

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
Instituto de Investigación y Desarrollo Educativo



D O C T O R A D O

“Metaevaluación de un sistema de evaluación docente en educación superior”

Tesis
Que para obtener el grado de
Doctora en Ciencias Educativas

Presenta

Alma Delia Hernández Villafaña

Enero, 2022

Ensenada, Baja California, México



Universidad Autónoma de Baja California
Instituto de Investigación y Desarrollo Educativo
Doctorado en Ciencias Educativas



**“Metaevaluación de un sistema de evaluación docente
en educación superior”**

TESIS

Que para obtener el grado de

DOCTOR(A) EN CIENCIAS EDUCATIVAS

Presenta

Alma Delia Hernández Villafaña

APROBADO POR:

Dra. Edna Luna Serrano
Directora de tesis

Dra. Graciela Cordero Arroyo
Sinodal

Dr. Luis Horacio Pedroza Zúñiga
Sinodal

Dr. Salvador Ponce Ceballos
Sinodal

Dr. Mario Rueda Beltrán
Sinodal



UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
Instituto de Investigación y Desarrollo Educativo

Ensenada, B.C., a 30 de noviembre de 2022

ASUNTO: Voto aprobatorio al trabajo
de tesis para el grado de Doctor(a) en Ciencias Educativas

Dr. Sergio Gerardo Malaga Villegas
Coordinador(a) de Investigación y Posgrado
Presente.

Después de haber efectuado una revisión minuciosa sobre el trabajo de tesis presentado por la C.
Alma Delia Hernández Villafaña, me permito comunicarle que he dado mi **VOTO**
APROBATORIO al mencionado trabajo.

Con base en lo anterior, dicho documento se considera listo para su defensa en el examen de grado
de Doctor en Ciencias Educativas, con el trabajo titulado:

“Metaevaluación de un sistema de evaluación docente en educación superior”.

Esperando reciba el presente de conformidad, quedo de usted.

Atentamente

Dra. Edna Luna Serrano



UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
Instituto de Investigación y Desarrollo Educativo

Ensenada, B.C., a 30 de noviembre de 2022

ASUNTO: Voto aprobatorio al trabajo de tesis para el grado de Doctor(a) en Ciencias Educativas

Dr. Sergio Gerardo Malaga Villegas
Coordinador(a) de Investigación y Posgrado
Presente.

Después de haber efectuado una revisión minuciosa sobre el trabajo de tesis presentado por la **C. Alma Delia Hernández Villafaña**, me permito comunicarle que he dado mi **VOTO APROBATORIO** al mencionado trabajo.

Con base en lo anterior, dicho documento se considera listo para su defensa en el examen de grado de Doctor en Ciencias Educativas, con el trabajo titulado:

“Metaevaluación de un sistema de evaluación docente en educación superior”.

Esperando reciba el presente de conformidad, quedo de usted.

Atentamente



Dra. Graciela Cordero Arroyo



UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
Instituto de Investigación y Desarrollo Educativo

Ensenada, B.C., a 30 de noviembre de 2022

ASUNTO: Voto aprobatorio al trabajo de tesis para el grado de Doctor(a) en Ciencias Educativas

Dr. Sergio Gerardo Malaga Villegas
Coordinador(a) de Investigación y Posgrado
Presente.

Después de haber efectuado una revisión minuciosa sobre el trabajo de tesis presentado por la **C. Alma Delia Hernández Villafaña**, me permito comunicarle que he dado mi **VOTO APROBATORIO** al mencionado trabajo.

Con base en lo anterior, dicho documento se considera listo para su defensa en el examen de grado de Doctor en Ciencias Educativas, con el trabajo titulado:

“Metaevaluación de un sistema de evaluación docente en educación superior”.

Esperando reciba el presente de conformidad, quedo de usted.

Atentamente



Dr. Luis Horacio Pedroza Zúñiga



UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
Instituto de Investigación y Desarrollo Educativo

Ensenada, B.C., a 30 de noviembre de 2022

ASUNTO: Voto aprobatorio al trabajo
de tesis para el grado de Doctor(a) en Ciencias Educativas

Dr. Sergio Gerardo Malaga Villegas
Coordinador(a) de Investigación y Posgrado
Presente.

Después de haber efectuado una revisión minuciosa sobre el trabajo de tesis presentado por la **C. Alma Delia Hernández Villafaña**, me permito comunicarle que he dado mi **VOTO APROBATORIO** al mencionado trabajo.

Con base en lo anterior, dicho documento se considera listo para su defensa en el examen de grado de Doctor en Ciencias Educativas, con el trabajo titulado:

“Metaevaluación de un sistema de evaluación docente en educación superior”.

Esperando reciba el presente de conformidad, quedo de usted.

Atentamente

Dr. Salvador Ponce Ceballos



UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
Instituto de Investigación y Desarrollo Educativo

Ensenada, B.C., a 30 de noviembre de 2022

ASUNTO: Voto aprobatorio al trabajo
de tesis para el grado de Doctor(a) en Ciencias Educativas

Dr. Sergio Gerardo Malaga Villegas
Coordinador(a) de Investigación y Posgrado
Presente.

Después de haber efectuado una revisión minuciosa sobre el trabajo de tesis presentado por la **C. Alma Delia Hernández Villafaña**, me permito comunicarle que he dado mi **VOTO APROBATORIO** al mencionado trabajo.

Con base en lo anterior, dicho documento se considera listo para su defensa en el examen de grado de Doctor en Ciencias Educativas, con el trabajo titulado:

“Metaevaluación de un sistema de evaluación docente en educación superior”.

Esperando reciba el presente de conformidad, quedo de usted.

Atentamente

Dr. Mario Rueda Beltrán

Dedicatoria

Para Daniela y Michelle siempre

Agradecimientos

A Dios por su perfecto amor.

A mi familia por su constancia y paciencia.

A la Dra. Edna por su invaluable instrucción, consejo y cuidado.

A mis amigas y compañeros del doctorado por su compañía y entusiasmo.

A mi hermana analítica y hermanos académicos por su revisión y respaldo.

A la Dra. Graciela Cordero, al Dr. Horacio Pedroza, al Dr. Mario Rueda y al Dr. Salvador por sus enseñanzas y aportaciones diligentes.

A los investigadores y personal administrativo del Instituto de Investigación y Desarrollo Educativo por el seguimiento y apoyo.

Al Consejo Nacional de Ciencia y Tecnología por el recurso otorgado.

Cimarrón mi corazón

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, el uso de los resultados, la infraestructura del SED y caracterizar las innovaciones propuestas al SED. El diseño correspondió a un estudio de caso múltiple integrado que adaptó elementos del modelo de evaluación responsiva, las muestras no probabilísticas fueron: 42 directivos, 1,431 docentes y 10,160 estudiantes, se realizaron entrevistas a dos directivos y el resto de los participantes contestaron un cuestionario en línea; los datos se procesaron con técnicas de análisis univariado, análisis cualitativo y análisis cualitativo del contenido (ACC). Los resultados dieron cuenta de la complejidad de la evaluación de la docencia; los hallazgos develaron que directivos, docentes y estudiantes priorizan el uso los puntajes del CECD por los estudiantes para la identificación de fortalezas y debilidades de la práctica docente, que el CECD cumple con los elementos para evaluar la docencia, pero proponen revisar su extensión e incluir elementos como uso de Tecnologías de la Información y la Comunicación (TIC) en el aula y puntualidad, mejorar la comunicación de los resultados a los docentes, incluir otras estrategias de evaluación y de forma global se valoró importante la utilidad del SED. 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 y la necesidad de mejorar la usabilidad de los resultados de la evaluación docente.

Palabras clave: educación superior, Competencia del docente, Evaluación del docente, cuestionarios de evaluación docente; retroalimentación.

Contenido

Planteamiento del problema	6
Preguntas y objetivos de investigación.....	8
Justificación	9
Características del contexto del estudio.....	10
Marco teórico.....	16
Fundamentos de la docencia.....	16
Evaluación educativa y evaluación docente	19
Fundamentos de la metaevaluación	25
Evaluación comprensiva.....	32
Antecedentes de investigación en metaevaluación.....	35
Método.....	48
Participantes	50
Materiales	55
Procedimiento.....	55
Fase 1. La opinión de coordinadores	58
Fase 2. La opinión de directores y subdirectores.....	59
Fase 3. La opinión de docentes.....	62
Fase 4. La opinión de los estudiantes.	65
Resultados.....	68
Unidad de análisis: Directivos.....	69
Coordinadores.....	69
Directores y subdirectores	75
Unidad de análisis: Docentes.....	81
Unidad de análisis: Estudiantes	93
Síntesis de los resultados: acuerdos y discrepancias	102
Discusión y conclusiones	107
Referencias	125
Apéndices	135

Índice de tablas

Tabla 1 <i>Preguntas y objetivos de investigación específicos</i>	9
Tabla 2 <i>Numeralia institucional de la UABC</i>	11
Tabla 3 <i>Tipos de reporte para las unidades académicas</i>	14
Tabla 4 <i>Tareas para la planificación de un proceso de metaevaluación</i>	28
Tabla 5 <i>Directrices para el diseño e implementación de programas de evaluación de la enseñanza</i>	30
Tabla 6 <i>Palabras clave (descriptores)</i>	35
Tabla 7 <i>Ecuaciones de búsqueda empleadas para la identificación de estudios</i>	36
Tabla 8 <i>Criterios de búsqueda</i>	36
Tabla 9 <i>Artículos resultantes de la búsqueda en las bases de datos</i>	37
Tabla 10 <i>Distribución de artículos según etiquetas</i>	38
Tabla 11 <i>Características generales de antecedentes de metaevaluación</i>	39
Tabla 12 <i>Dimensiones del programa DOCENTIA</i>	45
Tabla 13 <i>Características generales de los participantes del Subgrupo b) Docentes</i>	53
Tabla 14 <i>Características generales de los participantes del Subgrupo c) Estudiantes</i>	54
Tabla 15 <i>Participantes, técnicas de recolección de datos y materiales</i>	55
Tabla 16 <i>Dimensiones del cuestionario</i>	59
Tabla 17 <i>Número de datos por dimensión del cuestionario de directores y subdirectores</i> .	60
Tabla 18 <i>Contraste entre las dimensiones de la RIIED y las categorías de las agendas de códigos de directores y subdirectores</i>	61
Tabla 19 <i>Número de datos por dimensión del cuestionario de docentes</i>	63
Tabla 20 <i>Contraste entre las dimensiones de la RIIED y las categorías de las agendas de códigos de docentes</i>	63
Tabla 21 <i>Número de ítems de los cuestionarios de evaluación del SED</i>	66
Tabla 22 <i>Número de datos por dimensión del cuestionario de estudiantes</i>	67
Tabla 23 <i>Estructura de los resultados de la investigación</i>	68
Tabla 24 <i>Porcentaje de las respuestas sobre el uso de los resultados de evaluación según directores y subdirectores</i>	76
Tabla 25 <i>Puntajes de las respuestas sobre elementos del CECD según directores y subdirectores</i>	76
Tabla 26 <i>Estrategias de evaluación docente adicionales al CECD según directores y subdirectores</i>	77
Tabla 27 <i>Porcentaje de respuestas sobre el procedimiento de aplicación del CECD según directores y subdirectores</i>	78

Tabla 28 <i>Puntajes de las respuestas sobre el uso de los resultados de evaluación según Docentes</i>	82
Tabla 29 <i>Puntajes de las respuestas sobre elementos del CECD según Docentes</i>	82
Tabla 30 <i>Estrategias de evaluación docente adicionales al CECD según Docentes</i>	84
Tabla 31 <i>Puntajes de las respuestas sobre el procedimiento de aplicación del CECD según Docentes</i>	85
Tabla 32 <i>Puntajes de las respuestas sobre la apreciación del SED según Docentes</i>	86
Tabla 33 <i>Tasa promedio de respuesta de los ítems de respuesta abierta de Docentes</i>	86
Tabla 34 <i>Puntajes de las respuestas sobre el uso de los resultados de evaluación según Estudiantes</i>	93
Tabla 35 <i>Puntajes de las respuestas sobre la utilidad de los resultados según Estudiantes</i>	94
Tabla 36 <i>Puntajes de las respuestas sobre elementos del CECD según Estudiantes</i>	94
Tabla 37 <i>Puntajes de las respuestas sobre el procedimiento de aplicación del CECD según Estudiantes</i>	95
Tabla 38 <i>Puntajes de las respuestas sobre la valoración del SED según Estudiantes</i>	96
Tabla 39 <i>Tasas promedio de respuesta de los ítems de respuesta abierta de Estudiantes</i> ..	96

Índice de figuras

Figura 1 <i>Recomendaciones para la evaluación de la docencia</i>	24
Figura 2 <i>Lista de verificación de Scriven (2011)</i>	26
Figura 3 <i>Ecuación de la evaluación de programas</i>	29
Figura 4 <i>Eventos de una evaluación comprensiva</i>	333
Figura 5 <i>Estudio de caso múltiple integrado</i>	49
Figura 6 <i>Porcentaje de la muestra del Subgrupo a) Directores y subdirectores</i>	52
Figura 7 <i>Grupos de participantes</i>	54
Figura 8 <i>Procedimiento de la investigación</i>	56
Figura 9 <i>Distribución de los apartados del método</i>	57
Figura 10 <i>Propuestas de CECD para programas educativos particulares según Docentes</i>	84
Figura 11 <i>Relación entre categorías de las agendas de códigos y categorías analíticas</i>	87

Planteamiento del problema

La evaluación educativa es un elemento que favorece el diseño y consolidación de directrices y estrategias para la mejora. Scriven (1966, 2007) explica que la evaluación es un proceso, una actividad lógica que consiste en la recopilación y combinación de información para determinar el valor, mérito o importancia de un programa. Asimismo, la evaluación alude a una labor profesional realizada de forma objetiva y sistemática (Scriven, 2007), para dar evidencia de efectividad, confiabilidad, viabilidad e integridad (Stufflebeam y Coryn, 2014).

Se resalta también la función de la evaluación como sustento para tomar decisiones que pueden ser de tres tipos: mejora de un curso, decisiones sobre los estudiantes y regulación administrativa (Cronbach, 2000). En esta lógica, el punto central de la evaluación es apoyar la determinación sobre el valor de un programa educativo o apoyar la toma de decisiones (Phillips, 2018).

Scriven (2015) señaló que una especialidad en la evaluación es la metaevaluación, cuya finalidad es valorar la evaluación, de trascendencia para la disciplina evaluativa en virtud de que su desarrollo y práctica sustentan los beneficios de establecer y defender conclusiones para el entorno en el que se enmarca. La metaevaluación presta atención tanto al desarrollo como al producto del proceso evaluativo a partir del contexto y las circunstancias (Stake, 2017). Al respecto, Stufflebeam (2011) indicó que la metaevaluación puede tener una función formativa para ayudar a la mejora de la evaluación, y, una función sumativa en apoyo a la determinación del logro y valor de la evaluación. También, de forma práctica, se espera que la metaevaluación provea un procedimiento válido en dirección de resaltar los resultados especialmente importantes y dudosos (Stake, 2017).

Planteamiento del problema

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). Al respecto, algunas instituciones de educación superior han realizado procesos de metaevaluación de sus programas de evaluación de la docencia con la finalidad de generar información de calidad que posibilite mejorar la enseñanza (Cisneros-Cohernour y Patrón, 2017; Elizalde et al., 2008; Elizalde et al., 2017; Sánchez, 2019; Universidad Complutense de Madrid [UCM], 2016).

Cabe resaltar que, el concepto de metaevaluación no es nuevo, pero los sistemas de metaevaluación sí los son (Stake, 2017); por consiguiente, las investigaciones sobre metaevaluación de programas de evaluación de la docencia son escasas. 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). La UABC cuenta con un sistema que gestiona la evaluación de la docencia denominado Sistema de Evaluación Docente (SED), su propósito es administrar la aplicación y entrega de resultados de la evaluación docente. El SED recaba información por medio de cuestionarios de evaluación del desempeño docente contestados por los estudiantes. Para ello, desde 2011 se ofrecen reportes de evaluación a los docentes mediante un sistema en línea, y en 2014 fueron actualizadas las pantallas de captura de la evaluación para los estudiantes (UABC, 2017a y 2017b).

Los principales objetivos declarados del SED son la mejora y retroalimentación de la enseñanza, la promoción de docentes, y la asignación adecuada de los recursos (UABC,

Planteamiento del problema

2017a). Sin embargo, desde la creación del SED hasta ahora, no se ha realizado una evaluación al sistema, se desconoce la opinión de los directivos, docentes y estudiantes sobre el SED, sobre su funcionamiento y seguimiento. Por ello, se plantea la necesidad de llevar a cabo una metaevaluación del 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 (Scriven, 2015).

Preguntas y objetivos de investigación

En el marco anterior descrito, se plantea la siguiente pregunta de investigación ¿Cuáles son los aspectos a mejorar del SED? Así, el objetivo general correspondiente refiere realizar una metaevaluación al SED con el fin de identificar las fortalezas y los aspectos a mejorar en: el Cuestionario de Evaluación de la Competencia Docente (CECD), la administración del CECD, el sistema gestor del SED, el uso de los resultados y la infraestructura institucional. A continuación, en la Tabla 1 se presentan las preguntas específicas de investigación alineadas a los objetivos específicos.

Tabla 1*Preguntas y objetivos de investigación específicos*

Preguntas específicas	Objetivos específicos
¿Cuáles son los elementos y procesos que componen al SED?	Identificar las partes constitutivas y procesos que componen al SED.
¿Cuáles son las principales observaciones y sugerencias de mejora de los actores educativos (directivos, docentes y estudiantes) al CECD?	Caracterizar las fortalezas y debilidades de CECD.
¿Cuáles son las principales observaciones y sugerencias de mejora de los actores educativos (directivos, docentes y estudiantes) al sistema gestor del SED?	Delimitar los aspectos positivos y a mejorar del sistema gestor del SED.
¿Cuáles son las opiniones y recomendaciones de mejora al proceso que administra la aplicación del CECD?	Identificar las diversas opiniones sobre el proceso que administra la aplicación del CECD.
¿Cuáles son las opiniones de los actores educativos (directivos, docentes y estudiantes) en relación con el uso de los puntajes del CECD?	Identificar las diversas opiniones y argumentos sobre los usos de los puntajes del CECD.
¿El SED cuenta con la infraestructura institucional necesaria para su operación?	Valorar la infraestructura institucional disponible para la operación del SED.
¿Cuáles son las innovaciones que proponen los actores educativos?	Caracterizar las innovaciones propuestas al SED

Justificación

La metaevaluación implica una revisión y verificación sistemática de la evaluación, con la finalidad de describir y reconocer posibles problemas o dificultades (Stufflebeam, 2011), e instaurar o sustentar conclusiones en beneficio de los implicados en el proceso (Scriven, 2015). En este sentido, por medio de una metaevaluación al SED se busca identificar fortalezas y áreas de oportunidad que permitan sugerir y diseñar estrategias para la mejora del proceso de evaluación de la docencia en la universidad.

Planteamiento del problema

Un punto importante es que el estudio que se propone no tiene precedente en la UABC; así pues, a partir de la revisión, análisis especializado y sustento en la opinión de los actores educativos, la institución contará con una metaevaluación acorde a sus estándares de calidad. De esta forma, también se procura propiciar una cultura de la evaluación mayormente informada y apropiada por la comunidad universitaria.

Además, por medio del caso de la UABC se pretende abordar la complejidad de la metaevaluación de un sistema de evaluación de la docencia. Asimismo, se busca contribuir a la discusión sobre la evaluación en educación superior, y sobre la metaevaluación de sistemas de evaluación docente; así como brindar evidencia empírica en el escenario mexicano.

Características del contexto del estudio

La UABC se caracteriza por ser descentralizada en tres campus en los municipios de Mexicali, Tijuana y Ensenada del estado de Baja California, cuenta con siete áreas de conocimiento: Ingeniería y tecnología, Ciencias Agropecuarias, Ciencias de la Salud, Ciencias Naturales y Exactas, Educación y Humanidades, Ciencias Sociales y Ciencias Administrativas. La oferta 146 programas educativos de licenciatura y 69 de posgrado; las unidades académicas son 31 facultades, 8 institutos de investigación y una escuela. Asimismo, con base en la numeralia institucional (UABC, 2022), la matrícula de estudiantes de licenciatura es de 66,504, y los académicos (docentes de tiempo completo, de medio tiempo, de asignatura, técnicos académicos e investigadores) son 6,171 (UABC, 2022) (véase la Tabla 2).

Tabla 2*Numeralia institucional de la UABC*

Programas educativos			
Licenciatura		Posgrado	
146		69	
Unidad académica			
Escuela		1	
Facultades		31	
Institutos de investigación		8	
Estudiantes		Académicos	
Licenciatura	66,504	Docentes de tiempo completo	1,381
		Docentes de medio tiempo	12
Posgrado	2,117	Docentes de asignatura	4,359
		Técnicos académicos	419
Total	68,621	Total	6,171

Nota. De “Numeralia institucional” por Universidad Autónoma de Baja California, 2022. CC BY-NC-ND.

En la UABC los documentos institucionales reconocen a la evaluación docente como una tarea permanente y formativa (UABC, 2021c), realizada con el fin de brindar a los docentes retroalimentación de su desempeño (UABC, 2021d, sección “Evaluación Docente”). Al mismo tiempo, los puntajes tienen un peso en la puntuación global del programa de compensación salarial, es decir, los resultados de la evaluación son determinantes para la participación de los docentes en los Programas de Estímulos al Desempeño del Personal Docente (UABC, 2021d, sección “Evaluación Docente”). 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, 2021d, sección “Evaluación Docente”). De

Planteamiento del problema

esta forma, los resultados de la evaluación docente son utilizados tanto con fines formativos, de control administrativo y compensación salarial.

La UABC (2017a) destaca a la evaluación de la docencia como una labor sistemática y continua orientada a favorecer la práctica docente, la promoción del profesorado y la distribución conveniente de los recursos. Asimismo, uno de los cinco principios orientadores sustentados en la misión, visión y filosofía de la universidad, señala 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).

Sistema de Evaluación Docente (SED)

De acuerdo con los documentos oficiales el SED es dirigido por la Coordinación General de Formación Profesional conformada por el Coordinador(a) general y cuatro departamentos: Diseño curricular, Evaluación del aprendizaje, Formación y evaluación docente y Aseguramiento de la calidad educativa (UABC, 2021b). El SED tiene la finalidad de ofrecer información sobre la práctica docente con base en la opinión de los estudiantes (UABC, 2010), esta tarea se realiza por medio de la aplicación del CECD de manera automatizada y en línea (2017a), y está programada en el calendario escolar durante el último mes de cada periodo semestral.

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, 2021c). Dicha evaluación data del año 2000 (UABC, 2000, 2015) y en el 2005 se asoció con el proceso de reinscripción semestral del estudiante (UABC, 2005).

Planteamiento del problema

El SED atiende dos modalidades de programas educativos: licenciatura presencial y en línea. En el caso de la modalidad de licenciatura presencial, el SED dispone de dos cuestionarios: uno para las asignaturas teóricas de 25 reactivos, y otro para el componente práctico con 20 reactivos —este último solo para clases prácticas—. El estudiante evalúa a cada docente según las unidades de aprendizaje que haya cursado en el semestre que se encuentre vigente. La evaluación es realizada por medio del sistema gestor del SED, el estudiante ingresa con su correo institucional y su contraseña, ahí le son visibles las evaluaciones que le corresponden llevar a cabo. El estudiante puede ingresar al sistema gestor las veces necesarias para responder los cuestionarios de evaluación docente; la evaluación tiene un periodo aproximado de cuatro semanas (UABC, 2021c).

El sistema gestor del SED tiene tres módulos de acceso: para estudiantes, docentes y unidades académicas; el ingreso al sistema es a través del correo electrónico institucional. A partir de la información que recupera el SED se generan distintos reportes, para los docentes se producen dos tipos de reportes: Unidad de Aprendizaje del Docente y Promedios Globales del Docente, el primero se refiere a la evaluación por cada unidad de aprendizaje que el docente imparte, y el segundo es un reporte concentrado de los puntajes de todas las unidades académicas del docente en la universidad; a las unidades académicas se les proporcionan nueve tipos de reporte: No Realizadas, Todas las Evaluaciones, No Aplica, Alumnos Evaluados, Unidad de Aprendizaje del Docente, Unidad Académica, Promedios por Unidad Académica, Promedios Globales del Docente y Diagnóstico del Desempeño de los Profesores. Los reportes están dirigidos al docente, al director de la unidad académica y a la administración central de la UABC (UABC, 2010). En la Tabla 3 se describen dichos reportes.

Tabla 3*Tipos de reporte para las unidades académicas*

Tipo de reporte	Descripción
No realizadas	Relación de estudiantes que no han realizado la evaluación docente.
Todas las evaluaciones	Relación de estudiantes por programa educativo que realizaron y no realizaron la evaluación docente.
No aplica	Relación de docentes por unidad académica cuyos instrumentos de evaluación fueron eliminados.
Alumnos evaluados	Relación del avance de la evaluación a la fecha, respecto a la población vigente en la unidad académica.
Unidad de aprendizaje del docente	Relación de cada una de las unidades de aprendizaje de los docentes adscritos a la unidad académica.
Unidad académica	Relación de docentes de la unidad académica en orden descendente de acuerdo al puntaje en cada unidad de aprendizaje.
Promedios por unidad académica	Relación de docentes de la unidad académica en orden descendente, de acuerdo al promedio general de todas las unidades de aprendizaje impartidas en la institución.
Promedios globales del docente	Relación concentrada, especifica el puntaje en todas las unidades de aprendizaje que los docentes de la unidad académica imparten en la institución.
Diagnóstico del desempeño de los profesores	Relación de docentes de la unidad académica en cada una de las dimensiones evaluadas en el instrumento.

Nota. De “Sistema de Evaluación Docente (SED),” por Universidad Autónoma de Baja California, 2010. CC BY-NC-ND.

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 y para la mejora.

Planteamiento del problema

Ahora bien, aunque se han realizado esfuerzos constantes para mejorar el proceso de evaluación docente en la UABC —p. ej. en los CECD (UABC, 2021c) y en el sistema gestor del SED (UABC, 2000, 2015)—, se reconoce que la presente metaevaluación tendrá una incidencia favorable en la evaluación docente para las diferentes audiencias interesadas de la universidad, además, este trabajo de investigación podrá constituirse como un referente que aliente otros esfuerzos institucionales en la evaluación de la docencia en beneficio del aprendizaje de los estudiantes.

Marco teórico

En este apartado se integran los planteamientos teóricos sobre la perspectiva de evaluación y metaevaluación que sustentan este proyecto de investigación. Se compone por cuatro secciones: fundamentos de la docencia, evaluación educativa, fundamentos de la metaevaluación y evaluación comprensiva.

Fundamentos de la docencia

El docente se define como el profesional de la enseñanza (Marcelo y Vaillant, 2009), el cual perfecciona su labor por medio de la modificación de su desempeño en el aula, y de las estrategias y técnicas de enseñanza (Hativa, 2000). Por tanto, la enseñanza es considerada el núcleo que orienta al entorno educativo, renovando y revitalizando a la institución (Boyer, 1990). Además, la enseñanza se conforma por los objetivos de aprendizaje dirigidos a comprender y resolver problemas, propiciar el pensamiento crítico y creativo, y dominio de datos, normas o principios (Shulman, 1987, 2005).

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).

Asimismo, la docencia forma parte central del proceso educativo, constituida por diversos procesos y rasgos (Wood, 2017a), reconocida 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 enseñanza requiere que el docente posea una comprensión en dos sentidos: sobre el conocimiento disciplinar, y de los tipos de actividades de enseñanza (Bransford et al. 2000). Así pues, el docente experto cuenta con un entendimiento sólido de su conocimiento de la materia, reconoce las barreras conceptuales de los estudiantes al aprender sobre los contenidos, e identifica estrategias efectivas para la enseñanza (Bransford et al., 2000). En este sentido, se asume que el docente domina el contenido, y tiene la posibilidad de modificar sus habilidades y actitudes en escenarios pedagógicos (Shulman, 2005).

