.5944/ried.21.1.18816>
- Hamui-Sutton, A., Ortiz-García, A., Cejudo-Aparicio, L., Lavalle-Montalvo, C., y Vilar-Puig, P. (2017). La evaluación de los docentes desde la perspectiva de los médicos residentes del Plan Único de Especializaciones Médicas. *Educación Médica*, 18(2), 89-97.
<https://doi.org/10.1016/j.edumed.2016.06.010>
- Hativa, N. (2000). *Teaching for effective learning in higher education*. Springer Science & Business Media.
- Hornstein, A. (2017). Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance. *Cogent Education*, 4(1).
<https://doi.org/10.1080/2331186X.2017.1304016>
- Krippendorff, K. (2004). Reliability. In *Content analysis: An introduction to its methodology* (2a ed., pp. 211-256). SAGE.
- Jornet, J. M., Perales, M. J., y González-Such, J. (2020). El concepto de validez de los procesos de evaluación de la docencia. *Revista Española de Pedagogía*, 78(276), 233-252.
<https://doi.org/10.22550/REP78-2-2020-01>
- Luna, E. (2019). Evaluación formativa del modelo educativo en instituciones de educación superior en México. *Revista Mexicana de Investigación Educativa*, 24(83), 997-1026.
<https://www.scielo.org.mx/pdf/rmie/v24n83/1405-6666-rmie-24-83-997.pdf>
- Luna, E., y Torquemada, A. D. (2008). Los cuestionarios de evaluación de la docencia por los alumnos: Balance y perspectivas de su agenda. *Revista Electrónica de Investigación Educativa*, 10(Especial), 1-15. <https://redie.uabc.mx/redie/article/view/201>
- Linse, A. R. (2017). Interpreting and using student ratings data: Guidance for faculty serving as administrators and on evaluation committees. *Studies in Educational Evaluation*, 54, 94-106.
<http://dx.doi.org/10.1016/j.stueduc.2016.12.004>
- Mayring, P. (2014). *Qualitative content analysis: Theoretical foundation, basic procedures and software solution*.
<https://www.ssoar.info/ssoar/handle/document/39517>
- Nasser-Abu, F. (2017). Teaching in higher education: Good teaching through students' lens. *Studies in Educational Evaluation*, 54, 4-12. <http://dx.doi.org/10.1016/j.stueduc.2016.10.006>
- Penuel, W. R., y Shepard, L. A. (2016). Assessment and teaching. En D. H. Gitomer y C. A. Bell, *Handbook of research on teaching* (5a ed., pp. 787-850). American Educational Research Association.
- Red Iberoamericana de Investigadores de la Evaluación de la Docencia. (2008). Reflexiones sobre el diseño y puesta en marcha de programas de evaluación de la docencia. *Revista Iberoamericana de Evaluación Educativa*, 1(3), 163-168.
<https://revistas.uam.es/index.php/riee/article/view/4661/5098>

- Rueda, M. (2010). Reflexiones generales a considerar en el diseño y puesta en operación de programas de evaluación de la docencia. *Revista Iberoamericana de Evaluación Educativa*, 3(1), 344-350. <https://revistas.uam.es/index.php/riee/article/view/4552/4986>
- Rueda, M. (2021). Anotaciones para reorientar las prácticas de evaluación educativa. *Revista Mexicana de Investigación Educativa*, 26(88), 311-330. https://www.comie.org.mx/v5/sitio/wp-content/uploads/2021/01/RMIE_88.pdf
- Rueda, M., Luna, E., García, B., y Loredó, J. (2011). Resultados y conclusiones. En M. Rueda (Coord.), *¿Evaluar para controlar o para mejorar?* (pp. 197-222). Instituto de Investigaciones sobre la Universidad y la Educación-Educación superior contemporánea.
- Sánchez, R. (2019, 18-22 de noviembre). Meta-evaluación del desempeño docente en la UATx: El Caso de la Facultad de Ciencias de la Educación. Ponencia presentada en el Congreso Nacional de Investigación Educativa, México. <http://www.comie.org.mx/congreso/memoriaelectronica/v15/doc/1227.pdf>
- Scriven, M. (1991). *Evaluation thesaurus* (4a ed.). Sage publications.
- Scriven, M. (1996). Types of evaluation and types of evaluator. *Evaluation Practice*, 17(2), 151-161. <https://doi.org/10.1177/109821409601700207>
- Scriven, M. (2007, 6-9 de junio). *The logic of evaluation* [Sesión de conferencia]. Ontario Society for the Study of Argumentation Conference, Canadá. <https://scholar.uwindsor.ca/cgi/viewcontent.cgi?article=1390&context=ossaarchive>
- Scriven, M. (2011). *Evaluating evaluations: A meta-evaluation checklist* (9a ed.). Claremont Graduate University. http://michaelscriven.info/images/EVALUATING_EVALUATIONS_8.1.11.pdf
- Scriven, M. (2015). *The meta-evaluation checklist*. Claremont Evaluation Center.
- Smith, W. C., y Kubacka, K. (2017). The emphasis of student test scores in teacher appraisal systems. *Education Policy Analysis Archives*, 25(86). <https://epaa.asu.edu/ojs/article/view/2889>
- Spooren, P., Brockx, B., y Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research*, 83(4), 598-642. <https://doi.org/10.3102/0034654313496870>
- Spooren, P., Vandermoere, F., Vanderstraeten, R., y Pepermans, K. (2017). Exploring high impact scholarship in research on student's evaluation of teaching (SET). *Educational Research Review*, 22, 129-141. <https://doi.org/10.1016/j.edurev.2017.09.001>
- Stake, R. E. (1999). Introducción: Estudio intensivo de los métodos de investigación con estudio de casos. En *Investigación con estudio de casos* (2a ed., pp. 9-14). Ediciones Morata.
- Stake, R. E. (2011). Program evaluation particularly responsive evaluation. *Journal of MultiDisciplinary Evaluation*, 7(15), 180-201. http://journals.sfu.ca/jmde/index.php/jmde_1/article/view/303
- Stake, R. E. (2017). Meta-evaluation of systems for the evaluation of teaching. En E. Luna y M. Rueda, *Experiencias de evaluación de la docencia en Iberoamérica* (pp. 248-256). Universidad Autónoma de Baja California.
- Stufflebeam, D. L. (2011). Meta-evaluation. *Journal of MultiDisciplinary Evaluation*, 7(15), 99-158. http://journals.sfu.ca/jmde/index.php/jmde_1/article/view/300
- Stufflebeam, D. L., y Coryn, C. L. (2014). Overview of the evaluation field. En *Evaluation theory, models, and applications* (2a ed., pp. 3-43). John Wiley & Sons.
- Torquemada, D., Jardínez, L. (2019). La formación de competencias docentes universitarias a partir de la evaluación del desempeño del tutor. *Publicaciones*, 49(1), 39-52. <https://doi.org/10.30827/publicaciones.v49i1.9851>
- Universidad Autónoma de Baja California. (2020). *Indicadores institucionales: Numeralia institucional*. <http://www.uabc.mx/planeacion/numeralia/>

