

Evidencias de fiabilidad y validez de una escala para la autoevaluación de las prácticas de enseñanza en secundaria

NANCY NOEMÍ GUTIÉRREZ-ANGUIANO* | ALICIA ALELÍ CHAPARRO CASO LÓPEZ**

El objetivo del presente estudio fue desarrollar y validar una escala para la autoevaluación de las prácticas de enseñanza en secundaria, a partir de los indicadores del marco para la enseñanza de Danielson (2013). El desarrollo del instrumento implicó tres etapas: en la primera, se diseñó el instrumento; en la segunda, se establecieron evidencias de validez de contenido; en la tercera etapa se realizó un pilotaje con una muestra de 203 docentes de secundaria, a fin de generar evidencias de fiabilidad y validez relacionadas con el constructo. Como producto se obtuvo una escala de autoevaluación con 70 ítems integrados en cinco dominios: planeación, clima del aula, enseñanza, evaluación del aprendizaje y responsabilidades profesionales. Se identificó un índice de fiabilidad de .98 e índices aceptables de bondad de ajuste en el análisis factorial confirmatorio, lo que corrobora la congruencia de los ítems con el constructo.

The aim of this study was to develop and validate a scale for self-assessment of teaching practices in junior high school, based on the indicators of Danielson's Framework for teaching (2013). The instrument was developed in three stages: in the first, the instrument was designed; in the second, evidence of validity of content was established; and in the third, a pilot test was conducted with a sample of 203 junior high school teachers, to produce evidence of reliability and validity related to the construct. As product, we obtained a self-assessment scale with 70 items across five domains: planning, classroom atmosphere, teaching, evaluation of learning, and professional responsibilities. An index of reliability of .98 was identified, with acceptable indices of goodness of fit in the confirmatory factorial analysis, which corroborates the congruence of the items with the construct.

Palabras clave

Autoevaluación
Profesores
Práctica pedagógica
Educación secundaria
Validez

Keywords

Self-assessment
Teachers
Pedagogical practice
Junior high school
Validity

Recibido: 17 de noviembre de 2018 | Aceptado: 30 de agosto de 2019

DOI: <https://doi.org/10.22201/iisue.24486167e.2019.167.58854>

* Docente de la Universidad Autónoma de Baja California (México). Maestra en Ciencias Educativas. Líneas de investigación: evaluación educativa, eficacia escolar, enseñanza eficaz. CE: noemi.gutierrez@uabc.edu.mx

** Investigadora del Instituto de Investigación y Desarrollo Educativo de la Universidad Autónoma de Baja California (México). Doctora en Psicología Educativa. Líneas de investigación: eficacia escolar, enseñanza eficaz; convivencia escolar. CE: achaparro@uabc.edu.mx

INTRODUCCIÓN

La evaluación de la enseñanza ha ido cobrando cada vez más fuerza, no sólo en la investigación educativa, sino también en la implementación de políticas. Particularmente, en México han surgido esquemas de evaluación docente para la asignación y permanencia de plazas y remuneraciones, entre otros. Sin embargo, como señalan Luna *et al.* (2013), la evaluación docente debe, en esencia, servir a la mejora de los procesos educativos, como herramienta de monitoreo con fines formativos. Para ello, es necesario contar con instrumentos que permitan identificar las áreas susceptibles de mejora en la práctica cotidiana de la enseñanza, pero que, a su vez, muestren un sólido referente teórico-metodológico, además de elementos que demuestren su validez.

No obstante lo anterior, la Red Iberoamericana de Investigadores de la Evaluación de la Docencia (RIIED, 2008) ha destacado que, en la mayoría de los casos, los esfuerzos por realizar evaluaciones docentes no incluyen estos aspectos. Es así, que el presente estudio constituye un esfuerzo por generar una propuesta para la evaluación formativa de la enseñanza, con características de validez, y que se constituya como una guía para que los docentes puedan reflexionar sobre su propia práctica y sean ellos mismos quienes identifiquen las áreas de mejora.

Las prácticas de enseñanza hacen referencia al conjunto de tareas que configuran el quehacer del docente, es decir, las acciones que éste implementa para el tratamiento didáctico de los contenidos por enseñar (Cid-Sabucedo *et al.*, 2009; García-Cabrero *et al.*, 2008; Shulman, 1986). Anijovich y Mora (2009) señalan que las prácticas de enseñanza se emplean en tres momentos esenciales:

Planeación. Es considerado el componente principal para que una clase funcione adecuadamente, puesto que en este momento se construye la programación sobre el contenido

que se va a enseñar en el transcurso del ciclo escolar. Para ello, el docente deberá organizar sus estrategias de enseñanza (metodologías didácticas y técnicas por aplicar), así como los materiales y actividades que permitirán estructurar las experiencias de aprendizaje. Asimismo, en este momento se determinan las estrategias evaluativas para identificar el avance de los estudiantes respecto a la apropiación de los contenidos.

Acción de la enseñanza. Implica el proceso en el que se desarrollan las actividades de aprendizaje y las estrategias de enseñanza programadas por el docente, para que los estudiantes interactúen con los conocimientos previos y nuevos, con el fin de que logren transferirlos a distintas situaciones.

Evaluación del aprendizaje e implementación de las acciones establecidas. A este momento corresponde la reflexión acerca de los resultados obtenidos, es decir, acerca del logro de los aprendizajes esperados y los objetivos definidos para las sesiones de clase.

Evaluar las prácticas de enseñanza es un proceso de indudable importancia en todo sistema educativo, como un medio para generar insumos hacia la mejora del aprendizaje de los estudiantes y la calidad educativa (Benois *et al.*, 2016; Martínez-Rizo, 2016). Para evaluar las prácticas de enseñanza existen diversas estrategias como: cuestionarios de opinión para estudiantes (p. ej., Arregui *et al.*, 2018), observaciones en el aula (p. ej., Classroom Assessment Scoring System, CLASS; Pianta *et al.*, 2012), pruebas estandarizadas (p. ej., Encuesta internacional sobre docencia y aprendizaje, TALIS), y entrevistas a los directivos, entre otros. Sin embargo, como ha puntualizado la Organización para la Cooperación y el Desarrollo Económicos (OCDE, 2011), evaluar las prácticas de enseñanza debería ser un proceso formativo, a partir de una valoración reflexiva entre los evaluados, con

el fin de concientizarlos hacia la mejora de sus prácticas a corto plazo. Conocer las capacidades, habilidades y prácticas de enseñanza que desempeña el docente le permitirá identificar el potencial sobre las posibilidades de mejora en su labor (Cabrera y Ayala, 2010; Miranda y Martínez, 2015; Ravela, 2012). Para Nava y Rueda (2014), un proceso de evaluación docente debe tener credibilidad y utilidad para los profesores, puesto que con los resultados se espera que éstos identifiquen debilidades, las atiendan y mejoren su labor.

Dado lo anterior, la mejor estrategia para promover la reflexión en el docente sobre su práctica es la autoevaluación (Martínez-Rizo,

2012; Murillo *et al.*, 2007). No obstante, dicha estrategia no ha sido ampliamente desarrollada por la falta de uso en los sistemas educativos, así como la escasez de instrumentos que se adecuen al contexto en el que los docentes realizan su práctica (Díaz y Díaz, 2007; García, 2014; Martínez-Rizo, 2016). La autoevaluación docente ha primado en la educación superior, y son escasos los instrumentos que se identifican en educación básica. Por ejemplo, en una revisión exhaustiva de la literatura, sólo se lograron identificar ocho instrumentos para la autoevaluación docente en este nivel educativo, de los cuales únicamente uno está enfocado en secundaria (Cuadro 1).