Según Feldman (1976b), con base en la opinión de estudiantes, las características de una práctica docente ideal asociada a una buena enseñanza son: estimulación del interés por el curso, entusiasmo, dominio del conocimiento disciplinar, preparación y organización del curso, enseñanza clara y comprensible, habilidades de elocución, sensibilidad y preocupación por el nivel y el progreso de los estudiantes, claridad de los objetivos y requisitos del curso, utilidad y relevancia de los materiales, dificultad del curso —carga de trabajo—, equidad e imparcialidad de la evaluación, gestión del aula, retroalimentación a los estudiantes, fomento de la discusión, desafío intelectual, respeto, disponibilidad y amabilidad.

Cabe enfatizar la complejidad del trabajo docente en la educación universitaria, lo cual destaca la naturaleza multidimensional de sus prácticas (Van Lankveld et al., 2017; Wood, 2017a). Por ello, la función docente se opone a definiciones reduccionistas de la enseñanza y rechaza concepciones individualistas de esta tarea (Wood, 2017a). De esta forma, la enseñanza no debe ser un proceso definido de manera restringida (Wood, 2017b). Así pues, la docencia no se limita a presentar información, es adaptativa e interactiva y necesariamente compromete una evaluación (Penuel y Shepard, 2016).

La búsqueda de la excelencia de la labor docente universitaria es una inquietud mundial, por tanto, las universidades atienden cada vez más a la mejora de la docencia (Hativa et al., 2001). Las instituciones de educación superior están comprometidas a impulsar una buena enseñanza, primero, para dar evidencia de una educación de calidad y en atención a los interesados; y segundo, en respuesta a la creciente búsqueda de una enseñanza valiosa y destacada (Nasser-Abu, 2017b). Además, con el paso del tiempo, ha

crecido el interés en la enseñanza como un área de especial relevancia del trabajo de la academia (Wood, 2017b).

De esta forma, el fortalecimiento de la enseñanza se promueve y apoya en un ambiente de evaluación de la docencia que genere oportunidades, pues con frecuencia los docentes no cuentan con evidencias de éxito en ciertas ideas o proyectos que realizan, lo cual limita sus planes a emprender (Bransford et al., 2000).

En el caso de la UABC (2021c), las principales funciones que se reconocen del docente son: generar ambientes para la discusión, crear entornos para realizar investigación educativa, fomentar la comunicación, reflexión, trabajo colaborativo, entre otros, y que su desempeño contribuya a la sociedad del conocimiento. Por consiguiente, en esta investigación se reconocen los principios aquí expuestos como ejes vertebrales respecto a la concepción, características, labor y retos que compromete el trabajo docente.

Evaluación educativa y evaluación docente

La evaluación educativa es un elemento determinante para la discusión, reflexión e innovación de la educación. Cronbach (2000) definió a la evaluación educativa como la recopilación sistemática y uso de información con el propósito de tomar decisiones sobre un programa educativo, estas decisiones son de diferentes tipos y conlleva la búsqueda de una diversidad de información. Además, Cronbach refirió que ningún conjunto de principios aborda completamente todas las situaciones, de ahí que, la evaluación se considere una labor diversificada.

La evaluación es un cimiento aceptable y justificado para identificar buenos programas (Stufflebeam y Coryn, 2014). En este sentido, se resalta que el juicio de valor asumido por la evaluación descansa en un análisis racional de los resultados (Scriven,

1966). Por consiguiente, la evaluación brinda a los encargados de la toma de decisiones información necesaria en la selección sustentada para la mejora, y a los docentes retroalimentación de utilidad para favorecer su desempeño (Centra, 1993).

Las evaluaciones en educación son en sí mismas intervenciones centradas en una necesidad o problema y suponen diversas formas o tipos de evaluación (Phillips, 2018). Scriven (1966) indicó que la evaluación cumple con dos funciones: valoración del proceso y del resultado final de este proceso. Por su parte Cronbach (2000) señaló que, para la mejora de un programa de evaluación se requiere de nuevos hábitos de pensamiento, variedad de técnicas y descripción amplia de los resultados. Además, para llevar a cabo una evaluación es importante concentrar los esfuerzos en el problema o preocupación de los interesados, y en ese sentido, reunir información relevante para las decisiones que se planeen tomar (Phillips, 2018).

Shulha y Cousins (1997) enfatizaron la complejidad de realizar un proceso evaluativo, ya que el uso de la evaluación, las características del programa educativo y el contexto político influyen de forma directa. Los autores también resaltan el creciente interés de los evaluadores en determinar el tipo de uso de la evaluación que se busca por los involucrados del programa, de esta forma se manifiesta la evaluación educativa como un diálogo continuo de información, lo cual propicia una responsabilidad compartida de generar, transmitir y difundir el potencial de la evaluación.

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.

Aunado a lo anterior, la evaluación de la docencia ha cobrado mayor presencia y se ha convertido en una labor recurrente desempeñada de diferentes formas y con diversos fines (Nasser-Abu, 2017a). Así pues, para mejorar y reconocer la buena enseñanza se requieren mejores maneras de evaluar el desempeño docente, también, de descripciones de conocimiento y trabajo evaluativo que posibiliten fomentar vínculos con la enseñanza (Centra, 1993).

Adicionalmente, la evaluación de la docencia y el desarrollo del profesorado deben considerarse de manera conjunta, en dirección de una evaluación efectiva que incorpore una revisión de las fortalezas y debilidades de la práctica docente asociada con las características de los estudiantes, el contenido, el diseño y la instrucción (Theall, 2017).

Darling-Hammond (2012) mencionó que un sistema de evaluación docente requiere condiciones favorables para el aprendizaje y criterios institucionales que apunten a la mejora. Además, la autora resalta cinco elementos que componen un sistema de evaluación docente: 1) participación activa del cuerpo docente, 2) elaboración de lineamientos claros y accesibles para todos los docentes, 3) tomar en cuenta las necesidades que los estudiantes

requieren para realizar una evaluación adecuada de la enseñanza, 4) impulsar un ambiente de aprendizaje profesional de calidad y continuo, y 5) reunir diferentes evidencias de la eficacia pedagógica.

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 esta línea, la estrategia de evaluación de la docencia principalmente empleada por las universidades son los cuestionarios de opinión de los estudiantes con el fin de evaluar la calidad de la docencia (Smith y Kubacka, 2017). Asimismo, se cuenta con evidencias para asumir que estos cuestionarios tienden a ser estadísticamente confiables, válidos y parcialmente libres de sesgos (Benton y Cashin, 2012), además de, útiles para el docente como retroalimentación de su enseñanza, para los estudiantes en la elección de cursos y para los directivos en la toma de decisiones administrativas (Marsh, 2007). De manera generalizada, el uso apropiado de los puntajes es uno de los desafíos que enfrentan este tipo de evaluaciones (Linse, 2017). De aquí que, la recopilación, interpretación y

empleo correcto de los resultados de la evaluación docente en educación superior ha sido un tema de investigación estudiado durante varias décadas (Theall, 2017).

Una cuestión relevante a considerar, mencionada por Feldman (1978), son las características propias de cada curso que un docente imparte, debido a la variabilidad existente y el papel que tales características desempeñan en los puntajes que el estudiante asigna al evaluar el curso. También, el autor reconoció la relación de puntajes altos cuando: a) el estudiante evalúa de forma anónima, b) los resultados se relacionan con la promoción y cuestiones salariales (1979), c) el número de estudiantes en el curso es reducido, b) los cursos son de división superior —especializados y priorizan un pensamiento analítico, resolución de problemas y aplicaciones teóricas—, d) los cursos son elegibles, e) los cursos son de humanidades, bellas artes e idiomas (1978, Franklin y Theall, 1995), incluso los cursos relacionados con áreas de la salud (Franklin y Theall, 1995), y f) el tiempo que el estudiante dedica para prepararse para el curso —probablemente relacionado con un buen diseño e impartición del curso— (Franklin y Theall, 1995). Por tanto, reconocer lo anterior permite cierto grado de comparabilidad de los puntajes, con el fin de favorecer su utilidad para estudiantes, docentes y directivos (1978). Además, el interés que el docente manifiesta por el curso tiene un vínculo positivo con los resultados de la evaluación de su práctica, junto con las percepciones de los estudiantes sobre su rendimiento y grado de aprendizaje (1976a).

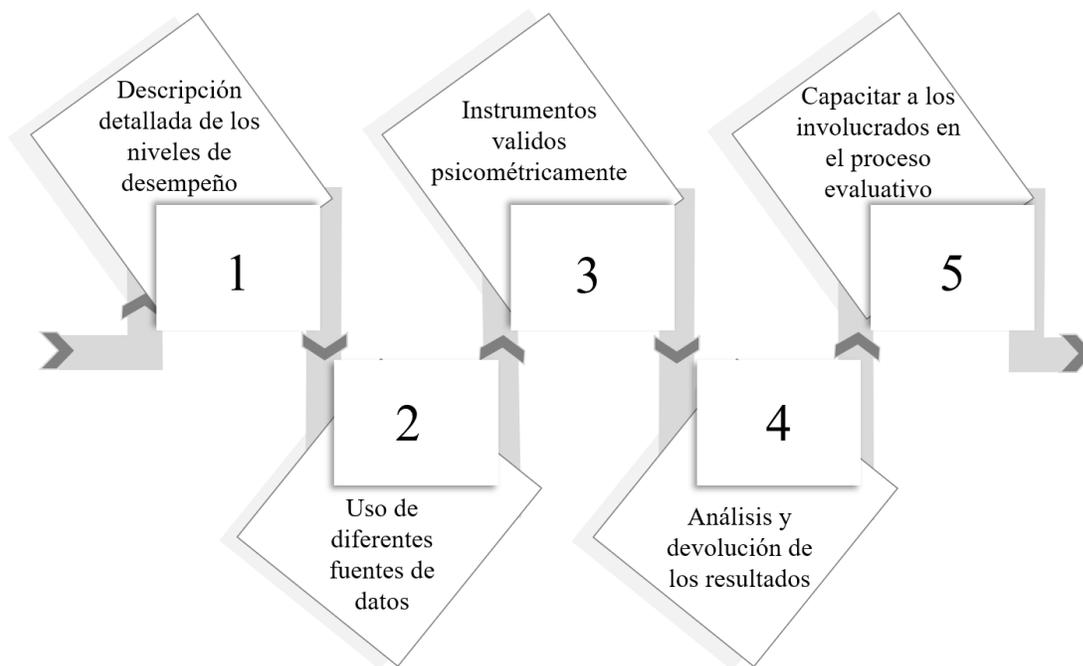
Adicionalmente, Theall (2017) destacó cuatro situaciones principales relacionadas a la tensión respecto al uso de los puntajes de la evaluación docente por los estudiantes: 1) los puntajes como único recurso —esto constituye un error metodológico y administrativo—; 2) la recolección, administración, análisis y comunicación de los puntajes realizada de

forma deficiente o equivocada; 3) malinterpretación y empleo incorrecto de los resultados de los puntajes; y 4) los puntajes considerados evaluaciones de la enseñanza. Lo anterior, según el autor, refiere a un error semántico y un problema debido a que los estudiantes no son los evaluadores, sino proveedores de un tipo de dato, los evaluadores son los docentes y administrativos que emiten juicios.

Dicho lo anterior, además de reconocer el amplio debate sobre los cuestionarios de evaluación docente para estudiantes, en relación a su utilidad y confiabilidad y validez de los puntajes, Theall subraya cinco recomendaciones para favorecer la evaluación de la docencia en la universidad (véase la Figura 1).

Figura 1

Recomendaciones para la evaluación de la docencia



Nota. De “MVP and faculty evaluation”, por M. Teall, 2017, *New Directions for Teaching and Learning*, (152), 91-98. CC BY-NC-ND.

En primer lugar, Theall recomienda describir de forma detallada los criterios de los niveles de desempeño, los objetivos y la extensión del trabajo previsto; segundo, emplear diversas fuentes de información provenientes de todas las instancias interesadas que se relacionen con las áreas de desempeño esperadas; tercero, aplicar a los estudiantes instrumentos con alto valor psicométrico; cuarto, analizar y comunicar los resultados de forma clara y comprensible a docentes y directivos para la toma de decisiones en evaluación; y la quinta recomendación es dotar de instrucción a los involucrados para la recopilación, interpretación y uso conveniente y equitativo de los resultados.

Al respecto, Jornet et al. (2020) reconocen 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. Así, se reitera que el sentido de la evaluación docente es para orientar la mejora (Rueda, 2021).

De forma específica, el presente trabajo de metaevaluación parte de los supuestos expresados en el reconocimiento de la evaluación docente universitaria como una tarea compleja influida por diversos elementos contextuales. Por tanto, esta investigación abonará a la delimitación de las fortalezas y áreas de oportunidad del proceso evaluativo docente en la UABC, de la voz de los diferentes actores educativos.

Fundamentos de la metaevaluación

Una evaluación adecuada y conveniente demanda que sus resultados sean valorados de forma sistemática, bajo esta consideración, Scriven (1991) resaltó a la metaevaluación como la evaluación de la evaluación. La metaevaluación se lleva a cabo a través de la

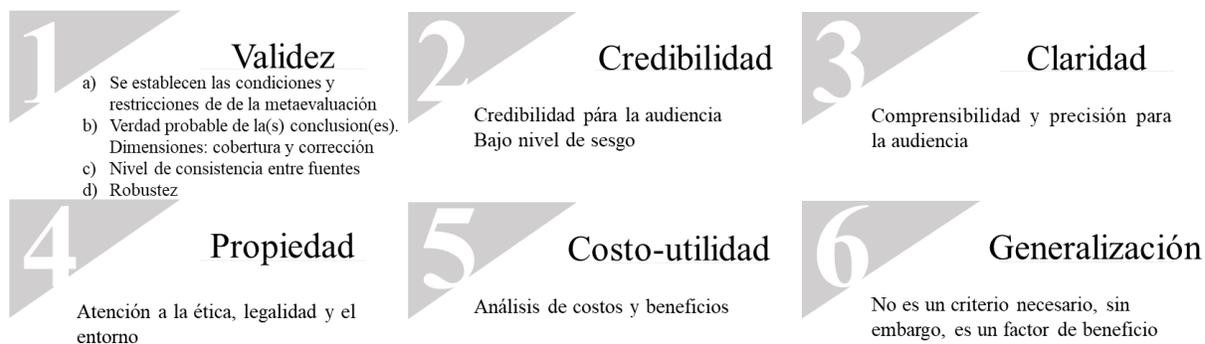
revisión y verificación de la evaluación, y de esta forma es posible detectar problemas tales como uso inadecuado, dificultades administrativas o errores técnicos (Stufflebeam, 2011).

Por tanto, la metaevaluación se define como un procedimiento con la finalidad de describir y juzgar una actividad de evaluación, con base en los supuestos que fundamentan una buena evaluación (Stufflebeam, 2011). A este respecto, la metaevaluación es una forma de evaluación de orden superior y compromete evaluaciones secundarias, terciarias, según las requeridas para identificar fortalezas y áreas de oportunidad.

Scriven (2011) indicó que para la realización de una metaevaluación se deben determinar los puntos de análisis que se llevarán a cabo. En este sentido, el autor propone una lista de verificación de seis elementos a tomar en cuenta para el desarrollo de una metaevaluación. Asimismo, se advierte que la utilidad de la lista depende del nivel de la consulta y de los asuntos específicos que se pretenden cubrir (véase la Figura 2).

Figura 2

Lista de verificación de Scriven (2011)



Nota. De “Evaluating evaluations: A meta-evaluation checklist,” por M. Scriven, 2011, *Claremont Graduate University* (9na ed., pp. 2-6), Claremont Graduate University. CC BY-NC-ND.

El primer elemento de la lista es la Validez, su tema fundamental es la verdad; este elemento está dividido en cuatro apartados, de los cuales los dos primeros son los más

dominantes: a) especificación de las condiciones y restricciones contextuales; b) afirmaciones factibles de la(s) conclusión(es); c) consistencia entre fuentes; y d) grado de dependencia de las conclusiones con el significado de las variables que representan. El primer apartado, sobre las especificaciones y restricciones contextuales, conduce a la precisión del enfoque de la evaluación requerida y del nivel de análisis; y el siguiente apartado se refiere a la verdad probable de la(s) conclusión(es) relacionada(s) con la cobertura y la adecuación de la evidencia.

El segundo elemento de la lista es la Credibilidad para los usuarios o interesados, relacionada con un bajo nivel de sesgo sobre los asuntos no cubiertos por consideraciones de validez. En tercer lugar, se encuentra la Claridad representada por la precisión y la comprensibilidad para los usuarios, de esta forma se reduce el esfuerzo y el número de errores de la interpretación que propicie mejorar la aceptación y puesta en marcha de la metaevaluación. El cuarto elemento es la Propiedad en función de la ética, legalidad y pertinencia cultural; aquí se incluyen las cuestiones contractuales, la privacidad y el consentimiento informado.

La quinta indicación es el Costo-utilidad concerniente al análisis de costos y beneficios con base al contexto, ganancias y pérdidas en general; para ello es necesario incluir un presupuesto detallado sobre la viabilidad y la efectividad de la metaevaluación. Como parte de la utilidad se incorporan beneficios de la eficiencia del trabajo, y ganancias específicas: conocimiento, la toma de decisiones, calidad del programa, mantenimiento de los recursos, entre otros.

Por último, el sexto elemento señala que, la Generalización es un componente benéfico más que un requisito definitorio, conlleva la utilidad del diseño de evaluación, los

procedimientos, resultados o conclusiones. En este sentido, la robustez cobra importancia en la medida que los resultados de la evaluación son libres de cambios o variaciones del programa relativos al contexto, a los valores de los datos o procesos inferenciales.

Por su parte, Stufflebeam (2001) presentó 11 tareas principales para tomar en cuenta en la planificación de un proceso de metaevaluación, son tareas generales que dependerán de la finalidad y contexto educativo, por ello no todas las tareas son aplicables o pueden requerirse tareas adicionales. En la Tabla 4 se muestran las tareas.

Tabla 4

Tareas para la planificación de un proceso de metaevaluación

No.	Tarea
1	Establecer acuerdos para la interacción con los involucrados
2	Contar con uno o más evaluadores calificados como parte del equipo de metaevaluación
3	Definición de las preguntas de metaevaluación
4	Acordar los estándares o criterios de valoración del sistema de evaluación
5	Desarrollar un memorando o contrato para regir la metaevaluación
6	Reunir y revisar la información pertinente disponible
7	Buscar información nueva (según sea necesario)
8	Analizar los datos cualitativos y cuantitativos
9	Valorar la consistencia de la evaluación a los estándares o criterios apropiados
10	Difundir los resultados de la metaevaluación por medio de informes, presentaciones, otros
11	Según sea necesario y factible, apoyar a los usuarios a interpretar y aplicar los resultados

Nota. De “The Metaevaluation Imperative,” por D. L. Stufflebeam, 2001, *American Journal of Evaluation*, 22(2), p. 191. CC BY-NC-ND.

La metaevaluación representa el conjunto de conocimiento evaluativo que incorpora los estándares para el desarrollo y práctica de un sistema de evaluación (Scriven, 2015).

Stufflebeam (2001) resaltó la necesidad de una metaevaluación para propiciar un mecanismo de garantía de la calidad que fortalezca los sistemas de evaluación, así como incitar a producir servicios defendibles, y estudiar la solidez y equidad de las evaluaciones. En este marco, la metaevaluación incorpora la justificación y crítica de la evaluación, por tanto, se reconoce a la metaevaluación como la conciencia de la evaluación, y esta a su vez de la sociedad (Scriven, 2009).

Sobre la evaluación de programas, Scriven (2011) destacó seis principios: 1) Apoyo a la toma de decisiones, responsabilidad y transparencia; 2) Mejora del evaluado; 3) Para el conocimiento; 4) Ejercer lo que se dispone; 5) Exigencia profesional; y 6) Exigencia ética. En este sentido, se resalta la relación de los valores mantenidos en el programa y las expectativas y criterios que se tienen de éste, así que, la función de la evaluación no es alcanzar una calificación resumida descriptiva, sino brindar información exhaustiva sobre las observaciones del programa, en alusión a la satisfacción e insatisfacción de los participantes (Stake, 2011) (véase la Figura 3).

Figura 3

Ecuación de la evaluación de programas



Nota. De “Program evaluation particularly responsive evaluation,” por R. E. Stake, 2011, *Journal of MultiDisciplinary Evaluation*, 7(15), .183-184. CC BY-NC-ND.

Desde otra perspectiva, la RIIED (2008) propuso directrices para orientar la reflexión en el diseño e implementación de programas de evaluación de la enseñanza. Estas directrices se

dividen en 5 dimensiones: Política de la Evaluación, Teórica, Metodológica-Procedimental, De Uso y De Evaluación de la Evaluación. En continuación a la propuesta de la RIIED, Rueda (2010) ahondó en el análisis de las dimensiones, y en la explicitación de una quinta dimensión Ética, relacionada con las buenas prácticas del proceso evaluativo, la postura teórica-ideológica, actitudes, valores, entre otras. A continuación, se enuncian y describen las dimensiones en la Tabla 5.

Tabla 5

Directrices para el diseño e implementación de programas de evaluación de la enseñanza

Dimensiones	Descripción
Política de la evaluación	La evaluación como práctica social alineada a la filosofía y el sistema de evaluación institucional. Claridad en las funciones de control y uso a partir de la evaluación. Participación de todas las áreas de la institución.
Teórica	Acuerdo de la comunidad escolar sobre el paradigma de evaluación de la institución.
Metodológica-procedimental	Vinculación del contexto educativo con la gestión académica y administrativa. Plan y seguimiento del sistema de evaluación. Grupo de expertos que respalden el proceso permanente del sistema de evaluación.
De uso	Demarcación detallada de los usos de los resultados de la evaluación, exposición de sus resultados. Propuesta de estrategias.
De evaluación de la evaluación	Seguimiento de las etapas del sistema de evaluación. Participación de grupos colegiados internos y externos en el sistema de evaluación. Desarrollo de estudios.
Ética ^a	Instauración y fortalecimiento de la confianza entre los diferentes actores. Propósitos y criterios de evaluación consensuados y conocidos. Involucramiento de todos los actores de la evaluación. Fortalecimiento de una actitud investigadora sobre la evaluación de la docencia.

Nota. De “Reflexiones sobre el diseño y puesta en marcha de programas de evaluación de la docencia,” por Red Iberoamericana de Investigadores de la Evaluación de la Docencia, 2008, 1(3), pp.164-167. CC BY-NC-ND. ^a La dimensión Ética fue una contribución posterior de “Reflexiones generales a considerar en el diseño y puesta en operación de

programas de evaluación de la docencia,” por M. Rueda, 2010, *Revista Iberoamericana de Evaluación Educativa*, 3(1), pp. 347-348. CC BY-NC-ND.

Estas directrices tienen la finalidad de contribuir a las buenas prácticas de la evaluación de la enseñanza concebida como un instrumento esencial para mejorar la educación, así como aportar elementos para la discusión de este tema y emprender acciones sobre las condiciones favorables para su desarrollo (RIIED, 2008). Por lo tanto, la propuesta de la RIIED es un referente que permite analizar y atender un sistema de evaluación institucional para mejorar su funcionamiento.

Se destaca la dimensión de evaluación de la evaluación en dirección de consolidar una agenda evaluativa que propicie, antes que otros fines, la mejora de la docencia (Rueda, 2010). Asimismo, resaltar la formación especializada y las condiciones institucionales requeridas para llevar a cabo la evaluación entendida como una actividad profesional (Rueda, 2010).

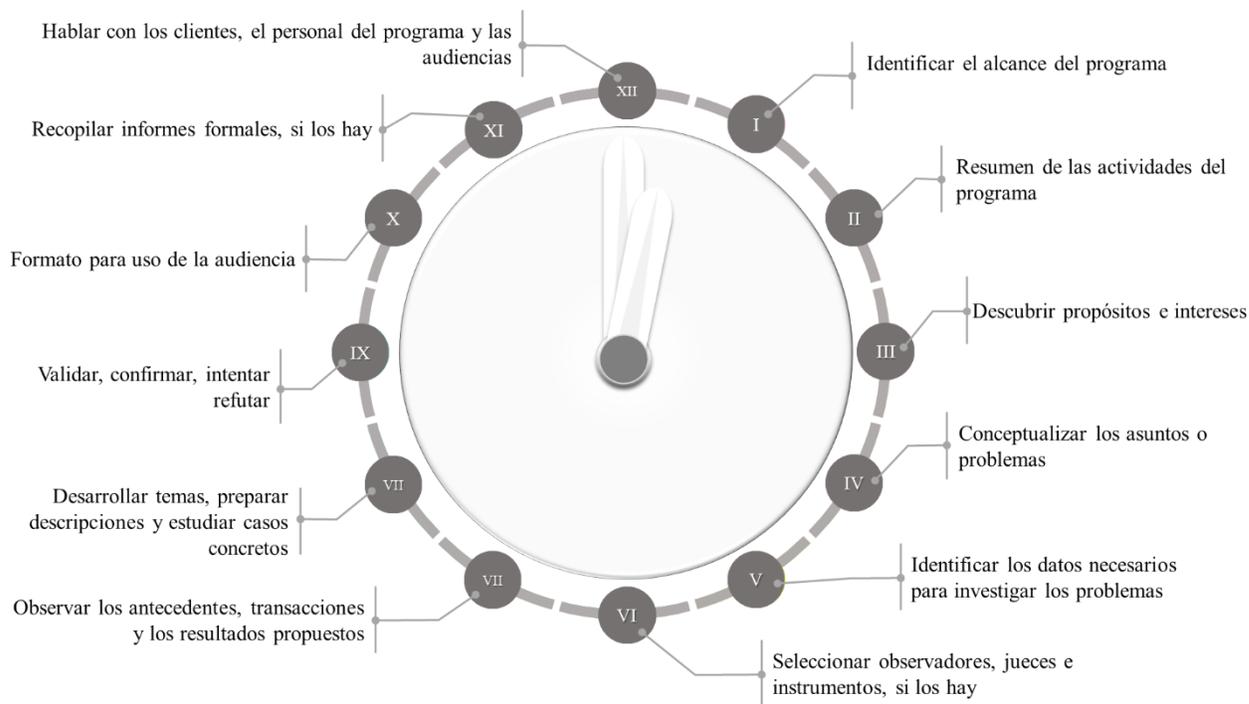
En suma, para la realización de esta metaevaluación del SED de la UABC, se parte de las propuestas de Scriven (2011) y Stufflebeam (2001) para hacer una metaevaluación, y de los planteamientos de Scriven (2011), RIIED (2008) y Rueda (2010) para el diseño e implementación de programas de evaluación docente. Específicamente, la validez respecto a las consideraciones contextuales, conclusiones viables y coherencia de los resultados, el bajo nivel de sesgo y la propiedad (Scriven, 2011); asimismo, la definición de las preguntas de metaevaluación, acuerdo de los criterios de valoración del SED, recabar y revisar la información disponible, analizar los datos y difundir los resultados (Stufflebeam, 2001).

Evaluación comprensiva

Dentro de las aproximaciones a la evaluación se encuentran distintos enfoques, en esta investigación se prioriza la evaluación comprensiva por centrar la atención en la particularidades y actividades del programa, así como a la diversidad social de sus participantes (Stake, 2003).

La evaluación comprensiva resalta: situaciones educativas más que objetivos o hipótesis; la observación directa e indirecta; también, la diversidad de normas expresadas por varios grupos de participantes, y se procura atender continuamente las necesidades de información de la audiencia (Stake, 1975, 1991). Por consiguiente, la evaluación comprensiva se enfoca en impulsar la resolución de un problema en un programa específico, con base en la experiencia de los involucrados y su contexto (Stake, 1976).

La estructura de un plan de evaluación comprensiva descansa en la observación y retroalimentación constantes, a tal efecto se reconocen doce eventos recurrentes que Stake (1975, 1991, 2011) representó con la forma de un reloj, sin embargo, las manecillas se mueven en cualquier dirección, esto es, los eventos no son subsecuentes ni dependen el uno del otro, entonces, es posible realizarlos en cualquier momento o incluso varios pueden ocurrir de forma simultánea (véase la Figura 4).

Figura 4*Eventos de una evaluación comprensiva*

Nota. De “Program evaluation particularly responsive evaluation,” por R. E. Stake, 2011, *Journal of MultiDisciplinary Evaluation*, 7(15), p. 189. CC BY-NC-ND.

