

CUESTIONARIO PARA VALORAR LAS PRÁCTICAS DE ENSEÑANZA EN SECUNDARIA DESDE LA PERCEPCIÓN DE LOS ESTUDIANTES

QUESTIONNAIRE FOR ASSESSING TEACHING PRACTICES IN HIGH SCHOOL FROM THE PERCEPTION OF STUDENTS

Irma Gloria Arregui Eaton¹
Alicia Alelí Chaparro Caso-López
Carlos David Díaz López

Universidad Autónoma de Baja California, México

RESUMEN

La calidad de las prácticas de enseñanza es uno de los principales factores escolares que influyen en el aprendizaje de los estudiantes. Por lo que es importante contar con herramientas de evaluación del ejercicio docente, que permitan generar insumos para fortalecer el trabajo en el aula. El presente estudio tuvo como propósito diseñar un instrumento para la valoración de las prácticas de enseñanza en educación secundaria, a través de la opinión de los estudiantes. El modelo conceptual que sustentó su diseño fue el Marco para la Enseñanza de Danielson, Edición 2013. El proceso de desarrollo implicó: (a) elaborar un banco de ítems y obtener las evidencias de validez de contenido; (b) reducir el número de ítems, mediante un proceso de optimización de la medida; (c) generar evidencias de validez de constructo a través de un análisis factorial confirmatorio. Como resultado se obtuvo un instrumento de 30 ítems a través de los cuales se valoran tres de las cuatro dimensiones del marco de Danielson: (1) Calidad de la planeación y preparación de las actividades docentes, (2) Clima de aula, y (3) Enseñanza o instrucción.

Palabras clave: prácticas de enseñanza; estudiantes; educación secundaria, evaluación docente, validez.

¹ Correspondencia: Irma Gloria Arregui Eaton, Instituto de Investigación y Desarrollo Educativo, Km. 103 Carr. Tijuana-Ensenada. Ensenada, Baja California, México C.P. 22830. Correo-e: irma.arregui@uabc.edu.mx.

ABSTRACT

The quality of teaching practices is one of the main school factors that influence the success or failure of student learning. To strengthen the work of teachers in classroom is necessary to assess their professional practice. This study was conducted with the purpose of designing an instrument for assessing the performance of teachers of secondary education, through the opinions of their students. The conceptual model used was Danielson's Teaching Framework, Edition 2013. The development process committed: (a) to elaborate an item pool and identify evidence of content validity; (b) to reduce the number of items using an optimization process measurement; (c) to obtain evidence of construct validity through confirmatory factor analysis. The product obtained was an instrument of 30 items, which allow us to evaluate three of the four dimensions of Danielson framework: (1) Quality Planning and preparation of teaching activities, (2) Classroom climate, and (3) Teaching or instruction.

Keywords: teaching practices; students; secondary education, teacher evaluation, validity.

Introducción

De acuerdo con la Red Iberoamericana de Investigadores de la Evaluación de la Docencia (RIIED, 2008), el proceso de enseñanza-aprendizaje ha sido el centro del diseño de políticas nacionales e internacionales, en las que se ha destacado la necesidad de precisar estrategias para la evaluación de la actividad docente. Asimismo, se ha evidenciado la importancia de la evaluación docente como estrategia de monitoreo con fines formativos, y de mejora de la calidad educativa (Luna, Elizalde, Torquemada, Castro y Cisneros-Cohernour, 2013). Sin embargo, la RIIED también ha destacado que la mayoría de las iniciativas emprendidas por las diversas instituciones educativas, carecen de sustento teórico y metodológico definido. Además, éstas obedecen más a cuestiones relacionadas con procesos de certificación y recursos económicos, que a la comprensión y desarrollo de los procesos con fines de mejora del ejercicio docente (RIIED, 2008). Por ello, resulta pertinente y necesario el fortalecimiento de los procesos evaluativos de la enseñanza, sobre todo a través de herramientas que demuestren su validez y utilidad, y en la que participen además de docentes y administradores educativos, estudiantes, quienes forman parte indiscutible de este proceso (Shulman, 1986), y por lo tanto tienen el derecho de participación (Román, 2010).

Conceptualización de la enseñanza

En la literatura especializada, la enseñanza es uno de los temas de mayor interés en la investigación educativa. Sin embargo, como lo señalaron Menges y Austin (2001), la enseñanza es una actividad compleja, por lo que determinar sus componentes o describir sus dimensiones es todo un reto. De acuerdo con Shulman (1986), los diferentes programas de investigación en el estudio de la enseñanza seleccionan diferentes partes del mapa para definirla, por lo que, en la configuración de sus dimensiones, existen diferentes aproximaciones. En su revisión, Shulman afirmó que el más influyente de los paradigmas para el estudio de la enseñanza es el de criterio de eficacia, el cual parte del modelo teórico de Proceso-producto. Citando el trabajo de Gage (1963) Shulman señaló que, en esta concepción de la enseñanza, existen cuatro elementos comunes "a) los procesos perceptuales y cognitivos del enseñante que dan como resultado b) elementos de acción por parte del profesor. Los actos del profesor van seguidos de c) procesos

perceptuales y cognitivos por parte del alumno, lo cual a su vez conduce a d) acciones por parte de los alumnos” (p. 8).

Dentro de este esquema, es que se incluyen los estudios de enseñanza eficaz, en los que se analiza de qué manera las prácticas de enseñanza (proceso) impactan el aprendizaje y desarrollo de los estudiantes (producto). Cappella, Aber y Kim (2016), en una revisión exhaustiva de la literatura, afirmaron que las características de las prácticas de enseñanza eficaz están integradas por al menos tres características esenciales. En primer término, la *Organización y manejo del aula*, las aulas con esta característica tienen reglas y consecuencias claras y consistentes, rutinas regulares y eficientes que maximizan el tiempo de aprendizaje, son proactivas y positivas, y aplican estrategias apropiadas para atender y prevenir los problemas de conducta. Las aulas bien organizadas incrementan la probabilidad de que los estudiantes desarrollen estrategias regulatorias que benefician las tareas del aula. Por el contrario, las aulas desorganizadas disminuyen las oportunidades de aprendizaje e incrementan la probabilidad de dificultades académicas en los estudiantes. Una segunda característica es el *Apoyo y clima emocional*, relacionado con el nivel de respeto, calidez y positivismo que impera en el aula, en el que se debe tener en cuenta que el docente es el modelo primario y agente de apoyo. Finalmente, la tercera característica es la *Implicación y métodos instruccionales que representen un reto para los estudiantes*. Esto se asocia con el conocimiento del docente acerca de las fortalezas y debilidades de sus estudiantes y la puesta en marcha de prácticas instruccionales alineadas al currículo, así como oportunidades de aprendizaje relevantes y variadas.

