



Páginas: 154-164
Recibido: 2022-07-01
Revisado: 2022-10-03
Aceptado: 2022-03-04
Preprint: 2023-21-04
Publicación Final: 2023-15-05

www.revistascientificas.us.es/index.php/fuentes/index

DOI: <https://doi.org/10.12795/revistafuentes.2023.21273>

Metaevaluación de un sistema de evaluación docente desde la perspectiva estudiantil

Meta-evaluation of a Teaching Evaluation System from students perspective



Alma Delia Hernández-Villafaña
Universidad Autónoma de Baja California (México)



Edna Luna Serrano
Instituto de Investigación y Desarrollo Educativo-Universidad Autónoma de Baja California (México)

Resumen

El objetivo de este estudio fue realizar una metaevaluación del Sistema de Evaluación Docente (SED) de la Universidad Autónoma de Baja California para identificar fortalezas y aspectos a mejorar en: el Cuestionario de Evaluación de la Competencia Docente (CECD), el sistema gestor del SED, la administración del CECD y el uso de los resultados. Corresponde a un estudio de caso que adaptó elementos del modelo de evaluación responsiva, la muestra no probabilística fue de 10,160 estudiantes que contestaron un cuestionario en línea, los datos se procesaron con técnicas de análisis univariado y análisis cualitativo del contenido. Se encontró una opinión dividida respecto al uso de los resultados, la estructura y contenido del CECD, y el procedimiento de aplicación; así como observaciones críticas hacia el SED. Pero la valoración global ponderó la utilidad del SED. Se concluye sobre la necesidad de mejorar la usabilidad de los resultados de la evaluación docente.

Abstract

The aim of this study was to carry out a meta-evaluation of the Teaching Evaluation System (TES) of the Autonomous University of Baja California to identify its strengths and aspects of the educational quality in: the Survey of Teaching Competence Evaluation (STCE), the management system of the TES, the administration of the STCE and how the results are used. The study corresponds to a case study that adapted elements of the responsive evaluation model. The non-probabilistic sample was of 10,160 students who answered an on-line questionnaire. The data was processed with univariate analysis techniques and qualitative analysis of the content. It was found a divided opinion with respect to the use of the results, the structure and content of the STCE, and the procedure of application, as well as with the critical remarks towards the TES. However, the global assessment weighted the usefulness of the TES. Conclusions relate to the necessity to better the usability of the results of the teaching evaluation.

Palabras clave / Keywords

Enseñanza superior, Personal académico docente, Competencia del docente, Evaluación del docente, Cuestionario, Estudio de caso, Retroalimentación, Calidad de la educación.
Higher education, Academic teaching personnel, Teacher qualifications, Teacher evaluation, Questionnaires, Case studies, Feedback, Educational quality.

1. Introducción

La evaluación de la docencia se asume como una práctica social con implicaciones de carácter público y privado, y repercusiones para la sociedad, las instituciones y los docentes; constituye un recurso fundamental en la mejora de los procesos educativos, siempre y cuando se tome como una oportunidad de reflexionar sobre las características que se consideran esenciales en los procesos de enseñar y aprender. Asimismo, requiere ser coherente con la filosofía institucional, con el modelo pedagógico que sustenta la práctica docente y con las condiciones particulares del contexto donde se realiza la evaluación, solo así, se logrará que los resultados de la evaluación sean útiles para mejorar la docencia (Red Iberoamericana de Investigadores de la Evaluación de la Docencia [RIIED], 2008).

La docencia se reconoce como una actividad compleja y multidimensional (Coll y Solé, 2014; Martínez et al., 2016), dada su complejidad no puede ser valorada a través de una estrategia única, por tanto, la recomendación generalizada es utilizar diversas fuentes de información que posibiliten la comprensión de la enseñanza (Darwin, 2017).

Entre las estrategias utilizadas para evaluar la docencia destacan los cuestionarios de evaluación docente con base en la opinión de los estudiantes. Estos 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; Marsh, 2007), persisten críticas respecto a su utilidad tanto en evaluaciones formativas como sumativas (Spooren et al., 2013). En este sentido cabe observar que, 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).

Ante el uso predominante de los puntajes en la evaluación de la docencia universitaria, la literatura especializada advierte sobre su mal uso y abuso que invalida la confiabilidad y validez de los instrumentos (Linse, 2017). Desde la década de los noventa se insistió en la necesidad de ubicarlos en el contexto de un sistema de evaluación docente (Theall y Franklin, 2002) que incluya múltiples fuentes de información, ya que la complejidad de la enseñanza y de los diferentes contextos educativos hace que los cuestionarios de apreciación estudiantil sean insuficientes para valorar la docencia.

Respecto al contenido de los cuestionarios se demanda que incorporen los cambios contextuales particulares de las situaciones educativas, en específico las nuevas prácticas instruccionales derivadas de los planes de estudio por competencias, aprendizaje colaborativo y la educación a distancia (Spooren et al., 2013). Actualmente puede hacerse referencia a los nuevos desarrollos de instrumentos, por ejemplo, los propios para el campo de las ciencias de la salud y el ámbito de la supervisión clínica en medicina (Flores-Hernández et al., 2017; Villavicencio-Martínez y Luna, 2018). Al reconocer el gran poder orientador que tienen los procesos de evaluación docente persiste la necesidad de instrumentos válidos, confiables y pertinentes a los distintos contextos educativos.