Stake (1972, 2006, 2011) destacó que la evaluación comprensiva presta atención a la audiencia, a la representación de las perspectivas vigentes y se favorece la comunicación. Lo anterior, se sustenta básicamente en el análisis de la experiencia personal de los participantes involucrados en el programa (Stake, 2006).

La evaluación comprensiva representa una indagación y fundamentación de la calidad de un programa (Stake, 2003), y en conjunto se enfatiza una pluralidad de opiniones (Stake, 2014). Adicionalmente, se distingue la labor de la evaluación comprensiva en dos momentos: en el proceso y en el resultado final, en el primero apoya el seguimiento de un programa, y en el segundo se orienta a comprender sus funciones, fortalezas y deficiencias (Stake, 1972, 2003, 2006, 2011). Además, la evaluación comprensiva pretende brindar una

descripción integral de las actividades del programa e interpretación de los hallazgos (Stake, 2014).

En este marco, la metaevaluación de un sistema de evaluación docente se considera como una actividad realizada de forma continua, que integra lo que se sabe sobre la evaluación y la medición de la competencia docente en función de la calidad; además, las opiniones de los diferentes usuarios son tomadas en cuenta (Stake, 2017). Por tanto, una metaevaluación aporta un control de calidad para las actividades de la evaluación (Stake, 1975). En este sentido, una buena metaevaluación revela una buena gestión de la evaluación que conlleva procesos minuciosos y empeño (Stake, 2014).

Así pues, la metaevaluación al SED retomó elementos de la evaluación comprensiva en función de la atención particular de los diferentes actores educativos involucrados en el proceso de evaluación docente para la UABC.

Antecedentes de investigación en metaevaluación

Este apartado integra los referentes empíricos sobre estudios de metaevaluación, en específico de metaevaluación de programas de evaluación de la docencia. Primero se especifica la realización de la búsqueda, después se abordan los estudios de metaevaluación, seguidos de estudios de metaevaluación de programas de evaluación de la docencia en el contexto internacional y por último se describen estudios de metaevaluación de programas de evaluación de la docencia en el contexto nacional.

Descripción de la búsqueda de estudios

Para la identificación de estudios, primero, se identificaron las palabras clave con base en el Índice de Revistas de Educación Superior e Investigación Educativa (IRESIE, 2015), el tesauruso de la Unesco y Eric (véase la Tabla 6). Cabe resaltar que la delimitación de las palabras se determinó a partir del objetivo general del presente estudio de investigación.

Tabla 6

Palabras clave (descriptores)

IRESIE		Tesauruso UNESCO		Tesauruso Eric	
Español	Ingles	Español	Ingles	Español	Ingles
Metaevaluación					<i>Meta-evaluation</i>
Educación superior		Enseñanza superior	<i>Higher education</i>		<i>Higher Education</i>
Evaluación del profesorado		Evaluación del docente	<i>Teacher evaluation</i>		<i>Teacher Evaluation</i>

Enseguida, se conformaron 3 ecuaciones de búsqueda con los operadores booleanos *OR* y *AND* (véase la Tabla 7).

Tabla 7

Ecuaciones de búsqueda empleadas para la identificación de estudios

Ecuaciones de búsqueda
Metaevaluación OR <i>Meta-evaluation</i>
Metaevaluación OR <i>Meta-evaluation</i> AND “Educación superior” OR “Enseñanza superior” OR “ <i>Higher education</i> ”
Metaevaluación OR <i>Meta-evaluation</i> AND “Educación superior” OR “Enseñanza superior” OR “ <i>Higher education</i> ” AND “Evaluación del profesorado” OR “Evaluación del docente” OR “ <i>Teacher evaluation</i> ”

La revisión de la producción científica se realizó en las bases de datos *Scopus*, *EBSCOhost*, y *Web of Science*. El acceso a dichas bases de datos fue posible por medio de los recursos bibliográficos digitales de información científica y tecnológica con los que cuenta la UABC. Asimismo, se definieron criterios de inclusión y exclusión para condicionar los resultados. Estos criterios fueron sometidos en cada base de datos. La Tabla 8 detalla los criterios de la búsqueda.

Tabla 8

Criterios de búsqueda

Criterio	Inclusión	Exclusión
Accesibilidad:	Acceso abierto.	No acceso abierto.
Periodo de publicación:	Enero de 2017 a septiembre de 2022.	Antes de 2017.
Tipo de documento:	Publicaciones académicas (arbitradas).	Artículos de información; tesis de licenciatura, maestría y doctorado; actas de congresos; ponencias; noticias; revisiones de libros.
Documento:	Documento completo.	Documento duplicado; sin resumen disponible.
Idioma:	Todos.	No aplica.
Área temática.	Educación.	No corresponde a Educación.

La primera revisión se llevó a cabo con la ecuación *Metaevaluación OR Meta-evaluation*; de esta búsqueda se hizo un preanálisis donde se identificaron los artículos que aluden de manera explícita la palabra de la ecuación en cualquiera de estos apartados: el título, el

resumen o las palabras clave. Asimismo, se establecieron dos normas de eliminación: artículos repetidos y sin resumen disponible. La localización de documentos se inició en *Scopus*, sin ningún interés particular, donde se ubicaron 34 documentos, le siguió *EBSCOhost* con 16 y *WoS* con 12 (véase la Tabla 9).

Tabla 9

Artículos resultantes de la búsqueda en las bases de datos

Ecuación de búsqueda: Metaevaluación OR Meta-evaluation	
Base de datos	No. Documentos
<i>Scopus</i>	33
<i>EBSCOhost</i>	16
<i>WoS</i>	12
Total de artículos para revisión	61

En la segunda búsqueda con la ecuación *Metaevaluación OR Meta-evaluation AND "Educación superior" OR "Enseñanza superior" OR "Higher education"* se localizaron 3 artículos en *Scopus* y 2 en *EBSCOhost*. Sin embargo, esos artículos se eliminaron porque eran repetidos de la primera búsqueda. Por último, con la ecuación *Metaevaluación OR Meta-evaluation AND "Educación superior" OR "Enseñanza superior" OR "Higher education" AND "Evaluación del profesorado" OR "Evaluación del docente" OR "Teacher evaluation"* no se encontraron documentos en ninguna de las tres bases de datos.

Para la gestión de los documentos se empleó la aplicación de escritorio de *Mendeley Reference Manager 2.77.0*, con el fin de almacenar, organizar y consultar todos los documentos en una sola biblioteca.

El paso siguiente consistió en identificar si los artículos encontrados correspondían a la realización de una metaevaluación declarada de forma explícita, para esto se asignaron etiquetas con el fin de sistematizar los hallazgos. De esta forma, en inicio, se asignaron dos

etiquetas: Aceptado y Descartado. La etiqueta Aceptado correspondió a una declaración directa y descriptiva de un ejercicio de metaevaluación, la etiqueta Descartado se asignó cuando se abordaba una metaevaluación, pero de forma indirecta o complementaria, no como parte del estudio que se presentaba, o si se indicaba la palabra metaevaluación sin aludir al desarrollo de la misma. Así, se descartaron 53 artículos, y restaron 8 en revisión (véase la Tabla 10).

Tabla 10

Distribución de artículos según etiquetas

Bases de datos	Etiquetas asignadas a los artículos	
	Aceptado	Descartado
<i>Scopus</i>	5	28
<i>EBSCOhost</i>	2	14
<i>WoS</i>	1	11
Total	8	53

Asimismo, fueron tomados en cuenta tres estudios más identificados de forma inductiva por consistir en una metaevaluación de programas de evaluación docente en educación superior, dos en el contexto nacional y uno internacional. Aunque dos de los estudios mencionados no cumplieron con el criterio de inclusión del periodo de publicación —uno es de 2008 y otro de 2016—, fueron considerados por su relevancia en la búsqueda de antecedentes para esta investigación. Los tres documentos se incluyen en la descripción que a continuación se presenta —los ocho artículos recuperados de las bases de datos y los tres apenas señalados, 11 en total— en orden cronológico del artículo más reciente al más antiguo en la Tabla 11. Posterior a la tabla se indican de forma general los principales elementos de cada estudio —objetivo, diseño, participantes y hallazgos según se identificó—.

Tabla 11

Características generales de antecedentes de metaevaluación

Año	Autore(s)	Título	País	Nivel educativo
2022	Patiño-Montero et al.	Actualización de la evaluación docente de posgrados en una universidad multicampus: experiencia desde la Universidad Santo Tomás (Colombia).	Colombia	Posgrado
2021	Bundi, Frey y Widmer	Does evaluation quality enhance evaluation use?	Suiza	Básico
2021	Chaverra-Fernández y Hernández-Álvarez	La metaevaluación en el profesorado de educación física	Colombia	Básico
2021	Gutiérrez-de-Rozas y Carpintero	Análisis de la evaluación de programas de intervención en motivación en Educación Secundaria	España	Básico
2020	Chen et al.	The construction of meta-evaluation indicators of Taiwan's university program evaluation: JCSEE program evaluation standards as a framework.	Taiwan	Superior
2019	Sánchez	Meta-evaluación del Desempeño Docente en la UATx: el Caso de la Facultad de Ciencias de la Educación	México	Superior
2019	Vieirai et al.	Avaliação do Relatório de Avaliação Pós-ocupação da Escola Municipal Albert Schweitzer: um estudo meta-avaliativo.	Brasil	Posgrado
2018	Mediell y Dionne	Un outil à visée pédagogique pour discuter de méthodologie	Canadá	Superior (impreciso)
2018	Rincon-Flores et al.	Strengthening an Educational Innovation Strategy: Processes to Improve Gamification in Calculus	México	Superior

Antecedentes de investigación en metaevaluación

Año	Autore(s)	Título	País	Nivel educativo
		Course through Performance Assessment and Meta-evaluation		
2016	Universidad Complutense de Madrid	Modelo de evaluación de la actividad docente del profesorado de la UCM: programa DOCENTIA-UCM	España	Superior
2008	Elizalde et al.	Metaevaluación del proceso de evaluación docente: Universidad Autónoma del Estado de Hidalgo	México	Superior

Patiño et al. (2022) realizaron un ejercicio de metaevaluación de la evaluación docente de posgrados de la Universidad Santo Tomás, en el período 2017-2020. El estudio fue cualitativo y correspondió a una metodología de investigación evaluativa, se contó con la participación de estudiantes, docentes, directoras y decanas de la facultad —muestra no probabilística—. El procedimiento consistió en ocho pasos: 1) Análisis contextual, 2) Revisión de referentes conceptuales e institucionales, 3) Análisis de implementación, 4) Evaluación de la batería de instrumentos utilizados —instrumento de evaluación de estudiantes; instrumento del director o líder del Programa Académico de Posgrado; instrumento del decano, solo se emplea cuando no se cuenta con director o líder del Programa Académico de Posgrado, e instrumento de autoevaluación Docente de Posgrado—, 5) Formulación de nueva batería de instrumentos, 6) Socialización y ajustes con los agentes involucrados, 7) Validación de métricas de los instrumentos y 8) Implementación. Los hallazgos principales reflejaron la falta de políticas institucionales efectivas, socialización de los resultados de la evaluación a los docentes, el derecho a réplica del profesorado. La metaevaluación conllevó a la consolidación y parametrización la actualización y mejoramiento de un aplicativo institucional.

El estudio de Bundi y Widmer (2021) tuvo el objetivo de investigar la relación de la calidad de una evaluación con la evaluación percibida y aceptada de los involucrados por medio de una metaevaluación. Fueron analizadas 34 evaluaciones externas de 2006 a 2014 y 307 participantes de estas escuelas (docentes, administradores y directivos), y se empleó una encuesta para recabar datos, estos fueron un análisis cualitativo de contenido (ACC). Los hallazgos mostraron que la calidad de la evaluación no se vincula con la opinión y el uso percibido de la evaluación, pero se relaciona con el impacto percibido de la evaluación. Por tanto, se concluyó que la calidad de la evaluación y su uso no necesariamente se asocian.

Otro referente de la metaevaluación se realizó con el fin de analizar las reflexiones de docentes después del proceso evaluativo de un curso—evaluación del aprendizaje— (Chaverra-Fernández y Hernández-Álvarez, 2021). El diseño de investigación correspondió a un estudio de caso múltiple, con el empleo de entrevistas semiestructuradas y un análisis documental. Participaron seis docentes de educación física seleccionados bajo criterios de conveniencia para el estudio. El análisis cualitativo de los datos reflejó la percepción positiva de los docentes sobre su práctica sustentada en el cumplimiento del programa del curso, y del aprendizaje que estiman de sus estudiantes. También los docentes reconocieron áreas de mejora de su desempeño, sin embargo, llegar a la conclusión del curso no conllevó necesariamente a una reflexión de los docentes sobre los procesos evaluativos que aplican, ni de los propósitos formativos de su evaluación, tampoco de aportes a una enseñanza de calidad.

Posteriormente, mediante la aplicación de estándares, se realizó una metaevaluación con el objetivo de analizar los procesos y resultados de evaluación de diferentes programas

de educación emocional y motivación para estudiantes de secundaria (Gutiérrez de Rozas y Carpintero, 2021). Para tal fin, se llevó a cabo una revisión de la literatura académica. Así, fueron estudiados 19 documentos sobre la evaluación de programas de intervención en motivación y en educación emocional. Los resultados arrojaron que existe una diversidad de procedimientos de evaluación y de formas de trabajar los resultados, donde sobresalieron evaluaciones de tipo cualitativo, cuantitativo y mixto. También, la mayoría de las evaluaciones fueron sumativas, centradas mayormente en los resultados que en el diseño e implementación del programa. Además, la mayoría de las evaluaciones se realizaron de manera interna.

Se continua con una investigación donde el objetivo fueron los indicadores de metaevaluación de la autoevaluación departamental universitaria de Taiwán a la luz de los estándares del Comité Conjunto sobre Estándares para la Evaluación Educativa (por su nombre en inglés Joint Committee on Standards for Educational Evaluation [JCSEE]) (Chen et al., 2020). La comparación de los cinco estándares del JCSEE (de utilidad, de factibilidad, de propiedad, de precisión y de responsabilidad de evaluación) con el apoyo de seis expertos y académicos develaron que los expertos y académicos consideraron los estándares de propiedad y de precisión son los más importantes para evaluaciones de programas universitarios, contrario al propio orden de importancia de los estándares del JCSEE.

En México, la Universidad Autónoma de Tlaxcala (UAT) llevó a cabo una metaevaluación al cuestionario de evaluación del desempeño docente que se evaluó por medio de la opinión del estudiante a partir de tres dimensiones: personal, disciplinar-pedagógica y profesional (Sánchez, 2019). En este estudio participaron 52 docentes y 204

estudiantes de la Facultad de Ciencias de la Educación. Algunos de los hallazgos indicaron que docentes y estudiantes priorizan evaluar en la dimensión personal: el respeto, dedicación de tiempo dentro y fuera del aula, y atención personalizada hacia el estudiante; en la dimensión disciplinar-pedagógica: planeación didáctica, claridad, diseño de estrategias enseñanza-aprendizaje y justicia de evaluación; y en la dimensión profesional: puntualidad y asistencia.

Enseguida, Vieirai et al. (2019) presentaron una metaevaluación del Informe de Evaluación Post-ocupación de la escuela municipal Albert Schweitzer con estudiantes de posgrado en Arquitectura. El propósito de la Evaluación Post-ocupación es obtener recursos para corregir fallas, valorar aciertos y delimitar nuevos proyectos; dicha evaluación es empleada por arquitectos y se aplica a edificios transcurrido un tiempo de su construcción. Este estudio verificó la concordancia entre el Informe de la Evaluación Post-ocupación y los estándares del JCSEE. Fueron delimitados 22 estándares relacionados con cuatro atributos del JCSEE. Los resultados mostraron que la Evaluación Post-ocupación cumplió con la mayoría de los estándares, con su propósito y expectativas de las partes interesadas.

El siguiente artículo propuso una matriz de metaevaluación formativa para la evaluación de programas centrada en los elementos metodológicos: planificación, implementación y comunicación (Mediell y Dionne, 2018). Se partió del fundamento de que la calidad de una evaluación depende, entre otras cuestiones, de una metodología demostrada que brinde las pruebas necesarias para realizar un juicio. Así, con base en un análisis de diversos criterios para la evaluación de programas (Patton, 2013; Sanders y Nafziger, 2011; Scriven, 2005, 2011; Stufflebeam, 1999, 2004; Widmer et al., 2000 como se citan en Mediell y Dionne, 2018) se construyó una propuesta que permitiera impulsar

una reflexión del evaluador y comprobar la calidad de las evaluaciones desde el diseño hasta la entrega del informe final. Los ejes de la matriz se relacionan con la descripción del proceso de evaluación, las fuentes de información y los procedimientos de muestreo, la descripción de las especificaciones de evaluación, la recolección de datos, la calidad y análisis de datos cuantitativos y limitaciones metodológicas.

Otra investigación realizada en México correspondió a la implementación de un proceso de mejora para una actividad de gamificación de un curso de cálculo para estudiantes universitarios (Rincon-Flores et al., 2018). El proceso contempló tres tipos de datos: 1) los resultados de 50 evaluaciones del desempeño de los estudiantes, 2) la metaevaluación de los efectos proactivos y costes de gamificación por parte de docentes del curso, 3) los datos sobre el nivel de satisfacción alcanzado en la actividad. También, fueron diseñados tres instrumentos: a) una rúbrica de evaluación del desempeño, b) un inventario para generar procesos de metaevaluación y c) un cuestionario de evaluación de la pertinencia de las actividades de gamificación. Se concluyó que la gamificación es una estrategia de alto nivel de innovación y aporta motivación y emoción para fomentar el aprendizaje.

Un referente de metaevaluación de un programa de evaluación docente en el escenario europeo es el Programa de apoyo para la evaluación de la actividad docente del profesorado universitario (DOCENTIA) de la Universidad Complutense de Madrid (UCM, 2016). Este programa fue conformado bajo la intervención de 3 organizaciones: la Agencia Nacional de Evaluación de la Calidad y Acreditación (ANECA), la Fundación para el Conocimiento Madrid y el Consejo de Gobierno de la UCM. La finalidad de DOCENTIA es mejorar la calidad de la enseñanza y generar información de utilidad para evaluar el

desempeño del profesorado. El programa se divide en 3 dimensiones: estratégica, metodológica y de resultados, revisión y mejora. En la tabla 12 se indican, de forma general, características de cada dimensión.

Tabla 12

Dimensiones del programa DOCENTIA

Estratégica	Metodológica	De resultados, revisión y mejora
Valoración sistemática del desempeño docente para la definición de objetivos. Finalidades de la evaluación docente.	Diversidad en las fuentes de información para evitar sesgos. Generar un espacio fiable y válido de la información. Fuentes de información: bases de datos de la institución, información aportada por los docentes y la información de las autoridades académicas. El periodo de evaluación es de 3 años.	Marco de la evaluación. Informes. Toma de decisiones.

Nota. De “Modelo de evaluación de la actividad docente del profesorado de la UCM: programa DOCENTIA-UCM,” por Universidad Complutense de Madrid, 2016, pp. 8, 13, 26. CC BY-NC-ND.

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).

La metaevaluación del programa DOCENTIA-UCM 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). Por tanto, los autores indicaron que se establecieron doce objetivos específicos, algunos de ellos señalan estudiar la dimensionalidad de los instrumentos, comparar las dimensiones

destacadas de la literatura sobre la calidad docente, 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 docentes.

En el contexto mexicano, la Universidad Autónoma del Estado de Hidalgo (UAEH) 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). Algunos de los aspectos que se evaluaron son: planeación, conducción, evaluación, materiales y técnicas didácticas, normatividad, asesorías y satisfacción del aprendizaje. Para obtener información sobre lo anterior se diseñaron instrumentos institucionales conforme al nivel y modalidad educativa (Elizalde et al., 2008).

Elizalde et al. (2008) reportó algunos de los hallazgos de la evaluación docente en la UAEH, entre ellos se destacó que el 92.7% de los docentes consideraron que el cuestionario de evaluación debe estar en estricto apego a las especificidades de cada programa educativo. Además, el 83.2% de los docentes expresaron que la autoevaluación es una práctica que favorece la reflexión de su trabajo. De igual forma, con un porcentaje semejante, los docentes indicaron que los resultados de su evaluación hacen posible que sus fortalezas y debilidades sean más claras; de la misma manera, valoraron esta información para tomar decisiones que beneficien su desempeño académico.

En el marco de la descripción general de los estudios, se reconoce la escasez de trabajos de investigación sobre metaevaluación de programas de evaluación docente en educación superior. Solamente cuatro de los estudios anteriores abordó una metaevaluación a un sistema o programa de evaluación docente (Elizalde et al., 2008; Patiño-Montero et al.,

2022; Sánchez, 2019; UCM, 2016). De tales estudios la descripción del proceso de metaevaluación es limitada. Así pues, contar con referentes empíricos es valioso, pero la poca información de una metaevaluación detallada propicia incertidumbre. También, se identificó que existe una variabilidad en cómo se nombra y discute una metaevaluación — meta-análisis, revisión, evaluación y hasta reflexión—, que distan de las precisiones de Scriven (2009, 2011). Esto dificulta, por una parte, la identificación de estudios de metaevaluación debido a la polisemia, y por otra parte a la apropiación errada del constructo. Esa fue una de las principales razones por las que los artículos revisados se redujeron, pues, aunque el documento incluía la palabra metaevaluación su correspondencia se ubicó mayormente como un meta-análisis. Se resalta la necesidad de estructuras de metaevaluación para su operacionalización que favorezcan los resultados y emisión de juicios (Rincon-Flores et al., 2018).

Dicho lo anterior, la presente metaevaluación abonará al conocimiento de este tema, y aportará evidencias de un contexto particular en educación superior, bajo una aproximación específica.

Método

Esta investigación se fundamenta en el paradigma naturalista por considerarse el más adecuado para guiar un estudio en evaluación (Guba y Lincoln, 1981, como se cita en Guba y Lincoln, 2012). Una investigación naturalista no se orienta a la generalización con base en un caso representativo, sino a desarrollar interpretaciones particulares que posibiliten profundizar en la experiencia humana (Armstrong, 2010). Asimismo, este tipo de enfoques se adapta bien al estudio de grupos sobre los que se sabe poco, y de los que se busca ahondar en su visión del entorno, y como ésta se refleja en sus prácticas, estructuras e interacciones (Armstrong, 2010).

La perspectiva epistemológica optada fue subjetivista, pues se resalta la participación del investigador en la situación de estudio para comprender e interpretar esa realidad (Tojar, 2006). En este sentido, se parte de una visión teórica interpretativa en función del desarrollo de interpretaciones del entorno educativo con una postura cultural e histórica (Sandín, 2003).

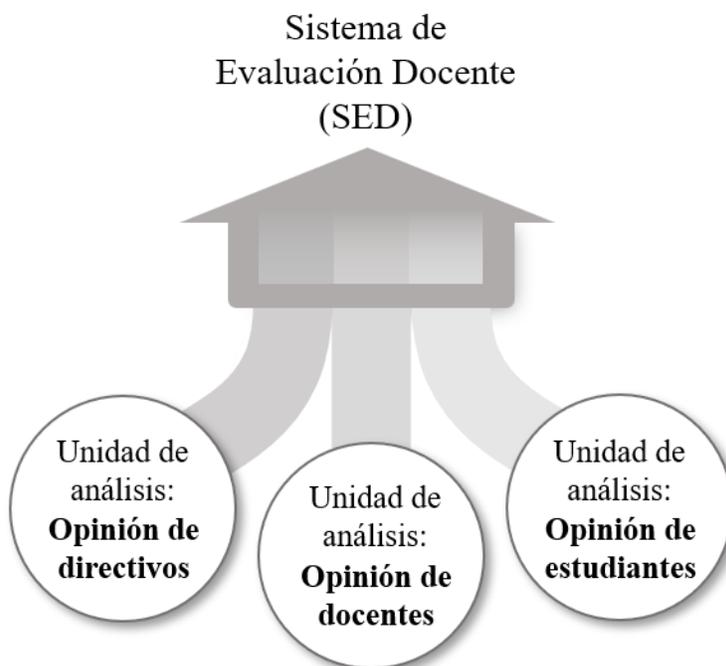
El diseño de investigación es un estudio de caso, en la perspectiva de atender el cómo o porqué de una condición actual (Yin, 2018) y ocuparse de la complejidad de una situación particular y de la interacción con su entorno (Stake, 1999). Yin (2017) menciona que este diseño puede incluir uno o múltiples casos. Así, el autor también afirma que, cuando un mismo estudio contiene más de un caso se le reconoce como estudio de caso múltiple, y estos a su vez están divididos en dos diseños: holístico e integrado. El diseño de caso múltiple holístico examina la naturaleza global del objeto de estudio; por su parte, el diseño de caso múltiple integrado involucra unidades de análisis (también llamadas

subunidades o casos de segundo nivel) dentro de un caso (igualmente denominado caso de primer nivel) (Yin, 2018). Además, como parte del diseño, fueron retomados elementos del modelo de evaluación responsiva, respecto a la atención de las condiciones del contexto y la recuperación de las distintas perspectivas de los diversos actores (Stake, 1972, 2006, 2011).

Por tanto, esta investigación se realizó bajo un diseño de estudio de caso múltiple integrado conformado por un caso de primer nivel y tres unidades de análisis. El caso de primer nivel correspondió a la metaevaluación del SED y las tres unidades de análisis fueron las opiniones de directivos, docentes y estudiantes (véase la Figura 5).

Figura 5

Estudio de caso múltiple integrado



Nota. De “Getting started: how to know whether and when to use the case study as a research method”, por R. K. Yin, *Case study research and applications: Design and methods* (6ta ed., pp. 31-56), 2018, Sage publications. CC BY-NC-ND.

Método

De esta forma, se abordó un análisis descriptivo minucioso y en profundidad (Tojar, 2006) que conllevó recoger y analizar datos tanto cualitativos como cuantitativos. Fueron empleadas técnicas de análisis cualitativo desde dos variantes, una variante en busca de una descripción densa inductiva-deductiva (Gibbs, 2012), y la otra variante fue el Análisis Cualitativo de Contenido (ACC) definido como un estudio empírico y metodológico de información desde su contexto de comunicación, con base en reglas y modelos paso a paso (Mayring, 2000). Asimismo, se realizó un análisis univariado para el estudio de los datos cuantitativos a fin de reportar estadísticos descriptivos y particularidades en la distribución de los datos (Afifi et al., 2020; Mujis, 2004).

Participantes

Se conformaron dos grupos de participantes: Grupo 1) Coordinadores de la administración central, y Grupo 2) Directores y subdirectores, docentes y estudiantes, ambos de la UABC; enseguida se describen sus características.

Grupo 1) Coordinadores de la administración central

Se eligieron dos participantes en virtud de la responsabilidad que tiene su labor con el SED. Los participantes fueron el Coordinador General de Formación Básica (CGFB) y el Coordinador de Planeación y Desarrollo Institucional (CPDI). Las funciones del CGFB tienen el objetivo de organizar y regular la pertinencia y reconocimiento de calidad de los programas educativos de licenciatura alineados al Modelo Educativo de la universidad (UABC, 2017a), además el SED es responsabilidad de esta coordinación. El segundo participante de este grupo, el CPDI desempeña diversas funciones, tales como elaborar, dar seguimiento y evaluar los avances del Plan de Desarrollo Institucional; realizar los informes anuales, parciales y de gestión rectoral; producir y publicar el Sistema Institucional de

Indicadores, así como generar la información estadística correspondiente; organizar y evaluar el cumplimiento de los estándares de calidad de los servicios de la universidad; y analizar el funcionamiento y la estructura académica y administrativa de la universidad (UABC, 2019).