Cuadro. 1. Instrumentos de autoevaluación de las prácticas de enseñanza

Instrumento	Objetivo	Dimensiones	Ejemplo ítem	Escala de respuesta	Propiedades psicométricas
Escala de autoevaluación del desempeño docente (Daoud, 2007)	Tener una mirada retrospectiva sobre la profesión docente	a) Didáctica b) Pedagogía c) Perfil psicológico d) Personalidad del docente	“17. Impulsa el trabajo en equipo”	Escala ordinal de cuatro puntos que va de nunca a siempre	No reportadas
Escala para la autorreflexión de la práctica pedagógica (Díaz y Díaz, 2007).	Analizar la acción docente, para profundizar sobre acciones o prácticas que se realizan adecuadamente en las aulas y cuáles son susceptibles de mejora	a) Programación de la enseñanza b) Metodología y uso de recursos c) Evaluación de los aprendizajes d) Formación y evaluación de la enseñanza e) Tutoría f) Atención a la diversidad g) Clima del aula	“5. Utilización de los recursos del medio”	Escala ordinal de cuatro puntos que va de mejorable a excelente	No reportadas
Cuestionario de autoevaluación de las competencias docentes (García-Cabreiro <i>et al.</i> , 2012)	Promover la reflexión, comprensión y análisis de las fortalezas y debilidades de la práctica docente	a) Planeación del curso b) Conducción de secuencias didácticas c) Evaluación de la progresión de los aprendizajes	“28. Enfaticé mediante el tono de voz la presentación de ideas importantes”	Escala ordinal de cuatro puntos que va de deficiente a muy satisfactorio	Cargas factoriales >.55. Fiabilidad de .87 (alfa de Cronbach)

Cuadro. 1. Instrumentos de autoevaluación de las prácticas de enseñanza

(continuación)

Instrumento	Objetivo	Dimensiones	Ejemplo ítem	Escala de respuesta	Propiedades psicométricas
Cuestionario para la autoevaluación de la práctica docente en escuelas primarias (Ramos et al., 2013)	Evaluar la práctica docente en escuelas primarias	a) Planeación b) Uso de los recursos c) Estrategias y actividades para promover aprendizajes significativos d) Estrategias y acciones para la evaluación e) Uso de los resultados de la evaluación f) Clima del aula	“2. Realizo una planeación escrita de mis clases”	Escala dicotómica con una valoración en sí y no	No reportadas
Instrumento de autoevaluación del docente (Akram y Zepeda, 2015)	Valorar las prácticas de enseñanza de docentes de secundaria	a) Conocimiento de la materia b) Planeación educativa y estrategias de instrucción c) Evaluación para el aprendizaje d) Entorno de aprendizaje e) Comunicación efectiva	“1. Demuestro un conocimiento preciso del tema”	Escala ordinal de cinco puntos que va de nunca a siempre	Fiabilidad de .94 e índices aceptables de bondad de ajuste (SRMR = .05, GFI = .86, CFI = .98, RMSEA = .056)
Rúbrica de autoevaluación del Modelo de desempeño docente de educación básica (Secretaría de Educación en Nuevo León, 2017)	Promover la mejora educativa y facilitar la implementación del modelo de desempeño docente en preescolar y primaria	a) Preparación del proceso enseñanza-aprendizaje b) Establecimiento de un ambiente favorable para el aprendizaje c) Desarrollo de las estrategias para el aprendizaje de los estudiantes d) Evaluación y mejora de la práctica de enseñanza-aprendizaje de los estudiantes e) Participación en la Ruta de Mejora Escolar y responsabilidades institucionales f) Resultados del desempeño docente g) Portafolio de evidencias	“2.1.1. Fomento de la comunicación”	Escala ordinal de cuatro puntos que va de insuficiente a excelente	No reportadas

Cuadro. 1. Instrumentos de autoevaluación de las prácticas de enseñanza

(continuación)

Instrumento	Objetivo	Dimensiones	Ejemplo ítem	Escala de respuesta	Propiedades psicométricas
Lista de cotejo para la autorreflexión de la práctica pedagógica (Kú y Pool, 2018)	Ser un elemento clave para la mejora y el desarrollo profesional en docentes de la educación básica	a) Conocimiento de los estudiantes sobre cómo aprenden b) Organización del trabajo educativo e intervención didáctica c) Apoyo a los estudiantes para aprender d) Características individuales, culturales y de contexto	“11. Establezco normas de comportamiento aceptables en conjunto con mis estudiantes”	Escala dicotómica con una valoración en sí y no	Cargas factoriales mayores a .54 y un índice de consistencia interna en .96 con el coeficiente alfa de Cronbach
Escala para la autoevaluación de las competencias docentes (Martínez-Izaguirre <i>et al.</i> , 2018).	Identificar debilidades y fortalezas de la enseñanza	a) Planificación y gestión educativa b) Gestión e implementación del currículo c) Evaluación educativa d) Orientación y tutoría e) Aprendizaje, investigación e innovación f) Ética y compromiso profesional g) Docente y trabajo en equipo en la comunidad educativa h) Gestión eficaz de las emociones y la creación de climas de confianza i) Comunicación	“4. Utiliza materiales didácticos que apoyen el aprendizaje”	Escala ordinal de cinco puntos, pero no se especifican las valoraciones	Consistencia interna de .97 (alfa de Cronbach)

Fuente: elaboración propia.

Se observa que los instrumentos identificados contienen un esquema y formato que se dirige más hacia un ejercicio de auto-calificación. También, dado el contenido acotado en los ítems, éstos no resultan descriptivos acerca de cómo debe de ser la práctica de enseñanza, ya que se perciben como más parecidos a una lista de cotejo (enunciados cortos y concretos dirigidos a determinar si se cumple o no con la tarea, y en algunos casos, en qué nivel se cumple) que a una autoevaluación. Por lo tanto, el diseño de estos instrumentos no logra del todo apoyar el verdadero propósito de una autoevaluación, como un recurso de reflexión en el que se identifican aspectos susceptibles de mejora. Además, pocos instrumentos reportan sus propiedades psicométricas; en

este caso, es cuestionable la validez de la interpretación de resultados que se realiza al utilizarlos. Desafortunadamente, como señalan Cisneros-Cohernour *et al.* (2012), en la literatura especializada en instrumentos de autoevaluación docente existe una escasez de información acerca de la validez e interpretación de los resultados. Aunado a ello, algunos instrumentos utilizan procedimientos de validez que no se ajustan a las características de tales instrumentos; por ejemplo, utilizar el coeficiente alfa de Cronbach para conjuntos de ítems que se encuentran en una escala ordinal, en cuyo caso debería aplicarse un alfa ordinal (Contreras y Novoa-Muñoz, 2018).

Si bien cualquier instrumento de evaluación debe ajustarse al contexto en el que éste

se utiliza, también es importante que los esquemas de evaluación permitan una reflexión que propicie realizar comparaciones y generalizaciones. Es menester que el instrumento retome los elementos esenciales de la enseñanza reconocidos por la literatura. En este sentido, varios autores han planteado que el marco para la enseñanza de Danielson (2013), es uno de los dos esquemas descriptivos de la enseñanza más referidos (Akram y Zepeda, 2015; Evans *et al.*, 2015; Leyva *et al.*, 2017), pues integra un marco universal para la evaluación de las prácticas que realiza el docente, y se adapta a cualquier contexto (Rowan y Raudenbush, 2016). Asimismo, el marco aporta mayor claridad sobre las acciones y comportamientos que distinguen la complejidad de la enseñanza y guían el logro del aprendizaje.

El marco para la enseñanza de Danielson (2013), fue el principal precursor de otros modelos; se rige bajo un enfoque constructivista de la enseñanza (Landolfi, 2016), es generalizable y constituye un referente teórico para cursos y programas de formación docente en países como Estados Unidos, Venezuela, Chile y Perú (Akram y Zepeda, 2015; Campoy y Xu, 2018; Colmenares, 2010; Evans *et al.*, 2015; Leyva *et al.*, 2017). Además, se reconoce que los elementos contenidos en este marco reflejan los criterios específicos y representativos para evaluar o autoevaluar las prácticas de enseñanza en cualquier nivel educativo (Rowan y Raudenbush, 2016). El marco de Danielson, por tanto, permite desarrollar procesos de autoevaluación y supervisión del desempeño docente (Vázquez *et al.*, 2014). Contempla cuatro dominios:

Planeación. Involucra aspectos que el docente debe considerar previamente a las sesiones de clase para organizar las experiencias de aprendizaje. Las prácticas que se revisan en este dominio son: conocimiento previo de los estudiantes, definición de estrategias de enseñanza y evaluación, recursos o materiales didácticos, entre otros.