En consonancia con estas características, Windschitl y Calabrese (2016) definieron las prácticas de enseñanza como el trabajo profesional y cotidiano del docente, dedicado a planificar, realizar y reflexionar sobre la instrucción. Puntualizaron que la enseñanza debe ser entendida como una práctica, en la que los procesos de participación y aprendizaje de los estudiantes, son mediados directamente por las decisiones que toman los docentes sobre los tipos de tarea, el lenguaje y las herramientas utilizadas en el aula. Por lo tanto, las diferencias entre los profesores en cuanto a la manera de organizar la instrucción, los métodos y materiales, así como la forma de interactuar con los alumnos, produce diferentes efectos en la forma en que éstos aprenden (Shulman, 1986). Es así, que, desde una visión de la enseñanza eficaz, se ha buscado integrar marcos, o estándares, que definan sus características.

Estándares para la evaluación de la enseñanza

Como señalaron Windschitl y Calabrese (2016), sin un marco común para describir y orientar la enseñanza, es difícil para los investigadores o los profesionales comunicar los problemas significativos en el aula; pero resulta particularmente difícil que el conocimiento profesional se comparta, evalúe y refine a través del tiempo, en ausencia de un marco definitorio. Los estándares de evaluación docente, constituyen, precisamente, marcos comunes sobre la definición de las características de la enseñanza. De acuerdo con Schmelkes y Mancera (2010), los estándares cumplen dos funciones esenciales: (a) sirven como un marco para conducir a los docentes hacia la mejora de su desempeño; y (b) permiten la valoración individual del ejercicio docente. Asimismo, en los marcos puede encontrarse la descripción precisa de lo que se considera una buena enseñanza, lo cual debe representar el común denominador en la construcción de nuevos modelos para la evaluación del desempeño docente. Por lo tanto, los estándares han sido considerados una base para la implementación de mecanismos de supervisión del profesorado, para orientar el desarrollo de sistemas de evaluación más integrales (Ávalos, 2008; Danielson, 2013; Elizalde y Reyes, 2008; Valdés, 2009; Valliant, 2008).

Cappella et al. (2016) en una revisión de diferentes procedimientos evaluativos de la enseñanza, reconocieron que uno de los estándares más referidos en la literatura es el Marco

para la Enseñanza de Danielson (2013). Rowan y Raudenbush (2016), señalaron que su utilidad estriba en que incorpora prácticas de enseñanza generalizadas, que pueden aplicarse a cualquier nivel y modelo de enseñanza. El marco de Danielson, cuenta con una estructura sólida, la cual ha ido evolucionando y perfeccionándose desde 1996, fecha en la que se publicó su primera versión; lo que ha contribuido a generar una descripción clara de indicadores de desempeño. Por lo tanto, se puede considerar como un estándar válido y confiable para la evaluación de la enseñanza. Este marco está dividido en cuatro dimensiones y 22 sub-dimensiones, las cuales hacen referencia a las diversas acciones que el docente realiza en los distintos momentos en los que el proceso de enseñanza-aprendizaje se desarrolla (ver tabla 1).

TABLA 1. Dimensiones y sub-dimensiones del marco para la enseñanza (Danielson, 2013)

<i>Dimensión</i>	<i>Sub-dimensiones</i>
1. Planeación y Preparación	1a Conocimiento del contenido y la pedagogía. 1b Conocimiento de los estudiantes. 1c Planeación de los resultados instruccionales. 1d Conocimiento de los recursos. 1e Diseño de una instrucción coherente. 1f Diseño de las evaluaciones de los estudiantes.
2. Clima de Aula	2a Creación de un ambiente de respeto y armonía. 2b Establecimiento de una cultura de aprendizaje. 2c Gestión de los procedimientos de clases. 2d Manejo de la conducta estudiantil. 2e Organización del espacio físico.
3. Enseñanza	3a Comunicación con los estudiantes. 3b Uso de técnicas de debate y discusión. 3c Involucrar a los estudiantes en el aprendizaje. 3d Uso de la evaluación en la instrucción. 3e Demostración de flexibilidad y capacidad de respuesta.
4. Responsabilidades Profesionales	4a Reflexiones sobre la enseñanza. 4b Mantenimiento de registros precisos. 4c Comunicación con las familias. 4d Participación con la comunidad profesional. 4e Crecimiento y desarrollo profesional. 4f Muestra de profesionalismo.

Fuente: Elaboración propia con base en Danielson (2013).

El Marco para la Enseñanza de Danielson (2013), por lo tanto, recupera en sus definiciones los aspectos que la literatura en el área ha referido como significativos en el proceso de enseñanza (Cappella et al., 2016). Incluye indicadores relativos a la organización y manejo del aula, que permiten maximizar los aprendizajes (Murillo, 2007; Murillo, Martínez y Hernández, 2011). Así también, describe elementos asociados al clima del aula, y a los aspectos relativos a el apoyo del docente, vinculados al desarrollo emocional y la autoestima del estudiante (Manota y Melendro, 2016; Martínez-Garrido y Murillo, 2016 ; Ríos, Bozzo, Marchant y Fernández, 2010); e incorpora información relativa a los métodos instruccionales y de compromiso profesional, ante los cuales los estudiantes responderán con implicación y entrega a las actividades y tareas escolares (Murillo, 2007; Murillo, Martínez y Hernández, 2011).