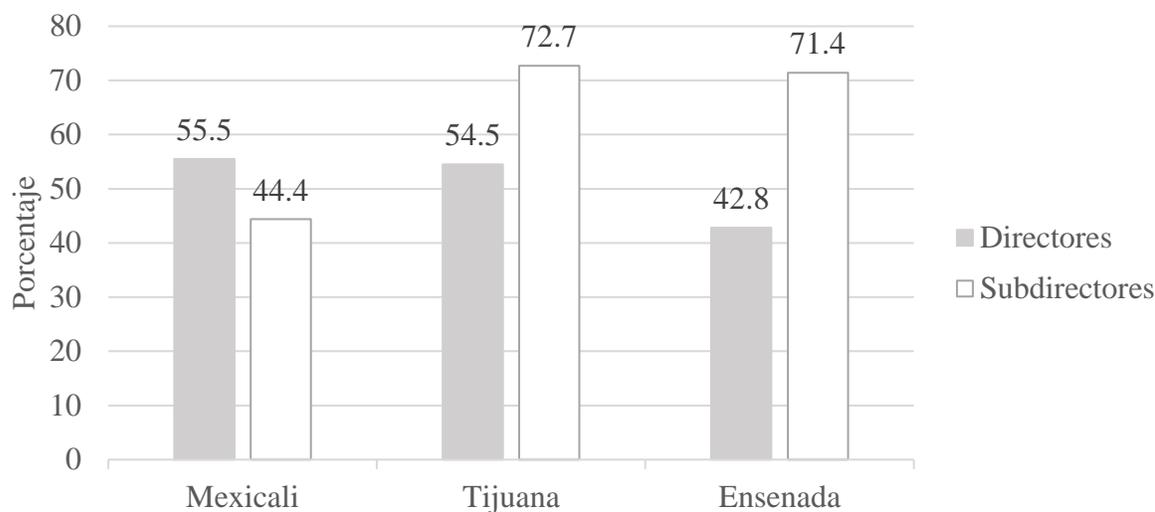
Grupo 2) Directores y subdirectores, docentes y estudiantes

En atención a los involucrados en el proceso evaluativo de la docencia en la UABC, el Grupo 2 se estructuró en tres subgrupos: *a) Directores y subdirectores, b) Docentes, y c) Estudiantes*. Para la conformación de las muestras se invitó a toda la población de cada subgrupo, esto con la finalidad de recabar la mayor información y brindar la oportunidad de participación a los actores educativos involucrados. De esta forma, se obtuvo una muestra no probabilística por autoselección en cada subgrupo.

En el caso del *Subgrupo a) Directores y subdirectores*, la muestra por autoselección fue de 40 participantes: 19 directores y 21 subdirectores, lo cual representó el 55.5% de participación del total de esta población. La Figura 6 muestra el porcentaje tanto de directores como de subdirectores por campus. Así, el porcentaje de colaboración específico de directores y subdirectores de Mexicali fue 50, 63.6% de Tijuana y 57.1% Ensenada.

Figura 6

Porcentaje de la muestra del Subgrupo a) Directores y subdirectores



En el *Subgrupo b) Docentes* se integró una muestra por autoselección de 1,431 docentes, esto representó el 24.60% del total de docentes de la universidad, 612 fueron docentes de carrera y 757 docentes de asignatura, el 4.3% no indicó el tipo de docente al que pertenecía. Los docentes involucrados correspondían a las siete áreas del conocimiento de la UABC (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), y de las tres modalidades educativas en las que se imparten clases (no presencial [a distancia], mixta [semipresencial] y presencial). Respecto a su antigüedad laboral, el 52.6% refirió contar con 8 o más años; 20.5% de 4 a 7 años y el 21.9% de 1 a 3 años. El 4.6% de los docentes no informó su tiempo de trabajo (véase la Tabla 13).

Tabla 13*Características generales de los participantes del Subgrupo b) Docentes*

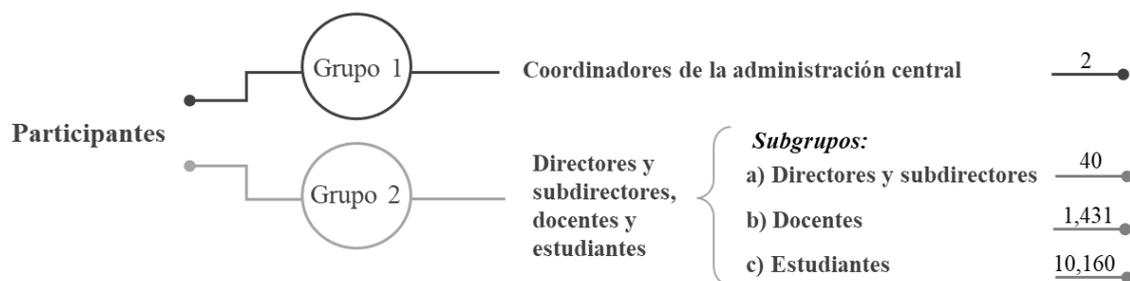
Ítem	Respuestas	Frecuencias	%
Tipo de académico	Docentes de carrera	612	42.7
	Docentes de asignatura	757	52.9
	Sin respuesta	62	4.3
Áreas del conocimiento	Ciencias de la Ingeniería y la Tecnología	322	22.5
	Ciencias de la Educación y Humanidades	289	20.2
	Ciencias de la Salud	260	18.1
	Ciencias Sociales	226	15.7
	Ciencias Económico Administrativas	209	14.6
	Ciencias Naturales y Exactas	91	6.3
	Ciencias Agropecuarias	34	2.3
Modalidad educativa	Presencial	1,285	89.7
	Semipresencial	70	4.9
	A distancia	76	5.3
Tiempo laboral	8 años o más	754	52.6
	4 a 7 años	296	20.6
	1 a 3 años	314	21.9
	Sin respuesta	67	4.6
Total de docentes de la muestra		1,431	

La muestra por autoselección del *Subgrupo c) Estudiantes* fue constituida por 10,160 estudiantes de licenciatura de las siete áreas del conocimiento de la universidad, lo cual significó el 15.9% del total de estudiantes matriculados en el periodo 2019-2. De estos, poco más de la mitad fueron de sexo femenino (52.1%) y casi tres cuartas partes eran estudiantes regulares (73.2%). Referente a las etapas de formación, cerca de la mitad se ubicó en etapa básica (42.5%) y el resto se distribuyó en las etapas disciplinar y terminal (46.6%), por lo tanto, se tuvo participación de los estudiantes de las distintas etapas de formación curricular. Y sobre la edad del estudiante, la mayoría contaba entre 19 y 20 años de edad (29.9%), le siguió entre 21 y 22 años con una cuarta parte (25.8%) y los demás se ubicaron entre 23 a 24 años (12.2%), 25 años o más (11.9%) y 17 a 18 años (10.1%) (véase la Tabla 14).

Tabla 14*Características generales de los participantes del Subgrupo c) Estudiantes*

Ítem	Respuestas	Frecuencias	%
Sexo	Femenino	5,298	52.1
	Masculino	3,740	36.8
	Sin respuesta	1,122	11.0
Tipo	Regular	7,439	73.2
	Irregular	1,596	15.7
	Sin respuesta	1,125	11.0
Etapa de formación	Básica	4,320	42.5
	Disciplinaria	2,884	28.3
	Terminal	1,859	18.3
	Sin respuesta	1,097	10.8
Edad	17-18 años	1,032	10.1
	19-20 años	3,046	29.9
	21-22 años	2,627	25.8
	23-24 años	1,242	12.2
	25 años o más	1,215	11.9
	Sin respuesta	998	9.8
Total de estudiantes de la muestra		10,160	

La Figura 7 presenta la distribución y número de participantes de cada grupo.

Figura 7*Grupos de participantes*

Materiales

Para obtener evidencia empírica se emplearon distintos materiales de acuerdo a cada grupo y subgrupo de participantes. Para el Grupo 1) Coordinadores de la administración central se utilizó un guion de entrevista. En el caso del Grupo 2) Directores y subdirectores, docentes y estudiantes se usó un cuestionario para cada subgrupo: *a) Directores y subdirectores, b) Docentes, y c) Estudiantes*. La Tabla 15 integra la organización de los grupos y subgrupos de participantes alineados a las técnicas de recolección de datos y los materiales dispuestos.

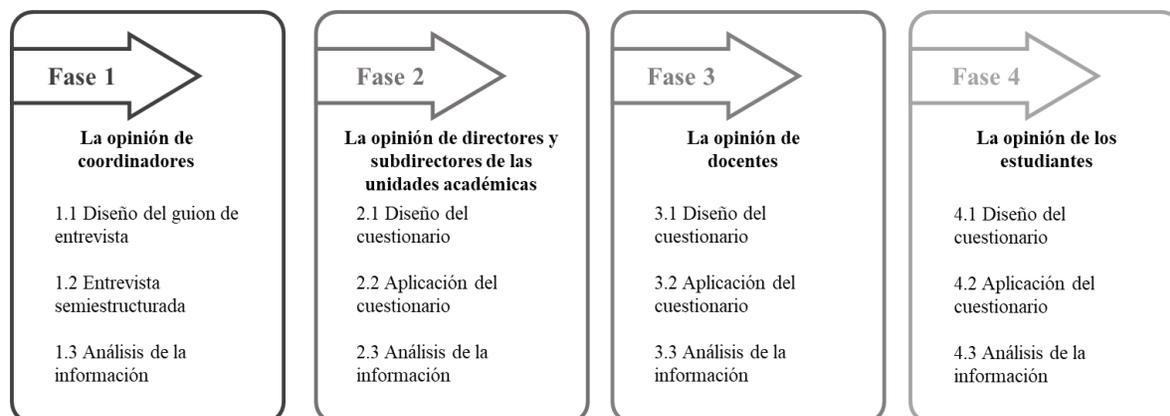
Tabla 15

Participantes, técnicas de recolección de datos y materiales

Grupos de participantes		Técnicas de recolección de datos	Materiales
Grupo 1) Coordinadores de la administración central		Entrevista semiestructurada	Guion de entrevista
Grupo 2) Directores y subdirectores, docentes y estudiantes		Subgrupos: <i>a) Directores y subdirectores</i> <i>b) Docentes</i> <i>c) Estudiantes</i>	Cuestionario

Procedimiento

El procedimiento para recolectar los datos se organizó en cuatro fases: Fase 1. La opinión de coordinadores; Fase 2. La opinión de directores y subdirectores; Fase 3. La opinión de docentes; y Fase 4. La opinión de los estudiantes. La Figura 8 se presenta la estructura de las fases.

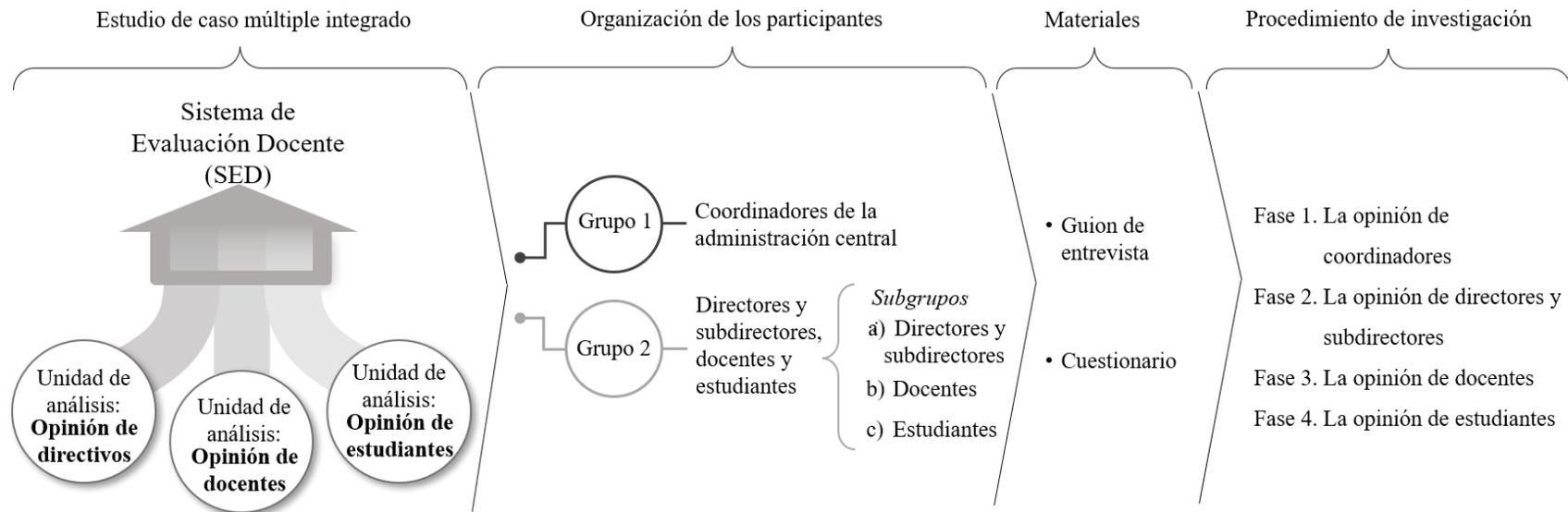
Figura 8*Procedimiento de la investigación*

En la Figura 9 se muestra la distribución y correspondencia de los apartados del método de esta investigación. La primera organización realizada fueron las unidades de análisis del estudio de caso múltiple integrado: Opinión de los Directivos, Opinión de los Docentes y Opinión de los Estudiantes; enseguida, la distribución de los dos grupos de participantes; posteriormente, se ubican las cuatro fases del procedimiento de la investigación.

Método

Figura 9

Distribución de los apartados del método



A continuación, se detallan las cuatro Fases del procedimiento.

Fase 1. La opinión de coordinadores

Etapa 1.1 Diseño del guion de entrevista. El guion de entrevista para el Grupo 1) Coordinadores de la administración central se conformó de un conjunto de preguntas guía con la finalidad de indagar en el proceso evaluativo que gestiona el SED, a partir de la función que desempeñan los coordinadores y su compromiso en la definición de las políticas y funcionamiento del SED. También, la construcción del guion se realizó a la luz del sustento teórico derivado de los planteamientos de la RIIED (2008) y las consideraciones referidas por Rueda (2010), por orientar la reflexión en el diseño e implementación de programas de evaluación de la docencia y su pertinencia al contexto mexicano (véase el Apéndice A).

Etapa 1.2 Entrevistas semiestructuradas. Las entrevistas se concertaron por medio de una invitación al correo electrónico institucional de cada coordinador, se les informó acerca del objetivo de la entrevista y el guion se les envió con antelación. Asimismo, se previó contar con el equipo tecnológico necesario para la entrevista (laptop o PC; grabadora; Internet). Las entrevistas se realizaron por medio de una videollamada en la aplicación *Meet* de *Google*. Además, se solicitó a cada coordinador su consentimiento para grabar el audio de la entrevista. La primera entrevista tuvo una duración de 50 minutos, y la segunda entrevista duró 47 minutos.

Etapa 1.3 Análisis de la información. Los audios de las entrevistas fueron transcritos para identificar los temas principales de análisis. La transcripción del audio de la primera entrevista fue de 12 cuartillas, y de la segunda fue de seis. Para el estudio de las transcripciones se realizó un análisis de contenido por categorías inductivas con la

intención de apoyar el desarrollo de los cuestionarios de las siguientes fases. El análisis de contenido refirió a un proceso de descubrimiento y organización de conceptos y relaciones de los datos, para sustentar las decisiones sobre dónde y cómo conseguir información adicional que favorezca el avance de la investigación (Strauss y Corbin, 2002).

Fase 2. La opinión de directores y subdirectores

Etapa 2.1 Diseño del cuestionario. Después del análisis de las entrevistas a los coordinadores, y con base en los referentes teóricos previstos en la Fase 1, se diseñó el cuestionario dirigido al subgrupo *a) Directores y subdirectores*. El propósito fue identificar su perspectiva sobre el sistema desde su función directiva. Para tal fin, el cuestionario se estructuró en cuatro dimensiones: Dimensión 1. Uso de los resultados de la evaluación; Dimensión 2. CECD; Dimensión 3. Aplicación del CECD; y Dimensión 4. Valoración del SED. Cada dimensión se orientó hacia un tema en particular en atención a los objetivos propuestos (véase la Tabla 16).

Tabla 16

Dimensiones del cuestionario

	Dimensiones	Finalidad
1	Uso de los resultados de la evaluación	Describir los usos de la evaluación.
2	CECD	Ahondar en la estructura, contenido y pertinencia del cuestionario.
3	Procedimiento de aplicación	Indagar en el procedimiento de aplicación y la generación de reportes de la evaluación.
4	Valoración del SED	Profundizar sobre la administración del sistema gestor del SED

El número de ítems del cuestionario fue de 24 y se contó con diversas opciones de respuesta: de escala, múltiple, dicotómica y abierta (véase el Apéndice B).

Etapa 2.2 Aplicación del cuestionario. El cuestionario se administró en el *software LimeSurvey* versión 1.92+ *Build* 120919. Los directores y subdirectores fueron invitados a participar por medio de un mensaje a su correo electrónico institucional. En el, se indicó el objetivo de la investigación, el enlace para acceder al cuestionario, el tiempo aproximado de duración y la información de contacto por considerarlo de interés de los participantes. El tiempo estimado para contestar el cuestionario fue de 10 minutos aproximadamente, y se brindó un margen de 15 días para acceder a contestarlo. Los resultados se monitorearon por medio del mismo *software*, después de 7 días se envió un recordatorio con la finalidad de obtener el mayor número posible de respuestas.

Etapa 2.3 Análisis de la información. Primero, se exportó la información del *software LimeSurvey* a un archivo de comandos de *SPSS* versión 26. Los datos de los ítems cerrados (opción múltiple y dicotómica) fueron integrados en una base de datos en dicho *software* para la realización de un análisis univariado en función de la distribución de frecuencia de las variables y descripción del caso o respuesta promedio (Mujis, 2004).

Después, los datos de los ítems abiertos del cuestionario se concentraron en un archivo para su análisis con el apoyo del *software Atlas.ti* versión 7.5.4. Los datos estudiados fueron 454; la Tabla 17 muestra el número de datos por dimensión.

Tabla 17

Número de datos por dimensión del cuestionario de directores y subdirectores

Dimensión	N.º de datos
1. Uso de los resultados de evaluación	101
2. CECD	115
3. Procedimiento de aplicación	34
4. Valoración del SED	204
Total	454

Nota. N.º= número.

La técnica empleada fue el ACC con el propósito de construir categorías con un enfoque de desarrollo inductivo-deductivo de categorías (Mayring, 2000). De esta forma, se realizó una agenda de códigos conformada por categorías de análisis para esclarecer y organizar la información, cada categoría se definió y compuso por códigos, reglas de codificación y ejemplos, de tal manera que se describiera con precisión bajo qué condiciones se codificarían los datos (Mayring, 2000). Para este fin, se consideraron algunas de las dimensiones propuestas por la RIIED (2008) para la planeación y funcionamiento de programas de evaluación del desempeño docente (véase la Tabla 18).

Tabla 18

Contraste entre las dimensiones de la RIIED y las categorías de las agendas de códigos de directores y subdirectores

Dimensiones de la RIIED	Categorías de las agendas de códigos
Política de la Evaluación	
Teórica	
Metodológica-Procedimental	Metodológica-Procedimental
De Uso	Uso de los resultados
De evaluación de la Evaluación	Valoración del SED

Nota. De “Reflexiones sobre el diseño y puesta en marcha de programas de evaluación de la docencia,” por Red Iberoamericana de Investigadores de la Evaluación de la Docencia, 2008, 1(3), pp.164-167. CC BY-NC-ND.

A partir del análisis, las categorías de la agenda de códigos fueron tres: Metodológica-procedimental, Uso de los resultados y Valoración del SED. La categoría Metodológica-procedimental concentró comentarios referentes al cuestionario y su aplicación. La categoría Uso de los resultados perteneció al empleo e implicaciones de los resultados de la evaluación, y la Valoración del SED integró información sobre la apreciación e infraestructura del SED.

Fase 3. La opinión de docentes.

Etapa 3.1 Diseño del cuestionario. El procedimiento general de construcción del cuestionario guardó semejanza con la Etapa 2.1 de la Fase 2, de esta forma, la realización de la Fase 2 retroalimentó el desarrollo de esta fase. En virtud del subgrupo *b) Docentes*, el cuestionario se mejoró a la luz de la información ya recogida en las fases previas y de las particularidades de la función docente en la universidad. Esto conllevó, la revisión de las cuatro dimensiones del cuestionario y el estudio detallado de los reactivos. El cuestionario se conformó de 29 ítems cerrados (opción múltiple y dicotómicos) y abiertos (véase el Apéndice C).

Etapa 3.2 Aplicación del cuestionario. El cuestionario para docentes se difundió y aplicó bajo las mismas condiciones de operación que el cuestionario dirigido a directores y subdirectores. La invitación para responder el cuestionario fue enviada al correo electrónico de todos los docentes de la universidad, asimismo, se sumó información sobre el propósito del estudio. La administración del cuestionario se realizó en el *software LimeSurvey*.

Etapa 3.3 Análisis de la información. Los datos se exportaron de la misma forma que en la Etapa 2.2 de la Fase 2. Igualmente, el análisis de los datos se desarrolló conforme al análisis realizado al cuestionario de los directores y subdirectores. Las respuestas de los ítems cerrados del cuestionario fueron organizadas en una base de datos que se trabajó en el *software SPSS* para realizar un análisis univariado, con la finalidad de visualizar cada una de las variables, pues esta es una de las mejores formas de identificar valores atípicos, frecuencias bajas y otras particularidades de los datos (Afifi et al., 2020).

La información derivada de los ítems abiertos se concentró en un archivo para el trabajo en el *software Atlas.ti*. y la técnica aplicada fue el ACC. Así, con el propósito de

Método

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; los datos analizados fueron 1,116 (véase la Tabla 19).

Tabla 19

Número de datos por dimensión del cuestionario de docentes

Dimensión	N.º de datos
1. Uso de los resultados de evaluación	443
2. CECD	220
3. Procedimiento de aplicación	199
4. Valoración del SED	254
Total	1,116

Nota. N.º= número.

Cada categoría de las agendas de códigos resultó de un solo principio de clasificación, por tanto, los niveles de análisis conceptualmente distintos se trabajaron por separado (Holsti, 1966). En consecuencia, las categorías emergieron del estudio continuo de los datos y de la identificación de referentes teóricos (RIIED, 2008; Rueda, 2010) considerados alineados a éstas (véase la Tabla 20).

Tabla 20

Contraste entre las dimensiones de la RIIED y las categorías de las agendas de códigos de docentes

Dimensiones de la RIIED	Categorías de las agendas de códigos
Política de la Evaluación Teórica	Política de la evaluación
Metodológica-Procedimental	Metodológica-Procedimental
De Uso	Uso de los resultados
De evaluación de la Evaluación	
Ética ^a	Ética
	Competencia docente

Nota. De “Reflexiones sobre el diseño y puesta en marcha de programas de evaluación de la docencia,” por Red Iberoamericana de Investigadores de la Evaluación de la Docencia, 2008, *Revista Iberoamericana de Evaluación Educativa*, 1(3), pp.164-167. CC BY-NC-ND. ^a La dimensión Ética fue una contribución posterior de “Reflexiones generales a considerar en el diseño y puesta en operación de programas de evaluación de la docencia,” por M. Rueda, 2010, *Revista Iberoamericana de Evaluación Educativa*, 3(1), pp. 347-348. CC BY-NC-ND.

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, surgió a partir de datos asociados con el dominio del docente relativo al conocimiento disciplinar, pedagógico y compromiso profesional.

Igualmente, se obtuvo la confiabilidad entre codificadores con el propósito de brindar un criterio de calidad analítico (Mayring, 2014). El tipo de confiabilidad fue la reproducibilidad, también denominada concordancia intersubjetiva y confiabilidad de formas paralelas, que alude a un proceso replicable por analistas diferentes, además de ser

considerada una medida sólida de confiabilidad (Krippendorff, 2004). La tarea se llevó a cabo con dos analistas: codificador 1 (c1, codificador titular) y codificador 2 (c2, codificador invitado), ambos codificaron un mismo material para después comparar sus resultados (Mayring, 2014). El coeficiente de confiabilidad se obtuvo de los acuerdos en la codificación de c1 y c2 y el número de codificaciones de c1 (Holsti, 1969, como se cita en Mayring, 2014). El proceso fue hecho en dos ocasiones con el objetivo de respaldar el ACC libre de sesgos o alteraciones intensionales o accidentales, asimismo, para sostener que el ACC mantiene una misma representación para quién decida efectuarlo (Krippendorff, 2004).

Fase 4. La opinión de los estudiantes.

Etapla 4.1 Diseño del cuestionario. En general, se siguió el procedimiento de la Etapa 2.1 de la Fase 2, y con mayor similitud de la Etapa 3.1 de la Fase 3. Por lo tanto, fueron conservadas las dimensiones de los cuestionarios anteriores, pero estudiadas con base en las características del subgrupo *c) Estudiantes*. Así, el diseño de los ítems se enriqueció a la luz de los cuestionarios previos. De esta forma, el cuestionario quedó conformado de 36 ítems de escala, opción múltiple, dicotómicos y abiertos (véase el Apéndice D). La Tabla 21 concentra el número de ítems por dimensión del cuestionario de cada subgrupo: *a) Directores y subdirectores, b) Docentes, y c) Estudiantes*.

Tabla 21

Número de ítems de los cuestionarios de evaluación del SED

Dimensiones del cuestionario	Directores y subdirectores	Docentes	Estudiantes
	Número de ítems en cada dimensión del cuestionario		
	Información general	4	11
1. Uso de los resultados de evaluación	5	4	4
2. CECD	10	12	9
3. Procedimiento de aplicación	4	5	10
4. Valoración del SED	5	5	2
Total	24	30	36

Etapa 4.2 Aplicación del cuestionario. La aplicación del cuestionario para los estudiantes se realizó de la misma forma que el cuestionario para directores y subdirectores y docentes; los lineamientos, recursos y condiciones fueron semejantes.

Etapa 4.3 Análisis de la información. La información fue recuperada de igual manera que en la Etapa 2.3 de la Fase 2, y el estudio de los datos guardó alineación con la Etapa 3.3 de la Fase 3. Los datos cuantitativos también se estudiaron con un análisis univariado a partir del estudio descriptivo de las variables de forma individual (Mujis, 2004), tal como se realizó en las etapas 2.3 y 3.3 de las fases correspondientes.

Como parte del ACC, el estudio de la información se efectuó por medio del *software Atlas.ti*. Los datos analizados fueron 2,132 (véase la Tabla 22).

Tabla 22*Número de datos por dimensión del cuestionario de estudiantes*

Dimensión	N.º de datos
1. Uso de los resultados de evaluación	1,032
2. CECD	371
3. Procedimiento de aplicación	314
4. Valoración del SED	415
Total	2,132

Nota. N.º= número.

Asimismo, se elaboraron cuatro agendas de códigos con base en las consideraciones teóricas señaladas en la Etapa 3.3 para la construcción de las categorías. Al respecto, se destaca el ACC como un análisis empírico y metodológico controlado, sujeto a lineamientos analíticos dentro de un contexto de comunicación (Mayring, 2000).

Además, se realizó un análisis de confiabilidad entre codificadores, esto conllevó una comparación de dos analistas que codifican un mismo contenido para generar una medida de objetividad (Mayring, 2014). En este sentido, Krippendorff (como se cita en Mayring, 2014) señala que la confiabilidad se refiere a la prueba de concordancia intra-codificador. El tipo de confiabilidad analizada fue la reproducibilidad que consiste en un proceso que puede ser reproducido por diferentes analistas (Krippendorff, 2004). Así, de igual forma, este análisis guardo similitud con la Etapa 3.3 de la Fase 3.

Para finalizar, la elección de los participantes, materiales y elaboración del procedimiento de esta investigación se considero con el fin de garantizar la obtención sistemática de la información en atención a los objetivos propuestos.

Resultados

Los hallazgos se presentan a partir las unidades de análisis del estudio de caso múltiple integrado (Yin, 2018) (véase la Figura 4). En relación al análisis univariado, cabe resaltar que, fue realizado con los datos de los ítems de respuesta cerrada, los cuales fueron contestados por el 100% de cada muestra. La Tabla 23 presenta la correspondencia de los estudios efectuados para cada unidad de análisis: Directivos, Docentes y Estudiantes¹, así como, las dimensiones o categorías atendidas.