Clima del aula. Comprende los aspectos socio-afectivos que intervienen en el proceso de enseñanza-aprendizaje. Algunas prácticas que se analizan son: gestión del aula, uso efectivo del tiempo de aprendizaje, manejo del comportamiento de los estudiantes, entre otros.

Enseñanza. Alude a la aplicación y regulación de los aspectos programados previamente en el momento de planeación. Las prácticas que se exploran son: promoción de la participación estudiantil en las actividades, claridad de las instrucciones y explicación del contenido, por mencionar algunos ejemplos.

Responsabilidades profesionales. Implica los aspectos presentes en la reflexión del proceso de enseñanza-aprendizaje. Algunas de las prácticas que se revisan son: la interacción con padres de familia y colegas, el progreso de los estudiantes, y el cumplimiento de actualización o capacitación profesional, entre otros.

Particularmente, en México se ha intentado diseñar un marco para la enseñanza a partir de los perfiles, parámetros e indicadores (PPI) —creados en el año 2014— para docentes y técnicos docentes en educación obligatoria, los cuales definen criterios generales para valorar el desempeño docente (SEP, 2017). Estos PPI están diseñados para cada uno de los niveles: preescolar, primaria, secundaria y educación media superior, aunque la mayoría tienen indicadores muy parecidos. La estructura de los PPI está organizada en cinco dimensiones que corresponden a las funciones generales que desempeñan los docentes en la educación obligatoria:

1. Un docente que conoce a sus alumnos, sabe cómo aprenden y lo que deben aprender.
2. Un docente que organiza y evalúa el trabajo educativo, y realiza una intervención didáctica pertinente.

3. Un docente que se reconoce como profesional que mejora continuamente para apoyar a los alumnos en su aprendizaje.
4. Un docente que asume las responsabilidades legales y éticas inherentes a su profesión para el bienestar de los alumnos.
5. Un docente que participa en el funcionamiento eficaz de la escuela y fomenta su vínculo con la comunidad para asegurar que todos los alumnos concluyan con éxito su escolaridad.

A partir de estas dimensiones se enuncian los parámetros e indicadores. Se espera que estos indicadores sirvan a los docentes para identificar los aspectos que deben desarrollar en su práctica cotidiana; sin embargo, la mayoría de tales indicadores son demasiado generales y no permiten al docente ubicar de

forma clara lo que se espera de él. Por ejemplo, en el Cuadro 2 se muestran los parámetros e indicadores correspondientes a la dimensión 1 de los PPI. Como puede verse, ninguno de los indicadores utiliza verbos específicos, sino verbos generales que dan pie a otros verbos (p. ej. identifica, domina, describe). En segundo lugar, incluyen términos que pueden dar cabida a diferentes interpretaciones o definiciones: por ejemplo “proceso de desarrollo”, “carácter formativo”, “principios pedagógicos”, entre otros. Por lo tanto, es probable que estas descripciones, tan poco específicas, obedezcan al interés de que el docente reflexione y derive, por sí mismo, lo que podría aplicarse para cada indicador de acuerdo a su contexto y realidad. No obstante, este propósito no se logra, y es común encontrar docentes que manifiestan la falta de certidumbre respecto a lo que deberían hacer para atender lo que marcan los PPI.

Cuadro 2. Ejemplo de parámetros e indicadores de los PPI en la dimensión 1

Parámetros	Indicadores
1.1 Conoce los procesos de desarrollo y de aprendizaje de los adolescentes	1.1.1 Identifica los procesos de desarrollo y aprendizaje de los adolescentes como referentes para conocer a los alumnos 1.1.2 Reconoce que en los procesos de desarrollo y aprendizaje de los adolescentes influyen factores familiares, sociales y culturales
1.2 Domina los propósitos educativos y los contenidos escolares de su asignatura en la educación secundaria	1.2.1 Explica el carácter formativo de los propósitos educativos del currículo vigente 1.2.2 Domina los contenidos de su asignatura en la educación secundaria 1.2.3 Describe la progresión de los contenidos educativos de su asignatura para favorecer el aprendizaje de los alumnos 1.2.4 Relaciona los contenidos de aprendizaje de su asignatura en la educación secundaria para el logro de los propósitos educativos
1.3 Explica los referentes pedagógicos y los enfoques didácticos del currículo vigente	1.3.1 Explica los principios pedagógicos que orientan la práctica docente en educación básica 1.3.2 Identifica las características del enfoque didáctico de su asignatura de la educación secundaria en actividades de aprendizaje

Fuente: elaboración propia a partir de SEP, 2017.

Además, los indicadores de los PPI no especifican las diferencias sobre las prácticas de enseñanza que caracterizan a cada nivel educativo; pareciera que en todos los niveles aplican los mismos parámetros. Autores como Córdoba (2016) y Pedroza (2016) afirman que

los PPI son complejos en su redacción y requieren precisar cuáles son los principios que los guían, y cómo es que el docente llega a tener un desempeño destacado para poder evaluarlo en cada nivel.

Con base en lo expuesto, se puede afirmar que existe una ausencia de instrumentos de autoevaluación de las prácticas de enseñanza en la educación secundaria; que los pocos reportados muestran escasa evidencia de sus propiedades psicométricas e inconsistencia de procedimientos de validez; y que es importante considerar que cualquier instrumento de autoevaluación de la práctica debe partir de esquemas generales de evaluación. Es así que se consideró pertinente el desarrollo de una propuesta evaluativa de las prácticas de enseñanza dirigida a docentes de secundaria, que atienda las limitaciones que se han identificado en la revisión de literatura. En consecuencia, el presente estudio tuvo como propósito desarrollar y validar una escala para la autoevaluación reflexiva de las prácticas de enseñanza en educación secundaria, sustentada en los indicadores de Danielson, los cuales, de acuerdo con lo referido en la literatura, incluyen prácticas de enseñanza de carácter universal.

ETAPAS DEL ESTUDIO

Dada la complejidad del constructo que se pretende evaluar, el procedimiento para el

desarrollo y generación de evidencias de validez requirió de tres etapas (Fig. 1). En la primera, se describe el proceso para el desarrollo del instrumento, que consistió en tres fases. Para la segunda etapa se generaron evidencias de validez de contenido por medio de un proceso iterativo, a partir de tres revisiones y ajustes a los ítems, a fin de asegurar que éstos fueran lo suficientemente descriptivos, pertinentes, congruentes y claros para un proceso de autoevaluación. En la tercera etapa se presentan evidencias relacionadas con la fiabilidad y estructura interna de la escala.

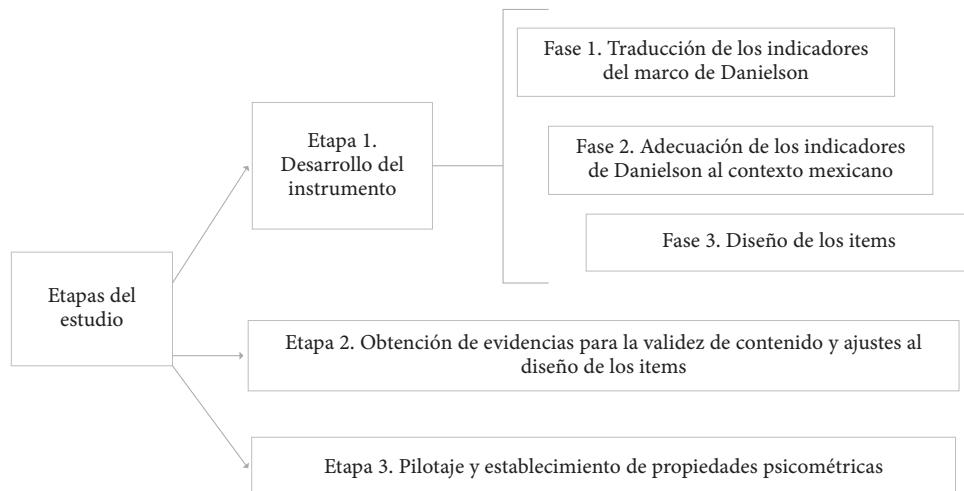
Etapa 1. Desarrollo del instrumento

Para desarrollar los ítems de la escala se siguió un proceso de construcción en tres fases: traducción de los indicadores del marco de Danielson (2013), adecuación de dichos indicadores al contexto mexicano y diseño de los ítems.