- Universidad Autónoma de Baja California. (2021a). *Anexo 1: Tabla de identificación de actividades (TIA)*. http://sriagral.uabc.mx/Secretaria_General/PPREDEPA/predepa2022-2023/TIA.pdf
- Universidad Autónoma de Baja California. (2021b). *Evaluación de la docencia de licenciatura en la UABC: Descripción general del proceso*. http://web.uabc.mx/formacionbasica/documentos/evaluacion_docencia_lic.pdf
- Universidad Autónoma de Baja California. (2021c, 28 de diciembre). *Formación docente*. <https://www.uabc.mx/formacion-docente/>
- Universidad Complutense de Madrid. (2016). *Modelo de evaluación de la actividad docente del profesorado de la UCM: programa DOCENTLA-UCM*. <https://www.ucm.es/data/cont/docs/3-2016-11-15-Docentia-noviembre2016.pdf>
- Wood, P. (2017). From teaching excellence to emergent pedagogies: a complex process alternative to understanding the role of teaching in higher education. En A. French y O'Leary M. (Eds.) *Teaching excellence in higher education: Challenges, changes and the teaching excellence framework* (pp. 39-74). Emerald Publishing Limited.
- Yin, R. K. (2018). Getting started: How to know whether and when to use the case study as a research method. In *Case study research and applications: Design and methods* (6a ed., pp. 31-56). Sage Publications.
- Zabalza, M. A. (2007). *Competencias docentes del profesorado universitario. Calidad y desarrollo profesional*. Narcea.

Sobre las Autoras

Alma Delia Hernández-Villafaña

Universidad Autónoma de Baja California
alma.delia.hernandez.villafana@uabc.edu.mx

Licenciada en ciencias educativas por la Universidad Autónoma de Baja California (UABC), maestra en ciencias de la educación por el Instituto de Investigación y Desarrollo Educativo de la UABC y doctoranda en la misma institución. Es docente en la Facultad de Ciencias Administrativas y Sociales de la UABC.

<https://orcid.org/0000-0003-4792-4350>

Edna Luna Serrano

Universidad Autónoma de Baja California
eluna@uabc.edu.mx

Licenciada y maestra en psicología por la Universidad Nacional Autónoma de México y doctora en educación por la Universidad Autónoma de Sinaloa (México). Su línea de investigación es evaluación educativa. Miembro del Sistema Nacional de Investigadores, Nivel II. Es coordinadora de la Red de Investigadores sobre Evaluación de la Docencia y directora editorial de la *Revista Electrónica de Investigación Educativa*.

<http://orcid.org/0000-0003-1496-548X>

archivos analíticos de políticas educativas

Volumen 31 Número 11

7 de febrero 2023

ISSN 1068-2341



Los/as lectores/as pueden copiar, mostrar, distribuir, y adaptar este artículo, siempre y cuando se de crédito y atribución al autor/es y a Archivos Analíticos de Políticas Educativas, los cambios se identifican y la misma licencia se aplica al trabajo derivada. Más detalles de la licencia de Creative Commons se encuentran en <https://creativecommons.org/licenses/by-sa/4.0/>. Cualquier otro uso debe ser aprobado en conjunto por el autor/es, o AAPE/EPAA. La sección en español para Sud América de AAPE/EPAA es publicada por el *Mary Lou Fulton Teachers College, Arizona State University* y la *Universidad de San Andrés* de Argentina. Los artículos que aparecen en AAPE son indexados en CIRC (Clasificación Integrada de Revistas Científicas, España) DIALNET (España), [Directory of Open Access Journals](#), EBSCO Education Research Complete, ERIC, Education Full Text (H.W. Wilson), PubMed, QUALIS A1 (Brazil), Redalyc, SCImago Journal Rank, SCOPUS, SOCOLAR (China).

Sobre el consejo editorial: <https://epaa.asu.edu/ojs/index.php/epaa/about/editorialTeam>

Por errores y sugerencias contacte a Fischman@asu.edu

Síguenos en EPAA's Facebook comunidad at <https://www.facebook.com/EPAAAPE> y en **Twitter feed** @epaa_aape.