Tabla 23

Estructura de los resultados de la investigación

Tipos de análisis de los datos	Unidades de análisis		
	Directivos	Docentes	Estudiantes
	Coordinadores	Directores y subdirectores	
Análisis cualitativo	<i>Categorías:</i> Metodológica-procedimental Uso de los resultados Cuestiones críticas Cuestiones favorables		
Análisis univariado^a	<i>Dimensiones:</i> 1. Uso de los resultados de evaluación 2. CECD 3. Procedimiento de aplicación 4. Valoración del SED		
ACC	<i>Categorías analíticas:</i> Metodológica-procedimental: Uso de los resultados		
Análisis de confiabilidad	Reproducibilidad		

¹ De aquí en adelante, se inició con letra mayúscula las palabras: Directivos, Docentes y Estudiantes cuando se alude a la opinión de tal grupo de participantes (nombre propio); de forma opuesta, se inició con minúscula docentes y estudiantes cuando los participantes (Directivos, Docentes y Estudiantes) expresaron opiniones de estos.

Nota. ^a En el caso de Directores y subdirectores, Docentes y Estudiantes se estudiaron las *Dimensiones* del Análisis univariado y las *Categorías analíticas* del Análisis univariado. Dada la naturaleza de la información obtenida, los resultados fueron estudiados bajo una visión de complementariedad analítica para favorecer los objetivos de investigación propuestos. Así, en este apartado los resultados se presentan por unidad de análisis — Directivos, Docentes y Estudiantes—, y cada unidad se encuentra organizada por tipo de análisis, primero el análisis univariado y después el ACC. Solo en el caso de Docentes y Estudiantes se incorpora antes de finalizar el análisis de confiabilidad. Enseguida, se describe la *Unidad de análisis: Directivos* dividida en *Coordinadores* y *Directores y subdirectores*.

Unidad de análisis: Directivos

Coordinadores

Este análisis concentró la participación de los coordinadores de la administración central de la UABC. Así, el estudio de la información se organizó en cuatro categorías: a) Metodológica-procedimental, relativa al procedimiento de aplicación y a la comunicación de los resultados de la evaluación docente; b) Uso de la información, acerca del empleo de los resultados de la evaluación; c) Cuestiones favorables, en relación a aspectos positivos de la evaluación docente y del SED; y d) Cuestiones críticas, respecto a elementos críticos de la evaluación docente y del SED.

En la categoría a) *Metodológica-procedimental* se puntualizó el número de docentes que el estudiante debe evaluar cada cierre de semestre, motivo por el cual esta tarea puede resultar agobiante:

el hecho de que todos los estudiantes evalúen a todos los profesores de un determinado semestre o periodo escolar, termina siendo una experiencia abrumadora, el estudiante tiene que evaluar a muchos profesores, la escala si la multiplicas por n profesores deja de ser chica. (2019_C2_7-10²)

Asimismo, se añadió la posibilidad de que el estudiante perciba la evaluación docente como una labor de poca importancia: “lo que genera que el estudiante lo tome con poca seriedad, y no todos están interesados en evaluar a los profesores y lo hacen de manera rápida” (2019_C2_12-13).

Sobre el periodo de aplicación del CECD, los coordinadores invitaron a considerar realizar la evaluación antes del fin del periodo semestral: “pensar en que la evaluación no sea tan al cierre, pero que si haya una estrategia para que los chicos la puedan contestar, para que (...) no lo dejen al último” (2019_C1_37-39), y de esta forma, tener la posibilidad de retroalimentar la enseñanza de manera conveniente “para que se tenga la oportunidad de crear estrategias en el semestre inmediato” (2019_C1_54-56).

En relación a la comunicación de los resultados, sobresalió la falta de información para conocer los resultados de la evaluación docente, tanto de las unidades académicas como de los docentes: “creo que hay mucha desinformación, porque a veces las unidades académicas nos piden reportes sobre la evaluación docente cuando el sistema tiene un reporte de evaluación docente para el directivo y para la unidad académica” (2019_C1_89-91).

^{2 2} Se empleó un código de identificación respecto al año, unidad de análisis, dimensión y número de línea de la cita establecido en la base de datos.

Así, los coordinadores ponen el acento en la entrega de los resultados, respecto a la falta de visibilidad y desconocimiento para conocer los resultados de la evaluación: “falta más difusión, hay maestros que dicen -a mí no me entregan mi evaluación-, y es que no se entrega, está en el sistema, tú la puedes consultar, he tenido maestros (...) que se sorprenden cuando les digo eso” (2019_C1_92-95). De igual manera, también se acentúa la necesidad de difundir a los estudiantes el valor e implicaciones de la evaluación: (...) un proceso de difusión en el que se sensibilice sobre la importancia, los usos y consecuencias de esta evaluación, [...] para que puedan los estudiantes ahora sí, (...) que sus valoraciones verdaderamente resulten significativas para los tomadores de decisiones (2019_C2_24-25, 26, 27).

Referente a la categoría b) *Uso de los resultados*, se enfatizó la relevancia que tienen los comentarios de los estudiantes para los directivos: “creo que a lo que más le ponen atención los directivos es a los comentarios (...) hay muchos comentarios ricos que ayudan a tener un perfil del profesor o su desempeño” (2019_C1_50-53). Asimismo, los coordinadores consideraron que los docentes emplean los resultados de su evaluación, pero desconocen en que medida:

los profesores yo creo que los usan (...) no sé cuáles son las acciones que ellos hagan a partir de esa evaluación (...) creo que como cualquier profesor que genera un ejercicio, seguramente ha de estar comprometido con las buenas prácticas de la docencia y hace esfuerzos por mejorar, pero no tengo información para saber si hay algo sistematizado. (2019_C1_114, 115, 117-120)

Por otra parte, fue destacado el uso formativo de los resultados de la evaluación docente: “los usos deberían de ser de naturaleza formativa y que toda la información que aquí se

genere verdaderamente alimente procesos de formación docente” (2019_C2_66-68), siendo de singular importancia la contribución formativa de la evaluación: “Con lo que se tiene se puede hacer un gran aporte formativo para los profesores” (2019_C1_100-101).

En relación a la categoría *Cuestiones favorables*, los coordinadores resaltaron como un acierto contar con el SED, con instrumentos de medición pertinentes apegados al modelo educativo de la universidad, asimismo, enfatizaron la utilidad del SED como espacio generador de información para favorecer diversos procesos:

tener un sistema de evaluación docente con instrumentos probados técnicamente (...), que tienen una fuerte relación con el modelo educativo (...), que está sistematizado, (...) en un sistema electrónico, que tiene muchas opciones de reporte (...) sencillo de entender [...] y que puede servir de gran insumo para varios procesos tanto en las unidades académicas como en la universidad, y además que es un sistema que tiene las bondades o posibilidades de esparcimiento (...) hay mucho que se puede hacer con el sistema y eso sin duda tiene que ver con el asunto tecnológico y de la calidad del instrumento. (2019_C1_128-132, 135-138)

De esta manera, la relevancia del SED reside en el trabajo de expertos: “la propuesta emerge de un grupo de especialistas, también ha habido espacios de validación de la propia propuesta que subyace al sistema; entonces, en el diseño o desarrollo del sistema hay muy buenas prácticas” (2019_C2_1, 3-5).

La característica virtual y sistemática del SED fue especialmente destacada: “está en línea, eso es positivo pues llega a un número importante de estudiantes, llega rápido a todos los profesores los reportes, en automático tienen sus resultados una vez que se termine el proceso de evaluación” (2019_C1_145-148). Así, el tema sobre contar con resultados

preliminares de la evaluación fue subrayado: “se puede tener un reporte preliminar, (...) en el momento que el sistema de evaluación se abre el director ya puede estar viendo los resultados parciales (...), el sistema tiene las bondades de brindar información en tiempo y forma” (2019_C1_149-151, 152).

Respecto a la categoría d) *Cuestiones desfavorables* se discutió la credibilidad de estudiantes y docentes hacia el SED:

creo que los estudiantes en muchos casos no creen tanto en la evaluación docente, en el sentido de que se tome en cuenta para algo, pasa lo mismo con ciertos maestros, pero yo también creo que no pasa con todos. (2019_C1_12-13)

Con base en lo anterior, los coordinadores estiman el desencadenamiento de una postura de indiferencia a la luz de la falta de hechos procedentes de la evaluación: “en algunos estudiantes hay una sensación de que contestarlo da lo mismo y no identifican mejoras en la docencia” (2019_C1_22-23). También, fue cuestionada la percepción de una naturaleza de castigo del SED: “creo que la concepción que se tiene actualmente del SED puede ser como muy punitivo” (2019_C1_16).

Además, se estimó un empleo limitado de la información derivada de la evaluación, aunado al escaso conocimiento sobre el valor de la evaluación por parte de los estudiantes: “no estamos haciendo buen uso a la información, y no estamos aprovechando la información, no hay una sensibilización en los estudiantes de la importancia de la evaluación” (2019_C1_154-156), “creo que también los estudiantes no han sido muy conscientes de la importancia de la evaluación docente” (2019_C1_9). A tal efecto, fue reconocida la carencia de información brindada en relación a este tema: “hay vacíos

importantes como la difusión en materia de la sensibilización de otros actores que deberían de estar apoyando o cooperando en esta evaluación” (2019_C2_29-31).

Igualmente, se cuestionó la puesta en marcha de estrategias de manera oportuna: “dudo que las estrategias de mejora se emprendan en tiempo” (2019_C1_53-54).

Asimismo, se discutió la concepción primaria del proceso evaluativo en la universidad: “estamos perdiendo de vista lo más importante que es crear esquemas de acompañamiento para que nuestros docentes verdaderamente se empoderen del desarrollo de habilidades, de ser el caso que no han alcanzado (...) a desarrollar o de fortalecer” (2019_C2_86-89).

Por último, los coordinadores realizaron sugerencias al SED, en inicio considerar sumar otro tipo de estrategias de evaluación con el fin de contribuir la puesta en marcha de acciones para la mejora de la práctica docente: “a partir de la evaluación docente y en complemento con otro tipo de evidencias podemos (...) orientar estrategias de mejora de grupos colegiados, desde mesas de detección hasta los programas de formación de la propia unidad” (2019_C1_108-111), “creo que un sistema de evaluación docente en la universidad podría incluir otros componentes ligados a la autoevaluación y a una evaluación con algún referente externo [...] con fines formativos” (2019_C2_52-53, 57). También, se sugirió una evaluación a partir de muestras aleatorias: “si logramos que el instrumento no lo contesten todos los estudiantes y que los estudiantes contesten nada más a un conjunto de profesores (...), eso ayudaría mucho, porque al estudiante en lugar de evaluar a siete evaluará a tres” (2019_C1_42-45).

Adicionalmente, se sugirió evaluar la docencia de ciertas disciplinas con base en instrumentos particulares: “hay áreas disciplinares muy distintas que a lo mejor tendrían que tener instrumentos más *ad hoc*, como explorar la práctica de la docencia en ciertas

áreas como artes o como deportes” (2019_C1_174-176). En esta línea, también fue considerada la modalidad educativa semipresencial en función de una evaluación acorde: “la modalidad semipresencial también habrá que explorarlo, (...) es importante estar abiertos a que también se puedan evaluar estas prácticas” (2019_C1_190-191, 193-194).

Otra propuesta aludió a entablar una relación directa de los resultados de la evaluación docente con el Programa Flexible de Formación y Desarrollo Docente de la universidad: “relacionar el resultado de la evaluación con la oferta de formación docente institucional, o algún punto de recomendaciones (2019_C1_205-206). Así, orientar los esfuerzos a un seguimiento formativo personalizado para los docentes: “ofreces al docente un paquete personalizado a partir de sus resultados, le dices al docente -hay estas opciones y la institución pone a tu disposición- (...) implicaría mucho tiempo pero bueno es como un tipo repositorio” (2019_C1_222-224). Y un punto importante señalado fue la desvinculación de la evaluación docente con el Programa de estímulos al desempeño del personal docente: “Hay que invertirle a la formación y hay que quitarle este elemento de consecuencia a nivel de estímulo asociado al desempeño de los profes” (2019_C2_76-77).

Directores y subdirectores

Esta unidad de análisis conllevó un análisis univariado y el ACC (véase la Tabla 20), y en ese orden serán descritos.

Análisis univariado. En la *Dimensión 1. Uso de los resultados de evaluación* los directores y subdirectores opinaron sobre el uso de los resultados, la mayor frecuencia de respuestas se obtuvo en retroalimentar la práctica docente (64.5%). Los directivos también señalaron emplear los resultados para asignar unidades de aprendizaje (56.4%), en los informes de acreditación (33.8%) y conferir horarios (19.3%). Además, los directivos

Resultados

estimaron que los docentes usan los resultados de su evaluación para identificar fortalezas y debilidades (62.9%) y, en contraste, con un porcentaje menor (6.4) consideraron que los docentes no realizan ningún uso con los resultados. (véase la Tabla 24).

Tabla 24

Porcentaje de las respuestas sobre el uso de los resultados de evaluación según directores y subdirectores

Ítem	Respuestas	%
Como director, ¿Qué uso da a los resultados de evaluación docente?	Retroalimentar la práctica docente	64.5
	Para la asignación de unidades de aprendizaje	56.4
	En los informes de acreditación	33.8
	Asignación de horarios	19.3
	Los resultados no son tomados en cuenta	0.0
Los profesores, ¿Qué uso dan a los resultados de evaluación docente?	Para identificar fortalezas y debilidades	62.9
	Los resultados no son tomados en cuenta	6.4

Referente a la *Dimensión 2. CECD*, se cuestionó si el cuestionario integraba los elementos importantes de la práctica docente, así, la suma de los puntajes de la escala de los valores 8, 9 y 10 fue 79.3 (en adelante, se presentarán los puntajes de la suma de los últimos tres valores de la escala —8, 9 y 10—). Lo anterior refleja, en apreciación de directores y subdirectores, que el CECD contiene los componentes suficientes para su función (véase la Tabla 25).

Tabla 25

Puntajes de las respuestas sobre elementos del CECD según directores y subdirectores

Ítem	Puntajes de los valores respecto a la escala									
	1	2	3	4	5	6	7	8	9	10
¿El CECD incluye los elementos importantes de la práctica docente?	0.0	0.0	0.0	2.9	2.9	2.9	8.8	29.4	23.5	26.4

Resultados

Además, ante el cuestionamiento de considerar necesario el desarrollo de otros CECD para los estudiantes dirigidos a programas educativos particulares, dos terceras partes de los directores y subdirectores indicaron no ser necesario (66.7%) y el resto (33.3%) señaló que sí. Del grupo de directivos que respondieron afirmativamente, la mayoría mencionó necesario un cuestionario de evaluación docente para Ciencias de la Salud (58.3%), le siguieron con el mismo porcentaje (16.6) Posgrado y por áreas del conocimiento, y también para la modalidad educativa semiescolarizada y a distancia (8.3%).

Adicionalmente, el 89.1% estimaron apropiado sumar otras estrategias de evaluación docente además del CECD. En esta línea, fue resaltada la pertinencia de incorporar otras estrategias de evaluación de la docencia, tales como autoevaluación (25.8%) y evaluación de pares académicos (14.5%) (véase la Tabla 26).

Tabla 26

Estrategias de evaluación docente adicionales al CECD según directores y subdirectores

Ítem	Respuestas	%
¿Considera pertinente utilizar otras estrategias de evaluación docente?	Autoevaluación	25.8
	Evaluación de directivos	24.2
	Evaluación de pares académicos	14.5
	No	8.1

En la *Dimensión 3. Procedimiento de aplicación*, poco más de la mitad de directores y subdirectores refirieron que las fechas de aplicación del CECD son apropiadas (51.6%), el resto indicó que no eran convenientes (48.4%). También, dos terceras partes (75%) consideraron pertinente la difusión para promover que los estudiantes realicen la evaluación docente, y la otra cuarta parte (25%) señalaron desfavorable la difusión. Además, tres cuartas partes expresaron estar de acuerdo con la estrategia de aplicación del CECD (75.8%) (véase la Tabla 27).

Tabla 27

Porcentaje de respuestas sobre el procedimiento de aplicación del CECD según directores y subdirectores

Ítem	Respuestas	%
¿Considera apropiadas las fechas de aplicación del CECD?	Sí	51.6
	No	48.4
¿Considera adecuada la difusión para promover que los estudiantes contesten el CECD?	Sí	75
	No	25
¿La estrategia de aplicación del CECD es adecuada?	Sí	75.8
	No	24.2

Respecto a la *Dimensión 4. Valoración del SED*, directores y subdirectores manifestaron que conocer el desempeño de los docentes (46.8%) y la opinión de los estudiantes (43.5%) son los aspectos positivos del SED, aunado a la retroalimentación de la enseñanza (33.9%).

Análisis cualitativo de contenido. Este análisis se realizó con las respuestas de los ítems abiertos, de los cuales hubo 19 obligatorios y cinco opcionales. Así, para el estudio de la información fueron delimitadas dos categorías analíticas: Metodológica-procedimental y Uso de los resultados, en estas fueron agrupadas las categorías de la agenda de códigos. La categoría analítica Metodológica-procedimental concentró temas en relación a la estructura, aplicación y comunicación de los resultados del CECD. La categoría analítica Uso de los resultados se ocupó del empleo de los resultados y la valoración del SED. De esta forma, la descripción de las categorías analíticas se expone en el orden mencionado.

Categoría analítica Metodológica-procedimental. Los directores y subdirectores resaltaron como necesidad del CECD revisar la estructura y redacción de los ítems en atención de la extensión del mismo: “evitar realizar demasiadas preguntas y poner las más relevantes” (2019_DyS_D2-35). Asimismo, se invitó a sumar o modificar aspectos en función de la pertinencia de conocimientos a evaluar: “el instrumento evalúa aspectos que

en ocasiones no aplican para todas las asignaturas por su naturaleza, me parece complicado evaluar con los mismos reactivos a todas las áreas de conocimiento” (2019_DyS_D2-42).

También, fue señalada la incorporación al cuestionario de elementos referentes a la puntualidad y al código de ética de la universidad: “pero podrían incluirse elementos relativos a el código de ética institucional” (2019_DyS_D2-3), “debería preguntar si el profesor es puntual, si les avisa oportunamente cuando se va a suspender o reprogramar la clase” (2019_DyS_D2-4).

Un tema expuesto sobre el diseño del CECD fue trabajar en la interfaz del SED: “el sistema no es amigable para su realización” (2019_DyS_D3-29), además de hacerlo compatible para dispositivos móviles: “se debe realizar el cuestionario para dispositivos móviles” (2019_DyS_D3-33).

Sobre la aplicación del CECD, fue mencionada la pertinencia del periodo de aplicación del cuestionario: “las fechas ahora son mejores” (2019_DyS_D3-17); pero, se sugirió retrasar la aplicación del CECD en función del término del semestre: “la apertura se hace prácticamente un mes y medio antes de la finalización del semestre, creo que debiera ser un poco más tarde 15 días antes de la finalización” (2019_DyS_D3-9). De igual forma, estimaron conveniente realizar la evaluación en dos etapas: “que se realice en dos momentos: al concluir el primer mes del ciclo escolar y al concluir el ciclo” (2019_DyS_D3-3).

En relación a la obligatoriedad de la evaluación, directores y subdirectores mostraron acuerdo en condicionar la reinscripción del estudiante a la evaluación docente: “que exista la restricción al alumno de no poder realizar su inscripción si no ha realizado la encuesta docente” (2019_DyS_D3-13). Además, se consideró reiterar al estudiante el

resguardo de identidad al realizar la evaluación, con el fin de generar mayor confianza y libertad de respuesta: “hacer énfasis en que los profesores no pueden ver de manera particular la evaluación y opinión del alumno y en consecuencia pueda hacer la evaluación con total confianza y libertad” (2019_DyS_D3-26).

Los directores y subdirectores estimaron que el estudiante hace la evaluación docente con desinterés: “apatía de los estudiantes por la evaluación” (2019_DyS_D4-108), “los alumnos se aburren con la encuesta” (2019_DyS_D4-109), “los alumnos contestan sin analizar” (2019_DyS_D4-128). Por tanto, suponen conveniente supervisar la realización de la evaluación: “que haya más supervisión en su llenado” (2019_DyS_D4-178), y enfatizar el propósito de la evaluación: “hacer mayor difusión de la importancia de la evaluación docente” (2019_DyS_D4-184).

Respecto a la entrega de los resultados, directores y subdirectores enunciaron la necesidad de contar con los resultados oportunamente: “los resultados de la encuesta tardan mucho en darse a conocer” (2019_DyS_D4-132). Encima, expusieron falta de claridad para interpretar los resultados: “las puntuaciones que se entregan en los resultados de la evaluación son confusas para algunos docentes” (2019_DyS_D4-137).

Categoría analítica Uso de los resultados. En este apartado, fueron mencionadas las implicaciones y acciones emprendidas por los directores y subdirectores con base en los resultados de la evaluación docente. Sobresalió el uso de los resultados para contratación del personal docente: “recontratación de docentes” (2019_DyS_D1-1), destacar buenas prácticas: “reconocimiento a los docentes mejor evaluados” (2019_DyS_D1-3), y detectar inconvenientes “identificación de problemáticas en clase” (2019_DyS_D1-5). Asimismo, se resaltó la importancia de la evaluación para conocer y fortalecer el trabajo docente: “se

detectan las principales áreas de oportunidad y establecen estrategias que permitan mejorar y fortalecer la práctica docente” (2019_DyS_D1-81). En cuanto al SED, se estimó su utilidad, principalmente, para conocer la opinión de los estudiantes: “conocer la opinión de los alumnos” (2019_DyS_D4-70).

La divulgación de los resultados también es asumida por los directores y subdirectores a través de un acercamiento con los docentes: “mediante charlas con el personal docente y en las sesiones de academias” (2019_DyS_D18), “se plática con el docente sobre las áreas en que fue mal evaluado, para buscar como apoyarlo a mejorar” (2019_DyS_D1-6). Así, con el objetivo de mejorar las prácticas, se le hacen sugerencias al docente para favorecer su desempeño: “se le recomienda atender los cursos que se ofertan en UABC” (2019_DyS_D1-7).

Con relación al uso que los docentes hacen a partir del conocimiento de sus resultados, los directores y subdirectores no cuentan con información al respecto, por tanto, suponen poco empleo de los mismos: “no todos los profesores revisan su evaluación y la toman en cuenta” (2019_DyS_D4-119), “no tenemos evidencia de cuantos los utilizan para identificar sus fortalezas y debilidades, algunos no toman en cuenta los resultados” (2019_DyS_D4-54). Pero, se reconoce que parte de los docentes usan los resultados para retroalimentar su labor: “algunos de ellos las contemplan para rediseñar sus estrategias docentes” (2019_DyS_D1-58).

Unidad de análisis: Docentes

Análisis univariado

A partir de la estructura del cuestionario, en la *Dimensión 1. Uso de los resultados de evaluación*, la mayor frecuencia de respuestas se obtuvo en: identificar fortalezas y

Resultados

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 bajo porcentaje que respondió no hacer uso de los puntajes (3.2%) (véase la Tabla 28).

Tabla 28

Puntajes de las respuestas sobre el uso de los resultados de evaluación según Docentes

Ítem	Respuesta	%
¿Qué uso da a los resultados de evaluación docente?	Para identificar fortalezas y debilidades de mi práctica docente	71.7
	Busco alternativas de mejora de mi práctica docente	60.9
	Selecciono cursos del Programa Flexible de Formación y Desarrollo Docente	23.6
	Comento con mis colegas vías de mejora	14.1
	Ninguno	3.2

Respecto a la *Dimensión 2. CECD*, los puntajes más altos se obtuvieron en incluir en el cuestionario elementos como respeto (94.6)³, puntualidad (86.9), y uso de las Tecnologías de la Información y la Comunicación (TIC) en el aula (80.5); en general, consideraron que el CECD incluye los elementos importantes de la práctica docente (71.7 puntos) (véase la Tabla 29). En este sentido, se valoró positivamente el contenido del CECD y se puntualizaron elementos a incorporar.

Tabla 29

Puntajes de las respuestas sobre elementos del CECD según Docentes

Ítem	Puntajes de los valores 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 la puntualidad?	2.5	.7	.9	.9	3.4	1.3	3.3	9.4	14.3	63.2

³ Tal como se indicó en la página 20, los puntajes señalados de las escalas refieren a la suma de los puntajes de los valores 8, 9 y 10. De igual manera serán expresadas la suma de esos puntajes en el resto del documento debido a que fueron los valores más altos.

Resultados

Ítem	Puntajes de los valores 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 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 casi a la par, por un lado, el 57.3% señalaron que no era indispensable, y por otra parte el 42.6% 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%) (véase la Figura 10).

Figura 10

Propuestas de CECD para programas educativos particulares según Docentes



También, el 75.0% consideró pertinente utilizar otras estrategias de evaluación docente junto con el CECD, principalmente una evaluación de directivos (78.8%), evaluación de pares académicos (67.6%) y autoevaluación (59.6%) (véase la Tabla 30). De esta forma, se indicó la pertinencia de incorporar otros instrumentos de evaluación de la docencia.

Tabla 30

Estrategias de evaluación docente adicionales al CECD según Docentes

Ítem	Respuestas	%
¿Considera pertinentes las siguientes estrategias de evaluación?	Evaluación de directivos	78.8
	Evaluación de pares académicos	67.6
	Autoevaluación	59.6

Resultados

Referente a la *Dimensión 3. Procedimiento de aplicación*, con los puntajes más altos, los Docentes estimaron adecuado responder el CECD de forma obligatoria (83.3). Asimismo, los Docentes apuntaron conveniente la difusión para promover que los estudiantes contesten el CECD (76.9 puntos), además de aprobar la manera en que la unidad académica realiza la aplicación del CECD (76.5 puntos) y estimaron oportunas las fechas de aplicación del cuestionario (58.7 puntos). De esta manera, se valoró positivamente el procedimiento de aplicación del CECD (véase la Tabla 31).

Tabla 31

Puntajes de las respuestas sobre el procedimiento de aplicación del CECD según Docentes

Ítem	Puntajes de los valores 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, los Docentes manifestaron que el SED es útil (72.5 puntos) y que el reporte de resultados de la evaluación es claro (67.3 puntos). Y con puntajes menores, los Docentes indicaron disponer de los resultados de la evaluación cuando los requieren (54.2 puntos) y devolución oportuna de los mismos (49 puntaje). De esta forma, si bien de manera general

Resultados

se valoró como útil el SED, se identificaron aspectos centrales a mejorar en favor de la usabilidad (véase la Tabla 32).

Tabla 32

Puntajes de las respuestas sobre la apreciación del SED según Docentes

Ítem	Puntajes de los valores 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 de 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 de forma voluntaria, estos fueron 1,116 comentarios. De forma global, la tasa promedio de respuesta fue de una quinta parte de la muestra (19.4%) (véase la Tabla 33).

Tabla 33

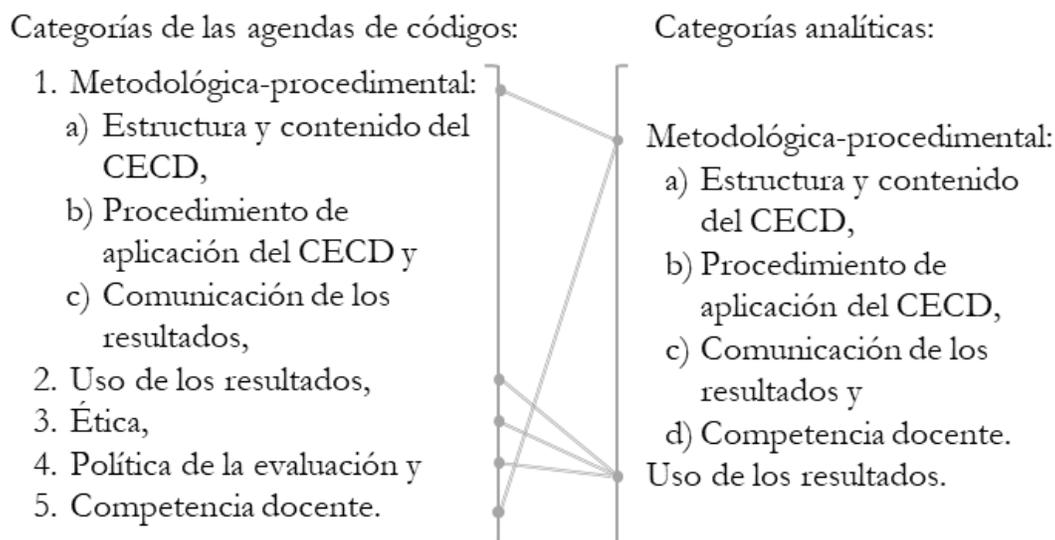
Tasa promedio de respuesta de los ítems de respuesta abierta de Docentes

Dimensión	%
1. Uso de los resultados de evaluación	30.9
2. CECD	15.3
3. Procedimiento de aplicación	13.9
4. Valoración del SED	17.7
Promedio global	19.4

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 la Figura 11).