Fase 1. Traducción de los indicadores del marco de Danielson

Los 104 indicadores del marco de Danielson se tradujeron al español y posteriormente fueron revisados por cinco especialistas en

Figura 1. Etapas para el desarrollo y validación de una escala para la autoevaluación de las prácticas de enseñanza en secundaria



Fuente: elaboración propia.

el constructo, todos bilingües. Acorde con las recomendaciones de la International Test Comission (2017), se realizó una evaluación directa por medio de un panel de expertos. Se hicieron modificaciones cuando la traducción no era fiel a la conceptualización de la versión original en inglés, o no era adaptable al lenguaje que se utiliza en el contexto educativo mexicano. Al revisarse la calidad de la traducción se identificó que todos los indicadores requerían de una modificación por convenciones culturales en algunos términos (p. ej., “lecciones” por “sesiones de clase”; “distrito” por “sistema educativo”). Asimismo, se realizaron algunos ajustes de redacción para asegurar una mayor claridad de la información (p. ej., “eventos cotidianos” por “situaciones de la vida diaria”; “adecuado para la lección” por “apropiado para la clase”).

Fase 2. Adecuación de los indicadores de Danielson al contexto mexicano

Para esta fase se requirió de una revisión directa de la traducción de los indicadores resultantes de la Fase 1, mediante una sesión plenaria con cuatro docentes de secundaria. Las modificaciones realizadas se relacionaban con inconsistencias en los términos culturales, ajustes en la redacción y eliminación de indicadores que no eran aplicables al contexto educativo mexicano. En esta etapa se hicieron modificaciones a los indicadores respecto de algunos términos que se emplean para el sistema educativo mexicano en secundaria (p. ej., “disciplina” por “materia”; “maestro” por “docente”). Además, se eliminaron cinco indicadores, dos por redundancia con otros y tres por no aplicar al contexto mexicano (p. ej., 40. Fomento de la proximidad física).

Fase 3. Diseño de los ítems

Los 99 indicadores revisados en la Fase 2 sirvieron como insumo para el desarrollo de los ítems. Dado que se pretendía que el instrumento fuera altamente descriptivo, a fin de que permitiera la autorreflexión del docente, se buscó que cada uno de los ítems se redactara tan específico como fuera posible. Así, se escribió un ítem por cada indicador. En el desarrollo de los ítems se consideró el apoyo de cuatro docentes de secundaria para incorporar ejemplos de actividades y materiales que se utilizan en las prácticas descritas en el indicador (p. ej., “para evaluar los aprendizajes esperados de los alumnos empleo diferentes estrategias, como: rúbricas, exámenes, cuestionarios, debates, ensayos, listas de cotejo, escalas de rango, entre otros”).

Etapa 2. Obtención de evidencias para la validez de contenido y ajuste al diseño de los ítems

El objetivo de esta fase fue evaluar la claridad, pertinencia y suficiencia de los ítems desarrollados. Dada la complejidad del constructo, se requirió de tres revisiones a fin de asegurar que los ítems fueran lo suficientemente descriptivos de las prácticas de enseñanza, pero que a su vez resultaran congruentes y pertinentes con el constructo a evaluar, así como claros y suficientes. En la primera revisión se envió un formato de validación, vía correo electrónico, a 10 docentes de secundaria, para que indicaran, en cada ítem, la claridad, la congruencia y la pertinencia con el constructo, así como la suficiencia de los 99 ítems para la autoevaluación de las prácticas de enseñanza en secundaria (producto final de la etapa 1). Como resultado, se determinó eliminar seis ítems que no cumplieron con el índice de validez de contenido (IVC) (Lawshe, 1975) al mostrar valores de -0.20 (Cuadro 3).

Cuadro 3. Resultados de la validación a partir del IVC de Lawshe en la revisión 1

Dominio	Ítems	IVC Lawshe
Planeación y preparación (ítems 1-32)	1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 18, 19, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32	1
	2, 16, 17	0.80
	20	0.60
Clima del aula (ítems 33-57)	33 al 57	1
Enseñanza (ítems 58-79)	58 al 79	1
Responsabilidades profesionales (ítems 80-99)	80, 81, 82, 82, 83, 85, 86, 87, 88, 89, 90, 92, 93, 94	1
	91, 96, 97, 98, 99	-0.20

Fuente: elaboración propia.

Sin embargo, como producto de esta revisión se identificó que los ítems no resultaban totalmente descriptivos, y requerían una mayor especificación para poder promover el autoanálisis de la práctica de enseñanza. Además, también se reconocieron ítems que llegaban a resultar redundantes, o incluso complementarios, y que, por lo tanto, deberían integrarse en uno solo. Esto dio pie a una segunda revisión.

Para esta segunda revisión se contó con la ayuda de cinco especialistas del constructo, quienes acordaron ajustar los 93 ítems (resultado de la revisión 1) de manera que todos manifestaran el qué, cómo y para qué. Es decir, que el ítem enunciara la práctica, describiera la forma de llevarla a cabo (ejemplos de estrategias, acciones, actividades, así como materiales que pueden utilizarse) y mencionara para qué (el propósito de la ejecución descrita). Producto de esta revisión, se determinó la necesidad de incorporar el dominio “evaluación del aprendizaje”, así como dos indicadores complementarios sobre la “organización de grupos en el aula” y “difusión de los criterios de evaluación”. Por último, en esta segunda revisión se eliminaron 25 ítems que se consideraron redundantes. Como producto de esta segunda revisión se derivó un total de 70 ítems, y dado este cambio fue necesario que docentes en servicio, nuevamente revisaran los ítems. Así se realizó la tercera revisión.

En ella se contó con el apoyo de 203 docentes de secundaria, de los cuales 97 por ciento afirmó que los ítems resultan útiles para autoevaluar las prácticas de enseñanza, debido a que les permiten reconocer fortalezas y aspectos que mejorar en su práctica. El 81 por ciento de los docentes mencionó que los ítems son aplicables en su contexto, y 88 por ciento que son suficientes para autoevaluar sus prácticas de enseñanza. Algunos comentarios vertidos en esta revisión de los ítems por los docentes fueron: “la escala abarca más ámbitos que las guías proporcionadas por la Secretaría de Educación” (CRE90-1); “permiten visualizar los aspectos que sí trabajamos y cuáles estamos dejando de utilizar” (ST4031-32 y CH40-5); “la escala es muy útil para todos los docentes que estamos frente a grupo” (ST1028-30), entre otros.

Etapa 3. Pilotaje y establecimiento de propiedades psicométricas

El propósito de esta etapa fue pilotear y generar evidencias de validez y fiabilidad de un conjunto de 70 ítems generados en la etapa 2.