Participación de los estudiantes en la evaluación docente

Como lo señaló Shulman (1986), “la enseñanza es concebida como una actividad que implica el trabajo conjunto de profesores y estudiantes. Involucra el ejercicio tanto del pensamiento como de la acción por parte de todos los participantes. En consecuencia, los enseñantes aprenden y los

estudiantes enseñan. Ambas funciones de cada actor pueden considerarse parte fundamental de la investigación” (p. 10). Por lo tanto, docentes y estudiantes deben participar en el proceso de evaluación de la enseñanza. Los estudiantes son informantes ideales y confiables para dar cuenta de lo que sucede dentro del aula (Darling-Hammond, 2012; Ferguson, 2012; Kane, Taylor, Tyler y Wooten, 2011). La participación de los estudiantes en la evaluación de la enseñanza, es esencial no sólo porque éstos tienen una perspectiva especial de la enseñanza y el aprendizaje, sino porque aportan información más sustancial, que cualquier otra forma de evaluación, pues sus valoraciones responden más a lo que observan en la práctica docente, que a los atributos de sus enseñantes (Menges y Austin, 2001).

Otro punto importante a tener en cuenta sobre la participación de los estudiantes en la evaluación docente, es la propia interpretación que ellos hacen de lo que experimentan en la enseñanza.

Como señaló Shulman (1986) en la interpretación del estudiante de los procesos del aula, se da una representación y construcción mental del contenido cognitivo que se está enseñando. Por lo tanto, como lo afirmó Román (2010): “El rol protagónico y prioritario del propio estudiante en el proceso de enseñanza y aprendizaje le confiere no solo el derecho, sino que la legitimidad, de ser un actor presente y válido a la hora de opinar y de juzgar la calidad de la enseñanza que recibe y, en tal sentido, analizar la práctica, el estilo y las actitudes de sus profesores” (p.5).

La validez de las evaluaciones realizadas por los estudiantes, constituye otro argumento a favor de su participación en la evaluación de las prácticas de enseñanza. La consistencia en los datos obtenidos en diversos grupos de estudiantes, es indicador de la confiabilidad como medida para realimentar el trabajo docente dentro del aula (Elizalde y Reyes, 2008; Luna y Torquemada, 2008). Además, desde hace varios años se han desarrollado investigaciones que han demostrado la validez y utilidad potencial de las evaluaciones docentes a partir de la opinión del estudiante (Menges y Austin, 2001). Como afirmaron Cappella et al. (2016) se cuenta actualmente varias medidas de evaluación docente basados en la valoración del estudiante, que han demostrado tener propiedades psicométricas que permiten evaluar diferentes aspectos de las prácticas de enseñanza.

Con base en la exposición realizada, se considera necesario y pertinente contar con instrumentos de medición de las prácticas de enseñanza basados en estándares, que incorporen la participación de los estudiantes. Sin embargo, en la literatura actual, de habla hispana, no se ubican instrumentos que cubran estas características en el nivel de secundaria. En consecuencia, el estudio que se presenta persiguió dos objetivos centrales. El primero fue desarrollar, a partir del *Marco para la enseñanza* de Danielson (2013), un instrumento para la evaluación de las prácticas de enseñanza en secundaria, con base en la percepción de los estudiantes. El segundo, fue aportar evidencias de validez del instrumento desarrollado.

Método

Muestra

En el estudio participaron 852 estudiantes de educación secundaria de los tres grados, con una media de 13 años de edad, y una desviación típica de 1.12, provenientes de seis escuelas secundarias de Baja California, México. De los cuales 434 (51%) fueron niñas y 417 (49%) niños; 590 (69%) asistían al turno matutino y 262 (31%) al vespertino; 329 (37%) cursaban el primer grado de secundaria, 92 (14%) el segundo grado y 431 (51%) el tercer grado. Es importante señalar que la muestra no fue representativa, y los participantes fueron seleccionados a

conveniencia, a partir de las escuelas que dieron autorización para la aplicación de los instrumentos.

Instrumento

Como se señaló anteriormente, el instrumento desarrollado se construyó a partir del *Marco para la enseñanza* de Danielson (2013). Al considerar a los estudiantes como informantes clave del desempeño docente, se realizó un análisis exhaustivo de los indicadores de cada dimensión, buscando identificar únicamente aquellos que eran susceptibles de ser evaluados por los estudiantes. De esta manera, se determinó que solo 89 de los 103 indicadores eran aptos para sustentar el diseño de cada uno de los ítems. De lo anterior, resultó un banco de 128 ítems, los cuales quedaron distribuidos en las cuatro dimensiones enunciadas por Danielson: Planeación y preparación (35 ítems); Clima de aula (44 ítems); Enseñanza o instrucción (30 ítems); Responsabilidades profesionales (19 ítems). Dado que el contenido de los ítems hacía alusión a prácticas docentes, se eligió una escala de respuesta ordinal de cuatro puntos, que diera cuenta de la frecuencia con la que se realizaban dichas prácticas: (a) siempre; (b) frecuentemente; (c) pocas veces y (d) nunca.

Para estimar la validez de contenido, el banco de 128 ítems se sometió a un proceso de validación de jueces, en el que participaron 13 docentes en servicio; quienes determinaron para cada ítem: (a) congruencia con la sub-dimensión a la que pertenecía; (b) congruencia con los indicadores; (c) determinación respecto a si cada uno de los ítems podía o no ser respondido por los estudiantes; y (d) sugerencias de cambio o eliminación de ítems que consideraran no aptos para formar parte de la estructura del instrumento.

A partir de las respuestas otorgadas por los jueces a cada ítem, y con base en el índice de validez de contenido (IVC) de Lawshe (1975) se identificaron aquellos ítems que resultaría necesario eliminar. Según las tablas de criterio del IVC, cuando participan 13 jueces el índice esperado es de al menos .54; por lo cual todos los ítems que tuvieron un índice menor, se eliminaron, conservándose un banco de 110 ítems distribuidos en las cuatro dimensiones de Danielson.

Procedimiento

Los 110 ítems se aplicaron a la muestra de 852 estudiantes, de forma grupal, en sus aulas. Todos los ítems se contestaron en la misma sesión. Junto con la aplicación se llevó a cabo un control de incidencias, lo que permitió asegurar que aquellos estudiantes que mostraban algún comportamiento que diera indicio de que no contestaban el instrumento con atención (es decir, que se mostraron distraídos, juguetones o apáticos) fueron descartados. Asimismo, el registro de incidencias permitió tomar nota de los ítems que los estudiantes reportaron que resultaban confusos.

Análisis de datos

Optimización de la medida. Dado que el diseño del instrumento que se propone está dirigido a estudiantes de secundaria, quienes frecuentemente muestran una atención dispersa y se distraen con facilidad (Heim, Ihssen, Hasselhom y Keil, 2013), se consideró que resultaba necesario optimizar la medida, pues 110 ítems resultaban excesivos para ser contestados por éstos, para ello, se realizó el procedimiento propuesto por Jornet, González-Such y Perales (2012), con el cual se buscó, por un lado, seleccionar aquellos ítems que aportaban información suficiente, con un elevado grado de confiabilidad y validez; y por otro, reducir la cantidad de ítems a partir de la

eliminación de aquellos que fueron redundantes o aportaban menos información. El proceso comprendió las siguientes fases:

1. Con el objetivo de asegurar que los 110 ítems fueran congruentes con el constructo, se realizó un análisis de consistencia interna utilizando Alfa ordinal. La razón por la que se eligió esta técnica se debe a la naturaleza ordinal de la escala de respuesta de los ítems.
2. Para evitar incluir ítems que no permitieran diferenciar respuestas de los estudiantes, se eliminaron aquellos en los que se observaba que más del 80% de los participantes se inclinaba hacia un extremo de la escala de respuesta. Después de eliminar los ítems, se obtuvo nuevamente el índice de consistencia interna (Alfa ordinal) y se comparó con el alfa de los 110 ítems.
3. Se analizó la calidad métrica de los ítems resultantes del procedimiento anterior, a través de la identificación de la correlación punto biserial y la correlación ítem total. En este caso, se eliminaron aquellos ítems que presentaban una correlación punto biserial inferior a .30, así como también aquellos que registraban una correlación inter-ítem mayor a .50. Este último criterio obedeció a la necesidad de eliminar ítems que fueran redundantes.
4. Al haber agotado los criterios métricos para la reducción del instrumento y tomando en consideración la cantidad de ítems, se realizó un análisis cualitativo, basado en el constructo de este instrumento: Marco para la enseñanza (Danielson, 2013).

Análisis Factorial Confirmatorio.

En esta etapa, el análisis se centró en establecer si el conjunto de 30 ítems derivados del proceso de optimización de la medida, se ajustaba métricamente con el constructo de interés, para lo cual se efectuó un análisis factorial confirmatorio con el programa *R project* (Culpepper y Aguinis, 2011). Dada la naturaleza ordinal de la escala de respuestas empleada en los ítems, se generó una matriz de correlaciones policóricas, a partir de la cual se efectuó este análisis.

Resultados

Optimización de la medida

Los resultados de cada una de las fases de este proceso se presentan en la tabla 2. En la fase 1, se determinó la consistencia interna del banco inicial de 110 ítems la que resultó de .98. Posteriormente, en la fase 2 se eliminaron 16 ítems en los que se observaba más del 80% de respuestas hacia un extremo de la escala. Esto dio como resultado un total de 94 ítems que conservaron un nivel de consistencia interna y una correlación con el total inicial de 110 ítems, superior a .98.

TABLA 2. Índices de consistencia interna y correlación por fase

<i>Fases</i>	<i>Total de ítems</i>	<i>Alfa</i>	<i>Correlación ítem total</i>
1	110	0.98	NA
2	94	0.98	0.997
3	71	0.98	0.993
4	35	0.96	0.981
5*	30	0.95	0.976

Fuente: Elaboración propia

Nota: * La fase 5 no formaba parte del procedimiento de optimización de la medida de Jornet et al. (2012), sin embargo, dado que implicaba también una tarea de optimización, se reporta en esta tabla. NA = no aplica.

En la fase 3, se eliminaron los ítems que mostraron una correlación punto biserial inferior a $<.30$, así como también aquellos que registraban una correlación inter-ítem superior a $.50$. De esta manera, se eliminaron 23 conservándose 71, los cuales mantenían tanto un nivel de consistencia interna, como una correlación con los 110 ítems iniciales, superior a $.98$.

Como resultado del análisis cualitativo, realizado en la fase 4, se eliminaron 36 ítems, ya que se encontró que presentaban redundancias o que evaluaban aspectos que no correspondían a la dimensión en la que se encontraban. De manera que al final de esta fase permanecieron 35 ítems, los cuales se observó que mantenían índices tanto de consistencia interna como de correlación con los 110 ítems iniciales, superiores a $.96$. Finalmente, dado que el instrumento se diseñó para aplicarse a estudiantes de educación secundaria, se determinó que los cinco ítems pertenecientes a la Dimensión 4. Responsabilidades profesionales, debían eliminarse a pesar de sus cualidades métricas, ya que, al realizar un análisis cualitativo de los mismos, se determinó que el contenido de éstos no podía ser evaluado a través de la percepción de los estudiantes.

Así al concluir las cinco fases del proceso de optimización de la medida, en la tabla 3 se puede observar que el instrumento quedó integrado por 30 ítems que se distribuyeron en tres dimensiones: Planeación y preparación (9 ítems); Clima de aula (12 ítems); Enseñanza o instrucción (9 ítems).

TABLA 3. Versión final del instrumento para la evaluación de las prácticas de enseñanza en educación secundaria

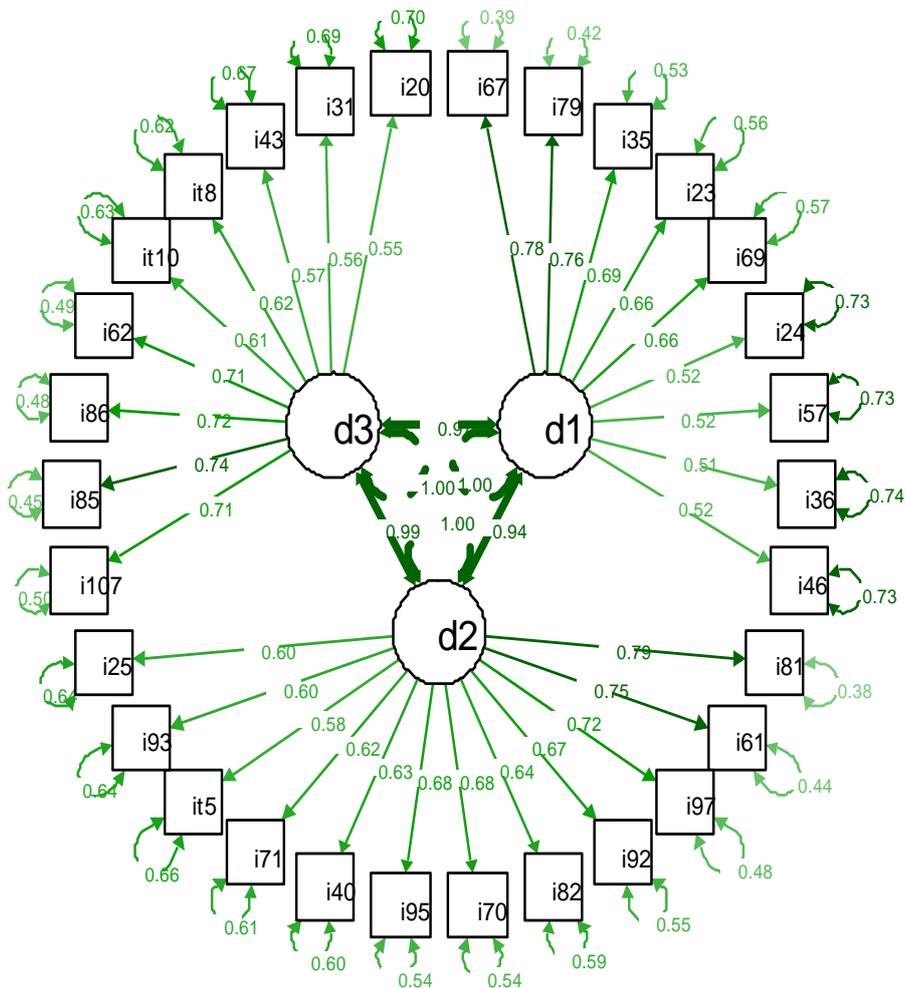
<i>Factores resultantes</i>		
<i>Factor 1. Planeación y preparación</i>	<i>Factor 2. Clima de aula</i>	<i>Factor 3. Enseñanza</i>
67. El profesor(a) utiliza ejemplos fáciles de comprender.	81. El profesor(a) reconoce el valor de nuestro esfuerzo en clase.	107. El profesor(a) utiliza nuestras participaciones para enseñarnos nuevas cosas.
79. El profesor(a) nos pone actividades de clase que nos motivan a pensar.	61. El profesor(a) nos felicita cuando nos portamos bien en clase.	85. El profesor(a) escucha nuestras ideas.
35. El profesor(a) nos ayuda a localizar diferentes textos que nos ayuden a comprender un tema.	97. El profesor(a) y los estudiantes trabajamos juntos durante la clase.	86. El profesor(a) nos hace preguntas para asegurarse que entendimos el tema.
23. El profesor(a) nos explica para qué nos servirá lo que aprenderemos en la clase.	92. El profesor(a) nos dice lo que espera de nosotros.	62. El profesor(a) nos explica los objetivos de aprendizaje que debemos alcanzar en cada tema.
69. El profesor(a) nos explica la forma de evaluar cada unidad.	82. Todos participamos en la organización de la clase.	10. El profesor(a) nos explica usando ejemplos de cosas que nos gustan.
24. El profesor(a) nos organiza en equipos de trabajo con compañeros diferentes.	70. El profesor(a) nos pide que participemos en clase.	8. El profesor(a) utiliza palabras fáciles de entender.
57. El profesor(a) explica la relación del tema de la clase con otros temas o materias.	95. El profesor(a) se asegura de que todos los compañeros podemos ver al pizarrón.	43. El profesor(a) nos explica usando ejemplos de cosas que suceden en la comunidad.
36. El profesor(a) emplea diferentes formas de evaluación.	40. El profesor(a) realiza los cambios necesarios para acomodar a los estudiantes con necesidades especiales.	31. El profesor(a) nos hace preguntas sobre cómo usaríamos en nuestra vida diaria lo que aprendemos en clase.
46. El profesor(a) utiliza diferentes materiales durante la clase.	71. El profesor(a) inicia la clase puntualmente.	20. El profesor(a) se pasea por el salón durante la clase, para supervisar nuestro trabajo.
	5. El profesor(a) se cerciora de que todos los estudiantes participemos en la realización de trabajos individuales o en equipo.	
	93. El profesor(a) termina la clase a tiempo.	
	25. El profesor(a) promueve el respeto de las diferentes costumbres y creencias entre compañeros.	

Fuente: Elaboración propia.

Evidencias de validez de constructo

Después del proceso de optimización de la medida, se realizó un análisis factorial confirmatorio a partir de matrices policóricas (ver figura 1) a los 30 ítems resultantes. Se puede observar que las cargas factoriales son, en todos los casos, superiores a .50, lo que demuestra la congruencia que cada conjunto de ítems tiene con la sub-dimensión.

FIGURA 1. Estructura factorial del instrumento para la evaluación de las prácticas de enseñanza en secundaria



Fuente: Elaboración propia a partir de los resultados del análisis factorial confirmatorio.

Nota: Los círculos representan las variables latentes o factores del instrumento (d1 = Planeación y preparación, d2 = Clima de aula, d3 = Enseñanza). Los cuadros simbolizan los ítems. En las flechas que salen de los círculos se observan las cargas factoriales de cada ítem.

En la tabla 4 se presentan los índices de bondad de ajuste de la estructura factorial presentada en la figura 1; los cuales mostraron valores aceptables según los criterios de referencia de Ruíz, Pardo y San Martín (2010).

TABLA 4. Estadísticos de bondad de ajuste y criterios de referencia

Tipo de ajuste	Prueba	Estadístico	Criterio*	RO
Ajuste absoluto	Razón Chi-cuadrado/grados de libertad	χ^2/gl	Menor que 3	1.579
Ajuste comparativo	Índice de bondad de ajuste	CFI	≥ 0.95	0.996
	Índice de Tucker-Lewis	TLI	≥ 0.95	0.996
	Índice de ajuste normalizado	NFI	≥ 0.95	0.990
Ajuste parsimonioso	Índice de parsimonia	NFI	Próximo a 1	0.990
	Índice de parsimonia corregido	PNFI	Próximo a 1	0.915
Otros	Índice de bondad de ajuste	CFI	≥ 0.95	0.996
	Índice de bondad de ajuste corregido	AGFI	≥ 0.95	0.992
	Raíz del residuo cuadrático promedio	RMR	Próximo a cero	0.040
	Raíz del residuo cuadrático promedio de aproximación	RMSEA	≤ 0.08	0.026

Fuente: Elaboración propia con base en * Ruíz, M., Pardo, A., y San Martín, R. (2010). Modelos de ecuaciones estructurales. *Papeles del Psicólogo*, 3 (1), 34-45. RO= Resultados obtenidos.

Conclusiones

Como objetivo central de este estudio, se planteó el desarrollo de un instrumento para la evaluación de las prácticas de enseñanza en educación secundaria basado en la opinión de los estudiantes y sustentado en el Marco para la Enseñanza de Danielson (2013). De acuerdo con los datos presentados, se puede afirmar que los objetivos se lograron de forma satisfactoria, ya que se obtuvo un banco de ítems, depurado y con propiedades métricas aceptables que se apega al constructo del que partió.

Se considera que una de las principales contribuciones de este estudio es haber desarrollado un instrumento para la evaluación de la enseñanza, con base en la opinión de los estudiantes. Pues, como se refirió anteriormente, los estudiantes forman parte del proceso de enseñanza aprendizaje (Shulman, 1986), tienen el derecho de opinar (Roman, 2010), y su opinión puede ser más sustancial que la de otros actores educativos (Menges y Austin, 2010). Asimismo, los estudiantes, son quienes tienen la información más acertada acerca del desempeño profesional de sus profesores (Darling-Hammond, 2012; Ferguson, 2012; Kane et al., 2011; Román, 2010). Como ya lo señalaron Román (2010) y Ferguson (2012), los estudiantes son quienes pasan más tiempo en el aula y, por consiguiente, son quienes pueden y tienen el derecho a valorar lo que en ella acontece. De la misma manera, se considera que la participación de los estudiantes en el proceso evaluativo del desempeño docente, es esencial (Vaillant, 2008).

Además, los resultados de la métrica del instrumento, reiteran, que la participación de los estudiantes como informantes de la enseñanza, puede considerarse válida (Cappella et al., 2016). Principalmente, se considera como una cualidad de este instrumento la participación del estudiante, pues, desde el enfoque de la eficacia escolar, la enseñanza tiene impacto directo en el desarrollo integral de éste, por lo que contar con su visión contribuye a identificar, en qué medida los esfuerzos del docente son percibidos por el estudiante. Como se mencionó, de acuerdo con Shulman (1986) y Menges y Austin (2001), los estudiantes reaccionan y responden al entorno a partir de la percepción que tienen de éste, por lo que, si la calidad del trabajo docente, no es percibida por el estudiante, no generará el impacto esperado. Por lo tanto, se tendrá mayor claridad respecto a los elementos que deben considerarse en la mejora (Schmelkes y Mancera, 2010).

A pesar de que los indicadores incluidos en las cuatro dimensiones que constituyen el Marco para la enseñanza de Danielson (2013) no representan una lista exhaustiva de acciones específicas que el docente en servicio debe realizar (Shulman, 1986), goza de reconocimiento y es ampliamente utilizado en la literatura de evaluación de la enseñanza, que recupera los aspectos significativos de ésta (Cappella et al., 2016; Rowan y Raundenbush, 2016). En consecuencia, considerando que el presente instrumento se fundamenta en este marco, se puede afirmar que es una herramienta que permite evaluar, de forma válida, las prácticas de enseñanza, que el docente de secundaria lleva a cabo.

Por lo tanto, la información que se obtenga a través del instrumento desarrollado en este estudio, podrá utilizarse como parte de una evaluación integral, que haga factible contrastar las opiniones sobre la práctica de enseñanza entre grupos como complemento de las valoraciones de los directores y supervisores de los planteles escolares Darling-Hammond, 2012; Ferguson, 2012; Kane et al., 2011). Además, es importante destacar que en la actualidad la implementación de estrategias para evaluar las prácticas de enseñanza, en las cuales se toma en cuenta la participación de los estudiantes, es llevada a cabo cada vez con mayor frecuencia (Cappella, et al., 2016; Román, 2010; Vaillant, 2008).

Asimismo, aunque el instrumento final, sólo incluyó tres de las cuatro dimensiones del marco de Danielson (2013), los factores finales guardan correspondencia con las características de la enseñanza eficaz enunciadas por Cappella et al. (2016): Planeación y preparación de la enseñanza, Desarrollo de un clima de aula propicio para el aprendizaje, y el momento mismo de la enseñanza (Windschitl y Calabrese, 2016). Así, por ejemplo, en el Factor 1, vinculado a la dimensión de Danielson (2013) de *Planeación y preparación*, se integraron ítems centrados en el grado de importancia que el docente otorga a la definición de los objetivos de la clase, en la especificación clara y precisa respecto a para qué les servirá a los estudiantes lo aprendido, así como las competencias y habilidades que deberían desarrollar; así como de organización y evaluación. Elementos que de acuerdo con la literatura de enseñanza eficaz se relacionan con la maximización de los aprendizajes de los estudiantes (Murillo, 2007; Murillo, Martínez y Hernández, 2011).

El Factor 2, recuperó ítems que integran la dimensión de *Clima de aula* en el marco de Danielson (2013), se conformó por ítems a partir de los cuales se obtendrá información sobre las expectativas del profesor sobre sus estudiantes, el reconocimiento de sus esfuerzos y el impacto de éste en los niveles de aprendizaje. También, se exploran acciones que permiten evidenciar si se genera una interacción respetuosa y empática que favorezca la convivencia y la participación e integración de todos los estudiantes, e incluso del docente, en las actividades que se han de realizar dentro del aula. Además, los ítems integrados en este factor, también indagan acerca de la capacidad del docente para adaptar y aprovechar al máximo los recursos materiales con los que cuenta para llevar a cabo su labor y lograr el éxito en el proceso de enseñanza. (Cappella et al., 2016; Windschitl y Calabrese, 2016). En consecuencia, la evaluación que se realice con base en los datos que arrojen estos ítems, permitirá identificar en qué medida el docente promueve el desarrollo emocional y la autoestima del estudiante (Manota y Melendro, 2016; Martínez-Garrido y Murillo, 2016; Ríos et al., 2010).

Por último, los ítems del Factor 3 se relacionaron con la dimensión que Danielson (2013) nombró como *Enseñanza*, captan información relacionada con la capacidad del docente para crear o adaptar escenarios de aprendizaje adecuados a las necesidades y capacidades de sus estudiantes, en los que se consideren sus conocimientos previos, sus participaciones y opiniones, así como los recursos y el tiempo disponibles (Windschitl y Calabrese, 2016). Además, en esta dimensión se da cuenta de la importancia de la habilidad del docente para diseñar estrategias de evaluación que le permitan el monitoreo constante del aprendizaje de sus estudiantes y al mismo tiempo, sea capaz de aprovechar esta información para realimentar su propio desempeño.

(Cappella et al., 2016). Estos aspectos, de acuerdo con la investigación en enseñanza eficaz, generan en el estudiante implicación, entrega y compromiso en las tareas escolares (Murillo, 2007; Murillo, Martínez y Hernández, 2011).

Otro punto destacable respecto a las aportaciones del presente estudio, es que el proceso de optimización de la medida que se siguió fue útil para lograr un instrumento parsimonioso válido y confiable, con ítems que tienen un alto nivel de relación con el constructo; y que resulta apropiado para aplicarse a estudiantes de nivel secundaria, que comúnmente muestran un alto nivel de distracción ante tareas como contestar un instrumento de amplia longitud (Heim et al., 2013). Además, se debe mencionar que el hecho de que se haya “descartado” una de las cuatro dimensiones que componen la estructura del marco, no implica una disminución en la calidad de la medida; sino por el contrario, representa una ventaja en la optimización de la misma, pues se debe tener presente que el Marco para la Enseñanza de Danielson (2013) aunque está integrado por cuatro dimensiones, éstas fueron diseñadas pensando en la observación directa y verificación de evidencias, por parte de directores, supervisores u otros docentes implicados en la supervisión de la práctica. De tal manera que los aspectos considerados en esta dimensión, no son necesariamente “accesibles” para el estudiante, y por ende, éste no puede dar cuenta de ellos. En consecuencia, la eliminación de la cuarta dimensión, se considera más una ganancia, que una pérdida.

En conclusión, con base en los resultados obtenidos, se considera que el instrumento diseñado constituye una herramienta útil para recoger información que permita la valoración de la calidad de la enseñanza desde la visión del estudiante, lo que permite establecer los elementos que deben tenerse en cuenta para generar un impacto mayor éste. Al ser desarrollado con base en un estándar descriptivo y ampliamente referido en la literatura, la evaluación que se realice con el presente instrumento, se puede utilizar como herramienta de diagnóstico y también de comparación y marco de referencia para la evaluación docente (Windschitl y Calabrese, 2016), así como para identificar de forma individual, los elementos para la mejora (Ávalos, 2008; Danielson, 2013; Elizalde y Reyes, 2008; Schmelkes y Mancera, 2010). Sin embargo, a pesar de que se reconoce la valía y utilidad de la evaluación realizada a través de la opinión del estudiante (Darling-Hammond, 2012; Ferguson, 2012; Kane et al., 2011; Menges y Austin, 2001).

Sin embargo, dada la complejidad del proceso de enseñanza, debe tenerse en cuenta que, para lograr una evaluación precisa y válida del trabajo docente, ésta no debe utilizarse como único medio de evaluación, sino que debe ser acompañada de otros instrumentos, y de la participación de otros informantes, además de los estudiantes, que ayuden a tener una visión integral del trabajo docente, sus necesidades de desarrollo profesional, y hasta su aporte a la evolución de sus estudiantes. Finalmente, debe reconocerse la importancia de llevar a cabo nuevos estudios que contribuyan a confirmar las propiedades métricas del instrumento (por ejemplo, la realización de análisis de invarianza factorial), que permitan recuperar evidencia sobre el comportamiento de los ítems en diferentes contextos, contribuyendo así a la verificación de su confiabilidad y validez, así como de los parámetros normativos para su aplicación, calificación e interpretación.

Referencias Bibliográficas

Cappella, E., Aber, J.L., & Kim, H.Y. (2016). Teaching beyond achievement test: Perspective from developmental and education science. En D. H. Gitomer, & C. A. Bell (Eds.), *Handbook of research on teaching* (pp. 249-349). Washington, DC: American Educational Research Association.

- Gage, N.L. (1963). Paradigms for research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching*. Chicago: Rand McNally. Citado en Shulman, L. (1986). Paradigmas y programas de investigación en el estudio de la enseñanza: Una perspectiva contemporánea. En M. C. Wittrock (Comp.), *La investigación de la enseñanza I: Enfoques, teorías y métodos* (pp. 9-84). México: Paidós.
- Luna, E., Elizalde, L., Torquemada, A., Castro, A., y Cisneros-Cohernour, E. (2013). Evaluación de la docencia en el sistema educativo nacional 2002-2012. En A. Maldonado (Comp.), *Educación y Ciencia: políticas y producción del conocimiento 2002-2011*. (pp. 399-454). México: ANUIES.
- Menges, R. & Austin, A.E. (2001) Teaching in higher education. En V. Richardson (Eds.), *Handbook of research on teaching* (4th Ed.) (pp. 1122-1158). Washington, D.C.: AERA
- Rowan, B, & Raudenbush, S.W. (2016). Teacher evaluation in american schools. En, D. H. Gitomer, & C. A. Bell (Eds.). *Handbook of research on teaching* (pp. 1159-1217). Washington, DC: American Educational Research Association.
- Shulman, L. (1986). Paradigmas y programas de investigación en el estudio de la enseñanza: Una perspectiva contemporánea. En M. C. Wittrock (Comp.), *La investigación de la enseñanza I: Enfoques, teorías y métodos* (pp. 9-84). México: Paidós.
- Windschitl, M. & Calabrese, A. (2016). Rigor and equity by design: Locating a set of core teaching practices for the science education community. En D. H. Gitomer, & C. A. Bell (Eds.). *Handbook of research on teaching* (pp. 1099- 1159). Washington, DC: American Educational Research Association.

Fuentes Electrónicas

- Ávalos, B. (2008). *Notas sobre la formulación y el uso de estándares para evaluar el desempeño docente*. Agencia de los Estados Unidos para el Desarrollo Internacional (USAID). Recuperado de https://pdf.usaid.gov/pdf_docs/Pnadr967.pdf
- Culpepper, S. & Aguinis, H. (2011). R is for revolution: A review of a cutting-edge, free, open source statistical package. *Organizational Research Methods*, 14(4), 735-740. Recuperado de <http://orm.sagepub.com/content/14/4/735>
- Danielson, C. (2013). *Framework for professional practice*. EEUU: The Danielson group. Recuperado de <http://www.danielsongroup.org/framework/>
- Darling-Hammond, L. (2012). *Creating a comprehensive system for evaluating and supporting effective teaching*. Stanford, CA: Stanford Center for Opportunity Policy in Education. Recuperado de <https://edpolicy.stanford.edu/sites/default/files/publications/creating-comprehensive-system-evaluating-and-supporting-effective-teaching.pdf>
- Elizalde, L. y Reyes, R. (2008). Elementos clave para la evaluación del desempeño de los docentes. *Revista Electrónica de Investigación Educativa*, 10,1-13. Recuperado de <https://redie.uabc.mx/redie/issue/view/23>
- Ferguson, R. (2012). Can student surveys measure teaching? *Kappan Magazine*, 9(3),24-28. Recuperado de <http://www.ciu10.org/cms/lib/PA06001249/Centricity/Domain/69/Ferguson%20in%20Kappan%20Nov%202012.pdf>
- Heim, S., Ihnssen, N., Hassekorn, M., & KEIL, A. (2013). Early adolescents show sustained susceptibility to cognitive interference by emotional distractors. *Cognition and Emotion*, 27(4), 696-706. doi: 10.1080/02699931.2012.736366

- Jornet, J. M., González-Such, J. y Perales, M. (2012). Diseño de cuestionarios de contexto para la evaluación de sistemas educativos: optimización de la medida de constructos complejos. *Bordón. Revista de Pedagogía*, 64(2), 89-110. Recuperado de http://www.uv.es/gem/gemeduco/publicaciones/Diseno_de_cuestionarios_de_contexto_para_la_evaluacion_de_sistemas.pdf
- Kane, T., Taylor, E., Tyler, J., & Wooten, A. (2011). Evaluating teacher effectiveness in Cincinnati public schools. *Education Next*, 11(3), 55-60. Recuperado de <http://educationnext.org/evaluating-teacher-effectiveness/>
- Lawshe, C. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563-575. Recuperado de <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1744-6570.1975.tb01393.x>
- Luna, E., y Torquemada, A. (2008). Los cuestionarios de evaluación de la docencia por los alumnos: balance y perspectivas de su agenda. *Revista Electrónica de Investigación Educativa*, 10, 2-15. Recuperado de <http://redie.uabc.mx/redie/article/view/201>
- Manota, M., y Melendro, M. (2016). Clima del aula y buenas prácticas docentes con adolescentes vulnerables: más allá de los contenidos académicos. *Contextos educativos*, 19, 55-74. doi: 10.18172/con.2756.
- Martínez-Garrido, C., y Murillo, F. J. (2016). Investigación iberoamericana sobre enseñanza eficaz. *Revista Mexicana de Investigación Educativa*, 21(69), 471-499. Recuperado de <http://www.scielo.org.mx/pdf/rmie/v21n69/1405-6666-rmie-21-69-00471.pdf>
- Murillo, F. J. (2007). *Investigación Iberoamericana sobre Eficacia Escolar*. Bogotá: Convenio Andrés Bello. Recuperado de <https://goo.gl/TeVLY>
- Murillo, F. J., Martínez, C., y Hernández, R. (2011). Decálogo para una enseñanza eficaz. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 9, 6-27. Recuperado de <http://www.redalyc.org/articulo.oa?id=55118790002>
- 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), 7-22. Recuperado de http://www.rinace.net/riee/numeros/vol1-num3_e/reflexiones.pdf
- Ríos, D., Bozzo, N., Marchant, J., y Fernández, P. (2010). Factores que inciden en el clima de aula universitario. *Revista Latinoamericana de Estudios Educativos*, 40(3-4), 105-126. Recuperado de <http://www.redalyc.org/pdf/270/27018888004.pdf>
- Román, M. (2010). *La voz ausente de estudiantes y padres en la evaluación del desempeño docente*. Serie Documentos No. 49. Chile: PREAL. Recuperado de http://www.cide.cl/documentos/preal_49_MR.pdf
- Ruiz, Miguel A., Pardo, A., y San Martín, R. (2010). Modelos de ecuaciones estructurales. *Papeles del Psicólogo*, 31(1), 34-45. Recuperado de <http://www.redalyc.org/articulo.oa?id=77812441004>
- Schmelkes, S., y Mancera, C. (2010). *Recomendaciones de políticas específicas sobre el desarrollo de un marco integral para la evaluación de maestros en servicio*. Instituto de Investigaciones para el Desarrollo de la Educación de la Universidad Iberoamericana, Ciudad de México. Recuperado de <http://www.inidedelauia.org/2012/02/recomendaciones-de-politicas.html>
- Vaillant, D. (2008). Algunos marcos referenciales para la evaluación del desempeño docente en América Latina. *Revista Iberoamericana de Evaluación Educativa*, 1(2), 7-22. Recuperado de <http://www.denisevaillant.com/PDFS/ARTICULOS/2008/RINACE2008.pdf>

Valdés, H. (2009). *Manual de buenas prácticas de evaluación del desempeño profesional de los docentes*. Lima Perú: Consejo Nacional de Educación. Recuperado de <http://repositorio.minedu.gob.pe/handle/123456789/5039>

Fecha de entrada: 27 abril 2017

Fecha de revisión: 17 mayo 2018

Fecha de aceptación: 20 mayo 2018