Figura 11

Relación entre categorías de las agendas de códigos y categorías analíticas



Enseguida, los resultados del ACC se presentan de acuerdo a las categorías analíticas señaladas.

Categoría analítica Metodológica-procedimental

Subcategoría Estructura y contenido del cuestionario. Una de las cuestiones mayormente mencionada por los Docentes fue la extensión del cuestionario, 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_D_D2-87). Asimismo, se mencionó que el cuestionario requería mejorarse: “hacer la redacción de la evaluación más amigable, menos técnica” (2019_D_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_D_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_D_D2-2.9[29]), “Artes Plásticas y Danza” (2019_D_D2-2.9[52]), “Posgrado” (2019_D_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_D_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_D_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_D_D3-86).

Asimismo, los Docentes 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_D_D2-25), “el objetivo se diluye al saber la cantidad de maestros a evaluar” (2019_D_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_D_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_D_D3-13), “continuamente el sistema no funciona adecuadamente y la evaluación no se concluye” (2019_D_D3-103). También, señalaron favorable realizar una mayor difusión para contestar el CECD: “considero que se debe hacer mayor difusión para que el alumno realice la evaluación” (2019_D_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_D_D1-789), “desconozco los resultados de la evaluación docente hasta el momento” (2019_D_D1-1014); seguido de la falta de

Resultados

información para acceder a estos: “desconozco la forma en que puedo acceder a los resultados de mi evaluación docente” (2019_D_D1-472), “no sé en dónde puedo consultar estos resultados” (2019_D_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_D_D1-1045), “los resultados deben estar más a la mano para los interesados ‘docentes’. Todavía hace un mes, no estaba disponibles en la plataforma los resultados de los cursos que impartí” (2019_D_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_D_D1-580).

Categoría analítica Uso de los resultados

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_D_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_D_D1-148).

Además, los Docentes indicaron que la evaluación carece de objetividad: “la evaluación docente resulta ser totalmente subjetiva por parte del alumno” (2019_D_D1-397), “no existe manera de comprobar que tan objetivos son los alumnos al contestar el cuestionario” (2019_D_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_D_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_D_D1-119), “creo que la evaluación es utilizada por algunos alumnos como venganza por la calificación obtenida” (2019_D_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_D_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_D_D1-2856), “no hay seguimiento ni estrategias para que la evaluación sea más integral y coadyuve al mejor desempeño docente” (2019_D_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 resaltaron 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_D_D1-49), “lo único que reviso son los comentarios, esto debido a que la calificación sale relativamente alta siempre” (2019_D_D1-1374).

Análisis de 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. El nivel aceptable de confiabilidad en el ámbito de las ciencias sociales es de .80 en adelante, pero no menos de .66, lo que representa la relación entre la comprensión lograda entre los codificadores y la claridad de las categorías involucradas (Krippendorff, 2004). Lo anterior

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).

Unidad de análisis: Estudiantes

Análisis univariado

En la *Dimensión 1. Uso de los resultados de evaluación*, la mitad de los Estudiantes señalaron que los docentes utilizan los resultados de la evaluación para identificar fortalezas y debilidades (53.5%). Sin embargo, poco más de una cuarta parte indicaron consideración nula de los resultados (27.8%), además, casi una cuarta parte manifestó desconocer el empleo de los resultados (véase la Tabla 34).

Tabla 34

Puntajes de las respuestas sobre el uso de los resultados de evaluación según Estudiantes

Ítem	Respuesta	%
En tu opinión, ¿qué uso dan los maestros 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
	Desconocido	18.6

Resultados

Asimismo, la utilidad de los resultados de la evaluación docente fue valorada con 71.8 puntos. La Tabla 35 muestra los puntajes de cada valor de la escala.

Tabla 35

Puntajes de las respuestas sobre la utilidad de los resultados según Estudiantes

Ítem	Puntajes de los valores respecto a la escala									
	1	2	3	4	5	6	7	8	9	10
Considero útiles los resultados derivados del CECD	1.7	1.0	2.0	1.6	8.49	4.5	9.8	26.0	18.1	27.0

Respecto a la *Dimensión 2. CECD*, los Estudiantes señalaron que el cuestionario incluye elementos importantes de la práctica docente (79.0 puntos); además, estimaron favorable incorporar al cuestionario elementos referentes a la puntualidad (82.5 puntos), cumplimiento del programa (82 puntos) y uso de las TIC (74.9 puntos) (véase la Tabla 36).

Tabla 36

Puntajes de las respuestas sobre elementos del CECD según Estudiantes

Ítem	Puntajes de los valores 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

Resultados

Con relación a la *Dimensión 3. Procedimiento de aplicación*, los Estudiantes indicaron que el ingreso al CECD es sencillo (66.5 puntos). Asimismo, consideraron pertinente responder el CECD de forma obligatoria (64.1 puntos), sumado a una adecuada aplicación del CECD por la unidad académica (64.9 puntos), además de una conveniente difusión para responder el cuestionario (61.8 puntos) conjuntamente con las fechas de aplicación del cuestionario (58 puntos) (véase la Tabla 37).

Tabla 37

Puntajes de las respuestas sobre el procedimiento de aplicación del CECD según Estudiantes

Ítem	Puntajes de los valores respecto a la escala									
	7	2	3	4	5	6	7	8	9	10
¿Te es fácil ingresar 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, los Estudiantes expresaron que los problemas para ingresar al cuestionario se han debido principalmente a dificultades con el sistema gestor del SED (51.7%), inconvenientes con el servicio de internet de la UABC (37.6%), y debido a no recordar la contraseña de acceso al SED (10.5%).

Resultados

Por último, sobre la *Dimensión 4. Valoración del SED*, los Estudiantes marcaron con 70.6 puntos la utilidad del SED (véase la Tabla 38).

Tabla 38

Puntajes de las respuestas sobre la valoración del SED según Estudiantes

Ítem	Puntajes de los valores respecto a la escala									
	7	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

Análisis Cualitativo de Contenido

En concordancia con el ACC de la Unidad de análisis: Docentes, el estudio de los datos se realizó con la información recabada de los ítems de respuesta abierta contestados de forma no obligada. Los datos analizados fueron 2,132, así, la tasa promedio de respuesta de la muestra de Estudiantes fue 5.2%. (véase la Tabla 39).

Tabla 39

Tasas promedio de respuesta de los ítems de respuesta abierta de Estudiantes

Dimensión	%
1. Uso de los resultados de evaluación	10.1
2. CECD	3.6
3. Procedimiento de aplicación	3.0
4. Valoración del SED	4.0
Promedio global	5.2

Igualmente, se retomaron las dos categorías analíticas: Metodológica-procedimental y Uso de los resultados, que concentraron las categorías de las agendas de códigos señaladas en la Figura 11.

Categoría analítica Metodológica-Procedimental

Subcategoría Estructura y contenido del cuestionario. Un tema sobresaliente por los Estudiantes fue la extensión del CECD referente al exceso en el número de preguntas que lo componen, por tanto, el propósito del cuestionario puede verse entorpecido: “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” (2019_E_D4-817). Además, algunos resaltaron que el cuestionario se compone por ítems 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 médica, las cuales son nuestro campo práctico” (2019_E_D2-2).

Por otro lado, los Estudiantes también expresaron la necesidad de introducir al CECD aspectos relacionados a la ética en la evaluación de la práctica docente: “también deberían de evaluar la ética profesional, responsabilidad, disposición, asistencia ya que una cosa es que firmen las listas de asistencia y otra que en realidad se presenten a dar clases” (2019_E_D2-360).

Subcategoría Procedimiento de aplicación. Respecto al periodo de aplicación, los Estudiantes expresaron que resulta inconveniente hacer la evaluación docente al final del semestre debido a la carga de compromisos académicos por culminar: “asimismo 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_E_D4-29). Por otra parte, un comentario recurrente señalado por los Estudiantes fue el tiempo limitado para responder el CECD:

El formato de uso de la evaluación docente tiene una deficiencia en cuanto al tiempo para contestarla, ya que cuando quieres agregar comentarios extras, al momento de querer guardar los resultados te recarga la página y tienes que contestar todas las casillas de nuevo. (2019E_E_D1-419)

A la par, los Estudiantes destacaron dificultades de acceso al sistema gestor del SED:

“nunca logre acceso a la página, marcaba error de página o error de usuario” (2019_E_D3-29). En esta línea, algunos Estudiantes estimaron que la dificultad para el acceso puede relacionarse a la saturación en el acceso al sistema gestor del SED:

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, siempre he realizado mis evaluaciones desde la fecha en que comienzan a anunciarlo y nunca he tenido problema alguno para contestarlas. (2019_E_D3-300)

Los Estudiantes 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 en 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_E_D1-322). En específico, consideran 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_E_D3-52). Por tanto, se propone que la evaluación sea de carácter opcional: “sería bueno que el cuestionario fuera opcional o que podamos solo comentar en que materia hubo algún problema ya sea con la clase o el maestro” (2019_E_D3-222).

En otro asunto, los Estudiantes hicieron referencia a la falta de una interfaz amigable, útil y funcional del sistema gestor del SED: “la interfaz del cuestionario de evaluación docente es muy molesta, difícil de llenar y en ocasiones se reinicia” (2019_E_D2-328); de igual forma, añadieron conveniente su manejo en dispositivos móviles:

creo que debería hacerse más óptimo para dispositivos móviles porque es un modo más rápido de hacer la encuesta ya que no todos tenemos siempre a la mano la computadora y es más sencillo y rápido hacerlo desde el teléfono celular.

(2019_E_D3-130)

Otra cuestión reconocida por los Estudiantes fue la realización de la evaluación de forma irreflexiva, solo por cumplir: “un problema muy grande es que los estudiantes responden de manera deshonesto para que sea más rápido terminar la evaluación docente” (2019_E_D1-260); asimismo, se reconoció que 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_E_D1-270). 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_E_D1-819).

Subcategoría Comunicación de los resultados. En este espacio, los Estudiantes comunicaron su interés sobre los resultados de la evaluación, con la finalidad de conocer el desempeño de sus docentes: “no lo sé. Supongo que, estaría bien tener el historial del

maestro a disposición. Así como los resultados de la evaluación docente, para poder seleccionar bien los profesores y las materias que deseo cursar” (2019_E_D1-1002).

Adicionalmente, los Estudiantes expresaron 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_E_D1-728).

Un aspecto que no se relaciona directamente con esta subcategoría, pero se estimó relevante en este espacio, es sobre los aspectos relacionados con el dominio disciplinar, pedagógico y profesional que algunos estudiantes señalan de sus docentes, por un lado, algunos Estudiantes expresaron conformidad en relación a sus docentes: “todos son buenos maestros” (2019_E_D1-226), en cambio, también se manifestaron 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_E_D2-80).

Categoría analítica Uso de los Resultados

Una estimación destacada por la mayoría de los Estudiantes fue que, los resultados de la evaluación no son tomados en cuenta, puesto que perciben poca o nula acción de cambio del docente para mejorar su práctica: “siento que no se toma 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_E_D1-89).

Adicionalmente, algunos Estudiantes valoran la oportunidad de expresarse: “agradezco que se nos tome en cuenta nuestras opiniones como alumnos” (2019_E_D1-731), pero solicitan que su opinión sea considerada por los tomadores de decisiones para

mejorar el desempeño docente: “el concepto de la evaluación es muy buena, pero hace falta que escuchen los resultados y las opiniones de los alumnos respecto a los docentes”

(2019_E_D1-182).

Análisis de confiabilidad

Semejante al análisis efectuado en la Unidad de análisis: Docentes, el proceso se realizó en dos ocasiones. En el primer análisis el coeficiente de confiabilidad fue de .35. A raíz del ejercicio, el c2 realizó comentarios sobre revisar la pertinencia y claridad de las categorías de las agendas, así como mayor especificidad en la definición de los códigos y correspondencia directa entre la regla de codificación y su ejemplo. Tales comentarios se concentraron principalmente para las categorías Ética y Política de la evaluación. De esta forma, fueron atendidas las observaciones con el objetivo de favorecer la claridad y comprensión de las agendas de códigos.

En el segundo análisis de confiabilidad se obtuvo un coeficiente de .73. No hay una respuesta establecida del nivel definido de confiabilidad en el campo de las ciencias sociales, sin embargo, los coeficientes no menores de .66 se han estimado aceptables (Krippendorff, 2004). Aquí, fue puntualizada la complejidad del estudio de los datos dada su naturaleza diversa y extensa; de forma específica se invitó atender la precisión de los códigos: Los resultados no se toman en cuenta (IRRE), El SED no es útil (SNOU) y Sugerencias de mejora de la evaluación (SMEV).

Síntesis de los resultados: acuerdos y discrepancias

La aproximación metodológica de este trabajo permitió recuperar las diferentes opiniones entre los participantes: Directivos (2 coordinadores, 19 directores y 21 subdirectores), Docentes (1,431) y Estudiantes (10,160), lo que denotó una diversidad de opinión respecto a un mismo sistema de evaluación. El propósito de este apartado fue elaborar una síntesis que represente los principales acuerdos y discrepancias señalados en los resultados. De esta manera, se presentan los consensos —destacados a través del análisis univariado—, seguidos de los aspectos opuestos —vertidos en el ACC que dio voz a los participantes—. Es preciso subrayar que, en el caso de los comentarios emitidos por los Docentes la tasa promedio de respuesta fue de 19.4% (véanse las Tablas 19 y 33) y de Estudiantes fue de 5.2% (véase las Tablas 22 y 39), esto representó el corpus de estudio del ACC de Docentes y Estudiantes. En relación con los Directivos, todos realizaron comentarios.

Referente a la categoría analítica *Metodológica-procedimental*, en el tema de la estructura y contenido del CECD, Docentes y Estudiantes mencionaron que el cuestionario integra los elementos importantes de la práctica docente (71.7 puntos y 79 puntos). Asimismo, los Directivos destacaron el CECD como un instrumento de medición de calidad apegado al modelo educativo de la universidad.

En cuanto a las discrepancias respecto al CECD, Directivos, Docentes y Estudiantes comentaron que el cuestionario es extenso en el número de ítems que integra. Además, los participantes sugirieron revisar la pertinencia de los ítems. Con relación a lo anterior, Directivos, Docentes y Estudiantes concordaron en incluir en el CECD aspectos como puntualidad (Docentes 86.9 puntos y Estudiantes 82.5 puntos) y uso de las TIC en el aula (Docentes 80.5 puntos y Estudiantes 82 puntos). También, Directivos y Estudiantes

propusieron incorporar elementos del código de ética de la universidad. A su vez, los Estudiantes indicaron añadir valorar el cumplimiento del programa (82 puntos).

Sobre el procedimiento de aplicación del CECD, la difusión para responder el cuestionario fue encontrada adecuada por Directivos (75%), Docentes (76.9 puntos) y Estudiantes (61.8 puntos). Correspondiente a la estrategia de aplicación del CECD, los participantes indicaron que era apropiada (Directivos 75.8%, Docentes 76.5 puntos y Estudiantes 64.9 puntos), al igual que las fechas de aplicación (Directivos 51.6%, Docentes 58.7 puntos y Estudiantes 58 puntos). También, los Estudiantes señalaron que es sencillo ingresar al SED (66.5 puntos), y el 40% especificó que accede en el primer intento. Además, fue apoyada la obligatoriedad de la evaluación (Directivos, Docentes 83.3 puntos y Estudiantes 64.1 puntos).

En contraste, se argumentó en oposición de la evaluación docente como una tarea obligatoria, debido a que Docentes (una proporción de la tasa promedio de 19.4%⁴) y Estudiantes (una proporción de la tasa promedio de 5,2%) señalaron que dicha condición pudiera propiciar concebir a la evaluación como un requisito condicionado a la reinscripción. Adicionalmente, los Estudiantes mencionaron de mayor conveniencia una evaluación opcional.

Por otro lado, comentarios de Directivos y Docentes expresaron desfavorable el número de cuestionarios que el estudiante debe responder, debido al tiempo que se invierte

⁴ Como se indicó al inicio de la síntesis, las posturas discrepantes señaladas por los participantes se alojaron en los comentarios, en el caso de los Directivos hubo aportaciones de todos, pero de Docentes y Estudiantes las observaciones representaron una porción de la muestra (tasa promedio de Docentes: 19.4% y de Estudiantes: 5.4%). Por tanto, es vital tomar en cuenta esta precisión.

para esta labor. Por ello, los Directivos sugirieron una evaluación a partir de muestras aleatorias.

En atención a la comunicación de los resultados, los Docentes indicaron disponer de los resultados cuando los requieren (54.2 puntos), así, señalaron que la devolución es oportuna (49 puntos). También, fue destacada la utilidad del SED (Directivos 72.1 puntos, Docentes 72.5 puntos y Estudiantes 70.6 puntos). En específico, los Directivos enfatizaron favorable que el SED es virtual y sistemático.

Inversamente, fue comentada por Directivos y Docentes la entrega de resultados fuera de tiempo. También, Directivos y Docentes mencionaron escasa información para acceder al sistema gestor del SED y conocer los resultados de la evaluación docente, además de una falta de claridad para interpretar los resultados. Asimismo, Directivos y Docentes estimaron una carencia de información con la que cuentan los estudiantes sobre el objetivo y valor de la evaluación docente. Este tema se identificó como parte de los sesgos de la evaluación.

Referente al sistema gestor del SED, se presentaron comentarios en función de fallas o dificultades del sistema expresadas por Docentes y Estudiantes (51.7%). Los Estudiantes consideraron como posible causa de fallas del SED la saturación debida a los intentos de ingreso al término del periodo de aplicación del CECD.

Por otra parte, Directivos (33.3%) y Docentes (42.6%) argumentaron sobre la conveniencia de incorporar otros instrumentos al SED. Al respecto, sobresalió el área de Ciencias de la Salud, donde se estimó apropiado un CECD específico (Directivos 58.3% y Docentes 26.7%). Además, Directivos (89.1%) y Docentes (75%) señalaron la importancia

de incorporar otras estrategias de evaluación, tales como: autoevaluación (Directivos, Docentes 42.7%) y evaluación de pares académicos (Directivos, Docentes 41%).

Adicionalmente, en relación a la competencia docente, los Estudiantes expresaron conformidad con la práctica de sus docentes. No obstante, algunos Docentes y Estudiantes manifestaron carencia en algunos docentes sobre el dominio pedagógico del contenido disciplinar y limitaciones didácticas.

Respecto a la categoría analítica *Uso de los resultados*, de manera general se destacó el empleo de los resultados de la evaluación para la identificación de fortalezas y debilidades (Directivos 62.9%, Docentes 71.7% y Estudiantes 53.5%). Adicionalmente, los Docentes indicaron emplear la evaluación para retroalimentar su práctica y seleccionar cursos del Programa Flexible de Formación y Desarrollo Docente (PFFDD). También, Directivos y Docentes aludieron a la relevancia de la opinión de los estudiantes, específicamente manifestada en los comentarios. De la misma manera, los Estudiantes valoraron la oportunidad de expresarse.

Con relación a observaciones críticas, estas se asociaron a los sesgos de la evaluación: a) la evaluación como una tarea irreflexiva, b) la evaluación con fines punitivos y c) desconocimiento o ausencia del uso de los resultados de la evaluación. De esta forma, Directivos, Docentes y Estudiantes mencionaron que responder el CECD es una acción realizada de manera irreflexiva, alejada del propósito de la evaluación. Por su parte, los Estudiantes reconocieron la posible realización de la evaluación con base en el agrado hacia el docente implicado.

Tocante a la evaluación con fines punitivos, Directivos, Docentes y Estudiantes coincidieron en señalar el uso de la evaluación como venganza por parte de los estudiantes

al responder el CECD. Al respecto, los Docentes también añadieron falta de objetividad de los estudiantes al llevar a cabo dicha tarea.

En cuanto al desconocimiento o falta de empleo de los resultados de la evaluación, el 3.2% de Docentes y el 27.8% de Estudiantes señalaron que no existe un uso de los resultados, y el 18.9% de Estudiantes indicó desconocer el uso de estos. Además, Directivos comentaron desconocer el alcance del uso de los resultados por los docentes. Se suma la percepción de Directivos respecto a una baja credibilidad de los docentes y estudiantes hacia el SED, posiblemente procedente de la falta de acciones derivadas de la evaluación docente. Por tanto, los Directivos consideraron que puede presentarse una postura de indiferencia por parte de los estudiantes hacia la evaluación. También, los Docentes estimaron que hay una carencia de información de los resultados de la evaluación y falta de seguimiento de los mismos.

A manera de sugerencia, Directivos, Docentes y Estudiantes coincidieron en enfatizar y difundir el propósito e implicaciones de la evaluación docente. Asimismo, los Directivos plantearon consolidar un mecanismo que asocie los resultados de la evaluación con estrategias de formación, así como, reconsiderar el vínculo entre la evaluación y el Programa de estímulos al desempeño del personal docente.

Discusión y conclusiones

El propósito general de esta investigación fue realizar una metaevaluación al SED con el fin de identificar las fortalezas y los aspectos a mejorar en: el CECD y su administración, el sistema gestor del SED, el uso de los resultados y la infraestructura institucional. Para abordar este trabajo se optó por un diseño de estudio de caso múltiple integrado que permitió la aproximación a un objeto de estudio complejo. Este diseño, es una variante de los estudios de casos, posibilitó atender casos particulares que forman parte de una misma situación y que recibieron un mismo tratamiento empírico (Yin, 2018). De esta forma, fue posible dar voz a los principales actores involucrados en la evaluación docente de la UABC.

Los resultados de la investigación se discutieron en función de su relación con estudios similares y referentes teóricos. De esta manera, la discusión se estructuró con base en tres categorías analíticas: *Metodológica-procedimental*, *Uso de los resultados* y *Ética*. La categoría *Metodológica-Procedimental* incluyó la estructura y contenido del CECD, su procedimiento de aplicación, la comunicación de los resultados y la competencia docente. La categoría *Uso de los resultados* correspondió al empleo de los resultados de la evaluación docente y política de la evaluación. La categoría *Ética* develó temas críticos que ponen en duda la confianza entre los actores educativos vital en un proceso de evaluación. También, al término de este capítulo se incorporaron las conclusiones y propuestas para investigaciones futuras.

Categoría Metodológica-procedimental

Los Docentes (71.7 puntos) y Estudiantes (79 puntos) indicaron que el CECD se conforma de los elementos sustantivos para evaluar la docencia, pero resaltaron que el cuestionario es extenso. Este hallazgo concuerda con otro estudio de metaevaluación, donde los estudiantes (64.8%) señalaron adecuado el cuestionario para evaluar a los docentes de forma integral (Elizalde et al. 2008).

Asimismo, Directivos, Docentes (86.9 puntos) y Estudiantes (82.5 puntos) coincidieron en incluir elementos como puntualidad, uso de las TIC en el aula (Directivos, Docentes 80.5 puntos y Estudiantes 82 puntos) y cumplimiento del programa (Estudiantes 82 puntos). Al respecto, es preciso atender de forma cuidadosa los elementos que se consideran evaluar de la práctica docente, estos requieren una interacción continua entre la teoría, la investigación y la práctica (Marsh, 2007).

Con relación a lo anterior, Feldman (1976b, 2007) destacó dos dimensiones mayormente asociadas a la práctica docente en educación superior. La primera dimensión integra un conjunto de actitudes, comportamientos y prácticas del docente en su tarea de participante y comunicador: estimulación del interés por el curso, enseñanza clara y comprensible, dominio del conocimiento disciplinar, preparación y organización del curso y entusiasmo del docente. La segunda dimensión comprende actividades reguladoras del docente como gestor: manejo del aula, equidad e imparcialidad de la evaluación, claridad de los objetivos y requisitos del curso, dificultad del curso —carga de trabajo— y retroalimentación a los estudiantes. Por tanto, el autor destaca que, el reconocimiento de las dos dimensiones antes señaladas, principalmente la primera, en la evaluación docente debe resultar alentador para los docentes que defienden la significancia de la opinión de los

estudiantes sobre su práctica. Así pues, en opinión de Directivos, Docentes y Estudiantes, es posible que se identifiquen otras características de tipo administrativo que formen parte de la tarea docente, pero, esas serán menores en importancia para la evaluación de la práctica docente. Dichas tareas pudieran valorarse por medio de otros mecanismos regulatorios.

En alusión a la aplicación del CECD, los participantes señalaron apropiada la difusión (Directivos 75%, Docentes 76.9 puntos y Estudiantes 61.8 puntos), la estrategia de aplicación (Directivos 75.8%, Docentes 76.5 puntos y Estudiantes 64.9 puntos) y el periodo de aplicación (Directivos 51.6%, Docentes 58.7 puntos y Estudiantes 58 puntos). Al respecto, en la UAEH existe una opinión diversa por algunos actores educativos (integrantes del Comité Institucional de Evaluación Docente, docentes y estudiantes) sobre el periodo de aplicación, por un lado, se considera un acierto la duración de 10 semanas en cada semestre; pero, por otra parte, dicho periodo puede resultar extenso debido a que se pierde el propósito de la difusión al mantener un periodo tan amplio (Elizalde et al., 2008).

En otro tema, Directivos, Docentes (83.3 puntos) y Estudiantes (64.1 puntos) asintieron conveniente realizar la evaluación de manera obligatoria. Sin embargo, comentarios de Docentes y Estudiantes mostraron oposición a tal obligatoriedad, asimismo, Directivos y Docentes puntualizaron desfavorable el número de cuestionarios a responder y tiempo invertido por los estudiantes en esta labor. Al respecto, comúnmente las instituciones condicionan la evaluación docente como un requisito para los estudiantes, lo que puede considerarse una carga debido a la cantidad de cursos por evaluar en cada periodo (Nasser-Abu, 2017a).

En otro asunto, hubo un acuerdo entre Directivos (89.1%) y Docentes (75%) en relación a la pertinencia de incorporar otras estrategias de evaluación como autoevaluación (Directivos, Docentes 42.7%) y evaluación por pares académicos (Directivos, Docentes 41%). De esta forma, se considera relevante diversificar las fuentes de información y fortalecer el SED, lo cual alude a la discusión referente a las limitaciones de los cuestionarios de apreciación estudiantil (Feldman, 1977; Hornstein, 2017), y la importancia de utilizar diversos instrumentos de evaluación para posibilitar una mejor comprensión de la actividad docente (Nasser-Abu, 2017b; Rueda, 2021; Spooren et al., 2013).

En relación con lo anterior, considerar a los cuestionarios de evaluación en opinión de los estudiantes como única herramienta para evaluar la docencia conduce a un ambiente de injusticia por parte de los docentes, también genera interrogantes sobre la validez y confiabilidad de los juicios emitidos por los estudiantes, así como, de la equidad de la recompensa y reconocimiento derivado de los resultados de la evaluación (Cook et al., 2021).

Respecto a otros instrumentos al SED, Directivos (33.3%) y Docentes (42.6%) apuntaron apropiado diseñar cuestionarios específicos en atención a las particularidades de la práctica docente, concretamente para evaluar a los docentes de las áreas de la salud (Directivos 58.3% y Docentes 26.7%). En este tema, particularmente, los componentes asociados a la evaluación docente en la carrera de medicina son considerados un elemento crítico por su relevancia social y responsabilidad de los estudiantes en la salud de la población (Flores-Hernández et al., 2017; Villavicencio-Martínez y Luna, 2018). Asimismo, con resultados semejantes, en el estudio de metaevaluación de la UAEH, los

docentes (92.7%) indicaron pertinente contar con cuestionarios específicos acordes con las particularidades de cada programa educativo (Elizalde et al., 2008).