Participantes

En el pilotaje participaron 203 docentes de secundaria: 56 desempeñaban su labor en secundarias de modalidad general, 125 en técnica

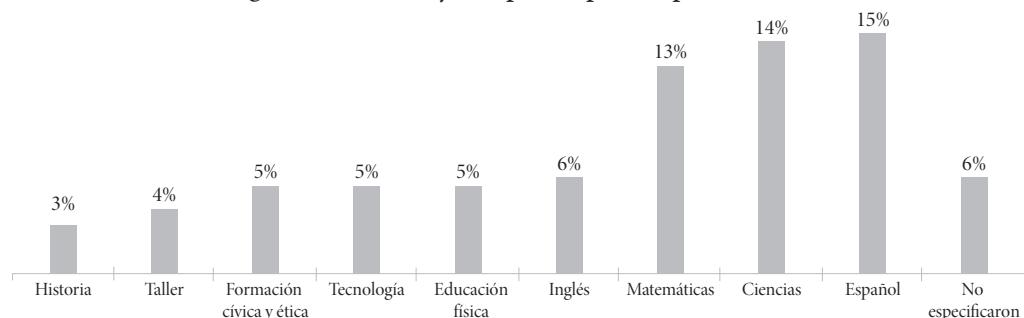
y 22 en privadas; al momento del estudio tenían de 1 a 33 años de experiencia docente en distintas materias de secundaria del sistema educativo mexicano. Entre los participantes, 3 por ciento impartía la materia de Tutorías, Computación o Asignatura estatal; 4 por ciento Artes o Geografía; y 16 por ciento estaban a cargo de dos y cuatro materias en distintos grados. Como se muestra en la Fig. 2, el mayor porcentaje de participación fueron docentes de Español, Matemáticas y Ciencias

(distribuidos entre las materias de Biología, Química y Física).

Instrumento

El instrumento conjuntó 70 ítems derivados de la etapa 2, distribuidos en cinco dominios, los cuales se presentan en el Cuadro 4. En la última columna se presenta un ejemplo de un ítem en cada dominio. Como se puede observar, cada ítem incluía el qué, el cómo y el para qué.

Figura 2. Porcentaje de participantes por materia



Fuente: elaboración propia.

Cuadro 4. Dominios e ítems en la escala de autoevaluación diseñada

Dominios	Ítems	k	Ejemplo de ítem
1. Planeación y preparación	1 al 13	13	4. Para adecuar las actividades, dentro o fuera del aula, a las características de los alumnos (capacidades, habilidades, estilos de aprendizaje, ritmos de trabajo, etc.) utilizo diferentes estrategias (modifico los materiales, la metodología de trabajo, los criterios de evaluación y los productos de la actividad) con el fin de promover la equidad y proporcionar a los alumnos lo que necesiten para el logro de los aprendizajes esperados.
2. Clima del aula	14 al 26	13	22. Con el propósito de mantener el orden y la disciplina dentro del aula, acuerdo con los alumnos las reglas de conducta que habremos de seguir durante el ciclo escolar (mantener el aula limpia y organizada, respetar a los compañeros y docente, solicitar turno para hablar, etc.).
3. Enseñanza	27 al 49	23	30. Para asegurarme de que los alumnos reciban instrucciones claras, describo de forma detallada y concreta lo que solicito, utilizando un lenguaje sencillo, ya sea oral o escrito, y ejemplifico lo que sí y lo que no deben realizar. Además, verifico que los alumnos comprenden la instrucción haciendo preguntas que me permitan identificar que la instrucción quedó clara.
4. Evaluación del aprendizaje	50 al 59	10	54. Llevo un registro (bitácora, guía de observación, portafolio de evidencias, escala de actitudes, diario de trabajo, etc.) con las actividades y adecuaciones curriculares necesarias que requieren los alumnos para ir valorando su avance.
5. Responsabilidades profesionales	60 al 70	11	67. Como parte de mi compromiso y desarrollo profesional, busco y participo en diversas iniciativas de formación continua (cursos, talleres, comunidades de aprendizaje, diplomados, etc.), para mantenerme actualizado y mejorar mi práctica docente.

Fuente: elaboración propia.

Para contestar cada uno de los ítems se incluyó una escala de cuatro puntos: a) nulo; b) básico; c) intermedio; y d) avanzado (Cuadro 5). Estos niveles representan el nivel de

aplicación de la práctica referida en cada ítem. El tiempo estimado para contestar el instrumento completo fue de 30 a 45 minutos.

Cuadro 5. Descripción de los niveles en la escala de respuesta

A. Nulo	Mi práctica de enseñanza refleja una mínima correspondencia con el contenido que se describe. Es decir, no aplico o realizo lo mencionado y, por lo tanto, no cuento con evidencias de esta práctica.
B. Básico	Mi práctica de enseñanza es similar a la descripción que se presenta. Es decir, en algunas ocasiones realizo lo mencionado y cuento con algunas evidencias de esta práctica.
C. Intermedio	Mi práctica de enseñanza tiene una alta correspondencia con el contenido que se describe. Es decir, realizo con frecuencia lo mencionado y cuento con varias evidencias de esta práctica.
D. Avanzado	Mi práctica de enseñanza se lleva a cabo tal como se describe, pero incorporo otros elementos que la enriquecen. Es decir, siempre realizo lo que se menciona y cuento con las evidencias suficientes para demostrarlo.

Fuente: elaboración propia.

Procedimiento

El pilotaje de la guía se aplicó individualmente tanto de forma presencial como virtual. Para la aplicación presencial se visitó a las escuelas y se solicitó autorización al directivo del plantel y a los propios docentes para que contestaran el instrumento. A los docentes se les proporcionó un cuadernillo de la escala para la autoevaluación, en el que se presentaba el objetivo y elementos del instrumento, las instrucciones y una lista de los ítems enumerados por dominio. Asimismo, se entregó una hoja de respuestas para solicitar los datos generales de los participantes y el listado con el número de ítem junto con la escala de respuesta (Fig. 3). En la mayoría de los casos, los docentes contestaron en un espacio fuera del salón de clases, mientras un miembro

del equipo de investigación o el prefecto de la escuela se quedaba con los estudiantes. En el caso de la aplicación virtual, el instrumento se registró temporalmente en formularios de Google, con una presentación similar al cuadernillo, y el enlace se compartió en grupos cerrados de enseñanza en secundaria, registrados en Facebook. Se hizo una invitación pública explicando el objetivo del instrumento y los elementos del mismo.

ANÁLISIS DE DATOS

Se llevaron a cabo dos procedimientos para analizar las propiedades psicométricas del conjunto de ítems.

- *Fiabilidad de las puntuaciones.* Se examinó el comportamiento y la estabili-

Figura 3. Ejemplo de cómo respondieron los docentes la escala de autoevaluación

55. Para evaluar los aprendizajes esperados de los alumnos empleo diferentes herramientas, tales como: rúbricas, exámenes, cuestionarios, debates, ensayos, listas de cotejo, escalas de rango, entre otros.

55 A B C D

Fuente: elaboración propia.

dad interna de las puntuaciones a partir del cálculo del coeficiente alfa ordinal. Se empleó debido a que los ítems se responden en una escala de valoración ordinal. Para estimar los coeficientes previamente mencionados se utilizó el paquete *psych* del programa estadístico *R*, y el cálculo se hizo por medio de matrices policóricas.

- *Análisis sobre la dimensionalidad de la escala diseñada.* Dado que el proceso de desarrollo de los ítems se realizó en apego a la estructura del marco de enseñanza de Danielson, se consideró pertinente realizar un análisis factorial confirmatorio (AFC); con ello se buscó reconocer la bondad de ajuste del conjunto de ítems con el constructo, e identificar el nivel de correlación de los ítems en cada factor (Campillo-Arias *et al.*, 2012). Para calcular el AFC se utilizó el método sobre la estimación de mínimos cuadrados ponderados

diagonalmente (WLSMV) aplicable para datos ordinales, como en este caso, utilizando el programa estadístico Mplus Editor.

RESULTADOS

Fiabilidad de las puntuaciones. A partir de matrices policóricas, se identificó una consistencia interna alfa ordinal de .98 en el conjunto de ítems. Asimismo, se calculó el alfa ordinal para cada uno de los factores (dominios de acuerdo al marco de Danielson), los cuales oscilaron entre .920 y .961 (Tabla 1).

Análisis sobre la dimensionalidad de la escala diseñada. Como se observa en la Tabla 1, se presenta un concentrado de las cargas factoriales y la fiabilidad de los resultados del AFC. Además, en la Fig. 4 se aprecia por completo el modelo teórico-hipotético de la escala de autoevaluación.