Dada la complejidad de los procesos involucrados en la evaluación de la docencia se requiere monitorear sus componentes y fases de forma sistemática, lo que conlleva a la metaevaluación. La metaevaluación refiere una especialidad de la evaluación, definida como evaluación de la evaluación (Scriven, 1991). La metaevaluación presta atención tanto al proceso como a los resultados de la evaluación a partir del entorno y las circunstancias (Stake, 2017), por tanto, una metaevaluación alude a una labor situada y contextualizada. Desde esta perspectiva, el objetivo del presente estudio fue realizar una metaevaluación del Sistema de Evaluación Docente (SED) de la Universidad Autónoma de Baja California (UABC) con base en la opinión de los estudiantes, para identificar fortalezas y aspectos a mejorar en: el Cuestionario de Evaluación de la Competencia Docente (CECD), el sistema gestor del SED, la administración del CECD y el uso de los resultados.

La metaevaluación tiene el objetivo de valorar la evaluación para fundamentar y defender resultados alineados al entorno (Scriven, 2011, 2015). Asimismo, es deseable que la metaevaluación brinde un medio válido para resaltar los hallazgos de mayor relevancia tanto como los dudosos (Stake, 2017). En congruencia con las funciones típicas de los procesos de evaluación, puede tener una función formativa para ayudar a la mejora de la evaluación, o una función sumativa en apoyo a la determinación del logro y valor de la evaluación (Stufflebeam, 2011).

La metaevaluación representa el conjunto de conocimiento evaluativo que incorpora los estándares para el desarrollo y práctica de un sistema de evaluación (Scriven, 2015). Se propone como un mecanismo de garantía de calidad para fortalecer los sistemas de evaluación, promover la producción de servicios defendibles, y estudiar la solidez y equidad de las evaluaciones (Stufflebeam, 2001). En este sentido, la

metaevaluación incorpora la justificación y crítica de la evaluación, por tanto, se reconoce como la conciencia de la evaluación (Scriven, 2009).

Un referente de metaevaluación en el escenario europeo es lo realizado al Programa de apoyo para la evaluación de la actividad docente del profesorado universitario (DOCENTIA) de la Universidad Complutense de Madrid. Esta metaevaluación se realizó con la finalidad de identificar las fortalezas y aspectos susceptibles de mejora del programa, además de proponer recomendaciones que favorezcan su eficiencia y utilidad (Gallego y Rojo, 2014). Entre los objetivos destacaron estudiar la dimensionalidad de los instrumentos, indagar la capacidad de discriminación entre los procedimientos de evaluación de los docentes, y plantear un sistema adecuado de producción de las puntuaciones para los mismos.

En el contexto mexicano, en la Universidad Autónoma del Estado de Hidalgo se realizó una metaevaluación del proceso de evaluación docente, con el propósito de sostener un proceso permanente de evaluación apegado al modelo educativo de la universidad, que apoye la calidad de la enseñanza y la formación docente (Elizalde et al., 2008; Elizalde et al., 2017).

Si bien es cierto que, la noción de metaevaluación no es nueva, 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.

1.1 Contexto del estudio

Este trabajo se desarrolló en la UABC en México. En el segundo periodo semestral de 2021 la matrícula de estudiantes de licenciatura fue de 66,029, y los académicos (docentes de tiempo completo, de medio tiempo, de asignatura, técnicos académicos e investigadores) eran 6,090 (UABC, 2021d).

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 por horas (UABC, 2021c, sección "Evaluación Docente").

La evaluación docente se realiza por medio del SED, que gestiona y almacena las respuestas sobre la evaluación del desempeño de los docentes con base en la opinión de los estudiantes (UABC, 2021b). 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 CECD se aplica de manera automatizada y programada en el calendario escolar durante el último mes de cada periodo semestral. Se generan tipos de reportes de acuerdo con los destinatarios: docentes, directivos de unidades académicas y directores de la administración central (UABC, 2021b).

2. Método

El diseño corresponde a un estudio de caso donde se consideraron elementos del modelo de evaluación responsiva. La evaluación responsiva atiende principalmente tres aspectos: 1) las actividades del programa no a las intenciones, 2) las exigencias de información de la audiencia, y 3) las diferentes opiniones de valor de las personas involucradas para informar sobre la aceptación del programa de evaluación (Stake, 1976, 2011).

2.1. Participantes

Se obtuvo una muestra no probabilística por autoselección de 10,160 estudiantes matriculados en el segundo periodo semestral de 2019, correspondiente del 15.9% del total de estudiantes. Los participantes de sexo femenino fueron 5,298 (52.1%) y 3,740 (36.8%) de sexo masculino; 1,122 no especificaron su sexo (11.0%). Asimismo, 7,439 correspondieron a estudiantes regulares (73.2%) y 1,596 (15.7%) irregulares; 1,125 (11.0%) no indicó su estatus escolar. Respecto a la etapa de formación curricular, 4,320 estudiantes se concentraron en formación básica (42.5%), 2,884 en disciplinar (28.3%) y 1,859 en terminal (18.3%); 1,097 no lo refirió

(10.8%); este dato implica que la muestra se conformó por estudiantes de todos los semestres escolares, con mayor concentración en los dos primeros años, seguidos del tercer año y por último del cuarto.

La distribución de los estudiantes por área de conocimiento fue: Ciencias Económico Administrativas 2,014 (19.8%), Ciencias de la Ingeniería y la Tecnología 1,974 (19.4%) y Ciencias Sociales 1,939 (19.1%), Ciencias de la Educación y Humanidades 1,358 (13.3%), y en Ciencias de la Salud el 1,358 (13.3%). Por último, Ciencias Naturales y Exactas concentró a 350 (3.4%) y Ciencias Agropecuarias 173 (1.7%).

Respecto a la edad, 3,046 estudiantes (29.9%) tenían entre 19 y 20 años; 2,627 (25.8%) entre 21 y 22 años; 1,242 (12.2%) entre 23 y 24 años; 1,215 (11.9%) tenían 25 años o más y 1,032 (10.1%) entre 17 y 18 años. 998 estudiantes (9.8%) no indicaron su edad.

También, se contó con la participación de dos coordinadores de la administración central, 60% de directores y subdirectores del total, y el 24.60% de docentes del total en el periodo 2019-2.

2.2. Procedimiento

Etapa 1. Diseño del cuestionario. La base teórica utilizada a lo largo de esta etapa fueron los planteamientos propuestos por la RIIED (2008) y Rueda (2010), respecto al diseño y puesta en marcha de este tipo de programas.

En principio, se realizaron entrevistas a los coordinadores de la administración central, con el fin de conocer e indagar los conocimientos, opiniones y expectativas que tenían sobre el SED. Enseguida, con base en la información derivada de las entrevistas y los referentes teóricos se elaboró un cuestionario dirigido a los directores y subdirectores de las unidades académicas; posteriormente, a la luz de la información recuperada fue diseñado el cuestionario para los docentes; por último, del discurso producido por los docentes, se retomaron los aspectos clave para desarrollar el cuestionario de los estudiantes, asimismo, un aspecto central que se consideró fue cuestionar únicamente los aspectos factibles de ser identificados por los estudiantes en relación con el SED.

De esta forma, el cuestionario se estructuró en cuatro dimensiones: *Dimensión 1. Uso de los resultados de evaluación docente* con cuatro ítems, *Dimensión 2. Cuestionario de evaluación de la competencia docente* con nueve ítems, *Dimensión 3. Procedimiento de aplicación* con 10 ítems, y *Dimensión 4. Valoración del SED* con dos ítems. En suma, el cuestionario se conformó de 36 ítems: 32 de respuesta cerrada (16 de escala, 12 de opción múltiple y cuatro dicotómicos) y cuatro de respuesta abierta (ver Tabla 1). Para los ítems de respuesta cerrada, se estableció una escala con valoración del 1 (la más baja) al 10 (la más alta), también, respuestas de opción múltiple alineadas a la naturaleza del ítem, y respuestas dicotómicas: sí y no. Respecto a los ítems abiertos, las respuestas fueron de extensión libre y opcionales en cada dimensión del cuestionario.

Tabla 1

Número de ítems respecto a su dimensión y tipo de respuesta

Dimensiones	Ítems				Totales
	Respuesta cerrada			Respuesta abierta	
	Escala del 1 al 10	Opción múltiple	Dicotómicos		
Información general		7	4		11
1. Uso de los resultados de evaluación docente	1	2		1	4
2. Cuestionario de evaluación de la competencia docente	8			1	9
3. Procedimiento de aplicación	6	3		1	10
4. Valoración del SED	1			1	2
N.º de ítems respuesta cerrada	16	12	4		
Total de ítems respuesta cerrada y abierta		32		4	
Total de ítems en el cuestionario					36

Etapa 2. Aplicación del cuestionario. El cuestionario fue aplicado y gestionado en línea en el *software Limesurvey versión 1.92*. El enlace de acceso al cuestionario fue enviado a todos los estudiantes de la UABC por medio de una invitación a su correo electrónico institucional, el tiempo estimado de respuesta del cuestionario fue de 10 minutos. Transcurridos 15 días, se les envió un recordatorio a los estudiantes que aún no respondían el cuestionario. El índice de respuestas se monitoreó a través del software.

Etapa 3. Análisis de la información. Los datos recuperados fueron separados para su análisis a partir de la fuente de respuesta de los ítems: los datos de las respuestas cerradas se concentraron en una base de datos en el *software SPSS versión 26*, a fin de realizar un análisis univariado para estudiar las variables de forma individual (Afifi et al., 2020), así, se generaron las frecuencias y porcentajes de cada ítem para conocer las características de los datos; los datos de las respuestas abiertas se concentraron en un archivo del *software Atlas.ti versión 7.5.4* y se analizaron con la técnica Análisis Cualitativo del Contenido (ACC), identificado como una indagación empírica y metodológica desde un entorno de comunicación, con el propósito de construir categorías de análisis a partir de un enfoque de desarrollo inductivo-deductivo (Mayring, 2000). En esta tarea, se hizo una depuración de los datos, se prescindió de respuestas con solo símbolos, letras aisladas, signos de puntuación y respuestas ajenas al tema.

3. Resultados

Los resultados del análisis univariado mostraron que, en la *Dimensión 1 Uso de los resultados* poco más de la mitad de los estudiantes consideraron que los docentes usan los puntajes para identificar fortalezas y debilidades de su práctica docente (53.5%), sin embargo, 27.8% señalaron que los docentes no toman en cuenta los puntajes, y 18.6% desconocen el uso que los docentes le dan a los mismos (ver Tabla 2).

Tabla 2

Porcentaje de las respuestas sobre el uso de los puntajes de evaluación docente

Ítem	Respuestas	%
En tu opinión, ¿qué uso dan los docentes a los resultados de evaluación docente?	Identificar fortalezas y debilidades de su práctica docente	53.5
	Los resultados no son tomados en cuenta	27.8
	No sé	18.6

En la *Dimensión 2. Cuestionario de evaluación de la competencia docente*, los estudiantes indicaron que el CECD incluye elementos importantes de la práctica docente con 66.3 puntos (resultado de la suma de los valores 8, 9 y 10; de igual manera, los siguientes enunciados de puntos alcanzados corresponden a la sumatoria de dichos valores); además otorgaron puntuación elevada a: incorporar aspectos referentes a la puntualidad 82.5, cumplimiento del programa 82 y uso de las Tecnologías de la Información y la Comunicación (TIC) 74.9 (ver Tabla 3).

Tabla 3

Puntaje de las respuestas sobre elementos del CECD

Ítem	Promedios de valoración respecto a la escala									
	1	2	3	4	5	6	7	8	9	10
En tu opinión, ¿el CECD incluye los elementos importantes de la práctica docente?	1.5	0.7	1.4	2.0	8.0	7.1	12.75	25.5	18.7	22.1
¿Qué tan de acuerdo estarías en incluir en el CECD valorar la puntualidad?	1.5	0.3	0.8	0.8	4.6	3.2	5.6	12.0	13.9	56.6
¿Qué tan de acuerdo estarías en incluir en el CECD valorar el cumplimiento del programa?	1.5	0.3	0.8	0.8	4.7	3.4	5.9	14.4	16.2	51.4
¿Qué tan de acuerdo estarías en incluir en el CECD valorar el uso de TIC en el aula?	2.0	0.7	1.2	1.4	6.5	4.6	8.1	16.3	15.3	43.3

Respecto a la *Dimensión 3. Procedimiento de aplicación*, los estudiantes informaron que el ingreso a la plataforma para contestar el CECD es fácil con 66.5 puntos, que la obligatoriedad para responder el CECD es adecuada con 59.8 puntos, la forma de aplicación del CECD en su unidad académica es favorable con

64.9 puntos, estimaron conveniente la difusión con 61.8 puntos, y propicias las fechas de aplicación del cuestionario con 58 puntos (ver Tabla 4).

Tabla 4

Puntaje de las respuestas sobre el procedimiento de aplicación del CECD

Ítem	Porcentaje de valoración respecto a la escala									
	1	2	3	4	5	6	7	8	9	10
¿Te es fácil ingresar a la plataforma para contestar al CECD?	5.5	1.8	2.4	2.8	6.4	5.3	8.5	13.1	13.1	40.3
¿Te parece adecuado que sea obligatorio responder el CECD?	9.9	2.2	2.8	2.9	9.0	5.3	7.6	13.7	11.9	34.2
¿Te parece adecuada la manera en que en tu unidad académica se realiza la aplicación del CECD?	4.2	1.6	2.1	2.4	8.8	5.7	9.8	18	16.2	30.7
¿Es adecuada la difusión para promover que contesten el CECD?	4.8	2.2	2.8	3.4	8.4	6.0	10.3	17.6	14.7	29.5
¿Son convenientes las fechas de aplicación del CECD?	8.5	2.2	2.8	3.1	7.9	6.8	10.3	17.2	13.9	26.9

De igual manera, el 40% de los estudiantes señalaron ingresar en una sola ocasión al CECD con éxito, el 22.2% dos veces, 14.8% tres veces, 5.7% cuatro veces y 17.1% cinco veces o más. Asimismo, expresaron que los problemas para ingresar al cuestionario se han debido principalmente a dificultades con el sistema gestor del SED (51.7%), seguido de inconvenientes con el servicio de internet de la UABC (37.6%), y a causa de no recordar la contraseña de acceso (10.5%) (ver Tabla 5).

Tabla 5

Porcentaje de las respuestas sobre el acceso al CECD

Ítem	Respuestas	%
¿Cuántas veces intentas ingresar al cuestionario hasta tener éxito?	1	40
	2	22.2
	3	14.8
	4	5.7
	5 o más	17.1
Tus problemas para ingresar al cuestionario se debieron a:	Problemas en el sistema gestor del SED (solo la página del cuestionario no funcionaba)	51.7
	Problemas de la Red UABC (por ejemplo, sin conexión a Internet)	37.6
	No recordé mi contraseña	10.5

Referente a la *Dimensión 4. Valoración del SED*, los estudiantes otorgaron 70.6 puntos en la valoración sobre la utilidad del SED (ver Tabla 6).

Tabla 6

Puntaje de las respuestas sobre la valoración del SED

Ítem	Promedios de valoración respecto a la escala									
	1	2	3	4	5	6	7	8	9	10
En tu opinión, ¿qué tan útil consideras el SED?	3.2	0.9	1.7	2.2	6.4	5.1	9.3	18.3	17.2	35.1

Con relación al ACC, cabe precisar que los resultados corresponden al corpus derivado de 1,424 cuestionarios (14% de la muestra) con respuestas a las preguntas abiertas. En total, se obtuvieron 2,132 datos, distribuidos por dimensión tal como se presenta en la Tabla 7.

Tabla 7*Número de datos por dimensión del cuestionario*

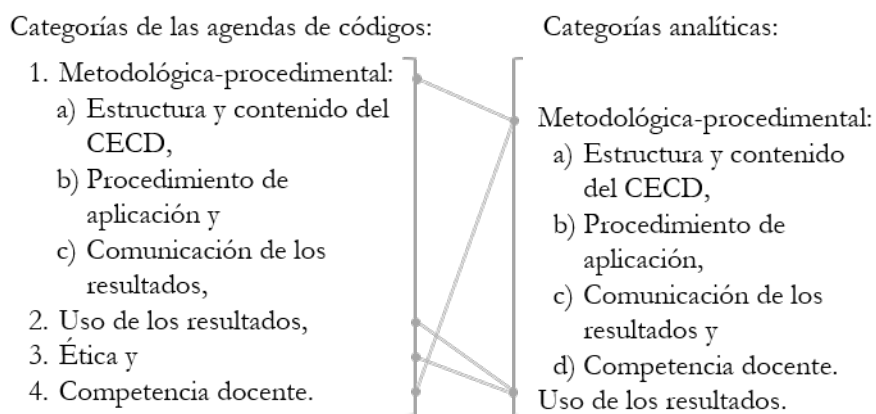
	Dimensiones	Número de datos
1.	Uso de los resultados de evaluación docente	1032
2.	Cuestionario de evaluación de la competencia docente	371
3.	Procedimiento de aplicación	314
4.	Valoración del SED	415
	Suma	2,132

Se elaboraron cuatro agendas de códigos, una por cada dimensión del cuestionario. La construcción de las agendas se basó en la definición de categorías establecidas bajo un término específico de clasificación (Holsti, 1966), integradas por códigos, reglas de codificación y ejemplos para representar de forma particular los criterios de codificación de los datos (Mayring, 2000). Las categorías fueron consolidadas con el análisis consecuente de los datos y las dimensiones propuestas por la RIIED (2008; Rueda, 2010).

Las agendas de códigos integraron cuatro categorías: Metodológica-procedimental, Uso de los resultados, Ética y Competencia docente. La categoría Metodológica-procedimental abordó aspectos del CECD y la generación de reportes (Rueda, 2010), con este fin se dividió en tres subcategorías: Estructura y contenido del CECD, Procedimiento de aplicación y Comunicación de los resultados. La categoría Uso de los resultados se refirió al empleo y utilidad de los resultados de la evaluación (RIIED, 2008; Rueda, 2010); la categoría Ética integró aspectos concernientes a la ética 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 Competencia docente incorporó las características relacionadas con el dominio disciplinar, pedagógico y compromiso profesional del docente.

Con base en las categorías de las agendas de códigos se derivaron 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 y cuestiones éticas (ver Figura 1).

**Figura 1. Relación entre categorías de las agendas de códigos y categorías analíticas**

Categoría analítica Metodológica-Procedimental. A continuación, los resultados en relación con cada subcategoría.

Subcategoría Estructura y contenido del cuestionario. Un tema recurrente fue la extensión del CECD cuando se evalúa una asignatura en su componente teórico y práctico, dado que al sumar el número de preguntas de los dos cuestionarios (teoría y práctica) resulta excesivo, "Me he dado cuenta de que la mayoría de las evaluaciones son extensas, lo que hace que muchos alumnos pierdan el interés por contestarla honestamente por los tiempos que les lleva leer el documento" (209_DV-817). Además, algunos estudiantes resaltaron que el cuestionario de práctica se compone por preguntas que no corresponden a la naturaleza de la unidad de aprendizaje o área del conocimiento en la que estudian: "La mayoría de los reactivos no concuerdan o no aplican para las rotaciones clínicas de medicina, las cuales son nuestro campo práctico" (2019_DC-2).

Por otro lado, los estudiantes también manifestaron la necesidad de introducir aspectos relacionados con la ética profesional docente: “También deberían de evaluar la ética profesional, responsabilidad, disposición y asistencia, ya que una cosa es que firmen las listas de asistencia y otra que en realidad se presenten a dar y clases” (2019_DC-360).

Subcategoría Procedimiento de aplicación. Un tema correspondió al periodo de aplicación del CECD, los estudiantes expresaron que resulta inconveniente realizar la evaluación docente al final del semestre debido a la carga de compromisos académicos por culminar: “El tiempo para realizar dichas evaluaciones son inapropiados ya que en esos tiempos el alumnado se encuentra en periodos de evaluación o entregas de proyectos finales” (2019_DV-29).

Por otra parte, se observaron aspectos alusivos al sistema gestor del SED como: limitaciones con el tiempo de respuesta, “El formato de la evaluación tiene una deficiencia en cuanto al tiempo para contestarla, ya que al agregar comentarios extras, y querer guardar los resultados, se recarga la página y tienes que contestar todas las casillas de nuevo” (2019_DU-419); dificultades de acceso, “Nunca logre acceso a la página, marcaba error de página o error de usuario” (2019_DP-29); y saturación del sistema gestor, “Considero que los problemas en cuanto al acceso a la evaluación son porque los alumnos siempre lo dejan al último y la red se satura” (2019_DP-300).

Asimismo, los estudiantes mencionaron la falta de una interfaz amigable, útil y funcional: “La interfaz del cuestionario es muy molesta, difícil de llenar y en ocasiones se reinicia” (2019_DC-328). Además, su manejo en dispositivos móviles “Debería optimizarse para dispositivos móviles ya que no todos tenemos siempre a la mano la computadora y es más sencillo y rápido hacerlo desde el teléfono celular” (2019_DP-130).

También hicieron referencia a su inconformidad por la obligatoriedad de la evaluación docente: “No debe de ser obligatorio el responder una evaluación, (...) es más fácil responder todo perfecto y terminar la evaluación que tomarte tu tiempo para poner lo que en realidad piensas. (...) Si nos quieren integrar en el proceso de evaluación está bien pero no debería de ser obligatorio” (2019_DU-322). En específico, consideraron desfavorable que la evaluación se asocie a la reinscripción: “Pero que no sea de índole obligatoria para la inscripción porque eso puede provocar cierto estrés” (2019_DP-52). Por tanto, se propone que la evaluación sea de carácter opcional: “Sería bueno que el cuestionario fuera opcional” (2019_DP-222).

Otra cuestión reconocida fueron los sesgos en la evaluación: “Un problema muy grande es que los estudiantes responden de manera deshonesta para que sea más rápido terminar la evaluación docente” (2019_DU-260); en algunos casos, la evaluación puede ser empleada por los estudiantes con base en el agrado: “Yo siento que muchos alumnos no califican de manera honesta y utilizan esta herramienta para calificar injustamente a profesores que no les cayeron bien (no en todos los casos)” (2019_DU-270). Por tanto, los estudiantes propusieron concientizar sobre la importancia y alcances de la evaluación docente: “Los resultados serían más provechosos si el alumnado realizara como es debido la evaluación. Hay que hacer más consciencia sobre la importancia de esta evaluación como una herramienta de ayuda a la mejora de la planta docente” (2019_DU-819).

Subcategoría Comunicación de los resultados. Los estudiantes expresaron su interés sobre los resultados de la evaluación con la finalidad de conocer el desempeño de sus docentes: “Supongo que, estaría bien tener el historial del maestro a disposición. Así como, para poder seleccionar bien los profesores y las materias que deseo cursar” (2019_DU-1002). Adicionalmente, compartieron su inquietud por conocer si los docentes fueron informados sobre la evaluación que ellos realizaron: “Que nos hagan saber si se les comunicó los resultados a los docentes y algún medio por donde nos informen la postura de ellos y los cambios que harán en su forma de manejar las clases y su contenido” (2019_DU-728).

Subcategoría Competencia docente. Por un lado, algunos estudiantes manifestaron conformidad en relación a sus docentes: “todos son buenos maestros” (2019_DU-226); y por otro, exteriorizaron comentarios dirigidos a la carencia de estrategias didácticas en el aula: “(...) hay maestros que no planean actividades e improvisan lo que harán en el salón de clases” (2019_DC-80).

Categoría analítica Uso de los Resultados. Se observó que los resultados de la evaluación no son considerados para generar cambios: “Siento que no se toman en cuenta los resultados obtenidos o que no toman cartas en el asunto respecto a comentarios que hacemos como alumnos para la mejora educativa de los maestros hacia nosotros” (2019_DU-89).

Además, algunos estudiantes valoraron la oportunidad de expresarse: “Agradezco que se tomen en cuenta nuestras opiniones como alumnos” (2019_DU-731), pero solicitan que su opinión sea considerada por los

tomadores de decisiones para mejorar el desempeño docente: “El concepto de la evaluación es muy bueno, pero hace falta que escuchen los resultados y las opiniones de los alumnos respecto a los docentes” (2019_DU-182).

En síntesis, los resultados derivados de las preguntas abiertas permitieron conocer las observaciones críticas de los estudiantes hacia el SED.

4. Discusión y conclusiones

Este trabajo reconoce la necesidad de realizar ejercicios de metaevaluación en los programas de evaluación de la docencia. Se desarrolló en la perspectiva de atender tanto al proceso como los resultados de la evaluación de forma situada y contextualizada (Stake, 2017). Asimismo, se propuso realizar un análisis detallado de los problemas reales encontrados en la evaluación de la práctica docente desde la perspectiva de los estudiantes. En consecuencia, acceder a la diversidad de opiniones sobre el SED.

En principio, dado que el SED utiliza únicamente un tipo de instrumento de evaluación (el CECD con base en la opinión de los alumnos), los resultados giran en torno a este tipo de estrategia de evaluación. El análisis univariado mostró una opinión dividida respecto al uso de los resultados, debido a que el 53.5% de los estudiantes opinaron que los docentes usan los puntajes para identificar fortalezas y debilidades de su práctica docente, y el 27.8% señalaron que los docentes no toman en cuenta los puntajes. Esta apreciación coincide con la discusión de la literatura en torno a que, los puntajes son utilizados por los docentes como fuente de información sobre las fortalezas y debilidades de su práctica docente (Nasser-Abu, 2017). Al mismo tiempo, persiste la discusión sobre los usuarios que no están convencidos de la utilidad de los puntajes, tanto con fines formativos como sumativos (Spooren et al., 2013); o el limitado uso por parte de los docentes (Spooren y Christiaens, 2017). Para lo cual se advierte en comunicación de los resultados su interés por que los puntajes sean comunicados a las distintas audiencias interesadas, en especial a los docentes.

En relación con el CECD se obtuvieron 66.3 puntos a favor de que incluye elementos importantes de la práctica docente. Además, propusieron incorporar aspectos referentes a la puntualidad, cumplimiento del programa y uso de las TIC; y observaron que en las asignaturas del campo clínico de medicina el cuestionario de práctica no corresponde a la naturaleza del proceso de enseñanza. Al respecto, cabe considerar la variabilidad en la literatura sobre los elementos que definen a la docencia efectiva o de calidad, dado que la buena docencia depende del contexto (Hativa, 2015); la necesidad de incorporar los cambios contextuales propios de cada situación educativa (Luna y Reyes, 2015); así como, los desarrollos en instrumentos propios para el campo de las ciencias de la salud y el ámbito de la supervisión clínica en medicina (Flores-Hernández et al., 2017; Villavicencio-Martínez y Luna, 2018).

Sobre el procedimiento de aplicación, si bien se informó que el ingreso a la plataforma para contestar el CECD es fácil 66.5 puntos, que la obligatoriedad para responder el CECD es adecuada con 59.8, así como las fechas de aplicación del cuestionario con 58 puntos. También se precisaron las discrepancias, como son las dificultades con el formato de respuesta; por ejemplo, en las asignaturas con un componente teórico y práctico, resulta excesivo el número de reactivos de evaluación, puesto que, al desplegarse como uno solo los dos cuestionarios (el de teoría y práctica) el número de reactivos se duplica.

Por otra parte, se advirtieron posibles sesgos en los puntajes derivados de la suposición de que los estudiantes contestan de forma deshonesta. En este tema existe suficiente evidencia empírica que apoya la confiabilidad de los puntajes en virtud de que los efectos de sesgo son muy pequeños (Spooren et al., 2017). Finalmente, en relación con la competencia docente se advirtió sobre el déficit en algunos docentes en la competencia de planeación didáctica.

En el balance general, sobre la valoración del SED los estudiantes ponderaron su utilidad. Este es un dato relevante en la consideración de que validar la evaluación es hacerla útil para los implicados (Jornet et al., 2020). Los resultados de este trabajo de metaevaluación derivan en recomendaciones puntuales orientadas hacia la mejora del SED. Entre las que destacan: 1) incorporar otras estrategias de evaluación que posibiliten acceder a la complejidad de la docencia; 2) mejorar la comunicación de los resultados a los docentes; 3) revisar el contenido del CECD a la luz de las demandas del contexto actual; 4) incluir cuestionarios de evaluación de la docencia por los estudiantes particulares para el ámbito de la enseñanza en ciencias de la salud y el campo clínico de medicina; 5) revisar los aspectos que atañen al sistema gestor del CECD.

Mejorar la calidad de los procesos de evaluación de la docencia involucra realizar ejercicios de metaevaluación que revelen tanto los aciertos como los problemas sentidos por la comunidad, con el fin de ajustar los componentes del sistema a la realidad institucional para la mejora de los procesos de enseñanza y aprendizaje.

Referencias

- Afifi, A., May, A., Donatello, R. A., y Clark, V. A. (2020). Data visualization. En *Practical Multivariate Analysis* (6ta. ed., pp. 37-58). Taylor & Francis Group.
- Benton, S. L., y Cashin, W. E. (2014). Student ratings of instruction in college and university courses. En Michael B. Paulsen (Ed.), *Higher Education: Handbook of Theory & Research* (vol. 29, pp. 279-326). Springer.
- Coll, C., y Solé, I. (2014). Enseñar y aprender en el contexto del aula. C. Coll, J., Palacios y Á. Marchesi (Comps.), *Desarrollo psicológico y educación: 2. Psicología de la educación escolar* (2da ed., pp. 357-386). Editorial Alianza.
- Darwin, S. (2017). What contemporary work are student ratings actually doing in higher education? *Studies in Educational Evaluation*, 54, 13-21. <http://dx.doi.org/10.1016/j.stueduc.2016.08.002>
- 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://bit.ly/3nqtXdB>
- Elizalde, L., Olvera, B. I., y Bezies, P. (2017). Metaevaluación: herramienta para la valoración integral del proceso evaluativo docente de la UAEH. En E. Luna y M. Rueda (Coords.), *Experiencias de evaluación de la docencia en Iberoamérica* (pp. 257-278). Universidad Autónoma de Baja California.
- 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://bit.ly/3nndFIJ>
- Hativa, N. (2015). ¿Qué dice la investigación sobre la buena enseñanza y los excelentes profesores? *Docencia en la academia*, 5, 42-75 [Hebreo]. <https://bit.ly/3noMdEJ>
- Holsti, O. R. (1966). Coding content data. En *Content Analysis for the social humanities* (pp. 94-126). Addison-Wesley Publishing Company.
- 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., y Reyes, É. (2015). Validación de constructo de un cuestionario de evaluación de la competencia docente. *Revista electrónica de investigación educativa*, 17(3), 13-27. <https://bit.ly/3a04GE2>
- 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. <https://doi.org/10.1016/j.stueduc.2016.12.004>
- Martínez, F., Taut, S., y Schaaf, K. (2016). Classroom observation for evaluating and improving teaching: An international perspective. *Studies in Educational Evaluation*, 49, 15-29. <http://dx.doi.org/10.1016/j.stueduc.2016.03.002>
- Marsh, H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and usefulness. En R. P. Perry y J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidencebased perspective* (pp. 319-384). Springer.
- Mayring, P. (2000). Qualitative content analysis. *Forum qualitative social research*, 1(2). <https://bit.ly/3ubZpAl>
- 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>
- 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://bit.ly/3OTxn4B>
- 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://bit.ly/3a2hutK>
- Scriven, M. (1991). *Evaluation thesaurus* (4ta. ed.). Sage publications.
- Scriven, M. (2009). Meta-Evaluation Revisited. *Journal of MultiDisciplinary Evaluation*, 6(11). <https://bit.ly/3npC8Ha>
- Scriven, M. (2011). Evaluating evaluations: A meta-evaluation checklist (9na ed.). Claremont Graduate University. <https://bit.ly/3yslhc1>
- Scriven, M. (2015). *The Meta-Evaluation Checklist*. Claremont Evaluation Center.
- 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., y Christiaens, W. (2017). I liked your course because I believe in (the power of) student evaluations of teaching (SET). Students' perceptions of a teaching evaluation process and their relationships with SET scores. *Studies in educational evaluation*, 54, 43-49. <http://dx.doi.org/10.1016/j.stueduc.2016.12.003>
- 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. (1976). A theoretical statement of responsive evaluation. *Studies in Educational Evaluation*, 2(1), 19-22. [https://doi.org/10.1016/0191-491X\(76\)90004-3](https://doi.org/10.1016/0191-491X(76)90004-3)
- Stake, R. E. (2011). Program evaluation particularly responsive evaluation. *Journal of MultiDisciplinary Evaluation*, 7(15), 180-201. <https://bit.ly/3OVHJru>
- 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. (2001). The Metaevaluation Imperative. *American Journal of Evaluation*, 22(2), 183-209. [https://doi.org/10.1016/S1098-2140\(01\)00127-8](https://doi.org/10.1016/S1098-2140(01)00127-8)
- Stufflebeam, D. L. (2011). Meta-evaluation. *Journal of MultiDisciplinary Evaluation*, 7(15), 99-158. <https://bit.ly/3Ab9v8h>
- Theall, M., y Franklin, J. (2002). Creating Responsive Student Rating Systems to Improve Evaluation Practice. *New Directions for Teaching and Learning*, 2000(83), 95-107. <https://doi.org/10.1002/tl.8308>
- Universidad Autónoma de Baja California. (2021a). *Anexo 1: Tabla de identificación de actividades (TIA)*. <https://bit.ly/39ZG9PD>
- Universidad Autónoma de Baja California. (2021b). *Evaluación de la docencia de licenciatura en la UABC: Descripción general del proceso*. <https://bit.ly/3Aan9IG>
- Universidad Autónoma de Baja California. (2021c, 28 de diciembre). *Formación docente*. <https://bit.ly/3QRPNV0>
- Universidad Autónoma de Baja California. (2021d, 15 de septiembre). *Sistema de indicadores y estadísticas institucionales*. <https://bit.ly/3xWGquy>
- Villavicencio-Martínez, R. A., y Luna, E. (2018). Diseño y validación de un cuestionario de evaluación de la supervisión clínica. *Revista Electrónica de Investigación y Evaluación Educativa*, 24(1). <http://doi.org/10.7203/relieve.24.1.9672>