En cuanto a la comunicación de los resultados, los Docentes expresaron contar con los resultados cuando los necesitan (54.2 puntos), de esta forma la devolución se valoró oportuna (49 puntos). Al respecto, los hallazgos en una metaevaluación de la evaluación docente en posgrado de la Universidad de Santo Tomas (Colombia) destacó casi nula la socialización de los resultados (Patiño-Montero et al., 2022). En cambio, la UAEH devuelve los resultados a los docentes de forma impresa, lo cual se valoró importante por los docentes (87.2%) (Elizalde et al., 2008).

Asimismo, Directivos y Docentes solicitaron que los reportes de resultados sean comprensibles y detallados, de tal forma que el profesorado pueda identificar sus fortalezas y aspectos a modificar. Con relación a lo anterior, en la UAEH, en algunos casos, la información con la que se cuenta es suficiente para explicar las dimensiones que conforman el cuestionario de evaluación docente (Elizalde et al., 2008). En este sentido, la credibilidad y utilidad de los resultados de la evaluación descansa en una entrega comprensible de los hallazgos, que describa sus alcances y restricciones (Jornet et al, 2020). De esta forma, una buena evaluación requiere proporcionar resultados claros, aunado a oportunidades de uso productivo de los resultados, incluso si el desempeño fuera deficiente (Theall, 2017).

Referente al SED, se enfatizó su utilidad (Directivos 72.1 puntos, Docentes 72.5 puntos y Estudiantes 70.6 puntos). En específico, los Directivos subrayaron favorable que el SED sea virtual y sistemático. Pero, de forma particular, Docentes y Estudiantes (51.7%) mencionaron fallas o dificultades en el manejo del sistema gestor del SED. Estos hallazgos concuerdan con los obtenidos en la metaevaluación de la UAEH, donde se considera un

acierto la aplicación electrónica de los cuestionarios, sin embargo, se resaltó la constante lentitud del sistema (Elizalde et al., 2008). Se agrega que, un buen sistema de evaluación debe ser capaz de brindar a los tomadores de decisiones fundamentos sólidos para el aprovechamiento de los datos disponibles, conjuntamente, el sistema requiere actualizaciones periódicas y revalidaciones estadísticas que sustenten las normas o cambios en los criterios de operación (Theall, 2017). Además, en relación a la validez y confiabilidad de los puntajes conviene enfatizar que las debilidades de los sistemas de evaluación poco se derivan de dichos elementos (Theall y Franklin, 1990).

Se advirtió también, por parte de algunos Docentes y Estudiantes, un déficit en ciertos docentes en relación al dominio pedagógico del contenido disciplinar y estrategias didácticas. En este tema, la literatura señala que los estudiantes, a partir de la percepción de su papel como evaluador de docentes, pueden discrepar en su concepción de competencia docente (Dunegan y Hrivnak, 2003; Nasser-Abu, 2017b; Spooren y Christiaens, 2017) y encontrar mayor acuerdo para determinar características asociadas a una enseñanza deficiente, así como presentar diversidad de opiniones al momento de caracterizar una práctica docente sobresaliente (Dunegan y Hrivnak, 2003).

Categoría Uso de los resultados

Antes de continuar, cabe resaltar que en la UABC la evaluación de la docencia en licenciatura solo se realiza por medio de la opinión de los estudiantes, y se cuenta con más de un cuestionario para esta tarea —teoría, práctica y para la modalidad en línea—. El SED es el sistema encargado de administrar y concentrar los resultados de la evaluación por medio del sistema gestor en línea. Bajo este contexto, se reconoce un uso formativo y

sumativo de los resultados de la evaluación docente por los diferentes actores de la universidad —Directivos, Docentes y Estudiantes—, pero sobresale que, las posibilidades de una evaluación formativa son escasas. A tal efecto, es preciso reiterar el valor del 13.3% otorgado a los puntajes de los resultados de la evaluación docente en el programa de compensación salarial de la UABC (UABC, 2021a), esta situación condiciona la obligatoriedad de la evaluación docente debido al peso de las consecuencias de alto impacto. Al mismo tiempo, existen implicaciones en relación a la toma de decisiones para la contratación de docentes por horas (UABC, 2021d, sección “Evaluación Docente”). De forma particular, los Directivos indicaron emplear los resultados para la asignación de unidades de aprendizaje y recontractación de docentes. Por lo tanto, el peso asignado a los puntajes dependerá de los intereses o enfoque institucional (Uttl et al., 2017). Sin embargo, es importante considerar que tales efectos pueden generar tensión para los docentes (Theall, 2017). Dicho lo anterior, se reconoce esencial el uso adecuado y coherente de los resultados de la evaluación (McKeachie, 1997; Theall y Franklin, 2001) para tomar decisiones respecto al personal docente —revisión posterior a la permanencia y para la contratación; aumentos por mérito, permanencia y promoción— (Linse, 2017).

En alusión a una evaluación con propósito formativo, Directivos (62.9%), Docentes (71.7%) y Estudiantes (53.5%) señalaron el empleo de los resultados para identificar fortalezas y debilidades de la práctica docente. También, en la UAEH los docentes (83.2%) indicaron utilizar los resultados para reconocer fortalezas y áreas de oportunidad de su práctica (Elizalde et al., 2008). Esta apreciación coincide con la discusión de la literatura en torno a que, la evaluación es utilizada por los docentes como fuente de información relacionadas a las fortalezas y debilidades de su práctica (Marsh, 2007, Nasser-Abu,

2017b). Además, Directivos y Docentes resaltaron el valor de los comentarios que los estudiantes expresan en el CECD. Lo anterior alude, según Nasser-Abu y Fresko (2009), a una preferencia de docentes por las observaciones escritas frente a resúmenes estadísticos, debido a que perciben un mayor provecho para su práctica. Por tanto, las autoras sugieren tomar medidas que propicien un aumento en el número de estudiantes que manifiestan su opinión con comentarios escritos como la calidad de los mismos. 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).

Con relación al seguimiento de los resultados, los Docentes expresaron falta de seguimiento y acciones que propicien la mejora. En este escenario cabe destacar la necesidad de vincular los puntajes de la evaluación con estrategias institucionales de mejoramiento profesional, ya que, solamente entregar los reportes al profesorado resulta insuficiente para incidir en el fortalecimiento de la práctica docente (Luna y Torquemada, 2008). Asimismo, el uso formativo y sumativo de la evaluación docente implica un equilibrio complejo y una variedad de consideraciones metodológicas y políticas (Martínez et al., 2016).

Respecto a lo anterior, Theall (2017) destaca el riesgo de considerar a la evaluación docente separada de un acompañamiento, dado que, en ausencia de un seguimiento docente formal la evaluación se vuelve punitiva y, por el contrario, el seguimiento sin evaluación es una suposición. Asimismo, el autor identifica que, cuando el efecto de la evaluación es únicamente una revisión negativa, el desempeño, motivación y voluntad del docente corren el riesgo de reducirse, junto con una afectación a la reputación y carrera del docente,

inclusive la deserción de los estudiantes puede incrementar. Por tanto, para contrarrestar tal escenario, Theall (2017) propone un acompañamiento que recopile y use de forma pertinente los resultados de la evaluación, incorporar actividades de desarrollo, trabajo colegiado, poner a prueba nuevos métodos de enseñanza y supervisar los avances.

Por otra parte, fue expuesto el desconocimiento o falta de uso de los resultados de la evaluación, Docentes (3.2%) y Estudiantes (27.8%) señalaron que los resultados no son tomados en cuenta, y los Estudiantes (18.9%) también indicaron ignorar su empleo. En este tema, los estudiantes (61.2%) de la UAEH mencionaron haber identificado mejoría en el desempeño de sus docentes producto de la evaluación; pero, si bien se reconoce el uso de los resultados para la toma de decisiones, también fue manifestado que los resultados no se han constituido a nivel institucional para la resolución y determinación de acciones —reategorización, contratación, promoción y estímulos— (Elizalde et al., 2008).

Adicionalmente, en un estudio reciente de metaevaluación en la UAEH, docentes expresaron, de forma jerárquica, como principal problemática que los resultados de la evaluación no son tomados en cuenta, y en el caso de los estudiantes revelaron dicha problemática en cuarto nivel (Elizalde et al., 2017). En este contexto, la UAEH, docentes (15.2%) subrayaron como un aspecto a mejorar el uso personal de los resultados, mientras que los estudiantes (18.4%) enfatizaron favorecer un uso institucional de los resultados (Elizalde et al., 2017). En alusión a otro referente, en la Universidad Santo Tomas se precisó que existe un alto porcentaje de docentes que no consultan sus resultados de la evaluación, y en otros casos se consulta como requisito para participar en convocatorias de promoción docente (Patiño et al., 2022).

Con base en la anterior situación, cabe retomar los planteamientos de Spooren y Christiaens (2017) sobre la opinión de estudiantes en relación al uso de los resultados de la evaluación docente, los autores indican que los estudiantes valoran y, en general, están dispuestos en participar en las evaluaciones de la docencia, sin embargo, la mayoría aprecia que sus docentes toman en cuenta de manera limitada sus comentarios. Por tanto, mientras no exista evidencia de cambio percibida por los estudiantes, fundamentada en sus aportes para mejorar la docencia, la evaluación de la docencia será vista como una tarea que no merece un esfuerzo consciente (Dunegan y Hrivnak, 2003).

Igualmente, los Directivos expresaron desconocimiento respecto al alcance del uso de los resultados por los docentes. En esta línea, Rueda (2019) alude a perspectivas distintas entre directivos y docentes, por un lado, los docentes consideran que los cursos de formación continua no se difunden bajo un trayecto regular, además de la falta de reconocimiento a directivos como autoridad académica por desconfianza. También, el autor añade, los directivos estiman que los docentes hacen caso omiso al ofrecimiento para participar en dichos cursos, e identifican una falta de comunicación con los docentes para que estos expresen sus dudas o comentarios al respecto. Por consiguiente, el uso conveniente de los resultados es crucial para generar un ambiente de enseñanza de alta calidad; por el contrario, el uso inadecuado produce desconfianza, desigualdades e incongruencias que desalientan a la institución (Linse, 2017).

En el marco de los planteamientos antes referidos, Directivos y Docentes priorizaron brindar información a los estudiantes acerca del propósito de la evaluación con el fin de sensibilizarlos sobre su importancia. De esta misma forma, en la UAEH los miembros del Comité Institucional de Evaluación Docente expresaron que hay mucho por

hacer en la labor informativa para promover una cultura de la evaluación docente (Elizalde et al., 2008). Así, los estudiantes que están mayormente informados y comprometidos con la evaluación de la práctica docente, realizarán esta tarea de manera favorable, pues se concebirán como un actor importante en la comunidad académica (Spooren y Christiaens, 2017).

Categoría Ética

En cuanto a temas críticos correspondientes a los sesgos de la evaluación, se identificó la evaluación como labor irreflexiva. En este punto, conviene retomar las palabras de Marsh (2007) al indicar la naturaleza multidimensional de los resultados de la evaluación docente por los estudiantes, además de ser confiables y estables, relativamente válidos, poco afectados de sesgos y útiles para la institución en su conjunto.

Específicamente, Directivos, Docentes y Estudiantes refirieron la realización de la evaluación docente de forma irreflexiva por los estudiantes. Además, los Estudiantes comentaron la posible influencia del agrado hacia el docente al momento de evaluar su práctica. Esta es una discusión recurrente en torno a los cuestionarios de apreciación estudiantil, al respecto, el afecto de estudiantes hacia los docentes parece ocupar un sesgo considerable en las evaluaciones docentes a diferencia del interés en los contenidos de la materia, por tanto, si se parte de que el agrado no tiene un vínculo conceptual con la calidad de la enseñanza, la validez puede verse comprometida en las evaluaciones por los estudiantes (Feistauera y Richter, 2018). También, es importante indicar que, los estudiantes frecuentemente desisten reflexionar de manera seria en el cumplimiento de los cuestionarios de evaluación, derivado del compromiso de realizar esta tarea para cada

curso, lo que implica ignorar las consecuencias inmediatas imperceptibles por los estudiantes de los resultados de la evaluación para la institución (Nasser-Abu, 2017a). En consecuencia, los resultados de la evaluación en la opinión de estudiantes deben usarse e interpretarse con cautela (Feistauera y Richter, 2018).

Otro tema expuesto fue el uso de la evaluación docente como venganza por los estudiantes, en voz de Directivos, Docentes y Estudiantes se destacó dicha condición comúnmente relacionada a la baja calificación de los estudiantes. En esta línea, se destaca que las principales preocupaciones de los estudiantes se alejan de una postura frívola, por el contrario, realizan la evaluación docente con base en una valoración del aprendizaje (Nasser-Abu y Fresko, 2009). Por tanto, aunque los docentes han asumido una relación entre las bajas calificaciones de sus estudiantes con la reducción de los puntajes en su evaluación, Feldman (1976a, 2007), señala que tal relación es poca o nula, pero, no se puede negar la presencia de un posible sesgo en este tema. Asimismo, el resultado de un proceso de evaluación docente en la opinión de los estudiantes justo está íntimamente vinculado con un trabajo docente imparcial y transparente (Tripp et al. 2018).

Conclusiones

Este estudio de metaevaluación demostró ser una estrategia valiosa para identificar fortalezas, cuestionamientos y propuestas de los Directivos, Docentes y Estudiantes 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 de las observaciones del objeto evaluado (Stake, 2011). En consecuencia, los resultados de este trabajo de metaevaluación derivaron en recomendaciones puntuales orientadas hacia la mejora del SED.

Adicionalmente, esta metaevaluación aportó información valiosa para orientar el diseño de mecanismos que ayuden a una mayor usabilidad de los resultados de evaluación. 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 educativa, con el fin de ajustar los componentes del sistema a la realidad institucional para la mejora de los procesos de enseñanza y aprendizaje.

En México, la evaluación se ha planteado como un recurso fundamental para mejorar el sistema educativo, al mismo tiempo, las políticas de evaluación de la docencia en educación superior han permitido la implementación de programas asociados a estímulos económicos. Entre tanto, 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).

La evaluación docente con base en la opinión de los estudiantes se realiza prácticamente en todas las IES del país (Rueda et al., 2016). Así, los resultados obtenidos pueden ser útiles para fines administrativos y de rendición de cuentas, pero insuficientes como único sustento para informar sobre la enseñanza (Nasser-Abu, 2017a). De esta manera, a pesar del amplio consenso sobre la relevancia de los cuestionarios para estudiantes, la evaluación docente se ha construido con tensión y conflictos durante años entre los partidarios —mayormente investigadores— de estos cuestionarios y los destinatarios de las evaluaciones —docentes— (Hammer, 2018). Por tanto, cabe recalcar

que la utilidad de la evaluación no descansa únicamente en contar con puntajes válidos y confiables, sino también en una política educativa clara, en la certidumbre y confianza de los procesos de recopilación de datos, en una comunicación de los resultados y en como estos pueden ser aprovechados (Theall y Franklin, 1990). Adicionalmente, es importante también incorporar otros instrumentos de evaluación que abonen al conocimiento de la práctica docente en la institución.

Dicho lo anterior, un sistema de evaluación docente que carece de un interés por atender a docentes, estudiantes y directivos hace un uso desfavorable del sistema y entorpece su propósito (Theall y Franklin, 1990, 2001). Estos sistemas requieren profesionales capacitados, directivos, docentes y estudiantes comprometidos, componentes técnicos y recursos continuos (Neumann, 2001, Theall y Franklin, 2001). Por tanto, se reconoce la complejidad y dinamismo de estos sistemas, así como el compromiso necesario de planificación a largo plazo, junto con la responsabilidad de las IES de trabajar los sistemas de manera adecuada y justa (Neumann, 2001).

Se agrega que, la falta de articulación entre el desarrollo profesional del profesorado y los sistemas de evaluación docente representan una de las problemáticas más graves (Smylie, 2014). Al respecto, las políticas que conducen los sistemas de evaluación, usualmente tienden a establecer acciones generales o imprecisas para el desarrollo profesional, en consecuencia, las oportunidades de calidad son débiles para mejorar la práctica docente. De ahí que, en ausencia de un seguimiento al desarrollo profesional, como complemento principal de la evaluación, es posible que los sistemas de evaluación docente no incidan a la mejora de las prácticas docentes tal como la literatura supone (Smylie, 2014).

Adicionalmente, de acuerdo con Dee y Wyckoff, (2013), los sistemas de evaluación docente vinculados a incentivos altos deberán comprometer múltiples indicadores del desempeño docente. Además, los autores mencionan que cualquier sistema de evaluación de docentes incurrirá en faltas cuestionables relacionadas a la forma de evaluar y a las consecuencias que los docentes enfrenten; así pues, las instituciones pueden reducir las faltas por medio de sistemas de evaluación más sofisticados, pero esto implicará atender desafíos de implementación y gastos de operación. Así, los formuladores de políticas educativas en evaluación docente deben ponderar los costos frente a los beneficios educativos, y como resultado a largo plazo los sistemas de evaluación docente pueden evitar la retención de docentes con un bajo desempeño (Dee y Wyckoff, 2013).

Por consiguiente, la evaluación de la docencia debe conducirse hacia una mejor comprensión de la enseñanza y la evaluación, a la búsqueda de formas en las que la investigación y la práctica puedan informarse de manera recíproca (Theall y Franklin, 1990). En este escenario, las IES enfrentan el gran reto de conocer y comprender mejor las condiciones y particularidades de la práctica docente, con el objetivo 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 estrategias de acción que favorezcan reconocer la complejidad de la docencia y su evaluación formativa.

Cabe enfatizar que, en palabras de Theall (2010), los docentes son el recurso humano más importante en la educación superior por el hecho de desempeñar la tarea educativa, representada en la creación, transferencia, aplicación e integración del conocimiento. Sin embargo, el autor reconoce una evolución desigual entre la política y la

práctica de la evaluación docente con el desarrollo profesional del profesorado, esto pone de manifiesto la necesidad de incrementar las oportunidades para la mejora de la práctica docente, en dirección de una enseñanza y aprendizaje eficaz y a una rendición de cuentas diligente para las IES.

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.

A manera de cierre, derivado de esta investigación, se realizan algunas recomendaciones puntuales para el SED en dirección de favorecer la tarea de la evaluación docente:

- Proyectar y debatir las condiciones y posibilidad de un vínculo institucionalizado de los resultados de la evaluación con el PFFDD.
- Repensar y discutir la relación de los resultados de la evaluación docente con programa de estímulos económicos.
- Promover y hacer más visible a la comunidad docente la consulta de sus resultados de la evaluación, indicar las formas de acceso al sistema gestor del SED y los periodos de consulta de sus últimos reportes de evaluación.
- Esclarecer, para las diferentes audiencias en la UABC, los objetivos y alcance de la evaluación docente.
- Considerar la aplicación del CECD a muestras aleatorias.

- Valorar la pertinencia y condiciones para incorporar al SED otras estrategias de evaluación —autoevaluación, coevaluación—, y otros CECD para programas específicos

Áreas de oportunidad y posibles líneas de indagación futuras

Stake (2017) señala que todo lo que se sabe sobre evaluación y medición de la competencia docente se aplica a la metaevaluación, puesto que se atiende el proceso y el producto, se identifica la evaluación formativa y sumativa, se considera la importancia del contexto y las circunstancias, existe una búsqueda de la calidad, la bondad, fortalezas, debilidades, negligencias y sesgos, cómo se ha evaluado y qué se ha encontrado. A pesar de lo anterior, Stake advierte que el actual sistema de evaluación obligatoria de la enseñanza se encuentra prácticamente estancado. Por tanto, con el objetivo de abonar al campo de conocimiento científico en evaluación de la docencia en educación superior, es necesario emprender estudios minuciosos y sistemáticos de metaevaluación que permitan contar con un conocimiento vigente sobre las expectativas y realidades de la evaluación para verdaderamente contribuir a la mejora.

De manera específica, a raíz del confinamiento derivado de la enfermedad por coronavirus 2019 (COVID-19) —causada por el virus del síndrome respiratorio agudo severo tipo-2 (SARS-CoV-2)—, en esta metaevaluación se dificultó el acercamiento a los diferentes actores educativos para el empleo de otras técnicas de recolección de datos, p. ej. entrevistas o grupos focales. Por tanto, resulta conveniente para próximos estudios considerar dichas técnicas para contar con un mayor y diverso bagaje de información.

Otro punto importante a considerar para estudios de metaevaluación en la UABC es la selección de una muestra aleatoria, aunque se considera valiosa la participación por autoselección de Directivos, Docentes y Estudiantes, sería conveniente para aminorar predilecciones o sesgos (Kerlinger y Lee, 2002).

Asimismo, dado que los cuestionarios de opinión de los estudiantes son la técnica mayormente empleada, se considera relevante investigar la relación de la actitud de los estudiantes frente a un proceso de evaluación de la docencia con su comportamiento al responder un cuestionario de evaluación docente (Spooren y Christiaens, 2017); y en como la experiencia de realizar un cuestionario de evaluación docente puede convertirse en una tarea educativa más clara y reflexiva para los estudiantes. También, estudiar el efecto de las emociones de los estudiantes en la finalización de la evaluación docente (Cook et al., 2021). Además de, replantear en qué medida se depende principalmente de los resultados de la evaluación docente en la opinión de los estudiantes para proponer seguimientos formativos y tomar decisiones sumativas.

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.
- Armstrong, J. (2010). Naturalistic inquiry. *Encyclopedia of research design*, 2, 880-885.
- Benton, S. L., y Cashin, W. E. (2012). Student ratings of teaching: A summary of research and literatura. *IDEA Papers Series*, 50.
https://ideacontent.blob.core.windows.net/content/sites/2/2020/01/PaperIDEA_50.pdf
- Boyer, E. L. (1990). The Campuses: Diversity with Dignity. En *Scholarship Reconsidered: Priorities of the professoriate* (pp.53-64). The Carnegie Foundation for the Advancement of Teaching.
- Bransford, J. D., Brown, A. L., y Cocking, R. R. (2000). *How people learn: brain, mind, experience, and school* (Edición ampliada). National academy press.
- Bundi, P., Frey, K., y Widmer, T. (2021). Does evaluation quality enhance evaluation use?. *Evidence & Policy*, 17(4), 661-687.
<https://doi.org/10.1332/174426421X16141794148067>
- Centra, J. A. (1993). Approaches to teaching and implications for evaluation. En *Reflective Faculty Evaluation: Enhancing Teaching and Determining Faculty Effectiveness* (pp. 20-46). Jossey-Bass Publishers.
- Chaverra-Fernández, B., y Hernández-Álvarez, J. L. (2021). La metaevaluación en el profesorado de educación física. *Revista Internacional de Medicina y Ciencias de la Actividad Física y del Deporte*, 21(83), 501-513.
<https://doi.org/10.15366/rimcafd2021.83.006>
- Chen, Y. N., Kang, J. L., y Guo, C. U. (2020). The construction of meta-evaluation indicators of Taiwan's university program evaluation: JCSEE program evaluation standards as a framework. *International Journal of Humanities, Arts and Social Sciences*, 6(2), 10-24. <https://dx.doi.org/10.20469/ijhss.6.20002-1>
- Cisneros-Cohernour, E. J., y Patrón, R. M. (2017). Metaevaluación de un sistema de evaluación docente en una universidad pública mexicana. En E. Luna y M. Rueda (Coords.), *Experiencias de evaluación de la docencia en Iberoamérica* (pp. 279-294). Universidad Autónoma de Baja California.
- 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* (2a ed., pp. 357-386). Editorial Alianza.
- Cook, C., Jones, J., y Al-Twal, A. (2021). Validity and fairness of utilising student evaluation of teaching (SET) as a primary performance measure. *Journal of Further*

- and Higher Education*, 46(2), 172-184.
<https://doi.org/10.1080/0309877X.2021.1895093>
- 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
- Darling-Hammond, L. (2012). Desarrollo de un enfoque sistémico para evaluar la docencia y fomentar una enseñanza eficaz. *Pensamiento Educativo, Revista de Investigación Latinoamericana*, 49(2), 1-20.
<http://pensamientoeducativo.uc.cl/index.php/pel/article/view/25775>
- Dee, T. S., y Wyckoff, J. (2015). Incentives, selection, and teacher performance: Evidence from IMPACT. *Journal of Policy Analysis and Management*, 34(2), 267-297.
<https://doi.org/10.1002/pam.21818>
- Dunegan, K. J., y Hrivnak, M. W. (2003). Characteristics of mindless teaching evaluations and the moderating effects of image compatibility. *Journal of management education*, 27(3), 280-303. <https://doi.org/10.1177/1052562903027003002>
- 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.
- 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>
- Feistauer, D., y Richter, T. (2018). Validity of students' evaluations of teaching: Biasing effects of likability and prior subject interest. *Studies in Educational Evaluation*, 59, 168-178. <https://doi.org/10.1016/j.stueduc.2018.07.009>
- Feldman, K. A. (1976a). Grades and college students' evaluations of their courses and teachers. *Research in Higher Education*, 4(1), 69-111.
<https://link.springer.com/article/10.1007/BF00991462>
- Feldman, K. A. (1976b). The superior college teacher from the students' view. *Research in higher education*, 5(3), 243-288.
<https://link.springer.com/article/10.1007/BF00991967>
- Feldman, K. A. (1977). Consistency and variability among college students in rating their teachers and courses: A review and analysis. *Research in Higher Education*, 6(3), 223-274. <https://link.springer.com/article/10.1007/BF00991288>
- Feldman, K. A. (1978). Course characteristics and college students' ratings of their teachers: What we know and what we don't. *Research in Higher Education*, 9(3), 199-242. <https://link.springer.com/content/pdf/10.1007/BF00976997.pdf>

- Feldman, K. A. (1979). The significance of circumstances for college students' ratings of their teachers and courses. *Research in Higher Education*, 10(2), 149-172. <https://link.springer.com/article/10.1007/BF00976227>
- Feldman, K. A. (2007). Identifying exemplary teachers and teaching: Evidence from student ratings. En *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 93-143). Springer. https://link.springer.com/chapter/10.1007/1-4020-5742-3_5
- Franklin, J., y Theall, M. (1995). The relationship of disciplinary differences and the value of class preparation time to student ratings of teaching. *New directions for teaching and learning*, (64), 41-48. <https://onlinelibrary.wiley.com/doi/abs/10.1002/tl.37219956407>
- Gallego, C., y Rojo, N. (2014). *Meta-evaluación del programa Docencia de la UCM*. Universidad Complutense de Madrid. <https://bit.ly/3DRqG0d>
- 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>
- Gibbs, G. (2012). La naturaleza del análisis cualitativo. En *El análisis de datos cualitativos en investigación cualitativa* (pp. 19-30). Morata.
- Guba, E., y Lincoln, Y. (2012). Controversias paradigmáticas, contradicciones y confluencias emergentes. En N. K. Denzin y Y. S. Lincoln (Eds.), *Paradigmas y perspectivas en disputa: Manual de metodología cualitativa* (2da, ed., pp. 38-78). Gedisa.
- Gutiérrez-de-Rozas, B., y Carpintero, E. (2021). Análisis de la evaluación de programas de intervención en motivación en Educación Secundaria. *Revista de Investigación Educativa*, 39(2), 503-525. <https://doi.org/10.6018/rie.442341>
- Hammer, R., Peer, E., y Babad, E. (2018). Faculty attitudes about student evaluations and their relations to self-image as teacher. *Social Psychology of Education*, 21(3), 517-537. <https://doi.org/10.1007/s11218-018-9426-1>
- Hamui-Sutton, A., Ortiz-García, A., Cejudo-Aparicio, L., Lavallo-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). Becoming a better teacher: A case of changing the pedagogical knowledge and beliefs of law professors. *Instructional Science*, 28(5), 491-523. <https://doi.org/10.1023/A:1026521725494>