Tabla 1. Cargas factoriales y fiabilidad del modelo teórico-hipotético

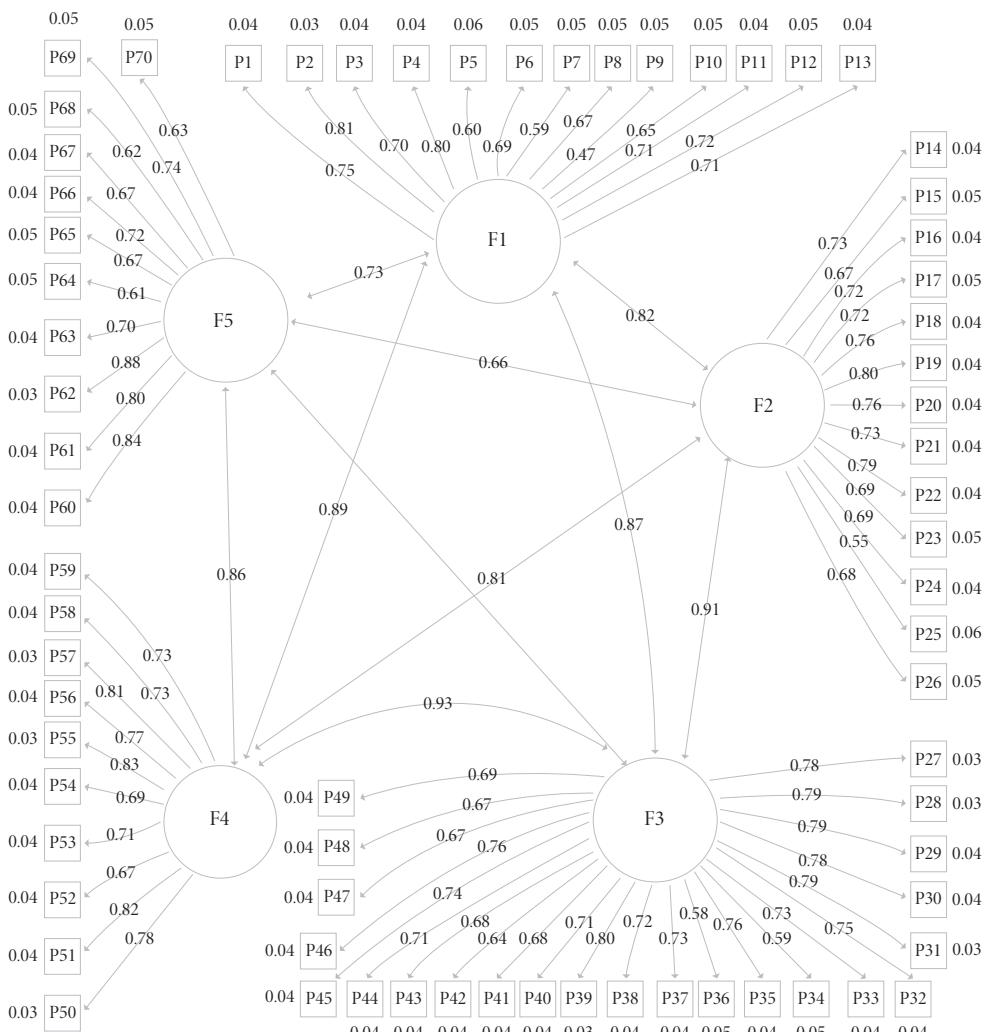
Factor	k	Ítems	Cargas factoriales	Fiabilidad
1. Planeación	13	1 al 13	.466 - .810	.920
2. Clima del aula	13	14 al 26	.549 - .796	.931
3. Enseñanza	23	27 al 49	.584 - .803	.961
4. Evaluación	10	50 al 59	.670 - .832	.929
5. Responsabilidades	11	60 al 70	.609 - .877	.921

Fuente: elaboración propia.

Asimismo, quedó confirmado que el conjunto de ítems diseñados para cada dominio guardó correspondencia con éste, lo que se

comprobó a partir de que los índices de bondad de ajuste resultaron en niveles aceptables (Tabla 2).

Figura 4. Modelo teórico-hipotético de la escala para la autoevaluación de las prácticas de enseñanza en secundaria



Fuente: elaboración propia.

Tabla. 2. Índices de bondad de ajuste del modelo propuesto

Prueba	Estadístico	Criterio*	Resultado
Raíz del residuo estandarizado medio	SRMR	Cercano a cero	.068
Error cuadrático medio de aproximación	RMSEA	<.05	.035
Índice de ajuste comparativo	CFI	> .95	.960
Índice de Tucker-Lewis	TLI	> .95	.960

Fuente: elaboración propia.

*Nota: los criterios adecuados para cada medida de bondad de ajuste se tomaron de Brown (2006).

DISCUSIÓN Y CONCLUSIONES

El propósito del estudio consistió en desarrollar y validar una escala para la autoevaluación de las prácticas de enseñanza en secundaria. Los resultados en el presente estudio demuestran que la escala diseñada constituye una herramienta válida, útil y pertinente para la valoración de dichas prácticas. Esta escala, a diferencia de otros instrumentos de autoevaluación, dispone de ítems con una descripción detallada sobre lo que el docente debe realizar para mejorar su desempeño y prácticas de enseñanza hacia el logro del aprendizaje; además, constituye una herramienta alternativa y complementaria a los PPI, pues integra elementos de la enseñanza de una forma descriptiva y clara, y guarda correspondencia con las dimensiones de éstos. Así, la escala permite que los docentes realicen una reflexión autocritica sobre aspectos susceptibles de mejora a corto plazo. Incluso, este instrumento es valorado por docentes de secundaria como una guía de estudio para los exámenes dedicados a la evaluación del desempeño docente; por lo tanto, es una herramienta que tiene credibilidad y utilidad para enriquecer la práctica de docentes de secundaria (Nava y Rueda, 2014).

La escala, además de ser una aportación importante a la investigación educativa, contribuye a los procesos de formación y evaluación docente, no sólo debido a las evidencias de validez presentadas, sino a la relevancia de los dominios que la integran, según lo que la literatura en el área señala. En consecuencia, a partir de la estructura del marco para la enseñanza de Danielson, pero integrando el contenido de los ítems que conformaron cada uno de los factores, el instrumento que se presenta permite valorar los siguientes dominios:

Planeación y preparación. Este dominio integra los aspectos que influyen en cómo el docente organiza sus sesiones de clase hacia el logro de los aprendizajes esperados. Los ítems que conforman el dominio hacen referencia

al conocimiento del docente sobre la materia y cómo adquiere información de los estudiantes para considerar sus conocimientos previos y adecuar las actividades. Además, se recuperan prácticas sobre el fomento y utilidad de recursos para el proceso de enseñanza-aprendizaje, así como el diseño de actividades pedagógicas coherentes para la materia. Las prácticas que se consideraron en este dominio concuerdan con las acciones que describen Anijovich y Mora (2009) en los tres momentos del proceso pedagógico, así como con la dimensión 1 de los PPI denominada “un docente que conoce a sus alumnos, sabe cómo aprenden y lo que deben aprender” (SEP, 2017).

Clima del aula. Contiene las prácticas y acciones que realiza el docente para propiciar un ambiente favorable en el proceso de enseñanza-aprendizaje hacia el logro de los objetivos planteados. En este dominio se agrupan ítems relacionados con la promoción de un entorno de respeto y armonía, mediante una interacción empática, equitativa, con comunicación y confianza entre los estudiantes y el docente. A través de estos ítems también se recupera información acerca de las acciones que permiten el establecimiento de una cultura de aprendizaje, por ejemplo: el reconocimiento de las expectativas y esfuerzo de los estudiantes, el manejo de grupo y la intervención en situaciones de mala conducta. Las prácticas descritas en este apartado guardan relación con los PPI en la dimensión 2: “Un docente que organiza y evalúa el trabajo educativo, y realiza una intervención didáctica pertinente” y 4: “Un docente que asume las responsabilidades legales y éticas inherentes a su profesión para el bienestar de los alumnos” (SEP, 2017).

Enseñanza. Agrupa los aspectos relacionados con la capacidad del docente para proporcionar explicaciones e instrucciones claras a los estudiantes y adecuar escenarios de aprendizaje hacia el logro de los objetivos de la materia. Los ítems contenidos en este dominio

dan cuenta de las habilidades que tiene el docente para involucrar a los estudiantes en el aprendizaje, gestionar los procesos del aula y utilizar actividades variadas que propicien la participación de los estudiantes, así como el aprendizaje significativo de los contenidos en la materia. Las prácticas que conforman este dominio concuerdan con los indicadores establecidos en los marcos para la buena enseñanza y los PPI (SEP, 2017). Además, este dominio coincide con las dimensiones e ítems de otros instrumentos para la autoevaluación de la enseñanza en educación básica (Akram y Zepeda, 2015; Daoud, 2007; García-Cabrero *et al.*, 2012; Kú y Pool, 2018; Ramos *et al.*, 2013).

Evaluación del aprendizaje. Este dominio es una aportación del presente estudio al marco de la enseñanza de Danielson (2013). En este nuevo dominio se concentraron todos los ítems dirigidos hacia la evaluación, seguimiento y retroalimentación del proceso de enseñanza-aprendizaje, los cuales estaban distribuidos entre los demás dominios. Dichos ítems recuperan información respecto a la definición, comunicación, diseño y adecuación de estrategias y acciones dirigidas a evaluar el aprendizaje de los alumnos. Asimismo, se presentan prácticas que el docente efectúa para el monitoreo y retroalimentación de la comprensión del estudiante. Con este nuevo dominio la escala concuerda con la estructura que se sigue en otros instrumentos de autoevaluación, como el de Kú y Pool (2018) y Ramos *et al.* (2013), en los que se recomienda colocar un dominio exclusivo para los aspectos de evaluación. Las prácticas de este apartado también tienen coincidencia con los aspectos que se abordan en los PPI, específicamente en la dimensión 2: “Un docente que organiza y evalúa el trabajo educativo, y realiza una intervención didáctica pertinente” (SEP, 2017).

Responsabilidades profesionales. Integra aspectos relacionados con las acciones o actividades que implementa el docente para mejorar sus prácticas de enseñanza y desarrollo profesional. Los ítems que conforman este dominio dan cuenta de cómo reflexiona el docente acerca del proceso de enseñanza-aprendizaje, qué estrategias utiliza para registrar el progreso de los estudiantes y cómo se comunica con los padres de familia. Estos ítems retoman aspectos acerca de la participación que tiene el docente en la comunidad profesional, en cursos de formación y proyectos de investigación educativa. El contenido de este dominio tiene similitud con los elementos que se valoran en la escala para la autoevaluación de las competencias docentes de Martínez-Izaguirre *et al.* (2018). Aunado a lo anterior, las prácticas que se consideraron en este dominio concuerdan con los indicadores propuestos en la dimensión 3 de los PPI: “Un docente que se reconoce como profesional que mejora continuamente para apoyar a los alumnos en su aprendizaje” (SEP, 2017).

Se puede concluir que la estructura que presenta el instrumento resulta útil para la autoevaluación de las prácticas de enseñanza y permite una reflexión descriptiva de los aspectos que pueden mejorarse. Además, si bien las prácticas incluidas en los ítems están estructuradas para el contexto mexicano, éstas parten de un marco de enseñanza universal, lo que hace factible su aplicación en otros contextos, así como la comparación con otras medidas del constructo. Sin embargo, se reconoce como limitación del estudio el tamaño y características de la muestra, que afectaron la posibilidad de realizar otros análisis, por ejemplo, acerca del funcionamiento diferencial de los ítems y de invarianza factorial. Por lo anterior, es necesario continuar con los estudios de la utilidad del instrumento en nuevas poblaciones en las que se determinen estos aspectos.

REFERENCIAS

- AKRAM, Muhammad y Sally Zepeda (2015), "Development and Validation of a Teacher Self-assessment Instrument", *Journal of Research and Reflections in Education*, vol. 9, núm. 2, pp. 134-148.
- ANIJOVICH, Rebeca y Silvia Mora (2009), *Estrategias de enseñanza. Otra mirada al quehacer en el aula*, Buenos Aires, Aique Educación.
- ARREGUI Eaton, Irma Gloria, Alicia Alelí Chaparro Caso-López y Carlos David Díaz López (2018), "Cuestionario para valorar las prácticas de enseñanza en secundaria desde la percepción de los estudiantes", *Revista Española de Orientación y Psicopedagogía*, vol. 29, núm. 2, pp. 55-70.
- BENOIS Muñoz, Norma, Fabiola Gómez Cauch, Maritza Briceño Caballero y Jessica Zumárraga Ávila (2016), "El Programa de Evaluación Docente en la Universidad Autónoma de Yucatán. Experiencia de la implementación", *Revista Iberoamericana de Evaluación Educativa*, vol. 9, núm. 1, pp. 161-179. DOI: <https://doi.org/10.15366/riee2016.9.1.010>
- BROWN, Timothy (2006), *Confirmatory Factor Analysis for Applied Research*, Nueva York, The Guilford Press.
- CABRERA Ortiz, Roberto y Eliseo Ayala Valdés (2010), "Evaluación del docente del siglo XXI", ponencia presentada en el "Foro Nacional de Ciencias Básicas. Selección y perfeccionamiento de profesores", México, UNAM, 13-15 de octubre de 2010, en: http://www.dcb.unam.mx/Eventos/Foro4/Memorias/Ponencia_51.pdf (consulta: 22 de abril de 2018).
- CAMPO-Arias, Adalberto, Edwin Herazo y Heidi Oviedo (2012), "Análisis de factores: fundamentos para la evaluación de instrumentos de medición en salud mental", *Revista Colombiana de Psiquiatría*, vol. 41, núm. 3, pp. 659-671.
- CAMPOY, Renee y Yuejin Xu (2018), "Using Teaching Videos to Integrate the Danielson Framework for Teaching into Secondary Certification Programs", *Southeastern Regional Association of Teacher Educators (SRATE) Journal*, vol. 27, núm. 2, pp. 1-10.
- CID-Sabucedo, Alfonso, Adolfo Pérez-Abellás y Miguel Zabalza (2009), "Las prácticas de enseñanza declaradas de los mejores profesores de la Universidad de Vigo", *Revista Electrónica de Investigación y Evaluación Educativa*, vol. 15, núm. 2, pp. 1-29.
- CISNEROS-Cohernor, Edith, María Jorquera Jaramillo y Ángel Aguilar Riverol (2012), "Validación de instrumentos de evaluación docente en el contexto de una universidad española", *Voces y Silencios: Revista Latinoamericana de Educación*, vol. 3, núm. 1, pp. 41-55.
- COLMENARES Gilly, Luis (2010), "Procesos de mejora y cambio en una escuela venezolana", *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, vol. 8, núm. 5, pp. 154-168.
- CONTRERAS Espinoza, Sergio y Francisco Novoa-Muñoz (2018), "Ventajas del alfa ordinal respecto al alfa de Cronbach ilustradas con la encuesta AUDIT-OMS", *Pan American Journal of Public Health*, vol. 42, núm. 65, pp. 1-8. DOI: <https://doi.org/10.26633/RPSP.2018.65>
- CÓRDOBA, Perla (2016), *Análisis de la alineación entre el "Perfil, parámetros e indicadores de desempeño docente" y el "Programa de estudios de preescolar 2011. Guía para la educadora"*, Tesis de Maestría, Ensenada, Universidad Autónoma de Baja California.
- DANIELSON, Charlotte (2013), *The Framework for Teaching Evaluation Instrument*, Princeton, The Danielson Group-Princeton University.
- DAOUD, Adrián (2007), "Propuesta de autoevaluación docente", *Revista de Informática Educativa y Medios Audiovisuales*, vol. 4, núm. 9, pp. 15-27.
- DÍAZ Alcaraz, Francisco y Juan Díaz Canals (2007), "Modelo para autoevaluar la práctica docente de los maestros de infantil y primaria", *Ensayos*, núm. 22, pp. 155-201.
- EVANS, Brian, Fran Wills y Megan Moretti (2015), "Editor and Section Editor's Perspective Article: A look at the Danielson framework for teacher evaluation", *Journal of the National Association for Alternative Certification*, vol. 10, núm. 1, pp. 21-26.
- GARCÍA, María (2014), *Modelos de evaluación docente en el mundo e instrumentos utilizados para evaluar: ventajas, desventajas y consideraciones para su elección*, Santiago de Chile, Pontificia Universidad Católica de Chile/MIDE.
- GARCÍA-Cabrero, Benilde, Javier Loredo Enríquez y Guadalupe Carranza Peña (2008), "Análisis de la práctica educativa de los docentes: pensamiento, interacción y reflexión", *Revista Electrónica de Investigación Educativa*, vol. 10, número especial, pp. 1-15, en: <http://redie.uabc.mx/NumEsp1/contenido-garcialoredocarranza.html> (consulta: 26 de julio de 2018).
- GARCÍA-Cabrero, Benilde, Alejandra Valencia y Vanía Pineda (2012), "Diseño y validación de un instrumento para la autoevaluación de las competencias docentes", *Revista Iberoamericana de Evaluación Educativa*, vol. 5, núm. 1, pp. 66-83.
- Gobierno de México-Secretaría de Educación Pública (SEP) (2017), *Perfil, parámetros e indicadores para docentes y técnicos docentes*, México, SEP.

- Gobierno de Nuevo León-Secretaría de Educación del Estado de Nuevo León (SEENL) (2017), *Autoevaluación del modelo de desempeño docente de educación básica*, México, SEENL.
- International Test Commission (ITC) (2017), *The ITC Guidelines for Translating and Adapting Tests*, Lincoln (EUA), ITC.
- KÚ Hernández, Omar y Wilson Pool Cibrián (2018), “Construcción y validación de instrumentos para la evaluación de la práctica pedagógica en educación básica del estado de Yucatán”, *Revista Iberoamericana de Evaluación Educativa*, vol. 11, núm. 1, pp. 23-41. DOI: <https://doi.org/10.15366/rie2018.11.1.002>
- LANDOLFI, Adrienne (2016), *An Exploration of Teachers and Administrators. Perspectives: The collaborative process using the Danielson framework for teaching model*, Tesis Doctoral, Fort Lauderdale, Keiser University.
- LAWSHÉ, Chad (1975), “A Quantitative Approach to Content Validity”, *Personnel Psychology*, vol. 28, núm. 4, pp. 563-575. DOI: <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>
- LEYVA Barajas, Yolanda, Moramay Guerra García, Mirna Santamaría Barreto, María Dander Flores, Blanca Jiménez Arteaga, Elvia Ortega Soriano, Jorge Lavin García, Susana Quinto Simón, Rosa León Manffer, Claudia Garduño Néstor, Alan López Alcántara, Olivia García de León, Karina Delgado Añorve, Diego Matus Ortega y Ulises Sandoval Velázquez (2017), “Caracterización de la práctica docente”, en Eduardo Backhoff y Gilberto Guevara (coords.), *La educación obligatoria en México. Informe 2017*, México, INEE, pp. 111-134.
- LUNA Serrano, Edna, Leticia Elizalde Lora, Alma Torquemada, Antelmo Castro López y Edith Cisneros-Cohernour (2013), “Evaluación de la docencia en el sistema educativo nacional 2002-2012”, en Alma Maldonado (comp.), *Educación y ciencia: políticas y producción del conocimiento 2002-2011*, México, ANUIES, pp. 399-454.
- MARTÍNEZ-Izaguirre, Miryam, Concepción Yániz-Álvarez y Lourdes Villardón-Gallegos (2018), “Autoevaluación y reflexión docente para la mejora de la competencia profesional del profesorado en la sociedad del conocimiento”, *Revista de Educación a Distancia*, vol. 10, núm. 56, pp. 1-30. DOI: <https://doi.org/10.6018/red/56/10>
- MARTÍNEZ-Rizo, Felipe (2012), “Procedimientos para el estudio de las prácticas docentes. Revisión de la literatura”, *Revista Electrónica de Investigación y Evaluación Educativa*, vol. 18, núm. 1, pp. 1-22.
- MARTÍNEZ-Rizo, Felipe (2016), *La evaluación de docentes de educación básica. Una revisión de la experiencia internacional*, México, INEE.
- MIRANDA López, Francisco y Arcelia Martínez Bordon (2015), “Dilemas y retos del INEE para la evaluación docente en México: algunos temas para la agencia institucional”, ponencia presentada en el “Segundo Congreso Latinoamericano de Medición y Evaluación Educativa”, Ciudad de México, marzo de 2015.
- MURILLO Torrecilla, Francisco Javier, Verónica González de Alba y Héctor Rizo Moreno (2007), *Evaluación del desempeño y carrera profesional docente: un estudio comparado entre 50 países de América y Europa*, Santiago de Chile, UNESCO.
- NAVA Amaya, Minerva y Mario Rueda Beltrán (2014), “La evaluación docente en la agenda pública”, *Revista Electrónica de Investigación Educativa*, vol. 16, núm. 1, pp. 1-11, en: <http://redie.uabc.mx/redie/article/view/662/898> (consulta: 17 de abril de 2017).
- Organización para la Cooperación y el Desarrollo Económicos (OCDE) (2011), *Establecimiento de un marco para la evaluación e incentivos docentes: consideraciones para México*, México, OCDE. DOI: <https://doi.org/10.1787/9789264059986-es>
- PEDROZA, Luis (2016), *Desarrollo de una estrategia de evaluación formativa para el profesorado de educación preescolar*, Tesis de Doctorado, Ensenada, Universidad Autónoma de Baja California (México).
- PIANTA, Robert, Bridget Hamre y Susan Mintz (2012), *Classroom Assessment Scoring System*, Charlottesville (USA), Teachstone.
- RAMOS, Gabriela, Berenice Ponce y Adriana Orozco (2013), *Autoevaluación de la práctica docente para profesores de educación primaria. Manual para la aplicación del cuestionario*, México, Instituto Nacional para la Evaluación de la Educación.
- RAVELA, Pedro (2012), “La evaluación del desempeño docente para el desarrollo de competencias profesionales”, en Elena Martín y Felipe Martínez-Rizo (coords.), *Avances y desafíos en la evaluación educativa*, Madrid, Organización de Estados Iberoamericanos/Fundación Santillana, pp. 113-126.
- Red Iberoamericana de Investigadores de la Evaluación de la Docencia (RIIED) (2008), “Reflexiones sobre el diseño y puesta en marcha de programas de evaluación de la docencia”, *Revista Iberoamericana de Evaluación Educativa*, vol. 1, núm. 3, pp. 7-22.
- ROWAN, Brian y Stephen Raudenbush (2016), “Teacher Evaluation in American Schools”, en Drew Gitomer y Courtney Bell (eds.), *Handbook of Research on Teaching*, Washington, DC American Educational Research Association, pp. 1159-1217.

SHULMAN, Lee (1986), “Paradigmas y programas de investigación en el estudio de la enseñanza: una perspectiva contemporánea”, en Merlin Wittrock (comp.), *La investigación de la enseñanza I: enfoques, teorías y métodos*, México, Paidós, pp. 9-84.

VÁZQUEZ Cruz, María del Ángel, Graciela Cordeiro Arroyo y Yolanda Leyva Barajas (2014), “Análisis comparativo de criterios de desempeño profesional para la enseñanza en cuatro países de América”, *Revista Electrónica Actualidades Investigativas en Educación*, vol. 14, núm. 3, pp. 1-20.