- Hativa, N., Barak, R., y Simhi, E. (2001). Exemplary university teachers: Knowledge and beliefs regarding effective teaching dimensions and strategies. *The journal of higher education*, 72(6), 699-729. <https://doi.org/10.1080/00221546.2001.11777122>
- Holsti, O. R. (1966). Coding content data. En *Content Analysis for the social humanities* (pp. 94-126). Addison-Wesley Publishing Company.
- 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>
- Índice de Revistas de Educación Superior e Investigación Educativa. (2015). *Vocabulario controlado*. Índice de Revistas de Educación Superior e Investigación Educativa. <http://132.248.9.1:8991/iresie/VocabularioControlado.pdf>
- 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>
- Kerlinger, F. N., y Lee, H. B. (2002). Muestreo y aleatoriedad. En *Investigación de comportamiento* (4ta. ed., pp. 147-198). McGraw-Hill.
- Krippendorff, K. (2004). Reliability. En *Content analysis an introduction to its methodology* (2da, ed., pp. 211-256). SAGE.
- 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>
- 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>
- Marcelo, C., y Vaillant, D. (2009). Nociones y alcances. En *Desarrollo profesional docente: ¿cómo se aprende a enseñar?* (pp. 75-82). Narcea Ediciones.
- Marsh, H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 319-383). Springer. https://link.springer.com/chapter/10.1007/1-4020-5742-3_9
- 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. <https://doi.org/10.1016/j.stueduc.2016.03.002>

- Mayring, P. (2000). Qualitative content analysis. *Forum qualitative social research*, 1(2). <http://www.qualitative-research.net/index.php/fqs/article/view/1089/2385>
- Mayring, P. (2014). *Qualitative content analysis: theoretical foundation, basic procedures and software solution*. <https://www.ssoar.info/ssoar/handle/document/39517>
- Mediell, M., y Dionne, E. (2018). Un outil à visée pédagogique pour discuter de méthodologie. *Canadian Journal of Program Evaluation*, 33(1), 90-113. <https://doi.org/10.3138/cjpe.31127>
- Mujis, D. (2004). Univariate statistics. En *Doing Quantitative Research in Education with SPSS* (pp. 91-112). Sage.
- Nasser-Abu, F. (2017a). Contemporary evaluation of teaching: Challenges and promises. *Studies in Educational Evaluation*, 54, 1-3. <http://dx.doi.org/10.1016/j.stueduc.2017.02.002>
- Nasser-Abu, F. (2017b). 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>
- Nasser-Abu, F., y Fresko, B. (2009). Student evaluation of instruction: what can be learned from students' written comments?. *Studies in Educational evaluation*, 35(1), 37-44. <https://doi.org/10.1016/j.stueduc.2009.01.002>
- Neumann, R. (2001). A decade of student evaluation of teaching: Case study of an evolving system. *Journal of Institutional Research*, 10(1), 96-111. <https://bit.ly/3TYFX53>
- Patiño-Montero, F., Godoy-Acosta, D. C., y Arias-Meza, D. C. (2022). Actualización de la evaluación docente de posgrados en una universidad multicampus: experiencia desde la Universidad Santo Tomás (Colombia). *Revista Educación*, 46(2), 457-474. <http://dx.doi.org/10.15517/revedu.v46i2.47955>
- Penuel, W. R., y Shepard, L. A. (2016). Assessment and teaching. En D. H. Gitomer y C. A. Bell, *Handbook of Research on Teaching* (5ta ed., pp. 787-850). American Educational Research Association.
- Phillips, D. C. (2018). The Many Functions of Evaluation in Education. *Analysis Archives*, 26(46). <http://dx.doi.org/10.14507/epaa.26.3811>
- 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>
- Rincon-Flores, E. G., Gallardo, K., y de la Fuente, J. M. (2018). Strengthening an Educational Innovation Strategy: Processes to Improve Gamification in Calculus Course through Performance Assessment and Meta-Evaluation. *International Electronic Journal of Mathematics Education*, 13(1), 1-11. <https://doi.org/10.12973/iejme/2692>

Referencias

- 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., 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.
- Rueda, M., Canales, A., y Leyva, Y. (2016). Desarrollo de la docencia: Universidad Nacional Autónoma de México. En *Prácticas y condiciones institucionales para el desarrollo de la docencia* (pp. 15-70). Universidad Nacional Autónoma de México
- Rueda, M. (2019). Prólogo. En *Programas y políticas de evaluación docente en educación básica (1993-2017)* (pp. 9-12). Universidad Nacional Autónoma de México.
- 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
- Sánchez, R. (2019). 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>
- Sandín, M. P. (2003). Perspectivas teórico-epistemológicas en la investigación educativa. En *Investigación Cualitativa en educación: Fundamentos y tradiciones* (pp. 45-71). McGrawHill.
- Scriven, M. (1966). The methodology of evaluation. *Social Science Education Consortium*, 110. <https://eric.ed.gov/?id=ED014001>
- Scriven, M. (1991). *Evaluation thesaurus* (4ta. 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). *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. (2009). Meta-Evaluation Revisited. *Journal of MultiDisciplinary Evaluation*, 6(11). https://journals.sfu.ca/jmde/index.php/jmde_1/article/view/220
- Scriven, M. (2011). *Evaluating evaluations: A meta-evaluation checklist* (9na ed.). Claremont Graduate University.
http://michaelscriven.info/images/EVALUATING_EVALUATIONS_8.1.11.pdf
- Scriven, M. (2015). *The Meta-Evaluation Checklist*. Claremont Evaluation Center.

Referencias

- Shulha, L. M., y Cousins, J. B. (1997). Evaluation use: Theory, research, and practice since 1986. *Evaluation practice*, 18(3), 195-208. [https://doi.org/10.1016/S0886-1633\(97\)90027-1](https://doi.org/10.1016/S0886-1633(97)90027-1)
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard educational review*, 57(1), 1-23.
- Shulman, L. S. (2005). Conocimiento y enseñanza: fundamentos de la nueva reforma. Profesorado. *Revista de currículum y formación del profesorado*, 9(2). <https://recyt.fecyt.es/index.php/profesorado/article/view/42675>
- 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>
- Smylie, M. A. (2014). Teacher evaluation and the problem of professional development. *Mid-Western Educational Researcher*, 26(2), 97-111. <https://bit.ly/3Wr0rYi>
- 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>
- Stake, R. E. (1972). *Responsive evaluation*. Document resume. <https://files.eric.ed.gov/fulltext/ED075487.pdf>
- Stake, R. E. (1975). *Program evaluation particularly responsive evaluation*. Center for Instructional Research and Curriculum Evaluation, 5. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.200.333&rep=rep1&type=pdf>
- 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. (1991). Excerpts from: " Program Evaluation, Particularly Responsive Evaluation". *Evaluation Practice*, 12(1), 63-76. [https://doi.org/10.1016/0886-1633\(91\)90025-S](https://doi.org/10.1016/0886-1633(91)90025-S)
- Stake, R. E. (1999). Introducción: Estudio intensivo de los métodos de investigación con estudio de casos. En *Investigación con estudio de casos* (2da ed., pp. 9-14). Ediciones Morata.
- Stake, R. E. (2003). Responsive Evaluation. En T. Kellaghan y D. L. Stufflebeam (Eds.), *International Handbook of Educational Evaluation* (pp. 63-68). Springer. https://doi.org/10.1007/978-94-010-0309-4_5

- Stake, R. E. (2006, 29 de septiembre). Standards-Based and Responsive Evaluation [Sesión de conferencia]. IPEN Development and Evaluation, Tbilisi, Georgia
- Stake, R. E. (2011). Program evaluation particularity 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. (2014). Information Science and Responsive Evaluation. *E-Learning and Digital Media*, 11(5), 443-450. <https://doi.org/10.2304/elea.2014.11.5.443>
- 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.
- Strauss, C., y Corbin, J. (2002). Descripción, ordenamiento conceptual y teorización. En *Bases de la investigación cualitativa. Técnicas y procedimientos para desarrollar la teoría fundamentada* (pp. 17-28). SAGE.
- 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. 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* (2da. Ed., pp. 3-43). John Wiley & Sons.
- Theall, M. (2010). Evaluating teaching: From reliability to accountability. *New Directions for Teaching and Learning*, (123), 85-95. <https://doi.org/10.1002/tl.412>
- Theall, M. (2017). MVP and faculty evaluation. *New Directions for Teaching and Learning*, (152), 91-98. <https://doi.org/10.1002/tl.20271>
- Theall, M., y Franklin, J. (1990). Student ratings in the context of complex evaluation systems. *New directions for teaching and learning*, (43), 17-34.
<https://doi.org/10.1002/tl.37219904304>
- Theall, M., y Franklin, J. (2001). Creating Responsive Student Rating Systems to Improve Evaluation Practice. *New Directions for Teaching and Learning*, 83, 95-107.
<https://doi.org/10.1002/tl.8308>
- Tojar, J. C. (2006). Disciplinas, paradigmas y tradiciones. En *Investigación cualitativa: comprender y actuar* (pp. 53-88). La Muralla.
- 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>
- Tripp, T. M., Jiang, L., Olson, K., y Graso, M. (2018). The fair process effect in the classroom: Reducing the influence of grades on student evaluations of teachers. *Journal of Marketing Education*, 41(3), 173-184.
<https://doi.org/10.1177/027347531877261>

Referencias

- Universidad Autónoma de Baja California. (2000). *Informe de Rectoría 1999-2000*. <https://bit.ly/3BY7Djc>
- Universidad Autónoma de Baja California. (2005). *Informe de Rectoría 2005*. <https://bit.ly/3YH9WB8>
- Universidad Autónoma de Baja California. (2010). *Sistema de Evaluación Docente*. <http://ed.uabc.mx/maestros/index.jsp>
- Universidad Autónoma de Baja California. (2015). *Plan de Desarrollo Institucional 2015-2019*. http://web.uabc.mx/planeacion/pdi/2015-2019/PDI_2015-2019.pdf
- Universidad Autónoma de Baja California. (2017b). *Informe de gestión 2009 2017*. http://pedagogia.mx1.uabc.mx/transparencia/InformeAnual/InformeActividadesFPIE_2010-2017.pdf
- Universidad Autónoma de Baja California. (2019). *Oficina de Planeación y Desarrollo Institucional*. <http://www.uabc.mx/planeacion/funciones.htm>
- Universidad Autónoma de Baja California. (2021a). *Anexo 1: Tabla de identificación de actividades (TIA)*. <https://bit.ly/39ZG9PD>
- Universidad Autónoma de Baja California. (2021b, 19 de mayo). *Coordinación General de Formación Profesional*. <http://www.uabc.mx/formacionbasica/index.html>
- Universidad Autónoma de Baja California. (2021c). *Evaluación de la docencia de licenciatura en la UABC: Descripción general del proceso*. <https://bit.ly/3Aan9IG>
- Universidad Autónoma de Baja California. (2021d, 28 de diciembre). *Formación docente*. <https://bit.ly/3QRPNV0>
- Universidad Autónoma de Baja California. (2022, 08 de enero). *Numeralia institucional*. <http://web.uabc.mx/planeacion/numeralia/#>
- Universidad Complutense de Madrid. (2016). *Modelo de evaluación de la actividad docente del profesorado de la UCM: programa DOCENTIA-UCM*. <https://www.ucm.es/data/cont/docs/3-2016-11-15-Docentia-noviembre2016.pdf>
- Uttl, B., White, C. A., y Gonzalez, D. W. (2017). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation*, 54, 22-42. <https://doi.org/10.1016/j.stueduc.2016.08.007>
- Van Lankveld, T., Schoonenboom, J., Volman, M., Croiset, G., y Beishuizen, J. (2017). Developing a teacher identity in the university context: A systematic review of the literature. *Higher Education Research & Development*, 36(2), 325-342. <https://doi.org/10.1080/07294360.2016.1208154>
- Vieirai, R., Storinoii, R., y Filho, O. (2019). Avaliação do Relatório de Avaliação Pós-ocupação da Escola Municipal Albert Schweitzer: um estudo meta-avaliativo. *Revista Meta: Avaliação*, 11(33), 638-663. <https://revistas.cesgranrio.org.br/index.php/metaavaliacao/article/view/1829>

- 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>
- Wood, P. (2017a). 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.
- Wood, P. (2017b). Holiploigy: navigating the complexity of teaching in higher education. *Journal of Learning Development in Higher Education*, 11. <https://journal.aldinhe.ac.uk/index.php/jldhe/article/view/421>
- Yin, R. K. (2017). Getting started: how to know whether and when to use the case study as a research method. En *Case study research and applications: Design and methods* (3ra ed., pp. 31-56). Sage publications.
- Yin, R. K. (2018). Getting started: how to know whether and when to use the case study as a research method. En *Case study research and applications: Design and methods* (6ta ed., pp. 31-56). Sage publications.
- Zabalza, M. A. (2007). *Competencias docentes del profesorado universitario. Calidad y desarrollo profesional*. Narcea.

Apéndices

Apéndice A. Guion de entrevista del Grupo 1) Coordinadores de la administración central

1. Cuál es su perspectiva sobre el SED en relación con:
 - a. *El procedimiento de aplicación*
 - i. *Calendario de aplicación*
 - ii. *Difusión*
 - iii. *Aplicación del cuestionario*
 - b. *Los instrumentos*
 - i. *Cuestionario de evaluación de la competencia docente (teoría)*
 - ii. *Cuestionario de evaluación de la competencia docente (práctica)*
 - iii. *Cuestionario de evaluación de la competencia docente en línea*
 - c. *La entrega de resultados*
 - d. *El uso de los resultados*
 - i. *Directivos*
 - ii. *Docentes*
 - iii. *Administración central (PREDEPA)*
2. ¿A su juicio cuales son los aspectos positivos del SED?
3. ¿Cuáles son los principales efectos (impactos) del SED?
4. ¿Cuáles son las problemáticas que enfrenta el SED?
5. ¿Cuáles son los cambios o aspectos a mejorar?
6. ¿Considera otras formas alternativas de evaluación docente?

Apéndice B. Cuestionario de evaluación del SED para Directores y subdirectores

Apreciable Director (a) y Subdirector (a):

El objetivo general de este estudio es realizar una Metaevaluación del Sistema de Evaluación Docente de la UABC, con el fin de identificar sus fortalezas y oportunidades de mejora desde la perspectiva de la comunidad académica involucrada en la función docente. Por tal motivo, este cuestionario tiene el propósito de conocer la opinión de los directivos sobre el Sistema de Evaluación Docente.

Aviso de Confidencialidad: Los datos proporcionados serán resguardados con base en los principios de licitud, calidad, consentimiento, información, finalidad, lealtad, proporcionalidad y responsabilidad consagrados en la Ley Federal de Protección de Datos Personales en Posesión de Particulares y su Reglamento. La información recabada será utilizada sólo con el fin de identificar las oportunidades de mejora de los programas institucionales.

Instrucciones: Las preguntas pueden tener más de una respuesta, elija o conteste las que considere convenientes.

Dimensión 1. Uso de los resultados de evaluación docente

1.1 Como director, ¿Qué uso da a los resultados de evaluación docente?

<input type="checkbox"/>	Retroalimentar la práctica docente
<input type="checkbox"/>	Para la asignación de asignaturas
<input type="checkbox"/>	En los informes de acreditación
<input type="checkbox"/>	Asignación de horarios
<input type="checkbox"/>	Los resultados no son tomados en cuenta

Otros, por favor especifique:

para los Alumnos–Teoría incluye los elementos importantes de la práctica docente?

2.2 ¿Qué elementos considera que deberían ser incluidos?

2.3 ¿Qué elementos considera que deberían ser modificados?

2.4 ¿Qué elementos sugiere eliminar?

Valore en una escala del 1 al 10 (siendo 1 la valoración más baja y 10 la valoración más alta)

2.5 En su opinión, ¿El Cuestionario de Evaluación de la Competencia Docente para los Alumnos- Práctica incluye los elementos importantes de la práctica docente?

1	2	3	4	5	6	7	8	9	10

2.6 ¿Qué elementos considera que deberían ser incluidos?

2.7 ¿Qué elementos considera que deberían ser modificados?

2.8 ¿Qué elementos sugiere eliminar?**2.9 ¿Considera necesario que se desarrollen otros Cuestionarios de Evaluación de la Competencia Docente para los Alumnos para programas educativos particulares?**

<input type="checkbox"/>	Sí
<input type="checkbox"/>	No

2.10 En caso afirmativo, ¿Para cuáles?**Dimensión 3. Procedimiento de aplicación**

3.1 ¿Considera apropiadas las fechas de aplicación del Cuestionario de Evaluación de la Competencia Docente para los Alumnos?

<input type="checkbox"/>	Sí
<input type="checkbox"/>	No

3.2 ¿Tiene alguna propuesta en particular?**3.3 ¿Considera adecuada la difusión para promover que los estudiantes contesten el Cuestionario de Evaluación de la Competencia Docente para los Alumnos?**

<input type="checkbox"/>	Sí
<input type="checkbox"/>	No

Comentarios:

3.4 ¿La estrategia de aplicación del Cuestionario de Evaluación de la Competencia Docente para los Alumnos es adecuada?

Sí

No

Comentarios

Dimensión 4. Valoración del Sistema de Evaluación Docente

4.1 A su juicio, ¿Cuáles son los aspectos positivos del Sistema de Evaluación Docente?

Conocer el desempeño de los docentes

Conocer la opinión de los estudiantes

Retroalimentar la enseñanza

Otros, por favor especifique:

4.2 En su opinión, el Sistema de Evaluación Docente es útil para:

1:

2

3

4.3 En su opinión, ¿Cuáles son las tres principales problemáticas que enfrenta el Sistema de Evaluación Docente?

1

2

3

4.4 En su opinión, ¿Cuáles son los cambios o aspectos a mejorar?

1

2

3

4.5 ¿Considera pertinente utilizar otras estrategias de evaluación docente?

No

Sí, Autoevaluación docente

Sí, Coevaluación docente

Sí, Evaluación de directivos

Otros, por favor especifique:

Apéndice C. Cuestionario de evaluación del SED para docentes

Apreciable docente:

El objetivo general de este estudio es realizar una metaevaluación del Sistema de Evaluación Docente (SED) de la UABC, con el fin de identificar sus fortalezas y oportunidades de mejora desde la perspectiva de la comunidad académica involucrada en la función docente.

Por tal motivo, este cuestionario tiene el propósito de conocer la opinión de los profesores sobre el SED. La información que proporcione será manejada de manera confidencial.

Aviso de Confidencialidad: Los datos proporcionados serán resguardados con base en los principios de licitud, calidad, consentimiento, información, finalidad, lealtad, proporcionalidad y responsabilidad consagrados en la Ley Federal de Protección de Datos Personales en Posesión de Particulares y su Reglamento. La información recabada será utilizada sólo con el fin de identificar las oportunidades de mejora de los programas institucionales.

Instrucciones: Las preguntas pueden tener más de una respuesta, elija o conteste las que considere convenientes.

Información general

1. Es usted personal académico de:

<input type="checkbox"/>	Asignatura
<input type="checkbox"/>	De carrera

2. ¿En qué área del conocimiento imparte clases?

<input type="checkbox"/>	Ciencias de la ingeniería y la tecnología
<input type="checkbox"/>	Ciencias sociales
<input type="checkbox"/>	Ciencias naturales y exactas
<input type="checkbox"/>	Ciencias de la educación y humanidades

- | | |
|--------------------------|------------------------------------|
| <input type="checkbox"/> | Ciencias de la salud |
| <input type="checkbox"/> | Ciencias agropecuarias |
| <input type="checkbox"/> | Ciencias económico administrativas |

3. ¿En qué modalidad educativa imparte clases? *

- | | |
|--------------------------|-----------------------------|
| <input type="checkbox"/> | Presencial |
| <input type="checkbox"/> | Mixta (semipresencial) |
| <input type="checkbox"/> | No presencial (a distancia) |

4. Tiempo de laborar en UABC:

- | | |
|--------------------------|--------------------|
| <input type="checkbox"/> | 1 a 3 años |
| <input type="checkbox"/> | 4 a 7 años |
| <input type="checkbox"/> | 8 años en adelante |

Dimensión 1. Uso de los resultados de evaluación docente.

Valore en una escala del 1 al 10 (siendo 1 la valoración más baja y 10 la valoración más alta)

1.1 Considero útiles los resultados derivados del Cuestionario de Evaluación de la Competencia Docente por los alumnos.

1	2	3	4	5	6	7	8	9	10
<input type="checkbox"/>									

1.2 ¿Qué uso da a los resultados de evaluación docente?

- | | |
|--------------------------|---|
| <input type="checkbox"/> | Para identificar fortalezas y debilidades |
| <input type="checkbox"/> | Busco alternativas de mejora de mi práctica docente |

2.4 ¿La redacción de los reactivos es clara?

--	--	--	--	--	--	--	--	--	--	--

¿Qué tan de acuerdo estaría en incluir en el cuestionario los siguientes aspectos?

1 2 3 4 5 6 7 8 9 10

2.5 Puntualidad

2.6 Cumplimiento del tiempo establecido en el programa

2.8 Uso de TIC en el aula

2.9 Considero necesario que se desarrollen otros Cuestionario de Evaluación de la Competencia Docente para los Alumnos para programas educativos particulares

<input type="checkbox"/>	Sí
<input type="checkbox"/>	No

En caso afirmativo, ¿en cuáles?:

2.10 Considero pertinente utilizar otras estrategias de evaluación docente

<input type="checkbox"/>	Sí
<input type="checkbox"/>	No

2.11 En caso afirmativo, ¿cuáles?

<input type="checkbox"/>	Autoevaluación
<input type="checkbox"/>	Evaluación de pares académicos
<input type="checkbox"/>	Evaluación de directivos

Otros (favor de especificar):

2.12 Otros comentarios:**Dimensión 3. Procedimiento de aplicación**

Valore en una escala del 1 al 10 (siendo 1 la valoración más baja y 10 la valoración más alta)

3.1 ¿Son oportunas las fechas de aplicación del Cuestionario de Evaluación de la Competencia Docente para los Alumnos?

1 2 3 4 5 6 7 8 9 10

--	--	--	--	--	--	--	--	--	--	--

3.2 ¿Es pertinente la difusión para promover que los estudiantes contesten el Cuestionario de Evaluación de la Competencia Docente para los Alumnos?

--	--	--	--	--	--	--	--	--	--	--

3.3 ¿Le parece adecuado que sea obligatorio para los estudiantes contestar el Cuestionario de Evaluación de la Competencia Docente?

--	--	--	--	--	--	--	--	--	--	--

3.4 ¿Considera adecuada la manera en que en su unidad académica se realiza la aplicación del Cuestionario de Evaluación de la Competencia Docente?

--	--	--	--	--	--	--	--	--	--	--

3.5 Otros comentarios:

Apéndice D. Cuestionario de evaluación del SED para Estudiantes

Apreciable estudiante:

El objetivo general de este estudio es realizar una evaluación del Sistema de Evaluación Docente (SED) de la UABC, con el fin de identificar sus fortalezas y oportunidades de mejora desde la perspectiva de la comunidad académica involucrada en la función docente.

Por tal motivo, este cuestionario tiene el propósito de conocer la opinión de los profesores sobre el SED.

Aviso de Confidencialidad: Los datos proporcionados serán resguardados con base en los principios de licitud, calidad, consentimiento, información, finalidad, lealtad, proporcionalidad y responsabilidad consagrados en la Ley Federal de Protección de Datos Personales en Posesión de Particulares y su Reglamento. La información recabada será utilizada sólo con el fin de identificar las oportunidades de mejora de los programas institucionales.

Información general

1. Sexo:

<input type="checkbox"/>	Masculino
<input type="checkbox"/>	Femenino

2. Edad:

<input type="checkbox"/>	17-18 años
<input type="checkbox"/>	19-20 años
<input type="checkbox"/>	21-22 años
<input type="checkbox"/>	23-24 años
<input type="checkbox"/>	25 años o más

3. Eres estudiante:

- | | |
|--------------------------|---|
| <input type="checkbox"/> | Regular |
| <input type="checkbox"/> | Irregular (adeudo actualmente una o más materias) |

4. Tu promedio general es de:

- | | |
|--------------------------|------------|
| <input type="checkbox"/> | 100-96 |
| <input type="checkbox"/> | 95-91 |
| <input type="checkbox"/> | 90-86 |
| <input type="checkbox"/> | 85-81 |
| <input type="checkbox"/> | 80-76 |
| <input type="checkbox"/> | 75-71 |
| <input type="checkbox"/> | 70-66 |
| <input type="checkbox"/> | 65-60 |
| <input type="checkbox"/> | 59 o menos |

5 ¿En qué área del conocimiento estudias?

- | | |
|--------------------------|---|
| <input type="checkbox"/> | Ciencias de la ingeniería y la tecnología |
| <input type="checkbox"/> | Ciencias sociales |
| <input type="checkbox"/> | Ciencias naturales y exactas |
| <input type="checkbox"/> | Ciencias de la educación y humanidades |
| <input type="checkbox"/> | Ciencias de la salud |
| <input type="checkbox"/> | Ciencias agropecuarias |
| <input type="checkbox"/> | Ciencias económico administrativas |

6. ¿En qué modalidad educativa estudias?

Presencial

Mixta (semipresencial)

7. ¿Has cursado asignaturas a distancia (en línea)?

Sí

No

8. ¿En qué etapa de formación te encuentras?

Etapa básica

Etapa disciplinaria

Etapa terminal

9. ¿Tienes conexión a Internet en tu casa?

Sí

No

10. ¿Tienes computadora personal? (Laptop o PC)

Sí

No

11. ¿Actualmente trabajas?

Sí

No

Dimensión 1. Uso de los resultados de evaluación docente

Valore en una escala del 1 al 10 (siendo 1 la valoración más baja y 10 la valoración más alta)

1.1 Considero útiles los resultados derivados del Cuestionario de Evaluación de la Competencia Docente por los alumnos.

1	2	3	4	5	6	7	8	9	10

1.2 ¿Para qué es útil la evaluación docente?

<input type="checkbox"/>	Para identificar fortalezas y debilidades
<input type="checkbox"/>	Seleccionar materias
<input type="checkbox"/>	Comentar con mis compañeros los resultados
<input type="checkbox"/>	Ninguno

Otros, por favor especifica:

1.3 En tu opinión, ¿qué uso dan los maestros a los resultados de evaluación docente?

<input type="checkbox"/>	Identificar fortalezas y debilidades de su práctica docente
<input type="checkbox"/>	Los resultados no son tomados en cuenta
<input type="checkbox"/>	No sé

Otros, por favor especifica:

1.3 Otros comentarios:

Dimensión 2. Cuestionario de Evaluación de la Competencia Docente

Valore en una escala del 1 al 10 (siendo 1 la valoración más baja y 10 la valoración más alta)

2.1 En tu opinión, ¿el Cuestionario de Evaluación de la Competencia Docente para los alumnos–Teoría incluye los elementos importantes de la práctica docente?

1	2	3	4	5	6	7	8	9	10

2.2 En tu opinión, ¿el Cuestionario de Evaluación de la Competencia Docente para los alumnos–Práctica incluye los elementos importantes de la práctica docente?

--	--	--	--	--	--	--	--	--	--

2.3 ¿El número de reactivos del cuestionario es suficiente?

--	--	--	--	--	--	--	--	--	--

2.4 ¿La redacción de los reactivos es clara?

--	--	--	--	--	--	--	--	--	--

¿Qué tan de acuerdo estarías en incluir en el cuestionario los siguientes aspectos?

2.5 Puntualidad

1	2	3	4	5	6	7	8	9	10

2.6 Cumplimiento del tiempo establecido en el programa

3.5 ¿Te parece necesario evaluar cada semestre a todos los maestros?

--	--	--	--	--	--	--	--	--	--

3.6 ¿Me es fácil ingresar a al cuestionario?

--	--	--	--	--	--	--	--	--	--

3.7 ¿Cuántas veces intentas ingresar al cuestionario hasta tener éxito?

	1
	2
	3
	4
	5 o más

3.8 Si respondiste dos o más veces indica el horario

	Mañana
	Mediodía
	Tarde
	Noche
	Todas las anteriores

3.9 Tus problemas para ingresar al cuestionario se debieron a:

	No recordé mi contraseña
	Problemas de la Red UABC (por ejemplo, sin conexión a Internet)
	Problemas en el servidor del cuestionario (solo la página del cuestionario no funcionaba)

3.5 Otros comentarios:**Dimensión 4. Valoración del Sistema de Evaluación Docente**

Valore en una escala del 1 al 10 (siendo 1 la valoración más baja y 10 la valoración más alta)

4.1 En tu opinión, ¿qué tan útil consideras el Sistema de Evaluación Docente?

--	--	--	--	--	--	--	--	--	--

4.5 Otros comentarios: