EVALUACIÓN EDUCATIVA

Experiencias de investigación en Posgrado

José Alfonso Jiménez Moreno
y Edna Luna Serrano
(Coordinadores)

Qartuppi

Esta obra se publica bajo una Licencia Creative Commons Atribución-NoComercial 4.0 Internacional.

https://creativecommons.org/licenses/by-nc/4.0/deed.es

EVALUACIÓN EDUCATIVA

Experiencias de investigación en Posgrado

José Alfonso Jiménez Moreno
y Edna Luna Serrano
(Coordinadores)

Qartuppi

La presente obra fue dictaminada por el método a doble ciego por dos evaluadores externos conforme a los criterios académicos vigentes.

Evaluación Educativa.
Experiencias de investigación en Posgrado

1era edición, diciembre 2018

ISBN 978-607-98270-3-8
DOI 10.29410/QTP.18.10

D.R. © 2018. Qartuppi, S. de R.L. de C.V.
Villa Turca 17, Villas del Mediterráneo
Hermosillo, Son. 83220 México
http://www.qartuppi.com

Diseño y Edición: Qartuppi, S. de R.L. de C.V.
Fotografía de portada: León Felipe Irigoyen Morales

CONTENIDO

7 Prólogo
Luis Lizasoain Hernández

15 Capítulo 1
La evaluación docente en la educación básica,
una mirada al contexto internacional y nacional
Antelmo Castro López, Edna Luna Serrano y Graciela Cordero Arroyo

37 Capítulo 2
Estándares de desempeño docente: México en el contexto internacional
Luis Horacio Pedroza Zúñiga, Edna Luna Serrano y Joaquín Caso Niebla

57 Capítulo 3
Caracterización de las motivaciones para ingresar a la Escuela Normal.
Resultados del estudio exploratorio en Baja California
Jihan Ruth García-Poyato Falcón, Graciela Cordero Arroyo,
Alejandra Priede Schubert y Edna Luna Serrano

81 Capítulo 4
Las técnicas de pensamiento en voz alta como herramienta para la obtención de evidencias de validez en los instrumentos de evaluación
Adán Moisés García Medina, Felipe Martínez Rizo, Graciela Cordero Arroyo y Joaquín Caso Niebla

103 Capítulo 5
Aproximaciones metodológicas para la adaptación de test en contextos educativos
Carlos David Díaz López, Joaquín Caso Niebla y José Alfonso Jiménez Moreno

119 Capítulo 6
La evaluación del egreso de la Licenciatura en Actividad Física y Deportes de la Universidad Autónoma de Baja California
José Alfonso Jiménez Moreno y Joaquín Caso Niebla

135 Capítulo 7
Apuntes sobre la perspectiva de calidad en la Educación Superior mexicana
Ana Patricia Gutiérrez Zavala y José Alfonso Jiménez Moreno

PRÓLOGO

Desde su fundación en 1990, el Instituto de Investigación y Desarrollo Educativo de la Universidad Autónoma de Baja California ha desarrollado una muy relevante actividad académica centrada en la investigación educativa, específicamente, en el ámbito de la evaluación.

Particularmente, son cuatro las líneas de investigación en la que el Instituto y los cuerpos académicos en él radicados, centran sus focos de interés: la evaluación educativa del docente y del estudiante, los sujetos educativos y sus prácticas discursivas, la medición, innovación y desarrollo en educación y, por último, la mediación tecnológica del proceso educativo.

Hace casi tres años, en el prólogo de otra obra de similares características a ésta, afirmaba que, por encima de enumeraciones más o menos detalladas de temáticas, líneas o de proyectos de trabajo e investigación, había una cuestión que, en mi opinión, resultaba aún de mayor relevancia, a saber: el enfoque, el modo y estilo de trabajo, las directrices que marcan y guían la actividad cotidiana de la institución y de los académicos que la conforman. Al respecto, citaba las siguientes cuatro cuestiones:

- Primero, el equilibrio logrado en la atención prestada tanto a las cuestiones teóricas como a las instrumentales y metodológicas.
- Segundo, una profunda preocupación por el impacto social de los resultados de dichas investigaciones de forma que las mismas puedan operar como eficaces herramientas y palancas de cambio y mejora educativa.
- Tercero, el rigor metodológico y la excelencia académica que se plasma en una preocupación constante por la formación al más alto nivel tanto de los investigadores y profesionales como de los estudiantes de maestría y doctorado.
- Cuarto, la importancia concedida al trabajo en redes y al fomento de la internacionalización del instituto.

Actualmente, con la perspectiva que da el paso del tiempo, creo que lo entonces afirmado cobra aún mayor vigencia y relevancia en la medida en que el instituto, sus académicos y estudiantes son una referencia académica y profesional tanto a nivel nacional como internacional.

Y este libro que ahora el lector tiene en sus manos, constituye una sólida evidencia de lo hasta aquí dicho. Se trata de una obra coordinada por el doctor Jiménez Moreno y la doctora Luna Serrano en la que se presentan las experiencias de investigación sobre evaluación educativa desarrolladas en el posgrado.

Como verán, un primer bloque de trabajos se configura alrededor de la figura central de cualquier sistema educativo, a saber: la de la persona que ejerce la docencia. Se tratan cuestiones relativas a la evaluación de su desempeño, los estándares para la misma y las motivaciones para elegir la profesión. Pero junto a esto, van a encontrarse con excelentes trabajos de corte más metodológico (una de las señas de identidad del IIDE), y también con muy notables aportaciones sobre la evaluación y la calidad en Educación Superior.

En la primera contribución, Castro, Luna y Cordero abordan el delicado y complejo asunto de la evaluación docente en la educación básica desde una perspectiva tanto internacional como centrada en el propio caso mexicano. El trabajo se inicia con una excelente síntesis de la evaluación docente en el contexto internacional presentando diferentes experiencias llevadas a cabo en Estados Unidos, Australia, Inglaterra, Finlandia, Chile, Colombia y Cuba. En todos los casos, el análisis se centra en considerar cuatro cuestiones básicas: la función de la evaluación, los criterios empleados en la misma, el procedimiento seguido y las consecuencias o impactos que de cada evaluación se derivan.

Una vez examinadas estas importantes experiencias internacionales, el foco se centra en la evaluación docente en México. Para ello, se describen los programas del escalafón, el de la carrera magisterial y el Concurso Nacional para el Otorgamiento de Plazas Docentes. Y a continuación se describen y analizan con detalle y rigor las características de los nuevos programas de evaluación instituidos desde 2013 a raíz de la Reforma Educativa de 2012 y de la promulgación de la Ley General del Servicio Profesional Docente (LGSPD). El trabajo finaliza con importantes conclusiones en las que se enfatiza la importancia de la orientación a la mejora y a los programas de formación docente, la de los principios éticos del propio proceso evaluativo y la participación activa de los propios docentes en el marco de una cultura evaluativa, y en el carácter integral que ha de tener este tipo de evaluaciones de forma que se valoren todas las prácticas vinculadas a los procesos de enseñanza-aprendizaje.

Siguiendo con cuestiones relativas a la evaluación de la actividad docente, en el segundo capítulo, Pedroza, Luna y Caso se centran en los estándares de desempeño docente analizando la situación en México desde una perspectiva internacional con el objetivo de analizar los distintos estándares utilizados en la evaluación del desempeño docente en diversos países y compararlos con los empleados en México. En concreto, se estudian los *Estándares Profesionales Nacionales para Docentes de Australia*, los del Consorcio Interestatal de EEUU (*InTASC*), los formulados por el *National Board for Professional Teaching Standards (NBPTS)* y, por último, para el caso de México, los establecidos en *Perfil, Parámetros e Indicadores (PPI)*. El trabajo se centra en analizar los fines y propósitos de cada uno de ellos tratando de detectar similitudes y diferencias. Dicho análisis se efectúa tomando en consideración los cuatro niveles en que se organizan: 1) los principios generales, 2) los dominios, 3) los propios estándares y 4) las elaboraciones.

El primer resultado relevante de este capítulo muestra un alto nivel de acuerdo entre los distintos países sobre lo que evalúan de la práctica docente y sus fines. Y, como afirman los autores, este nivel de consenso es por sí mismo relevante en la medida en que permite valorar con mayor certeza la adecuación de los contenidos de la evaluación docente. La comparación con los estándares mexicanos (PPI) les ha permitido detectar lagunas importantes en el nivel de los principios y en el de las elaboraciones descriptivas de los niveles de desempeño, ausencia que dificulta un adecuado proceso evaluativo. Junto a ello, el trabajo concluye señalando la importancia de la coherencia de los PPI con el modelo educativo y el currículo oficial.

El tercer capítulo está firmado por García-Poyato, Cordero, Priede y Luna; en el mismo se lleva a cabo una caracterización de las motivaciones que llevan a los estudiantes a ingresar en la Escuela Normal (EN). Este trabajo surge a raíz de constatar el preocupante hecho de la acusada disminución de aspirantes por ingresar en los programas educativos de las Escuelas Normales. En el caso del Estado de Baja California (BC), las cifras son contundentes: en el curso escolar 2011-2012 hubo un total de 1,818 aspirantes; en el 2017-2018 solo 316, lo que conlleva una reducción superior al 83%. Tal y como se afirma en el trabajo, estas cifras traslucen una falta de interés de los jóvenes en la carrera docente como opción personal y profesional y es un fenómeno sin precedentes en el sistema educativo mexicano que puede conllevar importantes repercusiones.

Partiendo de esta situación, el trabajo se marca como objetivo explorar y caracterizar las motivaciones de los estudiantes de nuevo ingreso en tres Escuelas Normales de BC. Y para ello, después de un examen del soporte teórico sobre esta cuestión, se presenta el método y resultados

del trabajo empírico realizado que es de corte cualitativo, exploratorio y transversal en el que se ha empleado como técnica de recogida de información la de los grupos focales. Los resultados apuntan a la presencia de factores motivacionales ya señalados en diferentes estudios a nivel internacional, pero también a otros específicos que atañen a la formación de docentes de primaria en México. Una conclusión muy relevante es que los jóvenes que egresan de bachillerato eligen su carrera profesional sin pasar por un proceso serio de elección. Asimismo, se señala que cuestiones del ámbito personal, como las experiencias previas de enseñanza o la influencia de familiares, tienen un impacto importante en el proceso de toma decisión por parte de quienes aspiran al ingreso en las EN. El trabajo finaliza con un conjunto de medidas o recomendaciones basadas en sólidas evidencias y tendentes a frenar este proceso de disminución de aspirantes.

Como acabamos de ver, las tres primeras contribuciones han abordado cuestiones relativas a los docentes como figura clave del sistema educativo. En los cuatro capítulos restantes se abordan otras temáticas igualmente relevantes.

Así, en el cuarto capítulo, García, Martínez, Cordero y Caso, desde una perspectiva más técnico-metodológica, se centran en el uso de las técnicas de pensamiento en voz alta para la obtención de evidencias de validez de los instrumentos evaluativos, entendido este proceso como un análisis del ajuste entre los procesos cognitivos que usan los evaluados al realizar una prueba y la correspondencia de dichos procesos con las tareas y demandas cognitivas que el constructo latente debe desencadenar.

Para ello, se analizan los protocolos y reportes verbales, y se presenta el ejemplo del estudio de un reactivo de una prueba de logro en matemáticas que en aplicaciones previas había mostrado funcionamiento diferencial a favor de estudiantes de alto nivel socioeconómico. A continuación, el trabajo presenta una completa y precisa síntesis de las seis fases que conlleva este tipo de estudios: diseño, pilotaje, realización, procesamiento de la información, análisis cognitivo de los ítems y, por último, el análisis del sesgo mediante el Diseño Universal de las Evaluaciones. El trabajo concluye señalando la utilidad de este tipo de técnicas no solo para conocer el proceso de respuesta de los sujetos sino también para identificar problemas de sesgo potencial. De aquí se deriva una importante recomendación centrada en que los diseñadores de pruebas organicen grupos de trabajo específicos que afronten y minimicen las fuentes de sesgos potenciales en determinadas poblaciones o grupos que muestren funcionamiento diferencial para así lograr procesos evaluativos más justos y equitativos.

El quinto capítulo se centra también en cuestiones metodológicas, en este caso en la adaptación de test y pruebas en contextos educativos, y está firmado por Díaz, Caso y Jiménez. Asumiendo el principio que el principal objetivo del proceso de adaptación de un test es garantizar la validez del mismo, especialmente en lo concerniente a la validez de constructo, de contenido y ecológica, los autores describen los tres principales enfoques del proceso de adaptación: la traducción o adopción, la adaptación, y el ensamblaje. Después de analizar los mismos, se describen con detalle las dificultades y los problemas que plantean las traducciones y se insiste en la importancia de realizar análisis de sesgo que garanticen la igualdad de significados.

En una segunda parte del trabajo se abordan los procedimientos de validación en el proceso de adaptación de una prueba: el estudio piloto y la prueba de campo, la validación del constructo dentro del contexto objetivo, y la validación del constructo entre el contexto original y objetivo. Con respecto al estudio piloto, se enfatiza la importancia y el papel de la entrevista cognitiva y del pensar en voz alta, tal y como en el trabajo anteriormente comentado se abordó. En lo relativo a la validación del constructo, el análisis se centra en el empleo del Análisis Factorial Confirmatorio (AFC) como estrategia analítica apropiada para la obtención de evidencias al respecto. Y, por último, en lo tocante a la validación de constructo entre la población original y la de destino, junto con el ya citado AFC, se presentan cuestiones relativas al análisis diferencial de ítems. Se trata, en consecuencia, de una valiosa aportación para los interesados en las cuestiones de adaptación de pruebas y de control de sesgos culturales y socioeconómicos lo que, al igual que antes se señaló, es requisito para llevar a cabo evaluaciones justas y equitativas, sobre todo en un contexto de diversidad cultural y lingüística como el mexicano.

En el sexto capítulo, los autores Jiménez y Caso presentan un caso de evaluación en el nivel de la educación superior como es el de la evaluación de egreso de la licenciatura en Actividad Física y Deportes de la Universidad Autónoma de Baja California (UABC). Se trata de uno de los programas de la UABC que no contaba con procedimientos específicos de evaluación tendentes a acreditar el nivel de egreso de los estudiantes. Por ello, el IIDE asumió la tarea de desarrollar una propuesta, la cual es presentada en este trabajo, y que asume una perspectiva comprensiva empleando dos instrumentos: uno centrado en la medición de las competencias establecidas en el plan de estudios y otro, un cuestionario, cuyo foco de interés son las competencias de orden personal. La perspectiva o el enfoque comprensivo asumido implica diversidad de perspectivas y, en este sentido, se presenta una síntesis de los 12 pasos propuestos por Stake y que deben ser con-

siderados al desarrollar una evaluación. De esta forma, el foco de atención de esta evaluación es el proporcionar información valiosa sobre la formación de los futuros profesionales de este campo asumiendo la diversidad de contextos.

En la segunda parte del trabajo se describe el procedimiento adoptado para la elaboración de los instrumentos, prestando una especial atención a la definición del marco de la evaluación para lo que se empleó una estrategia de definición de dimensiones del constructo en función de las asignaturas y del perfil de egreso. El trabajo concluye señalando la necesidad de generar evidencias robustas que doten de sentido a todo el método utilizado, de forma que se destaque el valor intrínseco del fenómeno por evaluar y, sobre todo, de elaborar una propuesta que no esté centrada exclusivamente en la mera medición del aprendizaje, sino que trate de proporcionar información válida y fiable que permita la toma de decisiones orientadas a la mejora del programa.

Por último, en el séptimo capítulo, y como cierre de esta importante obra colectiva, los autores Gutiérrez y Jiménez abordan el complejo asunto de la perspectiva de la calidad de la Educación Superior en México. Los autores comienzan presentando las dos posturas que al respecto emergen en los discursos académicos: una visión orientada o centrada en la perspectiva de los derechos humanos y, una segunda, que asume una perspectiva de corte más empresarial o economicista.

Con respecto a la primera, se enfatiza el hecho de que todas las personas tienen derecho a una educación de calidad, lo que lógicamente implica que la calidad sea un elemento inherente al propio proceso educativo. Bajo esta idea, el trabajo analiza y describe el sistema de las 4-A propuesto por Tomasevski: asequibilidad, acceso, aceptabilidad y adaptabilidad. Una vez analizadas estas cuestiones y la importancia de la responsabilidad social de las instituciones concernidas, la segunda parte del trabajo pasa revista a lo que los autores denominan la "otra cara de la moneda", la "empresarialización" de la educación superior. El análisis se centra en describir cómo en las citadas instituciones el discurso y los términos vinculados al modelo empresarial van poco a poco impregnando el discurso oficial: competencias, competitividad, mercado laboral, empleabilidad, rankings, etc. El trabajo finaliza con una brillante síntesis del concepto de calidad como un derivado del Estado Evaluador en el contexto de la globalización y con un conjunto de nuevas preguntas de investigación tan relevantes como si ante la hegemonía del enfoque empresarial y economicista es factible incorporar la perspectiva de los Derechos en la Educación Superior.

En definitiva, y como decía al inicio de estas líneas, esta obra constituye de nuevo una excelente muestra del alto nivel del quehacer docente e investigador tanto del claustro del IIDE como de los estudiantes inscritos en los programas de posgrado. La obra presentada no solo pretende hacer una contribución al estado de conocimiento relacionado con la conceptualización de la evaluación y sus procedimientos, sino que es un insumo relevante para el desarrollo de la docencia a nivel posgrado, dada la información incluida, que representa la oportunidad de conocer elementos relevantes para la formación en evaluación.

Creo que la mayor fortaleza del IIDE radica en una sabia y virtuosa combinación de las dos actividades básicas de la universidad: una investigación de primer nivel y una docencia (en este caso de posgrado) de alta calidad y siempre vinculada a la actividad investigadora. De esta forma, ambas se enriquecen mutuamente. Y esta afortunada síntesis de la labor formativa y de la investigadora al más alto nivel, se basa además en el rigor metodológico y la vinculación y compromiso con la realidad educativa y social del entorno en que se inserta. Los trabajos que constituyen esta obra son claras evidencias al respecto.

Así, solo me queda felicitar al IIDE y a los autores y recomendar vivamente la atenta lectura de los textos aquí compilados por los coordinadores.

Dr. Luis Lizasoain Hernández

Profesor de Métodos de Investigación
y Diagnóstico en Educación
Universidad del País Vasco
San Sebastián, España

Capítulo 1

La evaluación docente en la educación básica, una mirada al contexto internacional y nacional

Antelmo Castro López, Edna Luna Serrano y Graciela Cordero Arroyo

Introducción

La evaluación educativa se reconoce como necesaria en la medida que hace posible incrementar la comprensión sobre: los estudiantes, el profesorado, los programas y servicios que se ofrecen; así como identificar los resultados y actuar en consecuencia para mejorar el funcionamiento del sistema (Stufflebeam, 2001; Phillips, 2018).

En el ámbito educativo, la evaluación se ha generalizado desde finales del siglo pasado. Se evalúa a las escuelas, a los planes y programas de estudio, a los estudiantes y su aprendizaje, y con más énfasis a los docentes por ser los actores clave en el logro de los procesos de enseñanza y de aprendizaje (Schmelkes, 1995; Luna, Elizalde, Torquemada, Castro y Cisneros-Cohernour, 2013; Marcelo y Valliant, 2009).

La evaluación docente se ha incorporado en los sistemas de educación básica de los países latinoamericanos, incluidos los países miembros de la OCDE (Cuevas y Moreno, 2016). El fin último es mejorar la calidad de la docencia a través del desempeño de los maestros para favorecer el aprendizaje de sus alumnos. Se reconoce que en los sistemas de evaluación docente convergen tanto los propósitos asociados a compensación salarial o control administrativo como los orientados al mejoramiento a partir de la reflexión de la práctica (Isoré, 2010; Cordero, Luna y Patiño, 2013; Martínez-Rizo, 2016).

El propósito de este capítulo es describir el caso de México respecto a los programas de evaluación docente de educación básica. Para ello, se presentan las características de los programas

para el ingreso, el reconocimiento y la promoción operantes antes de la promulgación de la Ley del Servicio Profesional Docente. Así como los esquemas de evaluación que actualmente regulan la carrera docente en la educación básica. Con el fin de contextualizar el análisis, se revisan las funciones de la evaluación y las particularidades de algunos programas de evaluación docente de educación básica de sistemas educativos de otros países.

Las funciones de la evaluación docente

Se reconocen dos funciones críticas de la evaluación docente: evaluación para el control y evaluación para la reflexión. Scriven (1991) estableció el término evaluación sumativa para referirse a la valoración de un programa (producto o servicio), comprobar su eficiencia y tomar decisiones sobre su continuidad, acción que sitúa al final de un proceso. También introdujo el término evaluación formativa para referirse a la valoración de procesos con el objetivo de incidir en su mejoramiento.

En las prácticas de evaluación docente, la evaluación sumativa tiene un carácter de control y rendición de cuentas. Su fin último es aportar información que sirva de base para tomar una decisión (Tucker y Kindred, 1997; Jornet, Leyva y Sánchez, 2009; Isoré, 2010). Puede consistir en elaborar reportes de los conocimientos y aptitudes de un maestro regularmente medido a través de un examen que permita garantizar el cumplimiento de un grupo de criterios o estándares para reconocerlo (Isoré, 2010). Con esto, se toman decisiones de tipo: bien-mal evaluado; pasa-no pasa; aprobado–reprobado; certificado-no certificado; admitido-no admitido; acreditado-no acreditado; merece promoción-no la merece; merece incremento de salario-no lo merece (Jornet et al., 2009).

La función de la evaluación formativa es la identificación de los elementos deficitarios y de aquellos componentes que puedan dinamizarse para la mejora de aquél que se evalúa en cualquier proceso educativo, sea individual, colectivo o relativo a un programa, institución u organización educativa (Jornet et al., 2009).

Dentro de un sistema educativo en el que convergen múltiples elementos y actores, la evaluación como un proceso intencionado de reflexión de las prácticas docentes permite establecer brechas entre la realidad y las acciones deseables para lograr que los estudiantes aprendan los contenidos que el nivel o grado demanda (Madueño, Márquez, Valdez y Castro, 2009). Este tipo de evaluación sirve como un recurso que ofrece datos empíricos a los diferentes actores dentro de los sistemas y subsistemas educativos para la toma de decisiones sobre prácticas pedagógicas y en la formulación de políticas educativas. Así entendida la evaluación, contribuye a crear una cultura

en donde se vela por la calidad educativa a partir de las realidades que suceden dentro de un aula, una escuela o un sistema escolar.

Si la evaluación docente en la educación básica tiene el objetivo de la mejora de la enseñanza y los resultados de los estudiantes (OCDE, 2010; Isoré, 2010), la evaluación debería ser formativa, lo que implica brindar una retroalimentación apropiada a los docentes sobre las necesidades para realizar la mejora (Luna et al., 2013), y no abordarse como la emisión de un juicio al comparar los resultados de estudios con un parámetro establecido (De la Orden, 1997; Instituto Nacional para la Evaluación de la Educación [INEE], 2009). En este sentido, sin retroalimentación de la práctica difícilmente se contribuye a mejorar el trabajo pedagógico del profesor.

La evaluación docente en el contexto internacional

Durante la década de los ochenta del siglo pasado, en Estados Unidos inició la presión sobre la educación que invitaba a dar cuenta de la calidad de la enseñanza y del aprendizaje (Dwyer y Stufflebeam, 1996); como consecuencia surgieron reformas educativas que incluían procesos de evaluación del profesorado. En Tennessee se registraron las primeras iniciativas con el programa Tennessee Value Added Assessment System (TVAAS) que relacionaba el aprendizaje de los estudiantes con el desempeño docente, directivos y el distrito; en 1986, en Texas se implementó el programa Texas Examination of Current Administrators and Teachers (TECAT), que evaluaba los conocimientos sobre la enseñanza y la gestión de las escuelas a través de una prueba; en Carolina del Norte se aplicó el programa The National Board of Professional Teaching Standards (NBPTS), cuyo fin era la certificación de los docentes (Marcelo y Valliant, 2009). En la tabla 1 se describe la función de la evaluación de estos tres programas.

A partir de la década de los noventa del siglo pasado, países miembros de la OCDE y de Latinoamérica diseñaron e implementaron sistemas nacionales de evaluación docente; uno de los retos que enfrentaron fue establecer los fines de la evaluación.

Ante la complejidad de qué evaluar, en Australia se estableció un marco de referencia a partir de la definición de las buenas prácticas; en 1997, crearon The National Framework for Professional Standards for Teaching para asegurar una mayor calidad de la profesión docente (Australian Institute for Teaching and School Leadership, 2012).

Tabla 1. *Ejemplos de programas de evaluación docente surgidos en la década de 1980 en Estados Unidos*

Estado	Programa	Función de la evaluación
Tennessee	TVAAS	Evaluar los aportes del distrito, la escuela y el docente en la formación del estudiante. *Método:* Pruebas anuales que relacionaban el desempeño docente con los avances de los estudiantes en sus aprendizajes. La valoración se basaba en la mejora que lograba el estudiante durante todo el año escolar.
Texas	TECAT	Establecer los conocimientos y competencias que un docente debería poseer en cada etapa del ejercicio profesional. *Método:* Una evaluación anual basada en un plan de desarrollo profesional de cada docente era el requisito para la renovación del empleo.
Carolina del Norte	NBPTS	Evaluar y certificar el desempeño docente. *Método:* La evaluación se basaba en un portafolio presentado por cada docente en diversas pruebas. Los aspectos que adquirían mayor ponderación para la evaluación final eran las relaciones con las tareas del aula.

Fuente: Marcelo y Vaillant (2009).

Actualmente, Chile es un pionero en la evaluación docente que ha marcado un hito en las políticas de fortalecimiento del profesorado y el establecimiento de estándares de desempeño docente (plasmados en el Marco de la Buena Enseñanza) (Gobierno de Chile, 2011).

Con el propósito de brindar un panorama sobre el papel de la evaluación docente –y dilucidar sus usos en diferentes sistemas educativos– se presenta la tabla 2, que concentra la función de la evaluación en siete países, seleccionados en función de sus características y desarrollo educativo: Estados Unidos, Australia, Inglaterra, Finlandia, Chile, Colombia y Cuba.[1]

La función de la evaluación en la mayoría de los países representados en la tabla 2 atiende a una evaluación formativa que apoya la toma de decisiones para incidir en la calidad de la enseñanza y de los aprendizajes dentro del aula. Regularmente derivan en acciones de formación y actualización docente. Sin embargo, países como Estados Unidos (particularmente el Distrito de Cincinnati), Chile y Colombia toman la evaluación docente con fines de control para prescindir del servicio docente a aquellos maestros que no logran demostrar sus habilidades y conocimientos para ejercer su práctica.

1 *A partir de datos documentados en la primera década del siglo XX.*

Tabla 2. *Función de la evaluación en sistemas nacionales de evaluación docente*

Función de la evaluación	Criterios	Procedimiento	Consecuencias
Estados Unidos (Distrito de Cincinnati) [2] Evaluación sumativa vinculada con remuneración económica.	Sistema de evaluación basado en 17 estándares de desempeño derivados del Marco para la Enseñanza agrupados en cuatro dominios: (1) planificación y preparación; (2) generación de un ambiente conducente al aprendizaje; (3) enseñanza orientada al aprendizaje; y (4) profesionalismo.	*Instrumentos:* rúbricas, de acuerdo a dos fuentes de evidencia principales: seis observaciones en el aula y un portafolio preparado por el maestro. *Evaluadores:* Un docente, el director y el subdirector.	A los maestros principiantes con bajos puntajes en la evaluación global en dos periodos pueden anularles el contrato. Los docentes permanentes con puntajes altos pueden convertirse en un líder pedagógico; con puntajes bajos podría llevar a la asignación a un programa de asistencia entre pares y finalmente a un término del contrato.
Australia Evaluación formativa.	Evaluación del cumplimiento de siete normas contenidas en tres dominios de la enseñanza según el Marco Nacional de Estándares Profesionales de la Enseñanza: (1) conocimiento profesional, (2) práctica profesional y (3) compromiso profesional.	La evaluación se lleva a cabo en cuatro momentos en un periodo de 2 años: planificación preliminar, recolección de datos, entrevista y seguimiento.	A partir de los resultados se establecen planes futuros de desarrollo profesional. En el segundo año se realiza un seguimiento para analizar el apoyo que recibió el docente y su avance hacia el logro de metas establecidas.
Inglaterra Evaluación formativa como parte de políticas escolares.	La evaluación está diferenciada de acuerdo a las 5 etapas de la carrera profesional: (1) certificación como Maestro Calificado; (2) maestros incluidos en el escalafón principal; (3) maestros incluidos en el tramo superior de la escala de remuneraciones; (4) Maestros Excelentes; y (5) Maestros con Habilidades Avanzadas.	Los estándares apoyan a los docentes en la identificación de sus necesidades de desarrollo profesional. En cada etapa los estándares comprenden tres dominios: atributos profesionales, conocimientos profesionales y la comprensión de la profesión y habilidades profesionales.	Cuando los docentes desean avanzar a la siguiente etapa de la carrera profesional, los estándares del nivel siguiente ofrecen un punto de referencia; si no desean avanzar, los estándares también los apoyan en la identificación de las maneras de ampliar y profundizar su experticia en su actual etapa de la carrera profesional.

2 *En el caso de Estados Unidos, es importante precisar que las decisiones en materia educativa se toman por Distritos, a diferencia de los países latinoamericanos que suelen tener un currículum nacional.*

Función de la evaluación	Criterios	Procedimiento	Consecuencias
Finlandia Evaluación formativa como parte de políticas escolares.	El esquema de evaluación docente se caracteriza por el alto nivel de confianza asignado a la escuela y profesionalismo de los docentes como base para el mejoramiento de la calidad de la enseñanza.	La autoevaluación de los maestros es considerada uno de los principales medios de optimización profesional. Los directivos comprometen a los maestros en la auto-reflexión acerca de su propia práctica y en desarrollar una cultura de evaluación junto con ambiciosos objetivos, de acuerdo al contexto y los problemas de cada escuela.	La mayoría de las escuelas han implementado debates anuales entre los directivos escolares y los docentes para evaluar el cumplimiento de los objetivos personales establecidos durante el año anterior y para establecer nuevos objetivos personales que sean concordantes tanto con el análisis del maestro como con las necesidades de la escuela.
Chile Evaluación sumativa y formativa (mixta).	Se utilizan los estándares contenidos en el Marco para Buena Enseñanza. Contiene 20 criterios distribuidos en cuatro dominios: (1) preparación para la enseñanza, (2) creación de un ambiente propicio para el aprendizaje, (3) responsabilidades profesionales y (4) enseñanza para el aprendizaje de todos los estudiantes.	Se recogen evidencias mediante cuatro artefactos: (1) un portafolio con muestras del trabajo de los maestros y un video de una de sus lecciones; (2) un formulario de autoevaluación estructurado; (3) una entrevista estructurada con un evaluador par; y (4) un informe de las autoridades administrativas y pedagógicas del centro escolar. La evaluación se realiza cada cuatro años. Los evaluadores se capacitan y certifican.	Los docentes calificados en nivel 'básico' reciben oportunidades específicas de desarrollo profesional. Los calificados 'insatisfactorio' reciben oportunidades de desarrollo y se evalúan al año, si no mejoran se despiden. Los evaluados como 'competentes' o 'excepcionalmente competentes' tienen prioridad en oportunidades de ascenso y en las actividades de desarrollo profesional de su interés. También pueden postular a bonificaciones salariales siempre y cuando atiendan una prueba sobre conocimientos curriculares y pedagógicos.
Colombia	Se evalúan desempeños y competencias del docente en su periodo de prueba.	Evaluación obligatoria al término del ciclo escolar siempre y cuando el docente haya ejercido al menos cuatro meses.	Una calificación sobre el 60% de la prueba los docentes son inscritos en el escalafón docente según su título académico. Docentes por debajo de ese puntaje son retirados del servicio.

Función de la evaluación	Criterios	Procedimiento	Consecuencias
Colombia Evaluación sumativa y formativa en diferentes momentos de la carrera docente.	Evaluación anual para docentes en servicio que aborda: dominio de estrategias y habilidades pedagógicas y de evaluación, didácticas específicas, plan de estudio, actitudes hacia los alumnos, logro de resultados entre otros dominios.	Evaluación anual de carácter obligatorio.	Los resultados son insumos para programas de formación y desarrollo docente.
	Evaluación de 'competencias' para docentes con fines de promoción.	Carácter voluntario.	Los docentes con buenos resultados son candidatos a un nivel salarial superior o de obtener un ascenso en el escalafón.
Cuba Evaluación con propósitos formativos	Se evalúan prácticas pedagógicas orientadas a mejorar la calidad de los aprendizajes de los alumnos.	Evaluaciones anuales obligatorias al término del ciclo escolar.	Los docentes que presentan dificultades que afectan la calidad del aprendizaje de sus alumnos, asisten a curso dirigidos. Los docentes con adecuado desempeño profesional eligen opciones de formación (cursos, maestrías, diplomados, etc.) alineados a sus intereses laborales y personales.

Fuente: Elaboración propia a partir de Isoré (2010), Marcelo y Vaillant (2009), OCDE (2009), Manzi, González y Sun (2011), Díaz-Fuentes, 1997 y del sitio web de *Australian Institute for Teaching and School Leadership* (http://www.teacherstandards.aitsl.edu.au/).

La función de evaluación en la mayoría de los países representados en la tabla 2 atiende a una evaluación formativa que apoya la toma de decisiones para incidir en la calidad de la enseñanza y de los aprendizajes dentro del aula. Regularmente derivan en acciones de formación y actualización docente. Sin embargo, países como Estados Unidos (Distrito de Cincinnati), Chile y Colombia toman la evaluación docente con fines de control para prescindir del servicio docente de aquellos maestros que no demuestran sus habilidades y conocimientos para ejercer su práctica.

La evaluación docente en la educación básica en México

En México, hasta el 2013, la carrera docente no estaba normada por una ley, por el contrario, eran negociaciones particulares entre la Secretaría de Educación Pública (SEP) y el Sindicato Nacional de Trabajadores de la Educación (SNTE) para el ingreso y promoción de los docentes. Los programas de *Escalafón, Carrera Magisterial* y el *Concurso Nacional* eran, de alguna manera, la carrera docente; se crearon de forma independiente y cada uno de manera aislada tuvo un recorrido diferente. A continuación, se describen estos programas.

Escalafón, primer mecanismo de regulación: promoción vertical

Escalafón fue el primer programa de evaluación docente a gran escala que se aplicó en México. Fue un sistema organizado en la SEP para efectuar las promociones de ascenso[3] de los trabajadores de base y autorizar permutas. Su operación se basaba en el Reglamento de *Escalafón* de los Trabajadores al Servicio de la Secretaría de Educación Pública, publicado en el Diario Oficial de la Federación, el 14 de diciembre de 1973 (SEP, 1973).

El programa denominado concurso escalafonario fue el procedimiento mediante el cual la Comisión Nacional Mixta de Escalafón (CNME) otorgaba el ascenso (a la dirección o supervisión) con base en la calificación de los siguientes factores (SEP, 1973).

Tabla 3. *Factores que se evaluaban en el programa* Escalafón

Factor	Componente	%	Puntaje máximo	Documentación que acredite
Conocimientos	Preparación, mejoramiento profesional y cultural.	45	1080	Mayor grado académico. Constancias de mejoramiento.
Aptitud	Iniciativa, laboriosidad, eficiencia, etc.	25	600	Crédito escalafonario.
Antigüedad	Años de servicio.	20	480	Constancia.
Disciplina y puntualidad.	Orden, cumplimiento, etc.	10	240	Crédito escalafonario.
	Total	**100**	**2400**	

Fuente: Ortiz (2003).

3 *Para los fines escalafonarios, se considera como ascenso a todo cambio a una categoría superior (DOF, 1973).*

Escalafón no fue un mecanismo obligatorio, sino opcional de acuerdo a los intereses individuales donde podían participar todos los docentes o no docentes de base, con una antigüedad mínima de seis meses (SEP, 1973). El programa recibió críticas, se dijo que era un mecanismo de promoción ineficaz para los docentes en servicio (OCDE, 2010; Ortiz, 2003) y que no promovía la reflexión sobre el trabajo realizado con los estudiantes (Díaz-Barriga, 2001).

Carrera Magisterial, segundo mecanismo de regulación: promoción horizontal

En 1992, el Gobierno Federal, los gobiernos de las entidades federativas y el SNTE firmaron el *Acuerdo Nacional para la Modernización de la Educación Básica* (ANMEB), cuyo propósito central fue transformar el sistema de educación básica para asegurar a los niños y jóvenes una educación que los formara como ciudadanos de una comunidad democrática, que les proporcionara conocimientos y capacidad para elevar la productividad nacional, que ensanchara las oportunidades de movilidad social y promoción económica de los individuos, y que elevara los niveles de calidad de vida de los educandos y de la sociedad en su conjunto (ANMEB, 1992).

El ANMEB tuvo el propósito fundamental de elevar la calidad de la educación pública y obedeció a la necesidad de superar rezagos y ampliar la cobertura; se concentró en tres ejes principales: la reorganización del sistema educativo; la reformulación de los contenidos y materiales educativos; y la revaloración de la función magisterial.

En el Acuerdo se declaró que el maestro debe ser el protagonista de la transformación educativa de México, debe fomentar la curiosidad intelectual y debe ser ejemplo de superación personal (Presidencia de la República, 1992). Por ello, uno de los objetivos centrales de la transformación educativa fue revalorar la función del maestro en seis aspectos principales: la formación, su actualización, el salario profesional, su vivienda, la carrera magisterial y el aprecio social por su trabajo. En este contexto, en 1993 se creó el Programa Nacional de Carrera Magisterial (Sandoval, 2001).

Carrera Magisterial se concibió como un programa de reconocimiento de carácter no obligatorio e individual para profesores de educación básica (preescolar, primaria, secundaria y grupos afines) cuyo objetivo era incidir en la calidad educativa a través del reconocimiento y apoyo a los docentes, así como en el mejoramiento de sus condiciones de vida, laborales y educativas (SEP, 2010). El programa tuvo su sustento en un sistema de evaluación por medio del cual era posible determinar a quién se le otorgaba el estímulo económico. El programa evaluaba seis factores: antigüedad, grado académico, preparación profesional, cursos de actualización y superación

profesional, desempeño profesional y aprovechamiento escolar. En el factor preparación profesional se evaluaban conocimientos y habilidades que requiere el docente para desarrollar su función, en el factor desempeño profesional se evalúa la planeación y desarrollo del proceso enseñanza-aprendizaje y la participación en la escuela y su interacción con la comunidad (SEP, 2011).

Al igual que *Escalafón, Carrera Magisterial* no fue considerado un programa eficaz de evaluación docente porque el fin último eran los estímulos económicos, y no la reflexión del proceso de la enseñanza que permitiera su mejora de la práctica. En ambos programas se carecía de procesos sistemáticos y periódicos para una retroalimentación directa sobre los procesos pedagógicos dentro del salón de clases (SEP, 1995; Ortiz, 2003; Santibañez,et al., 2006; OCDE, 2010).

Concurso Nacional para el Otorgamiento de Plazas Docentes, tercer mecanismo de regulación: ingreso

En el 2008, la SEP y el SNTE firmaron la Alianza por la Calidad de la Educación (ACE) y concretaron el Concurso Nacional para que el ingreso y la promoción de plazas de nueva creación y de las vacantes definitivas fuera por medio de un concurso nacional público de oposición para así contribuir a la calidad educativa del país y fortalecer al profesorado mediante la contratación de los profesionales mejor calificados para el ejercicio docente; además de fortalecer la transparencia y la rendición de cuentas.

En las primeras ediciones del concurso, la ACE emitía dos tipos de convocatorias al año: (a) para maestros de nuevo ingreso egresados de instituciones de educación superior y formadoras de docentes como la Escuela Normal y Normal de Estudios Superiores, la Universidad Pedagógica Nacional e Instituciones de Educación Superior de las diversas especialidades, alineados a los perfiles de las diversas plazas que se ofertaban; y (b) para maestros en servicio adscritos a la SEP con diferentes tipos de contratación: maestros con nombramiento provisional, maestros con contrato por honorarios, aspirantes a obtener una plaza de base, maestros con plaza de jornada aspirantes a la doble plaza, y docentes con horas-semana-mes que deseaban incrementarlas hasta obtener la plaza inicial de 19 horas, o la plaza de tiempo completo de 48 horas.

Para la asignación de plazas del ciclo escolar 2008-2009 se utilizó un mecanismo transitorio acordado y supervisado bilateralmente por la Comisión Rectora SEP-SNTE y a partir del ciclo 2009-2010 la convocatoria y dictamen corrió a cargo del Órgano de Evaluación Independiente con Carácter Federalista (OEIF) integrado por un representante de la autoridad educativa y uno del

sindicato de cada entidad federativa (SEP-SNTE, 2008). Este órgano fue una instancia temporal y colegiada de especialistas con autonomía técnica en el ejercicio de sus funciones. Su objetivo era dictaminar técnica y académicamente las evaluaciones aplicadas y hacer recomendaciones generales sobre los temas que plantea la Comisión Rectora (Luna, Cordero, López y Castro, 2012).

Para evaluar a los aspirantes a una plaza docente se utilizaba el *Examen Nacional de*

Tabla 4. *Resultados del Concurso Nacional en sus primeras seis ediciones por desempeño y tipo de examen*

Edición Ciclo escolar	Tipo de convocatoria	Presentaron examen	Aceptable	Requiere nivelación académica	No acaptable
2008-2009	Nuevo ingreso	53,406	16,095	-	37,311
	En servicio	17,648	7,150	-	10,498
	Total	**71,054**	**23,245**	**-**	**47,809**
2009-2010	Nuevo Ingreso	81,490	18,661	59,277	3,552
	En Servicio	42,366	12,425	28,464	1,477
	Total	**123,856**	**31,086**	**87,741**	**5,029**
2010-2011	Nuevo Ingreso	84,207	21,753	60,162	2,292
	En Servicio	61,776	17,390	42,931	1455
	Total	**145,983**	**39,143**	**103,093**	**3,747**
2011-2012	Nuevo Ingreso	89,9001	79,081	-	1,820
	En Servicio	58,544	57,254	-	1,290
	Total	**139,445**	**136,335**	**-**	**3,110**
2012-2013	Nuevo Ingreso	81,317	78,374	-	2,943
	En Servicio	53,387	51,534	-	1,853
	Total	**134,704**	**129,908**	**-**	**4,796**
2013-2014	Nuevo Ingreso	75,517	74,398	-	2,119
	En Servicio	64,676	62,710	-	1,966
	Total	**140,193**	**129,908**	**-**	**4,085**
	Global	**615,042**	**359,717**	**190,834**	**137,152**

Fuente: Elaboración propia a partir de los resultados del Concurso en sus seis ediciones contenidos en el sitio: *http://concursonacionalalianza.sep.gob.mx*

Conocimientos y Habilidades Docentes integrado por cuatro áreas: (1) habilidades intelectuales específicas, (2) dominio de contenidos curriculares, (3) competencias didácticas y (4) normatividad, gestión y ética docente.

En las primeras seis ediciones del Concurso, 755,235 maestros presentaron el Examen Nacional en las dos convocatorias. La tabla 4 muestra el desempeño a nivel nacional por tipo de convocatoria. Cabe destacar que la asignación de plazas dependía de la disponibilidad de las mismas y no del resultado absoluto del examen, es decir, los docentes con resultado Aceptable no les garantizaba la obtención de una plaza.

Hasta 2013, las políticas relacionadas con la evaluación docente en México habían sido reguladas por comisiones SEP-SNTE. Los tres grandes programas atendían a propósitos diferentes que permitía a los docentes someterse a voluntad a procesos de evaluación que le posibilitara alcanzar sus metas en el sistema educativo nacional. En la tabla 5 se resume la función de la evaluación en estos tres programas.

A partir de la aprobación de la Reforma Educativa en México, con la promulgación de la Ley General del Servicio Profesional Docente (LGSPD, 2013) y la Ley del Instituto Nacional para Evaluación de la Educación en el 2013, se regula la carrera docente promoviendo cambios en los esquemas de ingreso, promoción y permanencia. En la tabla 6 se presentan los nuevos programas de evaluación instituidos desde 2013 y sus características.

Tres de los cuatro programas declarados en la LGSPD y descritos en la tabla 6, corresponden a una evaluación sumativa de alto impacto en la vida del profesorado de educación básica. En estos nuevos esquemas prevalece la función de la evaluación para el ingreso, promoción y reconocimiento como en los programas que los antecedieron, con la diferencia de contar con un fundamento y una estructura de meritocracia en relación a los esquemas burocráticos del siglo pasado asociados a reconocimiento y promoción docente en América Latina (Cuenca, 2015).

Tabla 5. *Función de la evaluación en sistemas nacionales de evaluación docente*

Función de la evaluación	Criterios	Procedimiento	Consecuencias
Escalafón Evaluación sumativa vinculada con el ascenso a una categoría superior. Surgió en 1973 con fines de promoción vertical.	Se califican cuatro factores: conocimientos, aptitud, antigüedad, disciplina y puntualidad.	Es una evaluación anual organizada y aplicada por la SEP y el SNTE que convergen en la Comisión Estatal Mixta de Escalafón*. La participación es voluntaria.	Los resultados se organizan en orden de prelación. A los docentes con puntajes más altos se le asignaban la plaza para la cual concursó. El docente podía aceptar o rechazar la plaza; en el segundo caso se asignaba al siguiente docente según el orden de prelación.
Carrera Magisterial Evaluación Sumativa vinculada con remuneración económica. Surgió en 1993 con fines de promoción horizontal, se modificó en 1998 y 2011.	Se evalúan cinco factores: 1. aprovechamiento escolar 2. actividades cocurriculares 3. formación continua 4. preparación profesional y 5. antigüedad.	Es una evaluación anual organizada y aplicada por la SEP y el SNTE. La participación es voluntaria.	Los docentes con puntajes más altos se promueven al siguiente nivel que se traduce en incremento salarial. Si no se ingresa, los docentes se mantienen en su mismo nivel, que deberán refrendar cada dos años, de lo contrario se les retira del programa.
Concurso Nacional para el otorgamiento de Plazas Docentes Evaluación Sumativa vinculada con el ingreso. Surgió en 2008 con fines de ingreso.	Se aplica el Examen Nacional de Conocimientos y Habilidades que evalúa cuatro áreas: 1. habilidades intelectuales específicas 2. dominio de contenidos curriculares 3. competencias didácticas y 4. normatividad, gestión y ética docente.	Es una evaluación anual que regula la SEP-SNTE a la que pueden participar docentes de nuevo ingreso y docentes en servicio que aspiran a doble plaza. Así como profesionales especializados en un área. La participación es obligatoria.	Los resultados se organizan en orden de prelación. A los docentes con puntajes más altos se les asigna la plaza para la cual concursaron. La asignación de plaza está en función de la oferta de la SEP.

Nota: *De acuerdo a la descentralización de la educación en México, en cada estado existían políticas y disposiciones para llevar a cabo el programa. Fuente: Elaboración propia.

Tabla 6. *Función de la evaluación en sistemas nacionales de evaluación docente*

Función de la evaluación	Criterios	Procedimiento	Consecuencias
Concurso de Oposición para la Educación Básica Evaluación sumativa. Garantizar la idoneidad de los conocimientos y capacidades de los aspirantes para el ingreso al servicio público educativo.	Se evalúan 5 dimensiones (perfil) en dos exámenes: Examen nacional de conocimientos y habilidades para la práctica docente: • Dimensión 1. Un docente que conoce a sus alumnos, sabe cómo aprenden y lo que deben aprender. • Dimensión 2. Un docente que organiza y evalúa el trabajo educativo, y realiza una intervención didáctica pertinente. Examen nacional de habilidades intelectuales y responsabilidades ético-profesionales: • Dimensión 3. Un docente que se reconoce como profesional que mejora continuamente para apoyar a los alumnos en su aprendizaje. • Dimensión 4. Un docente que asume las responsabilidades legales y éticas inherentes a su profesión para el bienestar de los alumnos. • Dimensión 5. Un docente que participa en el funcionamiento eficaz de la escuela y fomenta el vínculo con la comunidad para asegurar que todos los alumnos concluyan con éxito su escolaridad.	Es una evaluación anual organizada y aplicada por la Secretaría de Educación Pública. La participación es obligatoria.	Los resultados se organizan en orden de prelación. A los docentes con puntajes más altos y considerados idóneos, se les otorga un nombramiento provisional que queda sujeto a una segunda evaluación que se realizará a los dos años de haber ingresado.

Función de la evaluación	Criterios	Procedimiento	Consecuencias
Concursos de oposición para las promociones a cargos con funciones de dirección y supervisión, y a las funciones de asesoría técnica pedagógica en educación básica Evaluación sumativa. Acceso a una categoría o nivel docente superior al que se tiene, sin que ello implique necesariamente cambio de funciones, o ascenso a un puesto o función de mayor responsabilidad y nivel de ingresos.	Se aplican dos exámenes estandarizados: Examen de conocimientos y habilidades para la práctica profesional. • Evalúa aspectos relacionados con el conocimiento de la organización y el funcionamiento de la escuela, el trabajo en el aula, el ejercicio de una gestión escolar eficaz, el desarrollo profesional docente o la asesoría técnica pedagógica. Examen de habilidades intelectuales y responsabilidades ético profesionales. • Evalúa las habilidades intelectuales de Directores, Supervisores y Asesores Técnicos Pedagógicos, sus capacidades para el estudio, la reflexión y la mejora continua de su práctica, así como el conocimiento de los principios legales y filosóficos del sistema educativo mexicano y las actitudes necesarias para el mejoramiento de la calidad educativa, la gestión escolar y los vínculos con la comunidad.	Es una evaluación anual organizada y aplicada por la Secretaría de Educación Pública. La participación es voluntaria para quienes aspiran a cargos de dirección, supervisión y asesoría técnica pedagógica.	Las plazas de dirección, supervisión o asesoría técnica pedagógica, se asignarán en estricto orden de prelación. Únicamente se considerará a los aspirantes cuyos resultados sean idóneos.

Función de la evaluación	Criterios	Procedimiento	Consecuencias
Evaluación de Desempeño Evaluación formativa. Continuidad en el servicio educativo. Contribuye al "fortalecimiento de las prácticas de enseñanza porque sus resultados aportarán información que retroalimente a los docentes para mejorar su quehacer profesional, y permitirán orientar las políticas de formación continua dirigidas a los maestros en servicio" (SEP, 2018, p. 5).	Se lleva a cabo en tres etapas: Etapa 1. Informe de responsabilidades profesionales: • Prácticas de convivencia en la comunidad escolar • Prácticas de inclusión en la comunidad escolar • Cumplimiento de la normativa vigente • Vinculación con las familias e instituciones • Trabajo colegiado para la mejora escolar • Mejora de la práctica profesional docente y técnico docente Etapa 2. Proyecto de enseñanza: • Elaboración del diagnóstico y de la planeación didáctica • Intervención docente • Respuesta a tareas evaluativas diseñadas para llevar a cabo un proceso de reflexión Etapa 3. Examen de conocimientos didácticos y curriculares: Se evalúa el conocimiento curricular, disciplinar y sobre el aprendizaje que deben poseer los docentes o técnicos docentes, así como sus habilidades para la organización e intervención didáctica, considerados en las dimensiones del perfil.	Es una evaluación que realiza por lo menos cada cuatro años. Es una evaluación con carácter obligatorio.	Los docentes que obtengan un nivel No suficiente deberán participar en un programa de regularización por tutoría y formación continua. Los docentes que obtengan un nivel Suficiente logran una permanencia por 4 años en el servicio y son sujetos a otra evaluación dentro de 4 años.

Función de la evaluación	Criterios	Procedimiento	Consecuencias
Promoción en la Función por Incentivos Evaluación sumativa, asociada a estímulos económicos. "Otorgar niveles de incentivos al personal docente, técnico docente, de dirección, supervisión y asesoría técnica pedagógica, con resultados destacados y sobresalientes en sus evaluaciones, e impulsar su mejora continua" (SEP, 2015, p. 15).	La asignación de incentivos en esta promoción está sujeta a los resultados en los procesos de Evaluación del Desempeño y los criterios de evaluación descritos anteriormente. También atiende a una Evaluación Adicional para la Promoción en la Función por Incentivos que se aplicará de conformidad con lo que determine el INEE.	Es una evaluación anual que está sujeta a los tiempos de ingreso en el Servicio Profesional Docente. La participación en el programa es voluntaria e individual.	Quienes en la evaluación del Desempeño obtengan un resultado **Destacado**, participa en el Programa de Promoción en la Función e ingresa al primer nivel. Los que obtengan un resultado Incremento asciende en el Programa de Promoción en la Función. De acuerdo a los resultados, los incentivos son económicos, de carácter temporal o permanente, estos últimos los puede conservar el participante durante toda su vida laboral. Son independientes al sueldo, con repercusiones en aguinaldo y prima vacacional y se consideran para efectos de la seguridad social.

Fuente: Elaboración propia con base en los documentos oficiales.

Consideraciones finales

A partir de la función de la evaluación en los programas de evaluación docente en la educación básica en los sistemas educativos se derivan las siguientes reflexiones.

1. La evaluación en sí misma no genera cambio ni hace mejores docentes de manera automática (Martínez-Rizo, 2016), su alcance es aportar elementos, datos o información sobre la situación actual de contextos educativos, en particular de las prácticas asociadas al desempeño docente, por lo tanto, es importante la existencia de esquemas dentro de los sistemas educativos que permita una evaluación formativa.
2. La evaluación docente en la educación básica debe tener la intención de mejorar las prácticas que promueven el aprendizaje de los estudiantes y obtener resultados favorables según los objetivos o competencias que los planes establecen. A su vez, debe ayudar a identificar las fortalezas y deficiencias de los profesores para elaborar políticas o diseñar programas asociados a procesos de formación como lo sugiere Isoré (2010) y la OCDE (2010).
3. La evaluación docente debe ser confiable, instrumentada bajo principios éticos que generen en los docentes su participación dentro de una cultura evaluativa que obligue a reflexionar sobre las prácticas del salón de clases que promueven los aprendizajes y el dominio de competencias en los estudiantes.
4. En el caso de México, los nuevos programas de evaluación deben promover la confianza y participación por parte del gremio docente del sistema de educación básica, sobre todo de aquellos que vivieron procesos de evaluación con los programas que operaban antes de la implantación de la Ley del Servicio Profesional Docente, asociados a prácticas discrecionales para el ingreso, promoción y permanencia vinculadas (Castro, 2016).
5. La evaluación docente en sí misma es compleja, se deben considerar marcos de referencia o estándares de desempeño docente para valorar todas las actividades asociadas directamente con los procesos de enseñanza y de aprendizaje en un salón de clases (o en entornos virtuales) que abordan las prácticas del docente antes, durante y después de lo que ocurre en un 'episodio didáctico' que incluye su planeación, impartición de clases, revisión de trabajos, tareas y evaluación de los aprendizajes; la labor magisterial; y la actividad profesional. Se debe partir de conceptualizaciones que definan a un buen docente relacionado con resultados de aquel cuyos alumnos muestran altos niveles de aprendizaje, y relacionados con procesos cuyo trabajo en el aula consiste en buenas prácticas docentes (Martínez-Rizo, 2016).

Referencias

Acuerdo Nacional para la Modernización de la Educación Básica (ANMEB). (1992). *Diario Oficial de la Federación*, DOF 19-05-1992. Recuperado de http://dof.gob.mx/nota_detalle.php?codigo=4666809&fecha=19/05/1992

Australian Institute for Teaching and School Leadership (AITSL). (2011). *Australian Professional Standards for Teachers.* Australia: AITSL. Recuperado de https://www.aitsl.edu.au/teach/standards

Castro, A. (2016). *Los programas de evaluación docente de la educación básica en México: una mirada desde los actores participantes.* (Tesis de doctorado). Universidad Autónoma de Baja California, México.

Cordero, G., Luna, E., & Patiño, N. (2013). La evaluación docente en educación básica en México: panorama y agenda pendiente. *Sinéctica, 41*, 2-19. Recuperado de http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-109X2013000200008

Cuenca, R. (2015). *Las carreras docentes en América Latina. La acción meritocrática para el desarrollo profesional. Estrategia Regional Sobre Docentes.* Chile: Orealc-Unesco Santiago. Recuperado de http://unesdoc.unesco.org/images/0024/002440/244074s.pdf

Cuevas, C., & Moreno, T. (2016). Políticas de Evaluación Docente de la OCDE: Un Acercamiento a la Experiencia en la Educación Básica Mexicana. *Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas, 24*, 1-20.

De la Orden, A. (1997). Evaluación y optimización educativa. En H. Salmerón (coord.), *Evaluación educativa: teoría, metodología y aplicaciones en áreas de conocimiento* (pp. 13-28). España: Grupo Editorial Universitario.

Díaz-Fuentes, A. (1997). El sistema de formación inicial y continua del personal docente de Cuba. *Revista Educación y Pedagogía, 9*(17), 137-151.

Díaz-Barriga, A. (2001). Carrera magisterial y evaluación de profesores. Entre lo administrativo y lo pedagógico. En M. Rueda (coord.), *¿Hacia una nueva cultura de la evaluación de los académicos?* (pp. 84-102). México: Centro de Estudios Sobre la Universidad.

Dwyer, C.A., & Stufflebeam, D. (1966). Teacher evaluation. En D.C. Berliner y R.C. Calfee (eds.), *Handbook of educational psychology* (pp. 765-785). Estados Unidos: Prentice Hall International.

Gobierno de Chile. (2011). *Marco para la Buena Enseñanza.* Santiago: Gobierno de Chile.

Instituto Nacional para la Evaluación de la Educación (INEE). (2009). Los exámenes de la calidad y el logro educativos del INEE ¿Qué son y cómo se construyen? ¿Cómo saber si aprenden nuestros estudiantes? *Este País, 224*, 1-12. Recuperado de http://estepais.com/inicio/historicos/177/14_suplemento_como%20saber_inee.pdf

Isoré, M. (2010). *Evaluación docente: prácticas vigentes en los países de la OCDE y una revisión de la literatura.* Chile: Preal.

Jornet, J.M., Leyva, Y., & Sánchez, P. (2009). Dimensiones de la clasificación de los procesos de evaluación educativa. En J.M. Jornet & Y. Leyva (comps.), *Conceptos, metodología y profesionalización en la evaluación educativa* (pp. 66-75). México: Instituto Internacional de Investigación de Tecnología Educativa.

Ley General del Servicio Profesional Docente (2013). *Diario Oficial de la Federación*, DOF11-09-2013. Recuperado de https://www.dof.gob.mx/nota_detalle.php?codigo=5313843&fecha=11/09/2013

Luna, E., Cordero, G., López, G., & Castro, A. (2012). La evaluación del profesorado de educación básica en México: Políticas, programas e instrumentos. *Revista Iberoamericana de Evaluación Educativa, 5*(2), 231-244.

Luna, E., Elizalde, L., Torquemada, A.D., Castro, A., & Cisneros-Cohernour, E. (2013). La evaluación de la docencia en el Sistema Educativo Nacional 2002-2012. En A. Maldonado (coord.), *Educación y Ciencia: políticas y producción del conocimiento 2002-2011* (pp. 399-454). México: ANUIES-COMIE.

Madueño, M.L., Márquez, L., Valdez, A., & Castro, A. (2009). La evaluación de la docencia como elemento para la mejora de la calidad de la educación. *Apuntes y Aportaciones de Proyectos e Investigaciones en Educación*. México: Instituto Tecnológico de Sonora.

Manzi, J., González, R., & Sun, Y. (2011). *La evaluación docente en Chile*. Chile: Mide UC.

Marcelo, C., & Valliant, D. (2009). *Desarrollo profesional docente ¿cómo se aprende a enseñar?* Madrid, España: Narcea.

Martínez-Rizo, F. (2016). *La evaluación de docentes de educación básica. Una revisión de la experiencia internacional*. México: Instituto Nacional para la Evaluación de la Educación.

Organización para la Cooperación y el Desarrollo Económico (OCDE). (2009). *Análisis de las políticas para maestros de educación básica en México.* México: OCDE.

Organización para la Cooperación y el Desarrollo Económico (OCDE). (2010). *Mejorar las escuelas: estrategias para la acción en México.* México: OCDE.

Ortiz, M.B. (2003). *Carrera Magisterial, un proyecto de desarrollo profesional.* México: Secretaría de Educación Pública.

Phillips, D.C. (2018). The many functions of evaluation in education. *Education Policy Analysis Archives, 26*(46). http://doi.org/10.14507/epaa.26.3811

Sandoval, F. (2001). Ser maestro de secundaria en México: condiciones de trabajo y reformas educativas. *Revistas Iberoamericana de Educación, 25*, 83-103.

Santibáñez, L., Martínez, J.F., Datar, A., McEwan, P.J., Messan-Setodji, C., & Basurto-Dávila, R. (2006). *Haciendo camino. Análisis del sistema de evaluación y del impacto del programa de estímulos docentes Carrera Magisterial en México.* Estados Unidos: Rand Education.

Schmelkes. S. (1995). *Hacia una mejora de nuestras escuelas.* México: Secretaría de Educación Pública.

Scriven, M. (1991). *Evaluation thesaurus.* Estados Unidos: SAGE.

Secretaría de Educación Pública (SEP). (1973). Reglamento de escalafón de los trabajadores al servicio de la Secretaría de Educación Pública. *Diario Oficial de la Federación,* DOF 14-12-1973. Recuperado de https://www.sep.gob.mx/work/models/sep1/Resource/24e8c91d-d2fc-4977-ad19-dc572c3e4439/reglamento_escalafon_sep.pdf

Secretaría de Educación Pública (SEP). (1995). *Programa Nacional de Educación 1995-2000*. México: SEP.

Secretaría de Educación Pública (SEP). (2010). *Programa Nacional de Carrera Magisterial.* México: SEP. Recuperado de http://www.sep.gob.mx/es/sep1/cncm_pncm

Secretaría de Educación Pública (SEP). (2011). Comunicado 057-Firman SEP y SNETE Acuerdo para la Reforma de los Lineamientos Generales del Programa Nacional de Carrera Magisterial. *Sala de prensa.* Recuperado de http://www.sep.gob.mx/es/sep1/C0570511

Secretaría de Educación Pública (SEP). (2015). *Programa de promoción en la función por incentivos en educación básica.* México: SEP. Recuperado de https://www.sep.gob.mx/work/appsite/VBReglamento_final_2015.pdf

Secretaría de Educación Pública (SEP). (2018). E*tapas, aspectos, métodos e instrumentos proceso de evaluación del desempeño docente y docente y técnico docente en educación básica, ciclo 2018-2019.* México: SEP.

Secretaría de Educación Pública-Sindicato Nacional de los Trabajadores de la Educación (SEP-SNTE). (2008). *Alianza por la calidad de la educación.* México: SEP/SNTE.

Stufflebeam, D. (2001). Evaluation Models. *New Directions for Evaluation,* (89), 7-98.

Tucker, P.D., & Kindred, K.P. (1997). Legal considerations in designing teacher evaluation systems. En J. Stronge (edit.), *Evaluating teaching. A guide to current thinking and best practice* (pp. 59-90). Estados Unidos: Corwin Press.

Capítulo 2

Estándares de desempeño docente: México en el contexto internacional

Luis Horacio Pedroza Zúñiga, Edna Luna Serrano y Joaquín Caso Niebla

Introducción

En el ámbito nacional e internacional existe un interés creciente por realizar mejores evaluaciones de los docentes y las funciones que desempeñan. Este interés se encuentra influenciado por la preocupación por la calidad de la educación y el reconocimiento de que los docentes son la pieza más importante para mejorar los resultados educativos.

Uno de los aspectos más relevantes en cualquier evaluación es definir su contenido, es decir, seleccionar y acordar qué se evalúa. Los estándares representan el contenido de evaluación sobre la docencia. La práctica docente es una labor altamente compleja, por lo que llegar a definir lo que se tiene que evaluar es una tarea ardua. En años recientes, distintos países e instituciones han emitido estándares donde se establecen los contenidos de evaluación.

Con base en ello, el presente capítulo propone hacer un análisis de distintos estándares utilizados en la evaluación del desempeño docente en diversos países y compararlos con los empleados en México. Se pretende orientar la discusión sobre la evaluación docente en torno a dos de sus componentes: fines y contenido de la evaluación. Además, se busca hacer una valoración de si lo que se propone evaluar resulta congruente con la forma de diseñar y elaborar dichos estándares.

Los estándares, al igual que otras técnicas de evaluación, se utilizan para fines sumativos, conocer lo que los docentes conocen y son capaces de hacer. Pero al mismo tiempo los estándares también representan una herramienta poderosa para promover procesos de mejora y orientar al sistema educativo a mejores desempeños a través de procesos de formación y desarrollo profesional docente.

Estándares de desempeño

En castellano, *estándar* se define como un adjetivo "que sirve como tipo, modelo, norma, patrón o referencia", según la Real Academia Española (RAE, 2014). En lengua inglesa es un sustantivo con dos acepciones; como bandera o estandarte, y como especificación con la cual la calidad requerida de algo puede ser probada (Ingvarson y Kleinhenz, 2006). Estas definiciones son muy cercanas al uso que se le da al término en el contexto educativo. La acepción de estandarte se deriva de la orientación que se daba en las batallas a los soldados para alcanzar ciertas posiciones. Bajo este significado, los estándares intentan ser una inspiración, aspiración, una descripción de una visión de lo que podría ser. Se orientan hacia el futuro, pero capturan la esencia de lo mejor de una práctica, sustentada en lo que la investigación reciente señala como buenas prácticas.

Esta apreciación pone de manifiesto dos características de los estándares: su carácter aspiracionista y su fundamentación en la evidencia de la investigación científica. Si no se cumple con la primera característica y se incluye únicamente aquello que los docentes pueden lograr, es decir "se fija la vara muy abajo", con ello no se orienta a los docentes a alcanzar prácticas que se consideran mejores. Al respecto, Ingvarson (2013) señala que los estándares significan un puente entre la práctica y la investigación, pues son una forma de representar el lenguaje de la investigación con expectativas para la práctica docente. Asimismo, es necesario puntualizar que los estándares no son estáticos, puesto que la investigación y lo que se espera que los docentes hagan no lo son. Estos requieren de una revisión regular a la luz de nuevos hallazgos y expectativas sociales sobre la docencia y la educación.

En el terreno evaluativo, los estándares se refieren a normas, criterios de exigencia, o a la magnitud de una unidad de medida. Es decir, "es el nivel de desempeño respecto de los criterios que se evalúan" (Ingvarson, 2013, p. 28). Para ello, los estándares no solo necesitan describir lo que sabe y debe poder hacer un docente, también los distintos niveles de calidad.

Se pueden distinguir dos acciones relacionadas con los estándares: describir estándares y fijar estándares. La primera refiere a la descripción del conocimiento y habilidades que son valoradas, denominados como estándares de contenido. El proceso de fijar estándares implica llegar a acuerdos sobre las habilidades y conocimientos que se deben demostrar para alcanzar un nivel adecuado, identificado como fijar estándares de desempeño (Ingvarson, 2013). Este documento hace referencia a la primera de las acciones, la descripción de los estándares.

A continuación, se presenta una breve descripción de los estándares objeto de análisis en el presente documento:

Estándares Profesionales Nacionales para Docentes de Australia. Estos estándares son emitidos en 2011 por el propio ministerio de educación del país, para lo cual se tomaron en consideración las elaboraciones previas de sus estados y territorios. Investigadores educativos analizaron estándares disponibles en otros países y se promovió la participación de los docentes en la elaboración de los mismos (Australian Institute for Teaching and School Leadership [AITSL], 2012b).

Los estándares están inmersos dentro de un modelo de mejora de la enseñanza, el cual busca involucrar a toda la comunidad escolar en la creación de una cultura de desempeño y desarrollo. Esta cultura de desempeño y desarrollo está caracterizada por el claro propósito de la mejora de la enseñanza (AITSL, 2012a).

Interstate Teacher Assessment and Support Consortium (InTASC). Evaluación de los Estados Unidos, creada en 1987. En un principio sólo incluyó estándares para evaluar a los docentes que querían acreditarse en el ingreso a la profesión. Pero, en 2011 fueron sujeto de un proceso de reforma, a partir de la cual dejó de ser una evaluación exclusivamente para certificar el ingreso a la docencia, ampliando el rango a todo el ejercicio profesional docente (*Council of Chief State School Officers* [CCSSO], 2013). Otro cambio es su alineación a otras propuestas existentes en los estados de la unión americana, como los elaborados por el NBPTS, así como innovaciones en la descripción de los estándares para facilitar su uso con fines formativos.

National Board for Professional Teaching Standards (NBPTS). Es una asociación estadunidense sin fines de lucro, integrada principalmente por docentes. Su misión es promover el aprendizaje de los alumnos por medio del establecimiento de un sistema de estándares y certificaciones para docentes experimentados. Los estándares son específicos de acuerdo a la disciplina y el nivel de desarrollo de los alumnos. El sistema de estándares emite, en total, 25 certificados en 16 áreas de contenidos y cuatro niveles de desarrollo de los alumnos.

Es una certificación de gran prestigio. Inició sus trabajos en 1987, la creación de los estándares estuvo liderada por Lee Shulman y contaron con la participación de investigadores muy destacados como Cronbach. A la fecha, el NBPTS ha emitido más de cien mil certificados en 50 estados de EUA (NBPTS, s.f.).

Perfil, Parámetros e Indicadores (PPI). Este es el nombre que se le dio a los estándares para la evaluación docente en México. En el 2013 el sistema educativo mexicano inició una reforma educativa que implicó evaluar a los docentes del tramo de educación básica y media superior. Para ello, se promulgó la Ley General del Servicio Profesional Docente (LGSPD, 2013). La evaluación tiene como referente un conjunto de estándares de evaluación o marco de buena enseñanza denominados Perfil, Parámetros e Indicadores (PPI). La Ley propone tanto una evaluación sumativa: de ingreso, ascenso y permanencia, como una evaluación permanente y de carácter formativo, que tiene la finalidad de apoyar a los docentes en la evaluación y en su desarrollo profesional. La primera versión de los PPI se generó en el ciclo 2014-2015 y se han venido emitiendo cada ciclo escolar, con variaciones mínimas entre ellos (Luna, Pedroza, y Córdova, 2019).

Los estándares han sido utilizados con fines de certificación, selección por empleadores, evaluación, autoanálisis y reflexión sobre la práctica, así como para guiar el desarrollo profesional (Ingvarson y Kleinhenz, 2006). A continuación, se presentan los propósitos de los estándares analizados.

Los estándares australianos (AITSL, 2012b) tienen los siguientes propósitos:

1. Hacer explícitos los elementos de una enseñanza efectiva de alta calidad.
2. Ofrecer un marco para el aprendizaje profesional de los docentes, la autoevaluación y la reflexión sobre su práctica.
3. Contribuir a la profesionalización de los docentes y elevar el estatus de su profesión.
4. Establecer un modelo de rendición de cuentas profesional, ayudando a asegurar que los docentes tengan las competencias profesionales adecuadas.

El propósito de los estándares del InTASC es describir una nueva visión de la enseñanza en la cual se aspira a transformar el sistema educativo norteamericano para satisfacer las necesidades de los aprendices actuales, y servir como un recurso a disposición de distintos agentes educativos en el desarrollo de políticas y programas para preparar, certificar, apoyar, evaluar y recompensar a los docentes. Además, tienen tres funciones: servir como estandarte, ofreciendo una visión del lugar a donde se quiere ir; definir niveles del desempeño específico que debe ser alcanzado; y articular las actividades de aprendizaje profesional y los apoyos que deben proporcionarse, para garantizar que los docentes tengan las oportunidades para alcanzar los estándares (CCSSO, 2013).

Mientras que el propósito de los estándares del NBPTS es: elevar el estatus, la voz y el rol que cumplen los docentes y conformar a la docencia como una verdadera profesión; definir lo que los docentes expertos deben conocer y ser capaces de hacer; así como reconocer los logros de la enseñanza a través de un proceso de certificación riguroso, comparable con otras profesiones como medicina, ingeniería y derecho (NBPTS, s.f.).

Por su parte, el PPI de México retoma los propósitos de la LGSPD (SEP, s.f.), los cuales expresan lo siguiente:

1. Mejorar, en un marco de inclusión y diversidad, la calidad de la educación y el cumplimiento de sus fines para el desarrollo integral de los educandos y el progreso del país.
2. Mejorar la práctica profesional mediante la evaluación en las escuelas, el intercambio de experiencias y los apoyos que sean necesarios.
3. Asegurar, con base en la evaluación, la idoneidad de los conocimientos y las capacidades del personal docente.
4. Estimular el reconocimiento a la labor docente mediante opciones de desarrollo profesional.
5. Asegurar un nivel suficiente de desempeño en quienes realizan funciones de docencia.
6. Otorgar los apoyos necesarios para que el personal del Servicio Profesional Docente pueda, prioritariamente, desarrollar sus fortalezas y superar sus debilidades.
7. Garantizar la formación, capacitación y actualización continua a través de políticas, programas y acciones específicas.
8. Desarrollar un programa de estímulos e incentivos que favorezca el desempeño eficiente del servicio educativo y contribuya al reconocimiento escolar y social de la dignidad magisterial. (LGSPD, 2013, art. 13)

A partir de la exposición de los propósitos de cada uno de los estándares se pueden identificar algunas semejanzas y diferencias. Los cuatro modelos de estándares se utilizan como un mecanismo para favorecer la profesionalización de los docentes; establecen una guía de aquello que los docentes deben conocer y saber hacer, ofrecen un proceso de certificación de las competencias y representan un recurso para el desarrollo profesional, la autoevaluación y la reflexión sobre la práctica.

Una de las diferencias encontradas consiste en que el modelo desarrollado por el InTASC es el único que explícitamente propone el uso de los estándares para el diseño de políticas educativas.

Por otra parte, los estándares australianos son los únicos que declaran en sus propósitos utilizar los estándares con fines de rendición de cuentas.

Niveles de los estándares

Los modelos de estándares se estructuran en cuatro niveles: 1) los principios generales, 2) los dominios, 3) los estándares y 4) las elaboraciones. Éstos deben tener un orden que va de lo general a lo específico (Ingvarson, 2013):

- Nivel 1. Principios generales. Reflejan un aprendizaje y enseñanza de calidad. Son afirmaciones generales, buscan englobar los valores y fines educativos que debe guiar a los profesores.
- Nivel 2. Dominios. Son las categorías en que se dividen las funciones docentes y las afirmaciones que las delimitan.
- Nivel 3. Estándares. Estas afirmaciones describen lo que necesitan saber y hacer los docentes. Sirven como base para realizar las evaluaciones.
- Nivel 4. Elaboraciones. En este nivel las elaboraciones explican el significado de cada uno de los estándares. Se señalan acciones observables y adecuadas para el estándar. Estas elaboraciones proporcionan una base para usar rúbricas y valorar el desempeño del docente.

Nivel 1. Los principios

Los estándares de desempeño docente deben basarse en una visión coherente de lo que la sociedad considera una educación de calidad y en la función del profesor para ofrecer esa educación. Una visión de lo que significa el logro de un aprendizaje de calidad por los estudiantes y de aquello que los profesores deberían saber, creer y poder hacer para alcanzar ese aprendizaje de calidad (Ingvarson, 2013).

En la revisión de los estándares descritos anteriormente se encontró que algunos retoman una visión de educación de calidad de otros documentos más generales, incluso de orden normativo o legal. Por ejemplo, los estándares australianos retoman el concepto de aprendizaje de calidad de documentos oficiales, en ellos se especifican las metas de la escolarización, las cuales son: promover la calidad y la excelencia, y que todos los jóvenes lleguen a ser aprendices competentes, seguros en sus capacidades, creativos y unos ciudadanos activos e informados (AITSL, 2012b). Sobre la práctica docente se establece un continuo de experiencia o dominio de la profesión donde

se establecen niveles: profesor egresado, profesor competente, profesor con alto desempeño y profesor líder.

Por su parte, el InTASC expresa una visión en la cual cada niño debe aprender a tomar el control de su propio aprendizaje, no solo se enfatiza el aprendizaje sino la aplicación de los conocimientos y habilidades en la resolución de problemas de la vida real, se valoran las diferencias que cada aprendiz proporciona a la experiencia de aprendizaje, y se reconocen las posibilidades que ellos traen consigo para maximizar el aprendizaje de todos. El InTASC considera que los siguientes principios subyacen a la práctica docente (CCSSO, 2013):

- El aprendizaje y la enseñanza son complejos.
- La capacidad para la enseñanza puede ser aprendida, desarrollada en el tiempo y no es lineal.
- El crecimiento ocurre a través de la reflexión sobre la experiencia, la retroalimentación y las experiencias de aprendizaje grupales o individuales.
- El desarrollo profesional depende del contexto y del nivel de apoyo que se proporcione.
- El objetivo es la práctica docente y no sobre el docente en particular.

El NBPTS presenta una visión de una práctica docente experta, aquello que el docente debe conocer y ser capaz de hacer. Esta institución señala que a partir del concepto de aprendizaje se derivan principios o supuestos que guían la práctica docente. Además, indica que los requisitos fundamentales para una enseñanza de calidad están relativamente claros: conocimientos de los contenidos a enseñar, de las habilidades a desarrollar y del currículo; conocimiento general y específico de los métodos de enseñanza, así como de evaluación de los aprendizajes; conocimiento de los alumnos y del desarrollo humano; habilidades para la enseñanza de alumnos diversos; y habilidades, capacidades y disposiciones para utilizar ese conocimiento en favor de los alumnos (NBPTS, 1999 citado por National Research Council [NRC], 2008). Asimismo, el NBPTS hace explícito un modelo explicativo sobre el proceso que los docentes recorren para llegar a ser expertos donde, por ejemplo, alcanzar una práctica experta no es un proceso lineal ascendente sino más bien un proceso en espiral. Los siguientes puntos son los que el NBPTS (2013) estableció como los principios que guían la práctica profesional docente:

1. Los docentes están comprometidos con los estudiantes y su aprendizaje.
2. Los docentes saben los contenidos que ellos enseñan y cómo enseñarlos.
3. Los docentes son responsables de gestionar y monitorear el aprendizaje de los estudiantes.
4. Los docentes piensan sistemáticamente sobres su práctica y aprenden de la experiencia.
5. Los docentes son miembros de comunidades de aprendizaje.

Por su parte, la estructura del PPI organiza los distintos elementos en una lógica de lo general a lo específico. Las dimensiones del perfil son lo más general, seguida de los parámetros y como elemento más específico los indicadores. Sin embargo, carecen de un planteamiento general de principios sobre el aprendizaje y la función docente, pues no establecen explícitamente estos elementos y tampoco establecen algún vínculo con lo señalado en el currículo sobre este concepto. Aunque estos postulados podrían encontrase en el modelo educativo (SEP, 2017) y en los planes y programas de estudio, no se hace una explícita referencia a ellos.

Nivel 2. Los dominios

Los estándares de países presentan pocas variaciones en cuanto a los dominios o dimensiones que los organizan. Ingvarson (2013) identificó que los estándares de Australia, Chile, Inglaterra, Nueva Zelanda, Escocia y el InTASC de Estados Unidos se agrupan principalmente en tres dominios:

1. Conocimientos sobre el contenido, el proceso de aprendizaje y el desarrollo de los estudiantes.
2. Habilidades para favorecer experiencias de aprendizaje efectivas en alumnos diversos, para autoevaluarse y reflexionar sobre su práctica.
3. Compromisos profesionales que los docentes requieren para continuar con su desarrollo profesional de forma individual y en colectivo.

En la tabla 1, se presentan los dominios de los estándares australianos, el InTASC, las dimensiones del PPI, los cuales tienen una organización similar. Se puede ver que varios de los dominios se agrupan en las categorías de conocimientos, habilidades y compromisos. No obstante, algunos de estos modelos enfatizan ciertos aspectos. El InTASC, por ejemplo, tiene un dominio específico para el aprendiz y los procesos de aprendizaje, en el que se incluyen aspectos principalmente relacionados con el conocimiento de aprendiz y los procesos de desarrollo. Este dominio destaca a los alumnos como punto de partida de la intervención docente.

Tabla 1. *Tipos de dominio en que se clasifican los estándares de diferentes países*

Dominios	Australia	InTASC EUA	PPI México
Conocimientos	1. Conocimiento profesional	1. El aprendiz y el aprendizaje 2. Conocimiento del contenido	1. Un docente que conoce a sus alumnos, sabe cómo aprenden y lo que deben aprender.
Habilidades	2. Práctica profesional	3. Práctica instruccional	2. Un docente que organiza y evalúa el trabajo educativo, y realiza una intervención didáctica pertinente.
Compromisos	3. Compromisos profesionales	4. Responsabilidad profesional	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.

Nota: InTASC = *Interstate Teacher Assessment and Support Consortium*; PPI = Perfil, Parámetros e Indicadores.

Por su parte, el PPI tiene un dominio sobre el conocimiento que debe tener el docente sobre sus alumnos, los procesos de aprendizaje y el currículo. También tiene un dominio que agrupa los aspectos relacionados con las habilidades de intervención docente. En lo relacionado con los compromisos profesionales, el PPI destina a este aspecto tres de sus cinco dimensiones, lo que indica un énfasis del perfil en este ámbito.

Sin embargo, se reconoce que pueden existir distintas formas de organizar estos elementos, sin que esto signifique una deficiencia en los estándares. Lo relevante es que éstos tengan una organización que presente distintas áreas de la función docente y que no presente repeticiones en dominios distintos. De hecho, el NBPTS no establece dominios para agrupar a los estándares generalistas de preescolar.

Otra manera común de organizar los dominios es en función del proceso de enseñanza (Ingvarson, 2013). Un ejemplo de ello es el Marco para la Enseñanza de Danielson (2013). En este marco, el primer dominio se conceptualiza como la planificación y la preparación de la enseñanza,

donde ciertamente se incluyen los aspectos relacionados con el conocimiento, pero incluye otros más como la congruencia entre lo planificado y las necesidades de los alumnos. Algo similar sucede en el dominio de habilidades, donde uno se relaciona específicamente con la instrucción y otro con el ambiente del aula favorable para el aprendizaje. Aunque coincide con los demás estándares al incluir un dominio relacionado con las responsabilidades profesionales.

Nivel 3. Los estándares

Los estándares son las afirmaciones que describen lo que deben saber y hacer los docentes. Distintos autores especifican qué conocimientos, habilidades y compromisos profesionales deben desempeñar los profesores para favorecer el desarrollo y aprendizaje de los alumnos (Bransford, Darling-Hammond y LePage, 2005; Ingvarson, 2013; OCDE, 2013; Porter, Youngs y Odden, 2001).

En la tabla 2 se puede apreciar un comparativo de los cuatro modelos de estándares, se identifica que el modelo de Australia tiene siete estándares, mientras que el InTASC y el NBPTS tienen diez. Por su parte, el PPI tiene 17, una cantidad considerablemente mayor a los otros modelos.

Tabla 2. *Estándares básicos para los cuatro modelos*

AITSL	NBPTS	InTASC	PPI
1. Conocer a los estudiantes y cómo aprenden.	1. Usar el conocimiento del desarrollo infantil para entender al niño de forma integral.	1. El maestro entiende cómo los alumnos crecen y se desarrollan, reconociendo que los patrones de aprendizaje y desarrollo. Diseñan e implementan experiencias de aprendizaje desafiantes y apropiadas al desarrollo.	1.1. Describe las características y los procesos de desarrollo y de aprendizaje de los alumnos para su práctica docente.
2. Conocer el contenido y cómo enseñarlo.	4. Conocer los contenidos disciplinares para la enseñanza de niños.	1. El docente entiende los conceptos centrales, herramientas de indagación, y estructuras de la disciplina que enseña, y crea experiencias que hacen a la disciplina accesible y manejable para que los aprendices. 5. El docente entiende cómo conectar los conceptos y utiliza diversas perspectivas para involucrar a los aprendices en un pensamiento crítico, creativo y en la resolución de problemas colaborativos.	2. Un docente que organiza y evalúa el trabajo educativo, y realiza una intervención didáctica pertinente.

AITSL	NBPTS	InTASC	PPI
3.Planificación e implementación efectiva de la enseñanza y el aprendizaje	7. Planear para el desarrollo y aprendizaje. 8. Implementar la instrucción para el desarrollo y aprendizaje. 3. Fomentar la equidad, justicia y apreciación de la diversidad.	7. El docente planifica una instrucción que apoya a cada estudiante a alcanzar sus metas de aprendizaje. 8. El docente entiende y usa una variedad de estrategias didácticas para fomentar que los aprendices desarrollen un profundo entendimiento del contenido y sus conexiones, y construir habilidades para aplicar el conocimiento de forma significativa. 2. El profesor utiliza la comprensión de las diferencias individuales, las diversas culturas y las comunidades para garantizar ambientes de aprendizaje inclusivos que permitan a cada alumno alcanzar altos estándares.	2.1. Organiza su intervención docente para el aprendizaje de sus alumnos. 2.2. Desarrolla estrategias didácticas para que sus alumnos aprendan. 4.2. Establece un ambiente favorable para la sana convivencia y la inclusión educativa en su práctica docente.
4. Crear y mantener un ambiente seguro y de apoyo al aprendizaje.	6. Manejar el ambiente para el desarrollo y aprendizaje.	3. El maestro trabaja con otros para crear ambientes que apoyen el aprendizaje individual y colaborativo y animan la interacción social positiva, la participación activa en el aprendizaje y la auto-motivación.	2.4. Construye ambientes favorables para el aprendizaje. 4.3. Considera la integridad y seguridad de los alumnos en el aula y en la escuela en su práctica docente. 4.4. Demuestra altas expectativas sobre el aprendizaje de todos sus alumnos.
5. Evalúa, hace reportes y proporciona retroalimentación sobre el aprendizaje de los alumnos.	5. Evaluar el desarrollo y aprendizaje infantil.	6. El docente entiende y usa múltiples métodos de evaluación para atraer a los aprendices en su propio crecimiento, monitorear su progreso y guiar a docentes y aprendices en la toma de decisiones.	2.3. Utiliza la evaluación de los aprendizajes con fines de mejora.

AITSL	NBPTS	InTASC	PPI
6. Involucramiento en el aprendizaje profesional.	9. Reflexionar sobre la enseñanza de niños.	9. El docente emplea el aprendizaje profesional en marcha y usa las evidencias para continuamente evaluar su práctica, particularmente los efectos de sus decisiones y acciones en otros (aprendices, familias, otros profesionales y la comunidad), y adapta su práctica a las necesidades de cada aprendiz.	3.1 Reflexiona sistemáticamente sobre su práctica docente como medio para mejorarla. 3.2 Selecciona estrategias de estudio y aprendizaje para su desarrollo profesional. 3.3 Utiliza diferentes medios para enriquecer su desarrollo profesional.
7. Involucramiento profesional con colegas, padres y comunidad.	2. Asociarse con las familias y la comunidad. 10. Ser un ejemplo de profesionalismo y contribuir a la profesión.	10. El docente busca oportunidades y roles de liderazgo para tomar responsabilidad del aprendizaje del alumno, colaborar con los aprendices, familias, colegas, otros profesionales escolares y los miembros de la comunidad para garantizar el crecimiento del aprendiz y el progreso en su profesión.	5.1 Realiza acciones en la gestión escolar para contribuir a la calidad de los resultados educativos. 5.2. Propicia la colaboración de los padres de familia y de distintas instituciones para apoyar la tarea educativa de la escuela. 5.3 Considera las características culturales y lingüísticas de la comunidad en el trabajo del aula y de la escuela. 4.1 Considera los principios filosóficos, los fundamentos legales y las finalidades de la educación pública mexicana en el ejercicio de su función docente.

Nota: AITSL = Australian Professional Standards for Teachers. Australian Institute for Teaching and School Leadership. NBPTS = National Board for Professional Teaching Standards. InTASC = Interstate Teacher Assessment and Support Consortium. PPI = Perfil, Parámetros e Indicadores para los docentes de educación preescolar.

Las similitudes se observan en las primeras dimensiones del perfil, donde los conceptos que se incluyen son similares en todos los modelos. A continuación, se describen:

1. El conocimiento de los alumnos y cómo aprenden está expresado en un estándar en todos los modelos.
2. El conocimiento del contenido disciplinar y cómo enseñarlo está contenido en uno o dos estándares.
3. Las competencias docentes para planeación, intervención y la inclusión educativa están presentes como un estándar. Con excepción de los estándares australianos, los otros modelos, incluido el PPI, tienen un estándar para cada uno de estos aspectos.
4. La habilidad para crear y mantener un ambiente seguro y de apoyo al aprendizaje está presente con un estándar en todos los modelos, excepto en el PPI, donde se cubre con tres.
5. Todos destinan un estándar para la evaluación de los procesos de aprendizaje.
6. Todos los modelos abordan los aspectos relativos al aprendizaje profesional con énfasis en la reflexión sobre la práctica docente.

A propósito del PPI de México las diferencias principales con los otros modelos son las siguientes:

1. El PPI presenta un desglose más detallado de los parámetros en relación con los compromisos profesionales, que incluye el conocimiento de la normativa escolar y lo referente al desarrollo profesional. En este sentido, existe un desbalance en los parámetros que versan sobre los conocimientos y habilidades profesionales y los que están relacionados con las responsabilidades profesionales.
2. Mientras que los otros modelos agrupan sus estándares en los tres dominios mencionados por Ingvarson (2013), el PPI mezcla parámetros de distintos dominios. Ello es más visible en la dimensión cuatro, en ella se identifican parámetros que corresponden a distintos aspectos de las otras dimensiones. Por ejemplo, en la dimensión 4. Un docente que asume las responsabilidades legales y éticas inherentes a su profesión para el bienestar de los alumnos. Además de incluir parámetros relacionados con los compromisos profesionales, también se incluyen parámetros que tiene que ver con la práctica pedagógica. Tal es el caso de los parámetros 4.3, y 4.4, los cuales suponen una habilidad del docente para crear un ambiente favorable al aprendizaje, que bien podrían estar mejor ubicados dentro del parámetro 2.4 (tabla 2).

Nivel 4. Las elaboraciones

En este nivel, las elaboraciones hacen explícito lo que significan cada uno de los estándares. Pueden incluir elementos observables y justificaciones sobre el estándar. Estas elaboraciones proporcionan una base para desarrollar rúbricas y valorar el desempeño del docente. En este nivel se encuentran las diferencias más marcadas entre los estándares analizados. Enseguida se presentan algunos breves ejemplos.

Los estándares australianos desagregan los estándares en áreas de focalización, cada estándar tiene varias áreas. Las elaboraciones están dispuestas a manera de rúbrica con varios niveles de desempeño como se aprecia en la tabla 3, cada nivel significa una etapa en la carrera profesional docente, iniciando como profesor titulado y terminando como profesor líder. Las elaboraciones traducen a un nivel concreción mayor los principios de los estándares, donde se señala que la docencia es un proceso y pasa por varias etapas. Por otra parte, una disposición de las elaboraciones a manera de rúbrica tiene ventajas para una evaluación formativa.

Tabla 3. *Ejemplo de elaboraciones para una sección de los estándares nacionales de Australia*

Estándar 1 – Conocer a los estudiantes y cómo aprenden				
Área	Titulado	Competente	Experto	Líder
1.1 Desarrollo físico, social e intelectual y características de los estudiantes.	Demuestra conocimiento del desarrollo físico, social e intelectual, así como las características de los estudiantes y cómo esto afecta el aprendizaje.	Usa estrategias de enseñanza basadas en el conocimiento del desarrollo físico, social e intelectual y las características de los estudiantes para mejorar su aprendizaje.	Selecciona de un repertorio flexible y efectivo de estrategias para para adaptarse al desarrollo físico, social e intelectual y las características de los estudiantes.	Lidera a colegas para seleccionar y desarrollar estrategias de enseñanza para mejorar el aprendizaje de los estudiantes usando el conocimiento desarrollo físico, social e intelectual y las características de los estudiantes.

Fuente: AITSL (2012b).

El InTASC en las elaboraciones de los estándares incluye: desempeños, conocimientos esenciales y disposiciones críticas. También, se incluye una rúbrica con la descripción de desempeños en tres niveles, los niveles tienen solamente una etiqueta numérica: uno, dos, tres. Adicionalmente,

se explicitan algunas sugerencias que podría hacer el docente para el avance de un nivel al otro (CCSSO, 2013). El modelo de elaboración del InTASC es altamente estructurado y con una descripción abundante, con un claro énfasis en elementos necesarios para favorecer el desarrollo profesional docente.

El NBPTS tiene otra forma de realizar las elaboraciones de los estándares, es también una descripción detallada, pero con un formato de texto. La extensión de las descripciones puede ser de varias páginas, por ejemplo, para el primero de los estándares generalistas de preescolar, son ocho páginas.

Por su parte, el PPI muestra elaboraciones con descripciones menos extensas que los anteriormente expuestos. La tabla 4 presenta un ejemplo de la descripción completa que tiene un estándar o parámetro, la cual se realiza a través de los indicadores. Además de la escasa descripción, las elaboraciones carecen de grados o niveles. Aunque el Servicio Profesional Docente, genera dos tipos de PPI, uno para docentes de ingreso a la docencia y otros para las evaluaciones de permanencia y desempeño. El que no estén integrados en un mismo documento impide visualizar distintos grados de desempeño.

Tabla 4. *Ejemplo de elaboraciones de estándares para un parámetro del PPI*

Parámetros	Indicadores
1.1 Describe las características y los procesos de desarrollo y de aprendizaje de los alumnos para su práctica docente.	1.1.1 Describe las características del desarrollo y del aprendizaje de sus alumnos para organizar su intervención docente.
	1.1.2 Describe las características del desarrollo y del aprendizaje de sus alumnos para poner en práctica su intervención docente.
	1.1.3 Identifica las características del entorno familiar, social y cultural de sus alumnos para organizar su intervención docente.
	1.1.4 Identifica las características del entorno familiar, social y cultural de sus alumnos para poner en práctica su intervención docente.

Fuente: SEP (2016).

Conclusiones

El análisis realizado en este trabajo permitió identificar que existe consenso en los estándares de distintos países en relación con lo que se tendría que evaluar de los docentes, los propósitos que persiguen e incluso en la forma de organizarlos y presentarlos a la comunidad educativa. El hecho

de encontrar las coincidencias es relevante en sí mismo, pues permite tener mayor certeza de que los contenidos de evaluación docente son los adecuados. Desde el punto de vista de investigación se cuenta con una evidencia de validez en relación con otras variables o validez concurrente.

En todos los países analizados, los estándares son utilizados con fines de mejora y para promover el aprendizaje profesional, puesto que los estándares de desempeño docente bien elaborados son una herramienta de evaluación formativa. Lo que se espera de una herramienta de evaluación formativa como las rúbricas es que sean descriptivas, que tanto evaluadores como evaluados conozcan los criterios con los cuales se va a juzgar el desempeño. En estos casos, los evaluados intentan alcanzar los criterios de desempeños más altos, por lo que los estándares también son utilizados como herramienta para favorecer el desarrollo profesional.

No obstante, en los estándares mexicanos o PPI se encontró ausencia del nivel más general, el nivel de los Principios y del nivel de las elaboraciones o descripciones de los niveles de desempeño. Estos elementos faltantes resultan fundamentales para que los estándares cumplan con una de sus funciones, fungir como un referente para la práctica docente y orientar a los maestros hacia la mejora. Por ello, es necesario que el PPI de forma explícita incluya una visión sobre el aprendizaje de calidad y una manera de entender la enseñanza, así como los principios que guían la práctica pedagógica o cómo es que un docente llega a tener un desempeño destacado.

Los elementos ausentes en el PPI pueden repercutir en una falta de precisión en la comunicación hacia los docentes. La falta de principios rectores de la práctica pedagógica puede obstaculizar que los docentes tengan claridad sobre las metas y principios que guían su práctica y los medios para que los alumnos logren aprendizajes de calidad.

El ejemplo del NBPTS, InTASC y los estándares australianos son muestras remarcables que podrían direccionar los estándares mexicanos. En los tres casos se hace una amplia explicación sobre los principios que guían la práctica y la carrera docente. La docencia se concibe como un proceso, no como algo estático, y es a través del tiempo y de oportunidades de formación profesional de calidad que los docentes están en posibilidades de alcanzar los desempeños más altos. Tener estos elementos de forma explícita favorece una mejor comunicación de lo que significa una buena enseñanza y cómo se podría alcanzar.

La dimensión cuatro del Perfil, que aborda las responsabilidades legales y éticas de la profesión, tiene varias problemáticas. Por una parte, se presenta una cantidad excesiva de parámetros, y por otra, mezcla elementos que estarían mejor ubicados en otras dimensiones del perfil. En

comparación con los otros estándares de desempeño docente analizados, en esta dimensión se detecta una mayor cantidad de parámetros e indicadores. Ello podría repercutir en un sesgo sobre los aspectos a evaluar, puesto que se tendría un énfasis innecesario en aspectos burocráticos y normativos sobre otras de las funciones docentes. Si bien se reconoce a esta dimensión como relevante para incluirse en la evaluación, los otros modelos abordan esta dimensión con una cantidad menor de estándares y se hace más énfasis en las otras dimensiones, especialmente en las que están relacionadas con el conocimiento y las habilidades docentes para promover los aprendizajes en los niños: planeación, intervención y evaluación.

En una reforma educativa basada en estándares como la que propone la LGSPD, es crucial que el conjunto del sistema educativo transmita un mensaje coherente y unificado sobre los aspectos en los que una enseñanza efectiva debe concentrarse, y qué es lo más importante para enseñar y evaluar (Ingvarson, 2013; OCDE, 2013). Por ello, es indispensable que los PPI sean congruentes con el modelo educativo, currículo oficial (Planes y Programas), documentos normativos, de organización y funcionamiento escolar, así como con programas de formación del profesorado desde su etapa inicial hasta la formación continua.

Desarrollar estándares pertinentes y efectivos es una tarea altamente especializada, compleja y de largo plazo. A los sistemas analizados les ha tomado décadas construirlos. Los propuestos por el InTASC fueron reformulados en 2011, luego de un proceso de fundamentación y consenso. Por otra parte, el NBPTS somete constantemente a revisión su sistema de estándares, la mayoría de ellos tiene varias actualizaciones. Todo ello lleva a pensar que los estándares mexicanos deberán estar sujetos a continuos procesos de revisión, consenso y validación. Las instituciones que los desarrollan y validan deberían tomarse esta tarea como un proceso de largo plazo.

Referencias

Australian Institute for Teaching and School Leadership (AITSL). (2012a). *Australian Teacher Performance and Development Framework*. Melbourne, Brisbane, Canberra: AITSL.

Australian Institute for Teaching and School Leadership (AITSL). (2012b). *Australian Professional Standards for Teachers.* Melbourne, Brisbane, Canberra: AITSL.

Bransford, J., Darling-Hammond, L., & LePage, P. (2005). Chapter one. Introduction. En L. Darling-Hammond & J. Bransford (Eds.), *Preparing teachers for a changing world: what teachers should learn and be able to do.* San Francisco: Jossey-Bass.

Council of Chief State School Officers (CCSSO). (2013). *Interstate Teacher Assessment and Support Consortium InTASC Model Core Teaching Standards and Learning Progressions for Teachers 1.0: A Resource for Ongoing Teacher Development.* Washington, DC: CCSSO.

Danielson, C. (2013). *The Framework for Teaching. Evaluation Instrument.* Princeton: The Danielson Group. Recuperado de https://www.danielsongroup.org/framework/

Ingvarson, L. (2013). Estándares de egreso y certificación inicial docente: la experiencia internacional. *Calidad en la educación, 38,* 21-77.

Ingvarson, L., & Kleinhenz, E. (2006). *Standards for advanced teaching: A review of national and international developments.* Australia: ACER.

Ley General del Servicio Profesional Docente (LGSPD). (2013). *Diario Oficial de la Federación*, DOF 11-09-2013. Recuperado de https://www.dof.gob.mx/nota_detalle.php?codigo=5313843&fecha=11/09/2013

Luna, E., Pedroza, L.H., & Córdova, P. (2019). Alineamiento entre estándares de desempeño docente y el programa de estudios de preescolar mexicano. *Archivos Analíticos de Políticas Educativas, 27*(X). Recuperado de: http://epaa.asu.edu/ojs/article/view/3595

National Board for Professional Teaching Standards (NBPTS). (2013). *Early Childhood Generalist Standards.* Arlintong: NBPTS.

National Board for Professional Teaching Standards (NBPTS). (s.f.). *Mission & History.* Recuperado de http://www.nbpts.org/mission-history

National Research Council (NRC). (2008). *Assessing Accomplished Teaching: Advanced-Level Certification Programs.* Washington, DC: The National Academies Press.

Organización para la Cooperación y el Desarrollo Económicos (OCDE). (2013). *Learning standards, teaching standards and standards for school principals: a comparative study. OECD Education Working Papers No. 99*. París: OCDE.

Porter, A.C., Youngs, P., & Odden, A. (2001). Advances in Theacher Assessment and Their Uses. En V. Richardson (Ed.), *Handbook of research on teaching* (pp. 259-297). Washington, DC: American Educational Research Association.

Real Academia Española (RAE). (2014). *Diccionario de la lengua Española* [en línea]. Recuperado de http://dle.rae.es/?id=GltBvT9

Secretaría de Educación Pública (SEP). (s.f.). *Coordinación Nacional del Servicio Profesional Docente. Objetivos.* Recuperado de http://servicioprofesionaldocente.sep.gob.mx/cnspd/

Secretaría de Educación Pública (SEP). (2016). *Evaluación del Desempeño Docente. Ciclo Escolar 2015-2016. Perfil, parámetros e indicadores para docentes y técnicos docentes en educación básica.* Ciudad de México: SEP.

Secretaría de Educación Pública (SEP). (2017). *Modelo Educativo para la educación obligatoria.* Ciudad de México: SEP.

Capítulo 3

Caracterización de las motivaciones para ingresar a la Escuela Normal. Resultados del estudio exploratorio en Baja California

Jihan Ruth García-Poyato Falcón, Graciela Cordero Arroyo, Alejandra Priede Schubert y Edna Luna Serrano

Introducción

La carrera docente comprende fases clave consideradas en el diseño de políticas educativas: formación inicial, formación continua y desarrollo profesional (OREALC-UNESCO, 2013). En este sentido, es un trayecto que inicia con los estudios profesionales, pero no termina con ellos (Diker y Terigi, 1997).

El interés de esta investigación se enfoca en la primera fase, la formación inicial, definida "como el momento dentro de un continuo desarrollo profesional docente, curricularmente acotada, con espacios de inmersión-cualificación laboral, en un esquema de simultaneidad entre lo disciplinar y lo pedagógico" (Fortoul, Güemes, Martell y Reyes, 2013, p. 153). Por lo tanto, requiere del acercamiento progresivo a la práctica docente real, como parte de la preparación para ser profesor y realizarse en instituciones dedicadas exclusivamente a la formación de profesores (Marcelo, 1995).

En el caso mexicano, tradicionalmente esta institución específica ha sido la Escuela Normal, al menos para la formación de profesores de preescolar y primaria. Sin embargo, recientemente, ha enfrentado una importante reducción en la cantidad de aspirantes a ingresar en sus programas educativos. A nivel nacional, del ciclo escolar 2000-2001 al 2015-2016 la cantidad de aspirantes a ingresar ha disminuido en 44% (Medrano, Ángeles y Morales, 2017).

La disminución de aspirantes a ingresar a las Escuelas Normales se ha presentado en Baja California. A partir de los datos proporcionados por la Dirección de Educación Superior para Profesionales de la Educación (DESPE) de Baja California, en la figura 1 se presenta el descenso de 83% del ciclo escolar 2011-2012 al 2017-2018 en el número de aspirantes a la Licenciatura en Educación Primaria (LEPRIM).

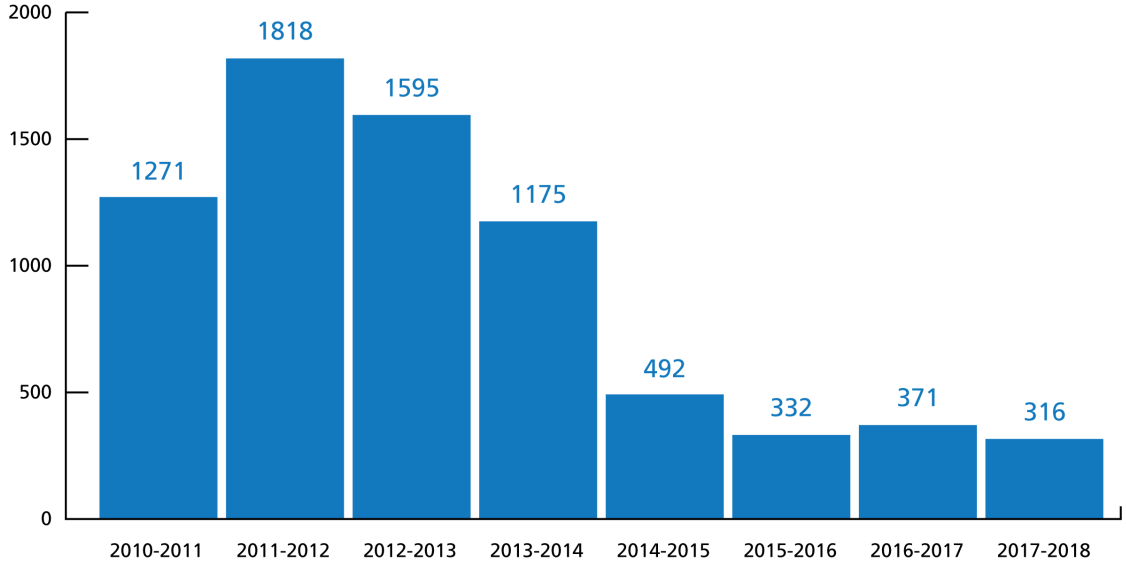

Figura 1. Cantidad de aspirantes a ingresar a la Licenciatura en Educación Primaria (LEPRIM) en Escuelas Normales públicas en Baja California, por ciclo escolar.

Esta situación puede deberse a diversas causas, algunas de ellas se describen a continuación. En 2010, la Organización para la Cooperación y el Desarrollo Económicos (OCDE) emitió una serie de recomendaciones a las autoridades mexicanas a partir del diagnóstico realizado al sistema educativo. Una de estas recomendaciones plantea la necesidad de fortalecer la formación inicial docente a través de la implementación de estándares rigurosos que permitan atraer a los mejores candidatos. En respuesta a dichas recomendaciones, ese mismo año, la Dirección General de Educación Superior para Profesionales de la Educación (DGESPE) determinó la expedición de una convocatoria única estatal para regular el proceso de selección de aspirantes a ingresar a las Escuelas Normales en cada entidad federativa. Además, se establecieron lineamientos generales en los requisitos de ingreso a la formación inicial en México: exigencia de un promedio mínimo de bachillerato, generalmente de 8, y la aplicación de un examen estandarizado de conocimientos,

definido por cada entidad.[1]

En segundo lugar, la Reforma Educativa vigente estipula en el Artículo 3º Constitucional que se garantizará la idoneidad de los docentes de educación obligatoria para lograr el máximo aprendizaje en los alumnos. Una de las acciones centrales de la Reforma Educativa fue la implementación en el año 2013 de la Ley General del Servicio Profesional Docente (LGSPD), que regula el ingreso, la promoción, el reconocimiento y la permanencia de los profesores de educación básica y educación media superior en todo el país. Con tales regulaciones, la situación laboral del magisterio cambió. Además de establecer los concursos de ingreso a la docencia, la LGSPD definió evaluaciones periódicas para asegurar la permanencia de los docentes en servicio. Esta regulación ha causado controversia entre el magisterio y la sociedad en general ante la incertidumbre de las nuevas disposiciones gubernamentales respecto a las prestaciones laborales y la permanencia vitalicia.

En tercer término, a partir de la LGSPD las Escuelas Normales perdieron su exclusividad en la formación de profesores de preescolar y primaria, ya que establece en su Artículo 24 transitorio, la participación en los concursos de oposición para el ingreso al servicio público de todas las personas con "formación docente pedagógica o áreas afines" (LGSPD, 2013), sin restringir a los egresados de cualquier Institución de Educación Superior (IES) a ocupar puestos como profesores de educación básica. Cabe señalar que dichas licenciaturas afines tienen competencias profesionales distintas a las licenciaturas de las Escuelas Normales.

Se identificó que, ante estos cambios, ciertos grupos resistentes han realizado protestas que impactan en gran medida la imagen social del maestro, lo cual genera opiniones encontradas ante la profesión y la función de las Escuelas Normales en el país. Dichas opiniones influyen en la decisión de pretender ser docente, ya sea en los jóvenes que pudieran estar interesados o en los padres de familia.

En quinto lugar, las condiciones de trabajo actuales para los docentes no son las ideales. Silva (2012) plantea como un factor negativo el número de alumnos por curso, donde México ocupa el quinto lugar de la lista de los 41 países latinoamericanos con mayor coeficiente de alumnos por docente. También destaca la disciplina del alumnado en los salones de clase y la rendición de cuentas a partir de evaluaciones estandarizadas como condiciones que los profesores pueden percibir desfavorables (Silva, 2012).

---

1 *A pesar de ser disposiciones nacionales, se registran variaciones entre los estados. Por ejemplo, el promedio solicitado en Michoacán en 2010-2011 fue de 6.0 y en Colima de 8.5 (Instituto Nacional para la Evaluación de la Educación [INEE], 2015).*

La disminución en el interés por cursar la formación inicial docente en una Escuela Normal provoca preocupación al preverse que las escuelas deban cerrar grupos por falta de aspirantes y, consecuentemente, el cierre total de las instituciones. Esta situación es grave, ya que las Escuelas Normales mexicanas cumplen una función social irremplazable consistente en la formación específica del profesorado, función que hasta el momento no ha sido suplida completamente por ninguna otra institución. Además, en el orden administrativo, el INEE (2015) plantea un "potencial problema de escasez cuando el volumen es menor o igual al número de docentes que deben ser reemplazados por retiro" (p. 47), ya sea por alcanzar la edad de jubilación o por abandono de la profesión.

En síntesis, a pesar de las condiciones adversas descritas, la demanda de aspirantes a la Escuela Normal, si bien ha descendido, no es inexistente. Por lo tanto, se considera relevante conocer los motivos que llevaron a los estudiantes de estas instituciones a ingresar a la LEPRIM, en un momento tan cambiante en la historia del magisterio mexicano. Se decidió enfocar esta investigación exclusivamente a los estudiantes de la LEPRIM ya que es el programa educativo con mayor matrícula a nivel estatal y nacional.

El objetivo del estudio es explorar y caracterizar las motivaciones de la población estudiantil de nuevo ingreso a la LEPRIM en tres Escuelas Normales públicas de Baja California (generación 2016-2020).

El tema de las motivaciones de los estudiantes para ingresar a dichas instituciones se menciona en documentos oficiales. En las publicaciones expedidas por la Secretaría de Educación Pública (SEP), a través de la DGESPE (2015), se afirma la falta de información respecto al impacto de las características de los estudiantes en su desempeño académico durante la formación inicial y establece la necesidad de definir un perfil de ingreso específico para cada licenciatura que se oferta en las Escuelas Normales. Se reconoce la necesidad de definir un perfil de ingreso específico a la docencia y estudios formales que aborden las trayectorias escolares durante formación inicial. El presente estudio exploratorio brinda insumos para el desarrollo de futuras investigaciones que faciliten información útil para los propósitos enunciados.

Referentes teóricos

Velasco y Meza (2013) aclaran que la elección de carrera, que pudiera considerarse en extremo personal y que tiene lugar "en la intimidad del hogar está influida por lo que hacen en torno a

su definición cotidiana los actores activos de dicha profesión" (p. 186). Es decir, la decisión del aspirante universitario se ve envuelta en influencias de carácter social, familiar y pedagógico, no solamente psicológico (Oliveira, Name y Leal, 2010; Mereshián y Calatayud, 2009).

Además, no todos los estudiantes pasan el mismo proceso al momento de elegir su carrera profesional, "para algunos la elección de una carrera es una etapa más o menos larga de toma de decisiones, para otros es producto del azar" (Sánchez, 2009, p. 137). Guzmán y Serrano (2011) indican que los factores asociados al proceso de elección son las características personales, educativas, socioeconómicas y culturales, por lo que es posible calificarlo como un proceso multidimensional.

A partir de la revisión de la literatura en el ámbito nacional e internacional, se identificó la Escala FIT-Choice de Watt y Richardson (2007), diseñada para medir la motivación para ser profesor en el contexto australiano. Este instrumento ha sido utilizado en más de 12 países de todos los continentes, y se ha aplicado a estudiantes de docencia y a egresados de distintos programas de formación docente a nivel internacional. Consiste en 18 factores agrupados en siete dimensiones que miden las razones para elegir la profesión docente, cinco de los factores miden las percepciones acerca de la enseñanza, uno la satisfacción con la elección y el resto las motivaciones de ingreso (Watt y Richardson, 2008).

El FIT-Choice se basa en la Teoría de la Expectativa del Valor de Wigfield y Eccles (2000), que clasifica las motivaciones en extrínsecas, intrínsecas y de utilidad social, además de tomar en cuenta las capacidades del individuo para lograr la acción por la cual está motivado.

También se revisó la Teoría Antropológica de la Acción Humana propuesta por Pérez-López (1991), que ha sido aplicada a estudios sobre las motivaciones para ser profesor en el contexto español (López-Jurado y Gratacós, 2013). Esta teoría propone clasificar la motivación en extrínseca, intrínseca y trascendente; sostiene el impulso del individuo hacia la acción debido al impacto que tiene en los otros. Debido a las características de la profesión docente, se considera que esta clasificación aporta un insumo valioso para analizar los motivos de ingreso a la carrera.

Método

El estudio se desarrolló bajo un enfoque de investigación cualitativo, de tipo exploratorio y transversal. A continuación, se presentan los participantes y la técnica de recolección de datos elegida en función del objetivo de la investigación.

En Baja California, la LEPRIM se ofrece en seis Escuelas Normales, tres en el municipio de Mexicali, dos en Ensenada y una en Tijuana. Se eligieron tres para participar en el estudio exploratorio, siguiendo la recomendación de la DESPE: Benemérita Escuela Normal Estatal, Profesor Jesús Prado Luna (BENEPJPL), Escuela Normal Experimental, Profesor Gregorio Torres Quintero (ENEGTQ) y Benemérita Escuela Normal Urbana Federal Fronteriza (BENUFF). En conjunto, su matrícula en primer semestre representa 44% de los alumnos de nuevo ingreso a nivel estatal inscritos en la LEPRIM en 2016, con 74 estudiantes. A continuación, se describen brevemente cada una de las instituciones participantes:

- **BENEPJPL.** Ubicada en Ensenada, se fundó en 1960 para la formación de profesores de educación primaria; posteriormente, se incluyó en la oferta académica la formación de profesores de educación preescolar. Actualmente, las dos licenciaturas que se imparten son la LEPRIM y la Licenciatura en Educación Preescolar, ambas cuentan con el reconocimiento vigente de los Comités Interinstitucionales para la Evaluación de la Educación Superior (CIEES). La institución labora en turno intermedio (10 a 19 horas).
- **ENEGTQ.** Ubicada en una localidad al sur del municipio de Ensenada, se fundó en 1977 bajo el programa de las escuelas experimentales. Su oferta educativa desde entonces había sido exclusivamente la LEPRIM, pero en la convocatoria del ciclo escolar 2018-2019 se incluyó la LEPRIM Intercultural Bilingüe. La institución labora en el turno matutino (7 a 14 horas).
- **BENUFF.** Ubicada en Mexicali, fue la primera Escuela Normal del estado de Baja California. Se fundó en 1949 para la formación de profesores de educación primaria. Posteriormente, se integró a la oferta educativa la Licenciatura en Educación Especial y la Licenciatura en Educación Secundaria, con distintas especialidades. La institución labora en el turno matutino (7 a 14 horas).

La selección de los participantes de este trabajo de investigación fue intencionada y diversificada al combinar los criterios: puntaje del examen de admisión, promedio de egreso del bachillerato, género y edad. Se seleccionaron en total 24 estudiantes, es decir, 32% de los alumnos inscritos en el primer semestre de la LEPRIM en las tres escuelas seleccionadas, ocho por cada escuela.

En la tabla 1, se describen los datos generales de los participantes: género, edad, estado civil, situación laboral, promedio de bachillerato y nivel de inglés. La mayoría son mujeres, tienen entre 17 y 20 años de edad, solteras, sin empleo, con promedio de egreso del bachillerato entre 9.0 y 9.9.

Tabla 1. *Datos generales de los estudiantes*

(n = 24)	General	BENEPJPL (Ensenada)	ENEGTQ (San Quintín)	BENUFF (Mexicali)
Género				
Femenino	19 (79%)	7	6	6
Masculino	5 (21%)	1	2	2
Categoria de edad				
Femenino	19 (79%)	7	6	6
Masculino	5 (21%)	1	2	2
Femenino	19 (79%)	7	6	6
Masculino	5 (21%)	1	2	2
Estado civil				
Soltero	18 (75%)	8	4	6
Casado	2 (8%)	0	1	1
Unión libre	3 (13%)	0	2	1
Viudo	1 (4%)	0	1	0
Trabajo previo				
Sí	11 (46%)	4	5	2
No	13 (54%)	4	3	6
Trabajo actual				
Sí	6 (25%)	0	4	2
No	18 (75%)	8	4	6
Rango promedio				
9.5-9.9	6 (25%)	2	1	3
9.0-9.4	8 (33%)	4	2	2
8.5-8.9	3 (13%)	0	2	1
8.0-8.4	7 (29%)	2	3	2
Nivel de inglés				
Alto	4 (17%)	1	0	3
Medio	5 (21%)	3	1	1
Bajo	15 (62%)	4	7	4

Fuente: Elaboración propia con base en el cuestionario de contexto.

Para la recolección de la información, se empleó la técnica de grupo focal con los estudiantes de nuevo ingreso seleccionados para participar. Esta técnica tiene un carácter colectivo, pues permite realizar una entrevista a todo un grupo de personas con la finalidad de obtener información y las percepciones sobre el tema a estudiar. Los grupos focales se realizaron durante el segundo mes del primer semestre que cursaron los estudiantes (2016-2).

El guión del grupo focal se diseñó con la intención de propiciar los comentarios libres de los participantes en torno a su decisión profesional. En este sentido, se incluyeron preguntas sobre su inclinación a la profesión, experiencias académicas, influencias familiares, expectativas de la formación y la carrera, así como la satisfacción con la elección.

Antes de iniciar cada una de las sesiones del grupo focal, se solicitó a los estudiantes que respondieran un cuestionario de contexto para recuperar sus datos generales (edad, estado civil, género y lugar de nacimiento), escolares (nombre del bachillerato, promedio de egreso, cantidad de extraordinarios presentados, nivel de inglés, becas y estudios universitarios previos) y familiares (escolaridad y ocupación de los padres, características de la vivienda y bienes materiales).

El punto de partida de la investigación consistió en la negociación de acceso a los participantes. A finales del mes de agosto de 2016 se presentó el tema de investigación al representante de la DESPE, quien brindó sugerencias para la implementación del estudio y autorizó su realización. Durante el mes de noviembre, se llevaron a cabo las tres sesiones de grupo focal en cada una de las escuelas seleccionadas, con una duración promedio de 67 minutos. Cada sesión fue audiograbada con la autorización de los participantes, quienes fueron informados del objetivo del estudio y de las características generales de su intervención. Todos consintieron su participación libremente y se aseguró la confidencialidad de la información que proporcionaron.

Para el análisis de la información recabada se empleó el análisis de contenido cualitativo, a través de la técnica de categorización y codificación deductiva e inductiva (Mayring, 2000). La categorización y la codificación inductiva-deductiva se construyó a partir del texto analizado, pero se tomó en cuenta la base teórica que sustenta el estudio y el objetivo de la investigación.

Una vez realizada la transcripción completa de los grupos focales, se convocó a un grupo de análisis con la finalidad de presentar el estudio y precisar las definiciones de las categorías y los códigos que se establecieron a partir de la primera lectura del material. Se efectuaron modificaciones en algunas categorías y códigos, lo cual favoreció a la organización y reducción de datos.

Resultados

A través del proceso de análisis de la información recuperada en el estudio se identificó la presencia de algunos de los factores motivacionales rescatados por la teoría a nivel internacional, además de otros específicos a la formación de profesores de primaria en las escuelas normales mexicanas. Estos factores motivacionales se integraron en el siguiente esquema de análisis (tabla 2), con definiciones teóricas y empíricas, validado a través de la doble codificación independiente.

Tabla 2. *Esquema de análisis deductivo-inductivo*

CATEGORÍAS	DEFINICIÓN	CÓDIGOS	DEFINICIÓN	F
Elección de la Escuela Normal (F=8)	Se refieren al proceso de elección realizado por los aspirantes para decidir ingresar a la formación inicial docente en una escuela normal (Sánchez, 2009).	*Prestigio de la institución*	Se refiere a los comentarios positivos que los aspirantes recibieron acerca de la institución educativa elegida.	5
		Condiciones institucionales	Se refiere a las características actuales de la institución normalista en relación a infraestructura, docencia, programas educativos y extracurriculares.	3
Influencias sociales para el ingreso a la carrera (F=55)	Se refieren a relaciones sociales y experiencias académicas previas que impactan la elección profesional (Watt y Richardson, 2007).	*Influencias familiares*	Se refieren a la influencia que ejerce la familia en la decisión profesional (Pariahuache, 2015).	18
		Experiencias de enseñanza	Se refieren a las situaciones que los aspirantes a profesor han vivido y que marcan su filiación a la docencia al ejecutar en algún momento ese rol (Pariahuache, 2015).	17
		Disuasión social	Se refiere a los comentarios negativos recibidos por los aspirantes en función de reorientar su elección profesional.	13
		Experiencias de aprendizaje	Se refieren a las situaciones que los aspirantes a profesor han vivido como alumnos durante su vida, que pueden asumirse de forma acrítica (Marcelo, 1995).	4

CATEGORÍAS	DEFINICIÓN	CÓDIGOS	DEFINICIÓN	F
Influencias sociales para el ingreso a la carrera (F=55)	Se refieren a relaciones sociales y experiencias académicas previas que impactan la elección profesional (Watt y Richardson, 2007).	*Influencias de otros*	Se refieren a las opiniones de personas cercanas al aspirante que influyen en la personalidad y, subsecuentemente, son un factor influyente en la elección de la carrera (Tomsik, 2016).	3
Capacidades para la docencia (F=29)	Se refieren a la opinión del individuo sobre sus habilidades, conocimientos y actitudes actuales en relación a una actividad determinada, en este caso, la docencia (Watt y Richardson, 2007).	*Rasgos de personalidad docente*	Se refieren a aquellas características de personalidad que los individuos consideran propicias para ejercer la docencia, además de identificarlas en ellos mismos (Roloff, Klusmann, Lüdtke y Trautwein, 2015).	17
		Habilidad para enseñar	Se refiere a las destrezas percibidas por el propio individuo como eficaces en relación a su función docente (Randall y Burley, 2015).	12
Valor intrínseco de la profesión docente (F=16)	Se refiere a los aspectos específicos de la profesión con el interés de ejercer la docencia sin un beneficio externo (Kyriacou y Coulthard, 2000).	*Trabajar con niños*	Se refiere a las condiciones de la labor docente que posibilitan la interacción con niños en las actividades cotidianas (Pariahuache, 2015).	11
		Valor intrínseco de la carrera	Se refiere a las características propias de la carrera docente como la satisfacción por aprender y el interés por la tarea (Cano, 2007).	5
Valor social de la profesión docente (F=14)	Se refiere a las condiciones de la profesión docente que se relacionan con la visión de la docencia como un trabajo importante y de valía social (Tang, Cheng y Cheng, 2014).	*Hacer una contribución social*	Se refiere a las acciones docentes orientadas a facilitar el avance educativo, pero no solamente de los niños en el aula, sino trascender para mejorar a la sociedad.	10
		Influir en el futuro de los niños	Se refiere a las acciones docentes orientadas a la formación académica y valoral de las generaciones de niños, con la intención de impactar favorablemente su futuro.	4

CATEGORÍAS	DEFINICIÓN	CÓDIGOS	DEFINICIÓN	F
Reconocimiento de la profesión docente (F=7)	Se refiere a los aspectos de la profesión que no son inherentes al individuo mismo, al ser de carácter externo y que influyen en la elección profesional (Kyriacou y Coulthard, 2000).	*Salario y seguridad laboral*	Se refiere a la remuneración económica que recibe el docente por el trabajo profesional que realiza, junto con las condiciones de trabajo que permiten tener continuidad laboral en la profesión docente (Morduchowicz, 2009, Pariahuache, 2015).	5
		Vacaciones y horario	Se refiere a los periodos de descanso pagados y a la carga horaria correspondientes al trabajo docente. Los primeros regularmente son mayores en comparación a otras profesiones, mientras que el horario es más reducido (Morduchowicz, 2009).	2

Nota: F = Frecuencia. *Fuente:* Elaboración propia.

Los resultados obtenidos en el estudio se describen a continuación en función de las categorías de análisis utilizadas. Al interior de la categoría se presenta la interpretación alcanzada en cada uno de los códigos, ordenados a partir de la frecuencia de aparición que tuvieron en los grupos focales.

Elección de la Escuela Normal

En esta categoría se ubicó información sobre las características de la escuela normal que llevaron a los participantes a elegirla como el espacio pertinente para llevar a cabo su formación inicial como profesores de primaria. Acerca del *prestigio de la institución*, los participantes de San Quintín y Mexicali afirmaron contar con referencias positivas. Ambas instituciones son consideradas como la mejor opción para formarse como docentes en la localidad respectiva.

En cuanto a la exigencia percibida por los participantes acerca de la formación inicial docente, en las tres escuelas se obtuvieron comentarios relacionados con una amplia carga académica, traducida en muchas materias curriculares que demandan tiempo para la realización de tareas. Los participantes coincidieron en la gran cantidad de lecturas y actividades relacionadas con su formación como profesores, por lo que califican la carrera como difícil. Sin embargo, aclaran que son actividades que consideran necesarias y pertinentes para su formación completa.

En relación a las condiciones institucionales, se obtuvieron comentarios de los participantes de la ENEGTQ, acerca de la cercanía del hogar de los estudiantes con la escuela y el horario matutino en que se imparten las clases. En los grupos focales realizados en Ensenada y Mexicali se recibieron comentarios sobre los docentes. En la BENEPJPL todos los participantes coinciden en calificar a los docentes como amables y con experiencia en grupos de primaria. Los participantes agregaron que cuentan con un tutor de grupo, quien les ha explicado de qué trata la carrera y orientado el diagnóstico de sus debilidades. En esta escuela fue en la única que se mencionó la figura del tutor grupal y los participantes reconocen su relevancia, al mencionar que sienten que los profesores realmente se interesan por ellos.

En cuanto a los programas institucionales, los alumnos de la BENEPJPL y ENEGTQ manifestaron tener información sobre las becas[2] a las que tienen derecho de aplicar. El otro programa que se mencionó fue el de movilidad académica estudiantil y sólo los participantes de Ensenada expresaron tener conocimiento general de sus características. De igual manera, fueron los únicos estudiantes que comentaron llevar clases de idioma extranjero, específicamente inglés. Los alumnos de la BENUFF sugirieron mejorar la organización de los aspectos administrativos del proceso de selección de aspirantes a ingresar, al evitar modificaciones con relación a las fechas y la falta de difusión de información oportuna a partir de medios electrónicos.

Influencias sociales para el ingreso a la carrera

Las influencias sociales se consideran un factor motivacional relevante en la decisión de ser docente ya que la elección profesional puede ser resultado de las interacciones que el estudiante ha tenido durante su vida (Mungarro y Zayas, 2009).

En cuanto a las *influencias familiares* se encontraron evidencias en los tres grupos de enfoque, pues 42% de los participantes compartieron que familiares cercanos (padre, madre, abuelos o tíos) los impulsaron a tomar la decisión de ingresar a la carrera porque ellos mismos han ejercido la profesión docente. Tomsik (2016) reconoce que la influencia de padres profesores tiene un peso importante en la decisión de ser maestro. En el caso de las madres de los participantes, la docencia es la respuesta de mayor frecuencia, luego de ama de casa; en cuanto a los padres también ocupa

2 *Al menos cinco tipos de becas federales (BAPISS, Proyecta 100000, PROSPERA); dos tipos de becas estatales ($300 al mes en efectivo, ISSSTECALI); y becas internas promovidas por cada escuela normal (alimentación, excelencia académica, movilidad).*

el primer lugar, pues cinco participantes los ubicaron en la profesión docente, mismo resultado que en la ocupación de comerciante.

Por otra parte, varios de los participantes mencionaron contar con *experiencias de enseñanza* exitosas. Algunos comentaron tener esa función en la organización religiosa a la que pertenecen o auxiliar la docencia de un familiar cercano que se desempeñaba como profesor de primaria cuando los participantes eran pequeños. Otros estudiantes expresaron diferentes acercamientos con la docencia. Por ejemplo, afirmaron que durante sus estudios de bachillerato fungían como tutores de sus compañeros o de familiares menores al explicar temas que les parecían difíciles. Dundar (2014) sostiene que las mujeres eligen la carrera docente porque en experiencias previas "se percibían a sí mismas como hábiles para enseñar" (p. 453). También se encontraron experiencias de enseñanza en el campo laboral (Mudavanhu, 2015). Dos participantes de la ENEGTQ compartieron tener experiencia como instructoras comunitarias del Consejo Nacional de Fomento Educativo (CONAFE), donde tuvieron la oportunidad de trabajar con niños migrantes en zonas rurales. De igual forma, se recuperaron experiencias de estudiantes al enseñar a niños de primaria, apoyar en una casa hogar con las tareas escolares y capacitar a otros en temas laborales.

Por el contrario, también se encontraron comentarios que buscaban desanimar a los estudiantes a ingresar a la carrera docente, los cuales se clasificaron bajo el código *disuasión social.* Destacan los relacionados a la falta de plazas en el servicio público, los bajos salarios o la percepción negativa que la sociedad tiene sobre los profesores. Los participantes confirmaron que los comentarios negativos provenían tanto de personas ajenas a la profesión como de maestros en servicio. Por ejemplo: *"Tenía una profesora de secundaria y me decía: a veces a los maestros ni siquiera nos pagan, cómo quieres ser maestra tú, busca otra cosa"* (GF2-03102016-16, 497-498)[3].

En cuanto a las *experiencias de aprendizaje*, los participantes comentaron que eligieron la profesión gracias al ejemplo de profesores de primaria que calificaron como buenos docentes, por ejemplo, una estudiante comentó: *"me gustaba cómo daba clases un profesor"* (GF2-03102016-3, 78-79). Este dato coincide con los resultados internacionales sobre el tema que reconocen la influencia de antiguos profesores como uno de los principales factores que atraen a los aspirantes a la formación docente (Malderez, Hobson, Tracey y Kerr, 2007).

---

3 La clave utilizada se conforma por el número de grupo focal, la fecha en que se realizó, la página y las líneas donde aparece la transcripción.

En menor frecuencia se identificaron *influencias de otras* personas cercanas a los estudiantes, como amigos y profesores, para elegir a la Escuela Normal como la mejor opción para su formación inicial como docentes.

Capacidades para la docencia

La autopercepción de las propias capacidades para ejercer la docencia cobra importancia, pues al elegir una carrera contraria a las características personales se corre el riesgo de abandonar los estudios posteriormente (Norzagaray, Maytorena y Montaño, 2011). Sobre los *rasgos de personalidad docente* que los estudiantes identifican en sí mismos se mencionaron: paciencia, responsabilidad, tolerancia, respeto, empatía y perseverancia. Por lo tanto, los elementos enlistados por los participantes son los que consideran necesarios para ejercer adecuadamente la docencia.

Además, los alumnos participantes perciben contar con las siguientes *habilidades para enseñar* relacionadas con la labor docente: facilidad para explicar a otros, creatividad, afinidad histriónica, así como aprendizaje permanente y autónomo.

En los casos en los que reconocen no poseer dichas habilidades o rasgos de personalidad, consideran que el cursar la formación inicial les permitirá desarrollarlos. Por ejemplo, uno de los participantes de San Quintín, mencionó no ser tan paciente al relacionarse con niños; ante esto, comenta que espera que sus estudios normalistas le permitan afinar este aspecto de su personalidad.

Valor intrínseco de la profesión docente

Se encontró que "los jóvenes no sólo eligen una carrera por motivos propios, sino que su decisión está mediada por múltiples factores tales como la personalidad, los elementos de su entorno, el deseo de satisfacer necesidades, entre otros" (Avendaño y González, 2012, p. 23). Además de los factores ya mencionados, se encuentran los factores motivacionales intrínsecos, los cuales la literatura afirma que tienen más posibilidad de permanecer en las personas que eligen la docencia como profesión (Sánchez, 2009).

Dentro de esta categoría, el código que se mencionó con mayor frecuencia en los tres grupos focales fue *trabajar con niños*. Los participantes afirmaron tener experiencias en las que se relacionaron con infantes, situación que les permitió reafirmar su gusto por la docencia. Sin embargo, no proporcionaron argumentos específicos que permitieran clarificar el significado del gusto referido, únicamente mencionaron que les parece atractivo enseñar a niños pequeños.

En general, los elementos identificados como *valor intrínseco* coinciden con los hallazgos de Hernández y Landa (2015) en un estudio realizado con estudiantes normalistas de Chihuahua. Por ejemplo, el gusto desde la infancia. De forma particular, esta razón fue constantemente referida por los estudiantes de la BENEPJPL. Comentaron que aun cuando tenían ese gusto desde pequeños, al momento de cursar el bachillerato presentaron inquietudes por otras profesiones.

Valor social de la profesión docente

El siguiente factor motivacional identificado guarda estrecha relación con el valor intrínseco, pero va más allá al relacionar la elección de la profesión docente con el efecto social trascendente que significa ser profesor (Pérez-López, 1991). Se identificó como un factor relevante en la decisión de ser profesor el hecho de *hacer una contribución social.* Este argumento va en el mismo sentido de facilitar el avance educativo, pero en este caso no solamente de los niños en el aula, sino trascender para mejorar a la sociedad. Uno de los participantes de Ensenada mencionó: *"creo que la docencia abarca eso en muchos aspectos, no nada más el conocimiento, sino en toda la formación del niño. Me gustaría mucho poder contribuir a que el niño avance o progrese de una manera positiva"* (GF1-23092016-13, 401-407).

Algunos de los participantes, de Ensenada y Mexicali, agregaron que el elegir la docencia como profesión les permitirá *influir en el futuro de los niños* que sean sus alumnos. Los estudiantes expresaron que al ser profesores podrán ayudar a sus alumnos a progresar y resolver las posibles problemáticas que enfrentan. En este sentido se recuperó que algunos estudiantes esperan que su labor sea en beneficio de otros.

Los estudiantes mencionaron el interés por trabajar en una escuela rural o de educación especial, por considerarlas poblaciones menos favorecidas en donde su función docente sería de ayuda. Una participante de la BENEPJPL mencionó ingresar a la formación docente con la intención de mejorar la equidad social al trabajar en escuelas rurales; comentó que a inicio de semestre realizaron una visita a escuelas en ese tipo de contexto y la experiencia le ayudó a ratificar su decisión profesional.

En general, los participantes perciben a la docencia como una labor que les permitirá contribuir con la mejora social a través de la formación académica y valoral de las futuras generaciones.

Reconocimiento de la profesión docente

Sánchez (2009) afirma la existencia de elementos externos a la profesión, como las condiciones del empleo que llegan a ser fuertes atractivos para que los jóvenes decidan ingresar a la formación inicial docente. Estos elementos son clasificados como el valor extrínseco de la docencia y fueron identificados con menor frecuencia en los participantes.

En esta clasificación se ubicó el código *salario y seguridad laboral* en la profesión docente. Dos de los participantes del estudio hicieron comentarios al respecto. Un estudiante de Ensenada mencionó: *"pensé en estudiar Derecho; pero pues, como también es una carrera que se batalla para el trabajo, dije: 'pues no, mejor a la Normal'"* (GF1-23092016-3, 96). Por su parte, un participante de BENUFF reconoció que actualmente se vive una situación compleja para el ingreso al Servicio Profesional Docente (SPD), sin embargo, afirma que los maestros siempre serán necesarios, aun en escuelas privadas. Comentó:

> *aunque no sea el gran trabajo que te va a dejar los millones, pero sí te va a dejar como que muchas experiencias, va a ser algo que a ti te guste hacer porque es lo que tú elegiste, no porque alguien te lo impuso.* (GF3-10102016-13, 425-427)

Dos estudiantes de la ENEGTQ fueron las únicas en comentar que elementos como *vacaciones y horario,* favorecen la conciliación familiar y los tomaron en cuenta al momento de decidir elegir la docencia como profesión. Ambas son madres de familia y afirmaron que decidieron estudiar para estar en posibilidades de brindar un mejor futuro a sus hijos.

En los grupos focales surgieron las opiniones de los participantes en relación a la imagen social del magisterio. Ellos expresaron comentarios positivos acerca de la profesión docente, sin embargo, reconocen que la imagen social actual de los profesores se encuentra dañada y que la carrera no es tan valorada socialmente como lo era en el pasado. En el grupo focal realizado en la BENEPJPL se recuperó que una posible causa para la disminución de aspirantes a ingresar a las escuelas normales radica en esta opinión negativa generalizada hacia el magisterio: *"si ya saben cómo están las cosas ahorita en México, y quieren aun así entrar, es porque les gusta; ya no es tanto por querer un trabajo o por el dinero, sino es porque les gusta"* (GF1-23092016-18, 591-594). Opiniones similares se recuperaron del resto de los grupos focales: *"ahorita, ser maestro es irse a caminar a un campo minado, es muy peligroso"* (GF3-10102016-14, 432-436).

Específicamente, se recuperaron comentarios en relación al impacto negativo que la serie de reacciones a la Reforma Educativa, como huelgas y marchas, ha tenido en la opinión que la sociedad tiene de los profesores. Al respecto, los participantes comentaron estar de acuerdo con las evaluaciones constantes del SPD, no así con el proceso de asignación de plazas, pues tienen sospechas de que no se maneja con total transparencia.

En las tres escuelas normales, surgió el tema de la dificultad existente en Baja California en relación al proceso de asignación de plazas docentes. En todos los casos los participantes compartieron anécdotas de familiares o amigos que han participado en el examen de ingreso al SPD y, a pesar de obtener un resultado satisfactorio, no logran conseguir trabajo. De igual forma, comentaron tener conocimiento sobre la tardanza en el pago a docentes interinos. Los participantes califican estas situaciones como condiciones laborales negativas. Sin embargo, se rescata:

> *prefiero hacer algo que en realidad me guste, algo que en realidad me llene y aunque gane poco, no tenga las mejores prestaciones o la permanencia laboral, a estar en un lugar haciendo algo que yo realmente…ahora sí que llegue a odiar y esté así toda mi vida. Prefiero estar haciendo lo que a mí realmente me gusta, me llena, que estar en otra carrera que a lo mejor me reditúa bien, tengo más prestaciones, más salario, pero que realmente no es lo mío, que no me llene y que incluso llegue a decir: 'ya no quiero trabajar'.*
> (GF3-10102016-13, 418-422)

Conclusiones

A pesar de las acciones de los sistemas educativos a nivel mundial, se identifica que en muchos casos los jóvenes que egresan de bachillerato eligen su carrera profesional sin pasar por un proceso serio de elección (Norzagaray et al., 2011). Ante esta situación y la importancia del papel del profesor de educación básica, se considera fundamental la implementación de estrategias para contrarrestar el descenso abrupto de aspirantes a las escuelas normales en México.

En este estudio exploratorio se presentaron los factores motivacionales que intervienen en la elección de la docencia como profesión. Se espera que su análisis brinde información para la futura elaboración de políticas de reclutamiento enfocadas a la atracción de más y mejores aspirantes a las escuelas normales. Watt y Richardson (2008) mencionan que:

> Históricamente, se ha esperado que los profesores vean a la docencia más como una vocación que un trabajo, pero en décadas recientes se han presentado cambios significativos en el mercado laboral con un potencial impacto en cómo aquellos que se inscriben en programas de formación docente ven a la docencia como una carrera. (p. 409)

En cuanto a la fase de elección profesional, se encontró que factores personales como las experiencias de enseñanza y la influencia de familiares profesores tienen un impacto importante en la decisión de los aspirantes. Sin embargo, también se recuperaron testimonios de estudiantes que fueron criticados por elegir ingresar a la Escuela Normal, quienes recibieron comentarios negativos que trataron de disuadirlos.

Resaltó también la presencia de estudiantes mayores a la edad regular para cursar los estudios universitarios. En el grupo de alumnos participantes, el 25% se ubica entre 21 y 31 años de edad, quienes compartieron que una de las razones para elegir la carrera fue la experiencia laboral con la que cuentan en áreas afines a la docencia; por ejemplo, la incursión como instructores comunitarios de CONAFE. Se considera que esta situación debe tomarse en cuenta en la reconfiguración del perfil del estudiante normalista.

Respecto a los elementos institucionales, la totalidad de los alumnos participantes consideraron que la Escuela Normal pública era la mejor opción para formarse como profesor de educación básica. Esta situación debe valorarse e impulsar a las escuelas normales a continuar con la mejora del servicio educativo que ofertan en Baja California.

Al momento de la recolección de datos (segundo mes del primer semestre de estudios), los alumnos participantes se manifestaron satisfechos con la decisión de ingresar a la Escuela Normal, pero mencionaron algunos aspectos que consideran áreas de oportunidad en sus escuelas, como el horario y la falta de cursos de inglés. Los alumnos de la BENEPJPL fueron los que mostraron mayor conocimiento de la dinámica institucional y de los programas o servicios con los que cuenta, como la movilidad estudiantil, las becas y la tutoría. También fueron los alumnos de esta escuela quienes manifestaron las desventajas del horario extendido en el que funciona.

Se encontró que "debido a las características peculiares que rodean la profesión de maestro, en la elección de la carrera intervienen tanto motivos intrínsecos como extrínsecos" (García-Garduño y Organista, 2006, p. 3). En los intrínsecos y trascendentes, la mayoría de los participantes

comentaron haber elegido la docencia para trabajar con niños y hacer una contribución social. En los rubros considerados extrínsecos, o de reconocimiento social, se encontró que los participantes valoran negativamente las condiciones laborales actuales de los profesores en servicio, el salario y la seguridad laboral. A pesar de esto, decidieron ingresar a la Escuela Normal.

Ante este escenario, se plantean como recomendaciones para atender la disminución de aspirantes en las escuelas normales públicas de Baja California las siguientes:

1. Seguir la indicación de las Normas Específicas de Control Escolar vigentes en relación a alinear las fechas de la convocatoria del proceso de selección de las Escuelas Normales con las convocatorias de otras IES. Esta medida pretende evitar que la licenciatura ofertada por la Escuela Normal sea considerada una carrera de respaldo.
2. Mejorar la organización del proceso de selección al facilitar los trámites burocráticos para impulsar la eficiencia de cada una de las fases del proceso. Asimismo, establecer medios de comunicación apropiados para que todos los aspirantes se mantengan informados de los requisitos y pasos a seguir durante el proceso de selección.
3. Una vez concluido el proceso de selección es fundamental dar seguimiento sistemático a aquellos aspirantes que, a pesar de haber sido admitidos, deciden no inscribirse. Es necesario indagar las razones por las cuales se toma esta decisión, para identificar posibles acciones de mejora institucional, si la causa está relacionada a factores académicos o de elección profesional.
4. La autoridad estatal argumentó que uno de los elementos centrales de su gestión es fortalecer la imagen de las Escuelas Normales de Baja California, por lo que propone el abandono a prácticas endógenas y promocionar con la comunidad el trabajo que se realiza cotidianamente en ellas. Sugiere que el intercambio con otras IES e impulsar la movilidad académica de estudiantes, son estrategias que impactarán positivamente la imagen que la sociedad tiene de las Escuelas Normales.
5. Implementar un programa estatal permanente de difusión de las actividades de las Escuelas Normales. A pesar de que cada institución realiza esfuerzos por socializar con la comunidad de la localidad donde se ubica sus labores cotidianas, es importante contar con una estrategia integral que presente a la sociedad una imagen unidad del normalismo bajacaliforniano.

A nivel internacional y nacional logró identificarse la escasa producción académica en relación a las motivaciones para la docencia, por lo que "conviene seguir explorando las relaciones entre actitudes y motivaciones hacia la docencia con las concepciones previas sobre los procesos y contextos educativos con vista a tener en cuenta tales aspectos en la mejora de la formación inicial" (Pontes, Ariza, Serrano y Sánchez, 2011, pp. 181-182).

Referencias

Avendaño, C., & González, R. (2012). Motivos para ingresar a las carreras de Pedagogía de los estudiantes de primer año de la Universidad de Concepción. *Estudios pedagógicos, XXXVIII,* 21-33.

Cano, M. (2007). Motivación y elección de carrera. *Revista Mexicana de Orientación Educativa, V*(13), 6-9.

Diker, G., & Terigi, F. (1997). *La formación de maestros y profesores: hoja de ruta.* Argentina: Paidós.

Dirección General de Educación Superior para Profesionales de la Educación (DGESPE). (2015). *Diagnóstico de las Escuelas Normales.* México: DGESPE.

Dundar, S. (2014). Reasons for choosing the teaching profession and beliefs about teaching: a study with elementary school teacher candidates. *College Student Journal, 48*(3), 445-460.

Fortoul, B., Güemes, C., Martell, F., & Reyes, E. (2013). Formación inicial de docentes para la Educación Básica. En P. Ducoing y B. Fortoul (Coords.), *Procesos de formación. Vol. I* (pp. 153-192). México: COMIE.

García-Garduño, J.M., & Organista, J. (2006). Motivación y expectativa para ingresar a la carrera de profesor de educación primaria: un estudio de tres generaciones de estudiantes normalistas mexicanos de primer ingreso. *Revista Electrónica de Investigación Educativa, 8*(2), 1-17.

Guzmán, C., & Serrano, O. (2011). Las puertas del ingreso a la educación superior: el caso del concurso de selección a la licenciatura de la UNAM. *Revista de la Educación Superior, XL*(1), 31-53.

Hernández, M., & Landa, T. (2015). *Intereses y competencias para iniciar la formación docente desde la mirada de los estudiantes. Licenciatura en Educación Primaria: Generación 2014-2015.* Trabajo presentado en el XIII Congreso Nacional de Investigación Educativa, Chihuahua, México.

Instituto Nacional para la Evaluación de la Educación (INEE). (2015). *Los docentes en México. Informe 2015.* México: INEE.

Kyriacou, C., & Coulthard, M. (2000). Undergraduates' views of teaching as a career choice. *Journal of Education for Teaching, 26*(2), 117-127.

Ley General del Servicio Profesional Docente (LGSPD). (2013). *Diario Oficial de la Federación*, DOF 11-09-2013. Recuperado de https://www.dof.gob.mx/nota_detalle.php?codigo=5313843&fecha=11/09/2013

López-Jurado, M., & Gratacós, G. (2013). Elegir enseñar: propuesta del modelo antropológico de la motivación de Pérez López aplicada al ámbito de la educación. *Estudios sobre educación, 24*, 125-147.

Malderez, A., Hobson, A., Tracey, L., & Kerr, K. (2007). Becoming a student teacher: core features of the experience. *European Journal of Teacher Education, 30*(3), 225-248. http://doi.org/10.1080/02619760701486068

Marcelo, C. (1995). *Formación del profesorado para el cambio educativo.* Barcelona: EUB.

Mayring, P. (2000). Qualitative Content Anlysis. *Forum: Qualitative Social Research, 1*(2). Recuperado de http://www.qualitative-research.net/index.php/fqs/article/view/1089/2385

Medrano, V., Ángeles, E., & Morales, M. (2017). *La educación normal en México. Elementos para su análisis*. México: INEE.

Mereshián, N., & Calatayud, P. (2009). Alternativas en la elección de la carrera docente. *Educación, lenguaje y sociedad, 6*(6), 251-264.

Morduchowicz, A. (2009). *La oferta, la demanda y el salario docente. Modelo para armar. Documento No. 45.* Argentina: PREAL.

Mudavanhu, Y. (2015). Contradictions and Tensions in Students' Motives of Enrolling in a Teacher Education Programme in Zimbabwe. *Journal of Education and Training Studies, 3*(2), 159-170. http://doi.org/10.11114/jets.v3i2.596

Mungarro, G., & Zayas, F. (2009). *Elección de carrera docente.* Trabajo presentado en el X Congreso Nacional de Investigación Educativa, Veracruz, México.

Norzagaray, C., Maytorena, E., & Montaño, A. (2011). Congruencia entre intereses, aptitudes y elección de carrera. *Revista Mexicana de Orientación Educativa, VIII*(21), 32-39.

Oficina Regional de Educación para América Latina y el Caribe/Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (OREALC/UNESCO). (2013). *Antecedentes y criterios para la elaboración de políticas docentes en América Latina y el Caribe.* Chile: OREAL/UNESCO.

Oliveira, V., Name, E., & Leal, L. (2010). Agentes de la Socialización: Influencia de la Familia y la Escuela en la Elección de Carrera. *Revista Mexicana de Orientación Educativa, 7*(19), 13-17.

Organización para la Cooperación y el Desarrollo Económicos (OCDE) (2010). *Mejorar las escuelas: estrategias para la acción en México. Resúmenes ejecutivos.* París: OCDE.

Pariahuache, M. (2015). *Identificación de Factores motivacionales influyentes en la elección de la carrera docente de educación en los y las estudiantes del primer año de la Facultad de Ciencias de la Educación de la Universidad de Piura* (Tesis de maestría). Universidad de Piura, Perú.

Pérez-López, J. (1991). *Teoría de la acción humana en las organizaciones. La acción personal.* Madrid: RIALP.

Pontes, A., Ariza, L., Serrano, R., & Sánchez, F. (2011). Interés por la docencia entre aspirantes a profesores de Ciencia y Tecnología al comenzar el proceso de formación inicial. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias, 8*(2), 180-195.

Randall, S., & Burley, H. (2015). Self-efficacy as a predictor of commitment to the teaching profession: A meta-analysis. *Educational Research Review, 15*,1-16.

Roloff, J., Klusmann, U., Lüdtke, O., & Trautwein, U. (2015). Who becomes a teacher? Challenging the "negative selection" hypothesis. *Learning and Instruction*, (36), 46-56.

Sánchez, E. (2009). Dos caras de la carrera docente: satisfacción y desmotivación. *Pedagogía Social. Revista Interuniversitaria, 16*, 135-148.

Silva, C. (2012). *Reclutamiento de docentes: Orientaciones para el diseño de las políticas en América Latina. Documento No. 62.* Chile: PREAL

Tang, S., Cheng, M., & Cheng, A. (2014). Shifts in teaching motivation and sense of self-as-teacher in initial teacher education. *Educational Review, 66*(4), 465-481. http://doi.org/10.1080/00131911.2013.812061

Tomsik, R. (2016). Choosing teaching as a career: importance of the type motivation in career choices. *TEM Journal. Technology Education Management Informatics, 5*(3), 396-400. doi: 10.18421/TEM53-21

Velasco, P., & Meza, J. (2013). Matrícula normalista y políticas públicas. En P. Ducoing (Coord.), *La Escuela Normal. Una mirada desde el otro* (pp. 157-189). México: IISUE-UNAM.

Watt, H., & Richardson, P. (2007). Motivational factors influencing teaching as a career choice: development and validation of the FIT-choice Scale. *The Journal of Experimental Education, 75*(3), 167-202.

Watt, H., & Richardson, P. (2008). Motivations, perceptions and aspirations concerning teaching as a career for different types of beginning teachers. *Learning and Instruction, 18,* 408-428.

Wigfield, A., & Eccles, J. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25,* 68-81.

Capítulo 4

Las técnicas de pensamiento en voz alta como herramienta para la obtención de evidencias de validez en los instrumentos de evaluación

Adán Moisés García Medina, Felipe Martínez Rizo,
Graciela Cordero Arroyo y Joaquín Caso Niebla

Introducción

Dentro de la teoría de la validación, en las últimas décadas del siglo XX se hizo cada vez más evidente que los procesos de respuesta eran el centro neurálgico en los errores de medición, y que debería de considerarse como una fuente de error, lo que finalmente recogió la versión de 1999 de los Estándares publicados por *American Educational Research Association, American Psychological Association* y *National Council on Measurement in Education* (AERA, APA y NCME, 1999). Desde entonces, los procesos de respuesta se consideran fuente de evidencia de validez. El supuesto es que los procesos cognitivos de quien responde conducen la respuesta que da al ítem, y entender la cognición es central para el diseño de reactivos y la comprensión y reducción de fuentes de error en las respuestas (Willis, 2005). La obtención de evidencias de validez sobre los procesos de respuesta consiste en un análisis empírico y teórico del ajuste que existe entre los procesos cognitivos que emplean los examinados para responder a las tareas que se le solicitan en determinado instrumento y su correspondencia con los procesos que se asume deben utilizar cuando se indaga el constructo medido (AERA, APA y NCME, 2014; Miller, Willson, Chepp y Padilla, 2014).

Existen varias técnicas para medir los procesos de respuesta, que pueden clasificarse en dos categorías: 1) Las que acceden directamente a los procesos cognitivos, como protocolos verbales, reportes verbales (estas dos conocidas como técnicas de pensamiento en voz alta), entrevistas

cognitivas y grupos focales (Schuman, 2001; Schwarz, 2001; Tanur, 2001); 2) Las que proveen indicadores indirectos y requieren una inferencia adicional, como los estudios del sendero de la vista y aquéllos que registran el tiempo de latencia de respuesta (Padilla y Benítez, 2014). Las que se han extendido más en el diseño de instrumentos de evaluación, han sido las de pensamiento en voz alta, protocolos y reportes verbales (Miller et al., 2014).

En el campo de la evaluación educativa en México son escasos los estudios que han utilizado las técnicas de pensamiento en voz alta (García-Medina, 2017; Perez-Morán, 2014), pese a su enorme potencial para mejorar instrumentos de medición, incluidos los test. Por ello, en este capítulo se ofrece un procedimiento detallado sobre cómo utilizarlas con un propósito específico: obtener evidencias de validez sobre los procesos de respuesta, para identificar posibles sesgos en reactivos de pruebas estandarizadas de logro. El procedimiento ha sido utilizado con resultados positivos en la investigación doctoral de la que se deriva este capítulo, cuyos objetivos fueron:

- Documentar los fundamentos teóricos de las evidencias de validez basadas en el aspecto de equidad de pruebas educativas.
- Evaluar el diseño de los ítems de una prueba aplicada en 2012, a través de las categorías de análisis del Modelo del Diseño Universal de las Evaluaciones (DUE).
- Identificar sesgo potencial en los ítems de la prueba aplicada en 2012 que mostraron un grado alto de funcionamiento diferencial, utilizando protocolos verbales y técnicas de pensamiento en voz alta (García-Medina, 2017).

Qué son los protocolos y los reportes verbales

En evaluación educativa, los protocolos verbales son una serie de indicaciones para promover que el sujeto, al tiempo que responde cada ítem, verbaliza los procesos cognitivos que pone en juego para dar respuesta a las tareas evaluativas (Ericsson y Simon, 1993). La función principal de estos protocolos es que el sujeto se mantenga verbalizando lo que piensa desde que inicia a leer el pie del reactivo hasta que elige una opción de respuesta (Leighton, 2009). El análisis de protocolos verbales se suele utilizar para identificar y medir el proceso de resolución de problemas, especialmente, para identificar los contenidos de la memoria de trabajo de los sustentantes (Ericsson y Simon, 1993), asociados con las funciones ejecutivas de la memoria.

Los protocolos verbales incluyen una serie de pautas para que el examinado verbalice mientras evoca los conocimientos y procesos cognitivos con los que se enfrenta a los ítems. Los

protocolos suelen incluir preguntas de sondeo (*probes*), para obtener información sobre aspectos específicos de aquellos ítems de los que se sospecha que tienen problemas en su formulación. Las preguntas de sondeo que se incluyen en el protocolo son redactadas para: corroborar o descubrir si un reactivo contiene términos que pueden resultar confusos o desconocidos para ciertos subgrupos; identificar si la estructura sintáctica dificulta la comprensión de la tarea; profundizar sobre las operaciones cognitivas que los individuos pusieron en juego; constatar si, para resolver determinado ítem, los sujetos tuvieron que recordar determinados conocimientos; o validar el modelo cognitivo del conjunto de reactivos de un test (Benítez y Padilla, 2014; Castillo-Díaz y Padilla, 2013).

Los reportes verbales, por su parte, propician que el sustentante vaya pensando en voz alta, pero a diferencia de los protocolos verbales, una vez que el examinado comienza a responder los ítems no se le dan instrucciones adicionales ni se le interrumpe durante el proceso de resolución, salvo para solicitar que continúe verbalizando su pensamiento. Con este método se suele indagar en la memoria de largo plazo de los examinados, la cual contiene los esquemas mentales más estables, con los que el entrevistado suele operar de forma cotidiana.

Cada técnica tiene ventajas y desventajas. La principal ventaja de los protocolos verbales es que permiten indagar un proceso cognitivo que no está siendo explicitado por el examinado, al tiempo que piensa en voz alta mientras responde a un reactivo, y eso permite conocer con bastante detalle cada paso que sigue el informante; por el contrario, tiene la desventaja de que el proceso cognitivo "natural" puede verse afectado, pues cuando el entrevistador plantea preguntas de sondeo incluidas en el protocolo, el entrevistado puede percatarse de algunas operaciones cognitivas que debió emplear y modificar la senda de su proceso cognitivo.

Los reportes verbales tienen la ventaja de que el examinado hace una introspección en su memoria de largo plazo y utiliza un proceso cognitivo natural; su principal desventaja es que el entrevistado podría no ser capaz de recordar con precisión el proceso cognitivo que empleó, y hacerlo saber o, en el peor de los casos, falsear la información (Willis, 2005; 2015).

Cuál es la utilidad de los protocolos y reportes verbales

Las técnicas de pensamiento en voz alta permiten explicitar el proceso de respuesta que se sigue con un instrumento de medición. En pruebas de logro, los protocolos y reportes verbales permiten mapear el proceso cognitivo utilizado por los sustentantes y usarlo con propósitos como la validación

del esquema cognitivo requerido para responder un reactivo, como insumo para generar ítems adaptativos al identificar con precisión la ruta cognitiva y el nivel de complejidad de ciertos ítems y, de forma especial, para identificar sesgo en los reactivos.

En el campo de las evaluaciones en gran escala, un ítem se considera sesgado cuando incluye contenido que, además de que podría ser ofensivo, penaliza a ciertos grupos de estudiantes debido a características personales como raza, género, etnicidad, discapacidad o estatus socioeconómico, que representan aspectos distintos a lo que se pretende medir y, por tanto, no debieran favorecer u obstaculizar a los sustentantes (AERA, APA y NCME, 2014). Se parte del supuesto de que algunos rasgos de una prueba pueden dificultar que algunos sustentantes muestren el manejo que tienen de lo que se quiere medir, por ejemplo, cuando el diseño tipográfico no permite que los sustentantes con alguna debilidad visual puedan leer los ítems de una prueba que no está destinada a medir la capacidad visual (Thompson, Johnstone y Thurlow, 2002).

Para ilustrar cómo las técnicas de pensamiento en voz alta pueden servir para identificar reactivos con sesgo potencial, se presenta el análisis de un reactivo analizado en el marco de la investigación doctoral, y mediante la cual se identificaron 13 reactivos con sesgo potencial de una prueba que se aplicó a estudiantes mexicanos de tercer grado de secundaria.

Análisis de un reactivo mediante protocolos y reportes verbales

El ítem SMB-01 (figura 1) es parte del banco de reactivos de una prueba de logro en Matemáticas, estandarizada y de gran escala. El ítem fue seleccionado para analizarse mediante reportes y protocolos verbales porque en un estudio previo (García-Medina, Martínez y Cordero, 2016) presentó un funcionamiento diferencial alto (0.98 lógitos) a favor de alumnos con nivel socioeconómico (NSE) alto, respecto a los de NSE bajo, lo que en términos simples significa que alumnos de NSE bajo, que tienen la habilidad matemática requerida para responderlo correctamente, lo responden de manera errónea por cuestiones ajenas a lo que el ítem pretende medir, lo que se puede señalar como una fuente de sesgo.

En el programa de estudios de Matemáticas, este ítem pertenece al eje *Sentido numérico y pensamiento algebraico*, y se ubica en el tema *Significado y uso de las operaciones*, y en el subtema *Problemas aditivos*. El reactivo se generó a partir de la especificación *Resolver problemas aditivos que impliquen el uso de números con signo* (Instituto Nacional para la Evaluación de la Educación [INEE], 2010a).

En el poblado donde vivo se registró una temperatura de 12°C el domingo a mediodía; durante las siguientes 18 horas la temperatura descendió 15°C. ¿Qué temperatura marcó el termómetro después de las 18 horas transcurridas?

- ○ -3°C *
- ○ 3°C
- ○ 27°C
- ○ -27°C

* Opción correcta

Figura 1. Reactivo SMB-01.

De acuerdo con su especificación, para resolver el reactivo se requiere que el sustentante domine las siguientes habilidades y tenga los siguientes conocimientos previos:

- Sumar y restar números naturales, fraccionarios y decimales positivos o negativos.
- Conocer la secuencia numérica ascendente y descendente a partir de cualquier número.
- Utilizar criterios, en forma intuitiva, para determinar si una situación problemática se resuelve con una suma o con una resta.
- Calcular sumas y restas apoyándose en la recta numérica.
- Resolver problemas aditivos de cambio y de composición (INEE, 2010a, p. 6).

Dos especialistas en enseñanza de las matemáticas, a las que se les pidió que anticiparan el proceso cognitivo que emplearían los sustentantes, señalaron que para responder este ítem se seguiría una ruta cognitiva como la presentada en la figura 2, que en un inicio tiene una parte común y, posteriormente, los sustentantes podrían optar por alguno de los dos procesos.

Un primer análisis, para identificar el potencial problema de definición imprecisa del constructo, consistió en contrastar los conocimientos y habilidades estipulados en la especificación con los procesos cognitivos anticipados por las especialistas, previo a la aplicación de los reportes y protocolos verbales. Este primer análisis apuntó a que este ítem tiene coherencia con el constructo que pretende medir. Cuando las especialistas anticiparon que, para enfrentarse a la tarea evaluativa, los sustentantes deberían reconocer que el reactivo les planteaba un problema aditivo que implicaba un descenso de temperatura, y que para operar con los valores de las temperaturas

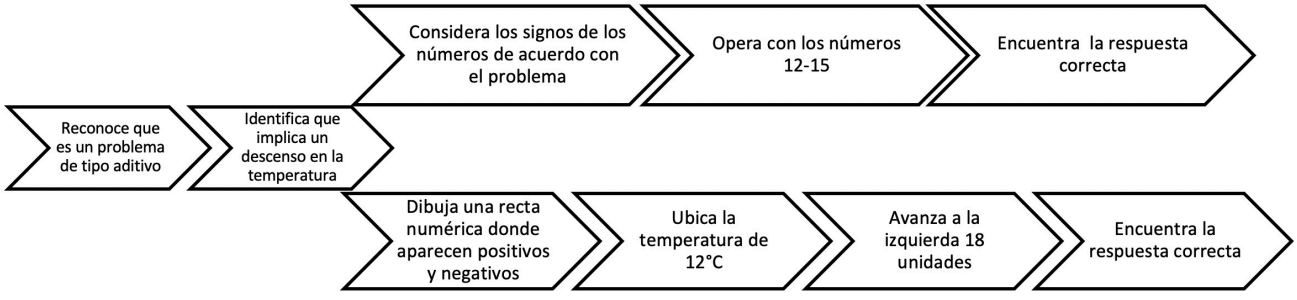

Figura 2. Proceso(s) cognitivo(s) anticipados por las especialistas en enseñanza de las matemáticas para el reactivo SMB-01.

deberían considerar los signos de los números, lo cual es coincidente con lo que pretende medir el ítem: si el sustentante es capaz de sumar y restar números naturales, fraccionarios y decimales positivos o negativos; usar intuitivamente criterios para determinar si una situación problemática se resuelve con una suma o una resta; calcular sumas y restas apoyándose en la recta numérica; o resolver problemas aditivos de cambio y de composición.

Un segundo análisis, a partir de los reportes y protocolos verbales, sugiere que los entrevistados, al enfrentarse a esta tarea, evocan los conocimientos señalados en la especificación del reactivo y emplean, con algunas variantes, las operaciones cognitivas anticipadas por las especialistas. La evidencia muestra, en efecto, que cuando los estudiantes responden al reactivo tienen que reconocer que implica un problema de tipo aditivo con números con signo; conocer la secuencia numérica ascendente y descendente a partir de cualquier número; utilizar intuitivamente criterios para determinar si una situación problemática se resuelve con una suma o una resta; y calcular la resta, apoyándose opcionalmente en la recta numérica.

Por ejemplo, al sondear con el protocolo verbal si uno de los alumnos que respondió de manera correcta el reactivo tiene claro lo que debe hacer, responde:

> *sí […] lo que el problema básicamente dice es que a las 12 del domingo estaba a 12 grados, durante las 18 horas descendió 15 [grados], entonces ¿qué temperatura marcó el termómetro después de las 18 horas transcurridas? Es decir, 12 grados menos 15 grados… no está tan difícil.* (E07)

Otro de los alumnos que respondió correctamente, verbaliza mientras lee el pie del ítem *"es una simple resta…"* (E05). En ninguno de los entrevistados se encontraron evidencias de que al resolver el SMB-01 evocaran conocimientos o habilidades distintos a los que el reactivo pretende medir, por tanto, el reactivo no parecer presentar el problema de una definición imprecisa del constructo.

Sin embargo, cuatro de los nueve sujetos a los que se les aplicaron los reportes y protocolos verbales (3 que respondieron correctamente y 1 que no seleccionó la respuesta correcta), aportaron elementos que sugieren que el reactivo SMB-01 tiene sesgo potencial para sustentantes según su NSE. Para el grupo de NSE bajo, la utilización de ciertos términos en el pie del ítem puede obstaculizar que los sustentantes muestren a cabalidad sus habilidades matemáticas.

El proceso empleado por uno de los estudiantes de telesecundaria es muy ilustrativo de cómo el desconocimiento de la palabra *descendió* ocasionó que eligiera una opción incorrecta. Este alumno, después de leer el pie del ítem, reconoce que se trata de un problema de adición y resuelve la operación 12+15=27, luego selecciona la opción correspondiente a 27°C, que es errónea. Este alumno no consideró de manera correcta el signo de las temperaturas, porque para él la palabra *descendió* significa un aumento en la temperatura; cuando verbalizaba su comprensión de la tarea evaluativa, dijo: *"Aquí observo que aumentó quince grados durante esas 18 horas"* (E01).

Durante el análisis de las opciones de respuesta el participante señala que, como la temperatura inicial es de 12°C, y luego aumenta 15°C, las opciones de respuesta con números negativos no pueden ser correctas; esto se ilustra cuando verbaliza: *"[…] menos tres grados centígrados no puede ser, si la temperatura está aumentando"* (E01). Su comprensión errónea de la palabra *descendió* se pudo corroborar en otras partes del protocolo verbal. Por ejemplo, cuando explica cómo descartó algunas opciones de respuesta, señala que, si el reactivo plantease "un incremento proporcional de 3°C por cada hora", algunos alumnos podrían elegir la opción de 3°C. El entrevistado, sin embargo, señala que eso no es lo que plantea el problema, por lo que no seleccionó la segunda opción.

La evidencia obtenida a través de las técnicas de pensamiento en voz alta permitió también conocer que el estudiante disponía de los conocimientos previos señalados en la especificación del reactivo (suma de números naturales, utilizar intuitivamente criterios para determinar si un problema se resuelve con una suma, resolución de problemas aditivos de cambio y de composición). Sin embargo, la inclusión en el reactivo de una palabra clave para la comprensión del problema, que era poco familiar para el estudiante, se convirtió en un obstáculo para escoger la opción correcta.

Por otro lado, al parecer el reactivo también se puede responder correctamente a partir de la información del contexto que tienen ciertos subgrupos poblacionales. Una alumna de NSE medio que respondió en forma correcta, pero no utilizó un procedimiento matemático para llegar a la solución, aplicó su conocimiento empírico sobre las temperaturas durante el día y la noche para acotar las opciones entre las que luego adivinó la respuesta. Cuando se le preguntó por qué seleccionó la opción A (correcta), la alumna respondió: *"cambió… al ser de día, va oscureciendo, va oscureciendo y pues va bajando más la temperatura, porque en el día hace más calor y cuando va anocheciendo baja la temperatura y hace más frío"* (E04).

Si se toma en cuenta que durante su proceso para la resolución del problema no realizó ninguna resta y, además, dudaba entre las dos opciones que tenían números negativos (-3° y -27°), se puede inferir que dedujo la respuesta correcta a partir de su conocimiento del medio. La alumna verbalizó *"yo creo que es ésta* [refiriéndose a la opción a)] *…pero también podría ser ésta* [opción d)]" (E04).

En síntesis, el desconocimiento o comprensión de una palabra como *descendió*, para ciertos grupos de población, denota que el reactivo, si bien mide conocimiento matemático, mezcla también cuestiones de vocabulario, que corresponde a la asignatura de Español (comprensión lectora o eventualmente reflexión sobre la lengua), lo que no es deseable.

Cómo realizar análisis cognitivos mediante reportes y protocolos verbales

El ejemplo de análisis de un reactivo mostrado en el apartado anterior condensa el resultado de un largo proceso de diseño, implementación y análisis de los reportes y protocolos verbales. Para guiar a otros investigadores educativos interesados en utilizar esta técnica, se presentan enseguida las seis fases que se siguieron en el estudio cognitivo que formó parte de la investigación doctoral.

Diseño

En la primera fase, que es fundamental, se desarrollaron once actividades:

A) Selección de ítems del estudio cognitivo

Del total de 100 reactivos de la prueba, se seleccionaron 18 para el estudio: los que presentaron un funcionamiento diferencial severo entre las subpoblaciones según su NSE (García-Medina et al.,

2016). Estos ítems y las especificaciones[1] a partir de las cuales se elaboraron, fueron facilitados por el INEE con condiciones de confidencialidad, por lo cual varios aspectos de dichos elementos no se presentan en este texto.

B) Identificación de los procesos cognitivos necesarios para responder a los ítems, declarados en el modelo teórico de la prueba

Se analizaron las especificaciones de los ítems, que están organizadas en seis apartados. El segundo (*Descripción del contenido curricular a evaluar*) incluye dos subapartados fundamentales:

- *Actividades cognoscitivas involucradas en la contestación correcta del ítem*, en que se establecen proceso(s) cognitivo(s) que los sustentantes tendrían que usar para responder correctamente. En una de las especificaciones para resolver problemas aditivos con números fraccionarios con distinto denominador se señala que las actividades cognoscitivas que se tendrán que poner en juego son: a) verificar que el problema presentado se resuelve mediante la suma o resta de números fraccionarios con distinto denominador; b) aplicar el algoritmo de equivalencia de fracciones en la adición o sustracción de dichos números; y c) resolver problemas con operación(es) que implique(n) el uso de números fraccionarios con distinto denominador, a partir de una situación (INEE, 2010b, s/p).
- *Conocimientos y habilidades previos requeridos para contestar el ítem correctamente.* Se establecen saberes matemáticos que son condición necesaria para que un alumno responda correctamente el ítem. Para la especificación anterior, se indica que los conocimientos y habilidades necesarios son: a) otorgar sentido a cantidades expresadas como fracción: b) interpretar cantidades fraccionarias en contextos de medición; c) expresar cantidades con fracciones; d) emplear algoritmos de adición y sustracción de números naturales mediante equivalencias de fracciones; e) resolver, con base en equivalencia de fracciones, problemas aditivos simples y combinados, directos e inversos que involucren números fraccionarios; y f) identificar y producir familias de fracciones equivalentes mediante la amplificación y simplificación, reconociendo que expresan una misma cantidad (INEE, 2010b, s/p).

1 *Las especificaciones de reactivos son indicaciones pormenorizadas a partir de las cuales se diseña un grupo de ítems con los que se pretende medir un contenido y/o habilidad en particular.*

C) Anticipación de los procesos cognitivos necesarios para responder a los ítems, con base en la opinión de especialistas

Tomando como referente las especificaciones, dos especialistas en enseñanza de matemáticas en educación básica definieron, para cada ítem, el proceso cognitivo principal, y, en su caso, sus variantes, por las que deberían transitar los alumnos para responder correctamente. Ayudaron también a identificar las distorsiones más comunes asociadas a ciertas respuestas incorrectas. Como resultado se obtuvo el o los procesos cognitivos que teóricamente utilizarían los entrevistados.

Se diseñó un formato en que las especialistas registraron: a) datos de identificación del ítem; b) los conocimientos y habilidades previos requeridos para contestarlo correctamente; c) las actividades cognoscitivas involucradas en la respuesta correcta; d) el proceso cognitivo que anticiparon que los entrevistados emplearían para enfrentar la tarea, para lo que hicieron un diagrama que representase el sendero cognitivo previsto (lineal, en paralelo, en espiral, lineal al inicio y luego una bifurcación, etc.); y e) el nivel de demanda cognitiva del ítem (conceptual, procedimental, resolución de problemas conocidos o típicos, y resolución de problemas desconocidos o transferencia).

D) Selección de la modalidad de técnica cognitiva a emplear

Como se ha señalado, hay varias formas de acceder al proceso de respuesta, que pueden englobarse en dos grandes tipos: reportes verbales y protocolos verbales.

A diferencia de los reportes verbales, que ocurren al mismo tiempo que se resuelve un ítem, los protocolos verbales pueden ser concurrentes o retrospectivos: concurrentes cuando poco tiempo después (incluso inmediatamente después) de que el examinado responde el ítem en cuestión, se le pide que explique, describa, profundice o haga explícito el proceso mental que siguió para hacerlo; retrospectivos si se llevan a cabo después de un tiempo considerable de que el sustentante se enfrentó a las tareas evaluativas.

En el estudio se emplearon reportes verbales y protocolo verbal concurrente. Aunque se privilegió identificar el modelo cognitivo empleado por los sustentantes a través de la verbalización de sus pensamientos mientras se enfrentaban a cada ítem, como estrategia complementaria se creyó conveniente el sondeo verbal de los protocolos, para identificar comprensiones de conceptos implicados en los ítems asociadas a aspectos culturales y del NSE, que fue el foco del trabajo.

E) Diseño de la guía para la realización de entrevistas concurrentes y la técnica de pensamiento en voz alta

La guía pautó la intervención del entrevistador durante el estudio. En ella se detallan los pasos que se debieron dar antes, durante y después de aplicar las técnicas de pensamiento en voz alta con cada alumno. En especial, se presentan las indicaciones que se dieron a los sujetos para entrenarlos para pensar en voz alta; se incluyeron las consignas a dar si dejaban de verbalizar lo que estaban pensando mientras resolvían los ítems; se definieron las posibles preguntas a utilizar en los sondeos verbales concurrentes (para precisar la ruta cognitiva utilizada por el estudiante y para identificar fuentes de sesgo asociadas al NSE del participante); y se precisaron aspectos particulares que el entrevistador debió cuidar con ciertos ítems.

F) Diseño del instrumento para recolectar los datos

Si bien las entrevistas de pensamiento en voz alta se videograbaron para tener un registro de la información obtenida, se planeó que también hubiera un instrumento que permitiese recolectar información básica de primera mano mientras se realizaba el levantamiento de datos. Para ello, se diseñó el *Formato de Recolección de Datos* que el entrevistador debía completar registrando, para cada ítem: palabras leídas de manera incorrecta por el entrevistado; términos en que manifestó confusión o desconocimiento; preguntas de sondeo concurrente adicionales a las previstas en el protocolo hechas por el entrevistador; si el entrevistado respondió correctamente el ítem; si el proceso para enfrentar la tarea fue apropiado, entre otras cuestiones. Este instrumento sirvió como un primer análisis de cada ítem, que se complementó con la revisión de las videograbaciones.

G) Definición de los instrumentos para el procesamiento de la información

Para procesar la información obtenida a través de la videograbación y del *Formato de Recolección de Datos* que completó el entrevistador, se diseñó el *Formato para concentrar y procesar la información de cada ítem-estudiante*, que se detalla en la fase del procesamiento de la información.

H) Selección de los participantes

La selección se hizo a partir de los siguientes criterios: el primero y fundamental fue su pertenencia a una de las tres subpoblaciones participantes de NSE (alto, medio o bajo) en las que se presumía sesgo en los ítems, como resultado de los análisis del funcionamiento diferencial de los ítems (DIF,

por sus siglas en inglés) llevados a cabo previamente; un segundo criterio fue que los participantes provinieran de distintas modalidades de la educación secundaria, lo que ayudó a aleatorizar las variables no consideradas como criterio de selección; el tercer criterio fue que no hubieran respondido antes la prueba; el cuarto fue su rendimiento en Matemáticas, según sus calificaciones: se pidió a sus maestros que escogieran un alumno con promedio de 9 o superior y otro alumno con promedio de 7 o menos, de tal forma que hubiera variedad en los niveles de logro, y se pudiera explorar la presencia de sesgo parcial o específico para ciertos grupos.

Se entrevistaron seis estudiantes de tercero de secundaria y sus tres docentes de matemáticas: dos alumnos de una Telesecundaria de NSE bajo y su profesor; dos de una secundaria general pública de NSE medio y su maestra; y dos de una secundaria privada de NSE alto y su profesor.

I) Definición de la cantidad de ítems por participante

Considerando que en cada grupo de NSE hubo sólo un estudiante de rendimiento alto y uno de rendimiento bajo, se definió que cada entrevistado respondiera los 18 ítems que mostraron DIF severo por NSE, para que ante una eventual situación donde los resultados requirieran explorar si el factor asociado a los resultados diferenciales fuera el rendimiento escolar, se tuviera la posibilidad de hacer dicha comparación. Así, al final se tuvieron 108 respuestas de alumnos y 54 respuestas de los profesores, siendo un total de 162 respuestas.

J) Preparación de cuadernillos para estudio cognitivo

Los 18 ítems que se aplicaron en el estudio cognitivo son sólo una muestra de la escala completa de la prueba, que consta de cien ítems. Por ello, se requirió preparar cuadernillos que tuvieran únicamente los 18 ítems a estudiar. Además, como se pretendió que no hubiera distractores que alteraran los resultados de los reportes y protocolos verbales, se cuidó que en cada hoja hubiera un solo ítem y que el cuadernillo estuviera armado de tal forma que el participante sólo pudiera trabajar en un ítem a la vez.

K) Adaptación del laboratorio cognitivo

Se adaptó un cubículo como laboratorio cognitivo: se dispuso de mobiliario (escritorio y sillas) para que cada entrevistado respondiera, y una videocámara de alta definición colocada en un trípode, con el receptor de un micrófono inalámbrico Lavalier, cuyo transmisor se puso a cada estudiante,

de manera que tanto sus verbalizaciones mientras respondía los ítems, como sus respuestas a las entrevistas concurrentes, fueran grabadas con toda claridad.

Pilotaje

La estrategia señalada en la fase del diseño se puso a prueba con una estudiante de tercero de secundaria. Los propósitos de esta fase fueron identificar y mejorar el entrenamiento que recibieron los participantes para las técnicas de pensamiento en voz alta, evaluar los instrumentos de obtención de información, y la operación del laboratorio cognitivo. A partir del piloto se hicieron ajustes a la *Guía para la realización de entrevistas* y al *Formato para recolección de información,* precisando algunas instrucciones, y anticipando la necesidad de disponer de materiales como borrador, sacapuntas, agua embotellada y un paquete de materiales como gratificación para los docentes.

Realización

La tercera fase consistió en llevar a cabo los reportes y protocolos verbales. Con apego a la *Guía para la realización de entrevistas mediante protocolos verbales y técnicas de pensamiento en voz alta*, en esta fase se realizaron los siguientes pasos:

1) Instalación del laboratorio cognitivo:

- Mobiliario: 1 escritorio y 2 sillas, una frente a otra, a no más de 1.5 m de distancia.
- Equipo: videocámara, trípode, micrófono Lavalier, memoria SD, cargador de la cámara, extensión eléctrica.
- Preparación de equipo de videograbación: en ubicación apropiada (evitando contraluz y que aparecieran en la toma tanto el entrevistado como el entrevistador); realizar las conexiones entre los componentes de manera que estuvieran listos para su utilización.

2) Revisión de materiales para la entrevista:

- Para entrevistado: cuadernillo con ítems, lápiz, borrador, sacapuntas, consentimiento informado, hojas blancas.
- Para entrevistador: guía de entrevista, formato de recolección de datos, ítems para entrenamiento, lápiz y sacapuntas.

3) Bienvenida:

- Breve presentación personal por parte del entrevistador.
- Agradecimiento al entrevistado por participar en el estudio.
- Breve presentación del estudio, sobre todo con participantes que no estuvieron en la negociación de participación.
- Descripción al entrevistado de las actividades que realizaría.

4) Firma de consentimiento informado (si no se hizo en la negociación de participación).
En concordancia con los criterios éticos, se pidió a los participantes su consentimiento para participar: se les informó sobre los propósitos del estudio; la manera en que se usarían los resultados; su propio papel en la investigación; y los cuidados previstos para salvaguardar la confidencialidad de la información que den. Se preparó una carta de consentimiento que los participantes debieron firmar como condición para participar.

5) Entrenamiento en las técnicas a emplear. Se leyó el siguiente mensaje a los entrevistados:

> *Estamos interesados en conocer cómo respondes algunas preguntas de opción múltiple de un examen de matemáticas. Para ello, voy a mostrarte unos ítems y te pediréque pienses en voz alta cómo los vas solucionando, después te podría hacer algunas preguntas. Pensar en voz alta significa que quiero que me digas todo lo que pasa por tu mente desde el primer momento en que veas la pregunta del examen hasta el final cuando elijas la opción que consideres correcta. Mientras revisas cada pregunta, es muy importante que te mantengas hablando, si en algún momento guardas silencio, te pediré que vuelvas a pensar en voz alta. Actúa como si estuvieras solo(a) en este salón y hablando para ti mismo(a). ¿Hasta ahora todo lo que te comenté es claro?*

Se dio a cada entrevistado un ejemplo de cómo pensar en voz alta: el entrevistador resolvió una resta o multiplicación de dos dígitos, verbalizando lo que pensaba en todo el proceso. Se hizo la primera práctica, resolviendo operaciones (no ítems todavía). Se leyó:

Ahora vamos a practicar un poco, para que te acostumbres a pensar en voz alta. Te voy a pedir que esta hoja blanca que te acabo de proporcionar multipliques dos números.Así que piensa en voz alta mientras multiplicas 15 por 25.

Como ejemplo de cómo resolver un ítem, el entrevistador respondió uno o dos reactivos, para que el participante pudiera observar el comportamiento y desempeño que se esperaba de él. Para esta etapa de entrenamiento se utilizaron algunos de los ítems de la prueba que no tuvieron DIF, para no distorsionar los procesos cognitivos que el estudiante empleaba en cada uno de los ítems que interesó analizar. El entrevistador leyó las siguientes instrucciones:

Ahora voy a pensar en voz alta mientras resuelvo este problema. Eso significa que voy a decir todo lo que pasa por mi mente, incluyendo cómo entendí el problema que tengo que resolver; si hay palabras que no comprendo, si me resultan claras las indicaciones, si el problema presenta los datos suficientes para resolverlo, si recuerdo algún procedimiento que pueda utilizar para resolverlo, si en las opciones de respuesta viene alguna palabra que no comprendo o tengo dudas de lo que significa, etc.

El entrenamiento terminó cuando el investigador observó que el participante presentaba un nivel suficiente de desautomatización[2] y conciencia de los procesos y recursos cognitivos que usaba para responder un ítem, de forma que en sus verbalizaciones se pudieran identificar de forma clara los pasos utilizados para responder y, de ser el caso, las dificultades que tuvo.

Una vez concluido el entrenamiento, se recordó al entrevistado que el objetivo por el que se le estaba entrevistando era conocer el proceso que empleaba para responder cada ítem. Al final, se le preguntó si tenía alguna duda, eventualmente se le resolvió y, ante la ausencia de dudas, se continuó con el siguiente paso.

2 *La desautomatización de las respuestas consiste en que el entrevistado se vuelva consciente del proceso cognitivo que emplea para resolver un problema y, posteriormente, sea capaz de expresar los pasos que siguió para responder a un ítem.*

6) Aplicación de técnica de pensamiento en voz alta y entrevista concurrente.

- Se inició la videograbación, cerciorándose que el micrófono estuviera funcionando.
- Se entregó el cuadernillo al estudiante y se le pidió que leyera las indicaciones que le indicaban que en cada página sólo encontraría un ítem, y que no podría cambiar de página hasta que el investigador se lo hiciera saber.
- Para cada ítem se realizó el siguiente ciclo, que se repitió 18 veces:
    a. Se pidió al estudiante que verbalizara sus pensamientos. Durante ese tiempo el entrevistador tuvo dos funciones: asegurarse que el entrevistado pensara en voz alta mientras respondía el ítem, e identificar lagunas de información para registrar en el Formato de recolección de datos las preguntas adicionales que formularía más tarde para subsanar dichos vacíos de información.
    b. Una vez que el alumno registró en el cuadernillo su respuesta a un ítem, se hizo la entrevista concurrente a partir de las preguntas eje del Formato de recolección de datos, y de preguntas adicionales que el investigador redactó durante la entrevista de pensamiento en voz alta, para detallar el proceso empleado e identificar términos que por la forma en que se entienden, pudieran alterar el resultado.

7) Agradecimiento. Además de un agradecimiento verbal, a estudiantes y maestros se le dio una batería portátil; a los docentes se les regaló además un paquete de materiales con temáticas para el mejoramiento de las prácticas de evaluación el aula.

8) Integración del expediente de cada entrevistado, que debía tener:

- Formato de recolección de datos debidamente complementado.
- Archivo con la videograbación, identificado con el número de folio del entrevistado.
- El cuadernillo del examen respondido por el entrevistado.
- Consentimiento informado firmado por el entrevistado.

Procesamiento de la información

Se utilizó el *Formato para concentrar y procesar la información de cada ítem-estudiante*. Como la información varía según el ítem, se diseñaron 18 formatos, cada uno con 12 puntos:

1. Datos de identificación del sustentante: folio interno, subpoblación a que pertenece, tipo de escuela de la que proviene: general, pública o privada, o telesecundaria.
2. Datos del ítem: folio, dominio al que pertenece, código de la especificación de la que se creó.
3. Conocimientos y habilidades previos requeridos para contestar el ítem correctamente, según la especificación correspondiente.
4. Actividades cognoscitivas implicadas en la respuesta correcta según la especificación.
5. Procesos cognitivos previstos por especialistas en enseñanza de matemáticas.
6. Nivel de demanda cognitiva previsto por dichos especialistas.
7. Procesos cognitivos empleados por los sustentantes, registrando el proceso que empleó cada entrevistado para responder a un ítem en particular.
8. ¿El entrevistado respondió correctamente el ítem? "Sí" o "No".
9. Adecuación del proceso de solución del problema.
10. Fluidez en la lectura del ítem.
11. Tipo de problema que presentó el ítem, registrando si hubo alguno de los siguientes, añadiendo una descripción en caso de respuesta afirmativa:
    a. Definición imprecisa del constructo
    b. Ítem no accesible o sesgado
    c. Procedimientos e instrucciones complejas, poco claras e intuitivas
    d. Incomprensibilidad o problemas de legibilidad del ítem
12. Notas adicionales: información sobre problemas distintos a los cuatro señalados en el punto anterior, o bien adelantando conclusiones preliminares.

Si bien, cada uno de estos puntos jugó un papel en el estudio, el N° 11 se consideró medular, pues se señaló si el entrevistado había mostrado evidencias de que un ítem presentaba alguno de los cuatro tipos de sesgo que se exploraron en este trabajo, definidos a partir de los elementos clave del Diseño Universal de las Evaluaciones (Thompson et al., 2002).

La evidencia se conformó a partir del pensamiento en voz alta que expresó cada entrevistado mientras respondía los ítems, y de sus respuestas a las preguntas que hizo el entrevistador a manera de sondeo concurrente, para explorar partes del proceso de respuesta no detalladas mientras el entrevistado verbalizaba lo que pensaba para enfrentarse a la tarea. No se hizo una transcripción

completa de las videograbaciones de las entrevistas, sino una selectiva, de manera que lo transcrito ilustrara la presencia o ausencia de los tipos de sesgo explorados.

En total, para procesar la información se llenaron 162 formatos, de los nueve entrevistados y su respuesta a cada uno de los 18 ítems. Para organizar los archivos se utilizó una nomenclatura en la que se concatenó el número del ítem y el folio del entrevistado. El archivo que contiene el procesamiento de la información del entrevistado N° 3 sobre el ítem SMA-05, por ejemplo, se identificó como SMA-05_F03.

Análisis cognitivo de cada ítem

Esta fase consistió en analizar la información procesada de los nueve entrevistados para cada uno de los 18 ítems, que fueron la unidad de análisis. El análisis de cada ítem incluyó:

- Identificación del ítem: el reactivo, señalando cuál opción de respuesta es la correcta; su ubicación en el programa de matemáticas (eje, tema y subtema); la especificación a partir de la cual fue diseñado; y el tipo de DIF (uniforme o no uniforme, en sentido positivo o negativo, y la magnitud en lógitos) identificado previamente.
- La resolución del ítem y adecuación del proceso empleado: cantidad de alumnos y docentes (desagregados por NSE) que respondieron de manera correcta e incorrecta el ítem, así como la adecuación del proceso empleado para su resolución.
- Tipo de problemas de varianza irrelevante al constructo que presenta el ítem: cantidad de entrevistados (desglosados por NSE) que aportaron evidencias para señalar que el ítem tiene alguno de los cuatro problemas de sesgo antes mencionados.

Es importante señalar que el foco del análisis no era identificar concordancia entre el modelo teórico de la prueba y un modelo identificado a partir de los procesos de respuesta de los examinados, sino identificar posibles fuentes de sesgo en función del NSE de los sujetos.

Análisis de sesgo mediante el Diseño Universal de las Evaluaciones (DUE)

La última fase consistió en un análisis de los resultados de la fase previa, tomando como eje cada una de las cuatro dimensiones del DUE que se exploraron en este trabajo (Thompson et al., 2002). Se hizo una síntesis para cada uno de los cuatro posibles problemas de sesgo que se exploraron: 1) Definición imprecisa del constructo; 2) Ítem no accesible o sesgado; 3) Procedimientos e instrucciones complejas, poco claras e intuitivas; y 4) Incomprensibilidad o problemas de legibilidad del ítem.

Para la realización de la síntesis por cada tipo de problema se contabilizó la cantidad de reactivos que lo presentaron, desglosándolo por el tipo de DIF con el que fue diagnosticado, y el tipo de entrevistado. Además, se incluyeron fragmentos representativos del problema en cuestión, que fueron obtenidos a partir de los reportes y protocolos verbales.

Conclusiones y recomendaciones

Las técnicas de pensamiento en voz alta son herramientas valiosas para conocer el proceso de respuesta cuando un sustentante se enfrenta a tareas evaluativas de un test, o responde preguntas de un cuestionario aplicado en una encuesta. La información conseguida mediante estas técnicas permite a los diseñadores de pruebas contrastar los esquemas cognitivos que deberían poner en juego los sustentantes al enfrentarse a las tareas evaluativas con los que en realidad utilizan, lo que permite identificar discrepancias entre las habilidades que los ítems pretenden medir y las que realmente evalúan.

El análisis de reactivos de una prueba mediante reportes y protocolos verbales permitió identificar problemas de sesgo potencial en ciertos reactivos que, de no haberse realizado este análisis cognitivo, difícilmente habrían conseguido tener suficiente evidencia de validez, reconociendo que los análisis DIF aportan elementos valiosos.

Las evidencias de validez obtenidas mediante los reportes y protocolos verbales han hecho patente la necesidad de que los diseñadores de pruebas formen equipos específicos para tratar de minimizar fuentes de sesgo de subpoblaciones que originalmente no se había considerado que pudieran tener un funcionamiento diferencial, como el caso de alumnos con cierto NSE. En el contexto mexicano, ante la escasez de trabajos de investigación que tengan por objetivo obtener evidencias de validez sobre los procesos de respuesta, este texto ofrece un ejemplo que puede utilizarse como guía para futuras investigaciones.

El proceso para utilizar las técnicas de pensamiento en voz alta requiere de sistematicidad y rigurosidad en cada una de sus fases y actividades. Es de vital importancia que, en el diseño del estudio cognitivo, su aplicación y el análisis de la información no pierdan de vista el propósito con el que se utilizan. El procedimiento descrito en este capítulo lo podrían seguir quienes estén interesados en obtener evidencias de validez sobre los procesos de respuesta, con propósitos específicos para identificar fuentes de sesgo para ciertas subpoblaciones. Si se desea emplear estas técnicas con propósitos diferentes deberán hacer ajustes tanto en su diseño como en los instrumentos para la recolección de datos.

Referencias

American Educational Research Association, American Psychological Association y National Council on Measurement in Education (AERA/APA/NCME). (1999). *Standards for educational and psychological testing.* Washington: AERA/APA/NCME.

American Educational Research Association, American Psychological Association y National Council on Measurement in Education (AERA/APA/NCME). (2014). *Standards for educational and psychological testing.* Washington: AERA/APA/NCME.

Benítez, I., & Padilla, J.L. (2014). Analysis of Nonequivalent Assessments Across Different Linguistic Groups Using a Mixed Methods Approach Understanding the Causes of Differential Item Functioning by Cognitive Interviewing. *Journal of Mixed Methods Research, 8*(1), 52–68. http://doi.org/10.1177/1558689813488245

Castillo-Díaz, M., & Padilla, J.L. (2013). How Cognitive Interviewing Can Provide Validity Evidence of the Response Processes to Scale Items. *Social Indicators Research, 114*(3), 963–975. http://doi.org/10.1007/s11205-012-0184-8

Ericsson, K.A., & Simon, H.A. (1993). *Protocol analysis. Verbal Reports as Data*. Cambridge: MIT Press. Recuperado de https://mitpress.mit.edu/books/protocol-analysis-revised-edition

García-Medina, A. (2017). *Equidad y análisis de sesgo en las evaluaciones. Los Excale de Matemáticas para tercero de secundaria*. Tesis doctoral. Instituto de Investigación y Desarrollo Educativo de la UABC, Ensenada, Baja California, México.

García-Medina, A., Martínez, F., & Cordero, G. (2016). Análisis del funcionamiento diferencial de los ítems del Excale de Matemáticas para tercero de secundaria. *Revista Mexicana de Investigación Educativa, 21*(71), 1191-1220.

Instituto Nacional para la Evaluación de la Educación (INEE). (2010b). *Especificaciones de los reactivos del Excale-09.* Documento interno. INEE.

Instituto Nacional para la Evaluación de la Educación (INEE). (2010a). *Excale 09. Especificación de reactivos M0910ABA007.* Documento interno.

Leighton, J.P. (2009). Two Types of Think Aloud Interviews for Educational Measurement. Ponencia. *Annual meeting of the National Council on Measurement in Education,* San Diego, CA. Recuperado de https://pdfs.semanticscholar.org/bc15/e348e8984540d7918bd284e6a959f4b2528b.pdf

Miller, K., Willson, S., Chepp, V., & Padilla, J.L. (Eds.). (2014). *Cognitive interviewing methodology*. Hoboken, New Jersey: Wiley.

Padilla, J.L., & Benítez, I. (2014). Validity evidence based on response processes. *Psicothema, 26*(1), 136-144. Recuperado de http://www.psicothema.com/pdf/4171.pdf

Perez-Morán, J.C. (2014). *Análisis del aspecto sustantivo de la validez de constructo de una prueba de habilidades cuantitativas.* Tesis doctoral. Instituto de Investigación y Desarrollo Educativo de la UABC, Ensenada, Baja California, México.

Schuman, H. (2001). Sample Surveys: Survey Design Issues and Strategies. En N. Smelser y P. Baltes (Ed.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 13470-13475). Amsterdam: Elsevier.

Schwarz, N. (2001). Questionnaires: Cognitive Approaches. En N. Smelser y P. Baltes (Ed.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 12667-12673). Amsterdam: Elsevier.

Tanur, J.M. (2001). Sample Surveys: Cognitive Aspects of Survey Design. En N. Smelser y P. Baltes (Ed.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 13448-13453). Amsterdam: Elsevier.

Thompson, S., Johnstone, C., & Thurlow, M.L. (2002). *Universal design applied to large scale assessments* (Synthesis Report 44). Minneapolis, MN: Universuty of Minnesota, National Center on Educational Outcomes. Recuperado de https://nceo.umn.edu/docs/OnlinePubs/Synth44.pdf

Willis, G.B. (2005). *Cognitive interviewing: a tool for improving questionnaire design.* Thousand Oaks: Sage.

Willis, G.B. (2015). *Analysis of the cognitive interview in questionnaire design.* Oxford: Oxford University Press.

Capítulo 5

Aproximaciones metodológicas para la adaptación de test en contextos educativos

Carlos David Díaz López, Joaquín Caso Niebla y José Alfonso Jiménez Moreno

Introducción

Cada vez es más común la aplicación e intercambio de instrumentos de evaluación psicológica o educativa de un país a otro y la utilización de exámenes traducidos a diversos idiomas, lo que ha exigido adaptar pruebas de rendimiento y aptitudes para su uso en múltiples culturas e idiomas (Hambleton, 2001; Klerk, 2008). La adaptación de instrumentos para uso en contextos lingüísticos y culturales diferentes a aquellos en donde fueron construidos, es una práctica que se remonta a los inicios del siglo XX (Muñiz, Elosua y Hambleton, 2013). Actualmente, la utilización de test adaptados se ha incrementado debido a la creciente globalización y a la facilidad y rapidez de uso de los medios de comunicación, pues aumenta la posibilidad de que los instrumentos generados en un país puedan utilizarse con prontitud en uno distinto rápidamente (Muñiz y Hambleton, 1996).

Lo anterior se ve reflejado en el creciente interés por los estudios comparativos internacionales y en investigaciones interculturales, pues en la mayoría de estos se utilizan instrumentos y otros métodos de recolección de información que requieren adaptarse para su uso en otras culturas (Byrne et al., 2009; Hambleton, Mereda y Spielberger, 2005). Tal es el caso del Programa Internacional de Evaluación de los Alumnos (PISA, por sus siglas en inglés), en el cual participan un gran número de países con idiomas, culturas y sistemas educativos diferentes. PISA evalúa las competencias en lectura, matemáticas y ciencia en estudiantes de 15 años de edad, a través de pruebas, encuestas, cuestionarios y manuales traducidos y adaptados en cada uno de los idiomas de instrucción de los países participantes (Organization for Economic Co-operation and Development [OECD],

2010; 2016). PISA, por la importancia y relevancia de sus resultados en el contexto internacional, probablemente represente el examen al que más recursos técnicos y presupuestales se le canalizan, lo que en cierta medida favorece la calidad técnica de sus resultados.

Así, la comunidad científica dedicada a la medición es cada vez más consciente de la necesidad de perfeccionar los métodos para traducir y adaptar estas pruebas. En lo particular, Solano-Flores, Backhoff y Contreras-Niño (2009) señalaron la importancia de comprender cómo estos instrumentos han sido desarrollados en un idioma, país y cultura, a fin de que puedan utilizarse de manera apropiada en otro contexto. No obstante, es común que los investigadores que utilizan test de otros países no se encuentren familiarizados con los métodos psicométricos necesarios para llevar a cabo un adecuado proceso de adaptación (Byrne et al., 2009).

La decisión de adaptar un test de una cultura a otra, supone que: (a) el constructo medido por el test guiará hacia las mismas, o similares, interpretaciones y predicciones en la nueva cultura; y (b) que la población donde se adaptará el test difiere de la población original en términos de cultura o idioma (Geisinger, 1994). De este modo, el proceso de adaptación de pruebas permite la obtención de mediciones válidas en cada contexto o cultura y facilita la realización de estudios comparativos entre los grupos involucrados, además, representa una reducción considerable del costo, pues resulta más oneroso crear un nuevo test que adaptar uno existente (Klerk, 2008). Para realizar una adaptación se deben considerar aspectos psicológicos, psicométricos, culturales, interculturales y lingüísticos, así como el conocimiento de todas las idiosincrasias específicas de un idioma o cultura que pueda causar sesgo en la evaluación (American Education Research Association, American Psychological Association, y National Council on Measurement in Education [AERA, APA y NCME], 2014).

El principal objetivo de la adaptación de un test es asegurar la validez de constructo, de contenido y ecológica, lo que proporciona sustento a las inferencias que se realizan en poblaciones específicas, independientemente de su contexto lingüístico y cultural. Cuando se utiliza un test en una cultura distinta a la cultura donde fue desarrollado y estandarizado, se asume que las características psicológicas, o al menos las raíces de estas características, son universales para todos los grupos involucrados (AERA, APA y NCME, 2014; Klerk, 2008). Así, para cuando se utiliza test en una población o cultura distinta a donde éste fue desarrollado se asume que: (a) el dominio conductual –habilidad o rasgo– medido por los ítems tiene, aproximadamente, el mismo significado en cada cultura; (b) la habilidad o el rasgo medido juega el mismo rol en la organización de la

conducta en cada cultura; (c) la puntuación tiene el mismo significado para quienes responden el test independiente de su contexto cultural; y (d) para referir la comparabilidad entre las puntuaciones del test de diferentes grupos culturales, es necesario valorar los términos sesgo y equivalencia de los puntajes (Klerk, 2008).

Para atender los supuestos anteriores es necesario realizar modificaciones y adecuar el instrumento al contexto en el cual se aplicará con el fin a establecer equivalencias y realizar comparaciones entre las poblaciones implicadas. Actualmente, se reconocen tres principales enfoques en el proceso de adaptación: (a) adopción o traducción; (b) adaptación; y (c) ensamblaje (He y van de Vijver, 2012; AERA, APA y NCME, 2014).

El primero de esos enfoques, llamado adopción o traducción, equivale a la traducción de un instrumento al idioma objetivo. Supone que el constructo subyacente que mide el instrumento es adecuado para cada grupo cultural y que la traducción directa es suficiente para obtener un instrumento que mide el mismo constructo en el grupo objetivo. Este enfoque es el más utilizado en la investigación empírica pues resulta simple de implementar, es económico y brinda la oportunidad de comparar las puntuaciones entre las poblaciones. Sin embargo, solo se puede utilizar cuando los ítems, en ambos idiomas o grupos culturales, tienen una cobertura adecuada del constructo medido en ambas poblaciones. El enfoque de adopción se debe considerar sí, y solo sí, las características del instrumento y del constructo son pertinentes en dichos grupos culturales (He y van de Vijver, 2012; van de Vijver y Tanzer, 2004). A este respecto, la APA (2013) ha cuestionado la contundencia y suficiencia del enfoque de traducción. Para dicho organismo la adaptación representa el enfoque ideal para la transformación de escalas de medición.

El enfoque de adaptación tiene como base la noción de que una adopción producirá instrumentos sesgados ya sea por razones lingüísticas, culturales o psicométricas. El proceso de adaptación requiere tanto de habilidades lingüísticas como del conocimiento de las culturas e idiomas involucrados para poder evaluar la relevancia y pertinencia del instrumento en otro contexto. Esta alternativa busca maximizar la equivalencia y posibilita la realización de comparaciones entre los grupos considerados (He y van de Vijver, 2012; van de Vijver y Tanzer, 2004).

El tercer enfoque, denominado ensamblaje, equivale a la compilación de un nuevo instrumento. Este enfoque maximiza la pertinencia cultural del instrumento, pues integra aspectos del constructo que son importantes para la cultura objetivo que no están siendo considerados por el instrumento original. Sin embargo, no permite la comparación de puntuaciones entre las culturas,

pues no garantiza la equivalencia de constructo ni de medición. Es una opción que emerge cuando la traducción o la adaptación de un instrumento no producen un test con la precisión lingüística, cultural y psicométrica que se requiere (He y van de Vijver, 2012; van de Vijver y Tanzer, 2004).

El término *adaptación de test* reemplazó al término *traducción de test* para hacer referencia al proceso de preparación de un test para su uso en otro idioma o cultura (Geisinger, 1994; Hambleton, 2001). La adaptación de test es un proceso más amplio y exhaustivo que refleja la manera más precisa de transformar y producir un test que sea igualmente válido en dos o más idiomas o contextos culturales (APA, 2013; Hambleton, 2001). A diferencia de la traducción de test, la adaptación estudia si la nueva versión del instrumento mide o no el mismo constructo, considerando aspectos del proceso de evaluación tales como la equivalencia de los procedimientos de aplicación y el formato utilizado en los ítems (Klerk, 2008). De acuerdo con Muñiz et al. (2013), la adaptación de un test no se trata meramente de una cuestión lingüística, exige la conjunción de aspectos culturales, conceptuales y métricos que han de acometerse desde perspectivas de análisis analítico-racionales y empíricas. El proceso de adaptación proporciona el medio más sistemático y exhaustivo para transformar con éxito una escala de medida. De acuerdo con la APA (2013), este proceso se apoya en tres fases enfocadas en asegurar la validez e interpretación de las puntuaciones en la población objetivo: traducción, validación y establecimiento de normas.

La traducción, validado previamente en una cultura e idioma específico, ha sido el procedimiento más comúnmente utilizado en la adaptación de un test (Casagrande, 1954). Así, la tarea del traductor es decodificar un mensaje presentado en un código (código A), y codificar el mensaje en un segundo código (código B), de tal forma que los dos mensajes sean aproximadamente equivalentes. La equivalencia perfecta, en el sentido de que los mensajes evoquen respuestas idénticas en los locutores de dos idiomas, es probablemente imposible de alcanzar (Casagrande, 1954).

Existen dos aproximaciones principales para la traducción de test: la traducción directa y la traducción inversa. En la traducción directa, un solo traductor o un grupo de traductores decodifican el idioma del instrumento original y lo codifican en el idioma de uso dentro de la nueva cultura. Otro grupo de traductores evalúa la equivalencia de estas dos versiones del instrumento y, si existen discrepancias, se realizan los cambios en la versión objetivo (APA, 2013; Byrne et al., 2009).

Por su parte, el enfoque de traducción inversa inicia con la traducción directa de la versión original a otro idioma y un traductor diferente traduce de nuevo la versión objetivo en el idioma original. Posteriormente, las versiones originales y traducidas se comparan por un grupo independiente de traductores, a fin de evaluar su equivalencia e identificar cualquier discrepancia entre las

versiones (APA, 2013; Byrne et al., 2009; Hambleton, 2005; Instituto Nacional para la Evaluación de la Educación [INEE], 2011).

Como ya se estableció, la traducción de instrumentos de medición para su uso en otra cultura o grupo lingüístico es más compleja que la simple traducción del texto en otro idioma debido, principalmente, a las fuentes de sesgo que pueden amenazar su precisión y grado de adecuación. En la traducción de un test, se deben considerar los significados dependientes de la cultura o las connotaciones de las palabras en cada contexto (International Test Commision, 2017; Klerk, 2008). Las diferencias entre los idiomas juegan un papel importante en la conformación del significado del entorno en que vivimos. Así, la presencia de factores lingüísticos específicos del contexto y la falta de correspondencia perfecta entre los idiomas, hace necesario evaluar la ejecución de los test con el fin de asegurar la correcta interpretación de las puntuaciones obtenidas (Elosua y López, 1999; INEE, 2011). Dado que una traducción lingüística precisa de los ítems es condición necesaria pero no suficiente para generar test equivalentes, se requiere de un proceso balanceado de consideraciones psicológicas, lingüísticas y culturales, que posteriormente permita realizar comparaciones interculturales válidas (Hambleton, 2001; van de Vijver y Tanzer, 2004).

En la actualidad, se reconoce que en general los procedimientos que se utilizan en los enfoques basados en la traducción no son suficientemente rigurosos para asegurar la validez de un instrumento (Byrne et al., 2009). En lo particular, tanto el enfoque de traducción directa como el de traducción inversa, presentan las siguientes limitaciones: (a) no consideran las pruebas de campo, lo cual reduce la posibilidad de identificar ítems defectuosos que puedan reducir la confiabilidad y validez del test (Hambleton, 2005); (b) no tratan con la comparabilidad del constructo subyacente que está siendo medido, por lo que solo pueden asumir que el constructo en ambas poblaciones es equivalente en cuanto al número de dimensiones consideradas o la cantidad y tipo de conductas consideradas en la medición del mismo (van de Vijver y Hambleton, 1996); y (c) no realizan análisis de la confiabilidad del instrumento traducido y la validez de las puntuaciones en la población objetivo (Mereda, 2005).

Recientemente, se han desarrollado enfoques complementarios que dan respuesta a las debilidades de la traducción directa y la traducción inversa. A este respecto, van de Vijver y Tanzer (2004) recomendaron el empleo de un enfoque de comité, donde un grupo de personas con diferentes tipos de especialización, evalúan la calidad de la traducción. A este comité se suma un segundo grupo de expertos quienes determinan si la traducción ha alterado el constructo que se desea medir con el test (INEE, 2011).

Lo anterior, sugiere que los responsables de conducir procesos de adaptación deben establecer y gestionar con éxito equipos con áreas de especialización en el constructo a evaluar, con especialistas en técnicas de construcción de test y psicometría, y con sensibilidad a las especificidades culturales, lingüísticas y contextuales de su propia cultura e idioma. Además, de acuerdo con Byrne et al. (2009), estos comités también deben incorporar, en medida de lo posible, a los autores originales del test.

En resumen, la traducción lingüística no garantiza la equivalencia métrica. Resulta indispensable realizar análisis del sesgo a través de procedimientos empíricos y analítico-racionales, que permitan confirmar la igualdad de significado de las puntuaciones obtenidas a través de la versión original y de la versión adaptada (Elosua y López, 1999; Muñiz et al., 2013).

Procedimientos de validación en la adaptación de un test

En el proceso de validación de un test que pretende adaptarse se identifican tres procedimientos empíricos independientes y lógicamente ordenados: (a) el estudio piloto y la prueba de campo, (b) la validación del constructo dentro del contexto objetivo, y (c) la validación del constructo entre el contexto original y objetivo (APA, 2013).

En el caso del estudio piloto y de la prueba de campo, estos se realizan una vez obtenida la versión traducida del test, con el propósito de recolectar información que permita conocer el funcionamiento del test y de sus ítems, de identificar los ítems que no están midiendo el constructo de la manera esperada (Geisinger, 1994), y de conocer el efecto de factores extraños en las puntuaciones de los ítems (APA, 2013). De esta manera, el estudio piloto considera la aplicación del test a una muestra pequeña de sujetos a manera de revisión inicial del funcionamiento del instrumento.

Dentro del estudio piloto se puede utilizar la técnica de entrevista cognitiva, para asegurar que el contenido de cada ítem del test tenga el mismo sentido para todos los sujetos, sea aplicable al contexto de vida de los sujetos, y su significado sea interpretado y comprendido de la manera esperada (Miller, Mont, Maitland, Altman y Madans, 2010). En este sentido, la técnica *pensar en voz alta*, permite conocer los procesos cognitivos involucrados en la respuesta al ítem (van de Vijver y Tanzer, 2004). Una vez identificados los ítems defectuosos, así como los factores que amenazan la validez de las puntuaciones del instrumento, se deben modificar y corregir para posteriormente realizar la prueba de campo, donde el test se aplica a una muestra más amplia y semejante a aquella con la que fue validado originalmente (APA, 2013).

Validación del constructo en la población objetivo

El test adaptado requiere de un proceso de validación de constructo para asegurar que el mismo constructo está siendo medido de la misma manera que en el caso del test original. La validación de constructo analiza la estructura factorial del instrumento adaptado para evaluar la estructura hipotética con respecto al número de factores subyacentes, las cargas factoriales de los ítems asignados a cada dimensión y a la relación entre los factores. De esta manera, en la medida en que el número de factores subyacentes y las cargas factoriales sean iguales entre la muestra original y la muestra objetivo, ello constituye una evidencia de validez de constructo (APA, 2013).

El análisis factorial confirmatorio (AFC), es, hasta el momento, el método más apropiado y estadísticamente más riguroso para obtener evidencias de validez de constructo de un test. Considera la estructura factorial postulada por el instrumento original como la estructura hipotética y la contrasta con la estructura observada en la muestra objetivo. Los resultados son descritos en términos de bondad de ajuste entre la estructura hipotética del instrumento y el conjunto de datos obtenidos en la muestra (puntuaciones de las respuestas a los ítems), donde altos índices de bondad de ajuste sugieren una fuente confiable de evidencia de validez de constructo del instrumento adaptado (APA, 2013).

En contraparte, cuando se obtienen niveles bajos de bondad de ajuste puede deberse a: (a) que los ítems de un test pertenezcan a más de un factor subyacente, (b) que el contenido de un ítem se sobrepone con el de otro ítem, o (c) que las correlaciones entre los factores subyacentes son inconsistentes con la estructura hipotética (APA, 2013).

Validación del constructo entre la población original y la población objetivo

Una vez que se hayan obtenido evidencias suficientes que den cuenta de la validez del constructo del instrumento adaptado en la población objetivo, se debe determinar el grado en el cual la estructura factorial y las puntuaciones del instrumento adaptado son equivalentes con las obtenidas por la versión original. Tanto el AFC como el análisis del funcionamiento diferencial de ítems (DIF, por sus siglas en inglés) son enfoques empleados para detectar evidencias de sesgo en un ítem y para estudiar la equivalencia del instrumento adaptado con relación al instrumento original (APA, 2013; Geisinger, 1994; He y van de Vijver, 2012). Ambos procedimientos se comentan a continuación.

Análisis factorial confirmatorio

El empleo del AFC para el estudio de la equivalencia de instrumentos de medición tiene como objetivo determinar el grado en el cual los ítems están operando de manera similar en las poblaciones consideradas, y estudiar la estructura subyacente del instrumento (APA, 2013). De manera específica, se busca conocer si: (a) el contenido del ítem es interpretado y percibido exactamente de la misma manera; (b) los ítems miden el constructo subyacente en el mismo grado; (c) el formato de los ítems opera de la misma manera; (d) el constructo medido por cada dimensión (del test) tiene el mismo significado; (e) el constructo tiene la misma estructura dimensional; y si (f) las relaciones entre las dimensiones son las mismas.

Si el modelo del AFC muestra un ajuste aceptable (tabla 1), significa que la estructura factorial asumida no puede rechazarse y que pueden establecerse diferentes niveles de equivalencia (He y Van de Vijver, 2012).

Tabla 1. *Clasificación de los índices de ajuste en el análisis factorial confirmatorio*

Inferencial / Descriptivo	Absoluto / comparativo	Índice de ajuste	Criterios de interpretación
Interferencial	Ajuste absoluto	Estadístico X^2	Se mantiene el modelo si $p \geq 0.05$; no se recomienda.
	Ajuste comparativo	Comparación de modelos anidados (X^2)	Los modelos no difieren si $p \geq 0.05$
Descriptivo	Ajuste absoluto	X^2/gl	Ratios menores que 2 o 3 suelen considerarse indicadores de ajuste aceptable; no se recomienda.
		RMR	Cuanto más pequeño, mejor ajuste; 0 indica ajuste perfecto; no se recomienda porque es difícil de interpretar.
		SRMR	≤ 0.08, el modelo se ajusta.
		RMSEA	≤ 0.06, el modelo se ajusta.
		GFI/AGFI	≥ 0.95, el modelo se ajusta; no se recomienda.
	Ajuste comparativo	NFI	≥ 0.95, el modelo se ajusta. Tiende a sobrestimarse en muestras pequeñas
		NNFI o TLI	≥ 0.95 (para variables continuas o ≥ 0.096 (para variables categóricas): el modelo se ajusta.
		NNFI o TLI	≥ 0.95 (para variables continuas o ≥ 0.096 (para variables categóricas): el modelo se ajusta.

Nota: Adaptado de Medición de ciencias sociales y de la salud por F. Abad, J. Olea, V. Ponsoda y C. García, 2011.

Para presentar evidencia de los diferentes niveles de equivalencia (equivalencia de constructo, equivalencia de medición y equivalencia de puntuaciones), se puede implementar un modelo de AFC multigrupo para contrastar las hipótesis respectivas. He y van de Vijver (2012) propusieron cinco modelos que brindan indicadores importantes sobre la equivalencia, su operacionalización e interpretación (tabla 2).

Tabla 2. *Modelos anidados en Análisis Factorial confirmatorio multigrupo*

Modelos	Operacionalización	Interpretación del nivel de equivalencia
Invarianza de la configuración	Mismo patrón de constructos observables y latentes	El mismo constructo latente es medido utilizando los mismos indicadores (no equivalencia métrica)
Pesos de medición	Las cargas factoriales en las partes de medición de cada grupo cultural son idénticas	Los mismos factores latentes son medidos entre los grupos, indicando equivalencia de constructo y equivalencia métrica
Invarianza de intercepto	Ítems tienen el mismo intercepto (media latente) entre las culturas	Todos los grupos representan la misma diferencia entre grupos, indicando estar libre de sesgo de ítem y una equivalencia completa en la puntuación
Residuales estructurales	La varianza de error del factor latente es idéntica	El rango de puntuaciones en los factores latentes no varía entre las culturas, lo cual indica equivalencia completa de puntuación
Residuales de medición	La varianza de error de los ítems observados es idéntica	Los grupos utilizan el mismo rango del continuo del constructo, lo cual indica equivalencia completa de puntuación

Nota: Adaptado de "Bias and Equivalence in Cross-Cultural Research" por J. He y F. van de Vijver, 2012, *Online Readings in Psychology and Culture, 2*(2).

Dentro de esta técnica, los estudios de validez analizan la invarianza de los parámetros calculados a través de los grupos. Es decir, se busca comprobar que las puntuaciones, en los grupos considerados, tienen el mismo significado o están libres de sesgo (Abad et al., 2011). De acuerdo con Geisinger (1994), cuando se identifica que los instrumentos adaptados miden las mismas dimensiones entre los grupos culturales o lingüísticos, pero las cargas de los ítems difieren, es un indicador de que efectivamente dichas dimensiones están presentes en ambas poblaciones, pero que las dimensiones difieren interculturalmente de alguna manera.

Además de la utilización del AFC, distintos autores recomiendan la implementación de otras técnicas que ayudan en el estudio de validez y de la equivalencia entre los instrumentos considerados (Byrne et al., 2009; Geisinger, 1994; van de Vijver y Watkins, 2006). Así, el modelamiento de ecuaciones estructurales (SEM, por sus siglas en inglés) y las técnicas factoriales analíticas de máxima verosimilitud permiten desarrollar pruebas de hipótesis sobre si el instrumento de medición mide las mismas cualidades, en las mismas proporciones y con la misma estructura factorial que el instrumento original. Como parte de otros diseños de validación de constructo, resulta conveniente utilizar la matriz multi-rasgo multi-método (MTMM), el estudio del cambio ante manipulaciones experimentales, la implementación del análisis de conglomerados y el escalamiento multidimensional.

Análisis del funcionamiento diferencial de ítems

De acuerdo a Elosua y López (1999), dentro de cualquier proceso de adaptación es imprescindible realizar un análisis detallado de las relaciones entre cada uno de los ítems y el rasgo o habilidad que pretenden medir. Un test bien construido (o adaptado) no debería penalizar a cierto grupo de personas por razones ajenas al atributo que se mide. Por lo tanto, dentro del proceso de validación de cualquier instrumento de medición, debe comprobarse que las puntuaciones que reflejan distintos niveles del rasgo que se pretender medir, no están determinadas por variables ajenas al constructo bajo estudio (Abad et al., 2011).

Para el análisis de sesgo, resulta relevante comprender la diferencia entre los términos: i*mpacto del ítem, funcionamiento diferencial del ítem* y *sesgo de ítem*. En este sentido, el impacto del ítem es evidente cuando los examinados de grupos diferentes presentan diferentes probabilidades de adscribirse (o responder correctamente) a un ítem, debido a que existen diferencias verdaderas entre los grupos en el rasgo subyacente medido por el ítem. Por su parte, el DIF tiene lugar cuando los examinados de distintos grupos presentan diferentes probabilidades de adscribirse (o responder correctamente) a un ítem, después de equipararse en función del nivel de rasgo que se pretende medir. En tanto, el sesgo de ítem ocurre cuando los examinados de un grupo tienen menos probabilidades de adscribirse (o responder correctamente) a un ítem a causa de algunas características del o de las condiciones de la aplicación, que no sea relevante para los propósitos del test (Zumbo, 1999).

El DIF es una condición necesaria pero no suficiente para identificar sesgo en los ítems, por lo que la presencia de DIF no es suficiente para señalar sesgo en un ítem y se deben implementar estudios de sesgo posteriores (análisis de contenido, entrevistas cognitivas, etc.) para poder determinar la presencia de sesgo (Zumbo, 1999).

Existen distintos procedimientos y análisis estadísticos para evaluar el DIF (Delta, chi-cuadrada, Mantel-Haenszel, comparación de parámetros de las curvas característica de los ítems, modelos log-lineales y regresión logística) (Elosua y López, 1999; Elosua, 2003; He y van de Vijver, 2012). En lo general, estos ayudan a identificar aquellos ítems que presentan DIF, mientras que otros permiten conocer si el funcionamiento diferencial es de tipo uniforme (cuando el efecto es constante) o no uniforme (si el efecto varía en función del nivel de rasgo) (figura 1). Algunos de estos procedimientos emplean el análisis de regresión logística (Zumbo, 1999), permitiendo conocer la magnitud del funcionamiento diferencial y valorar si ésta es insignificante, moderada o severa. La utilidad de estos procedimientos ha quedado demostrada al permitir determinar la medida en que los ítems son equivalentes entre los grupos de sujetos bajo estudio (APA, 2013; Klerk, 2008).

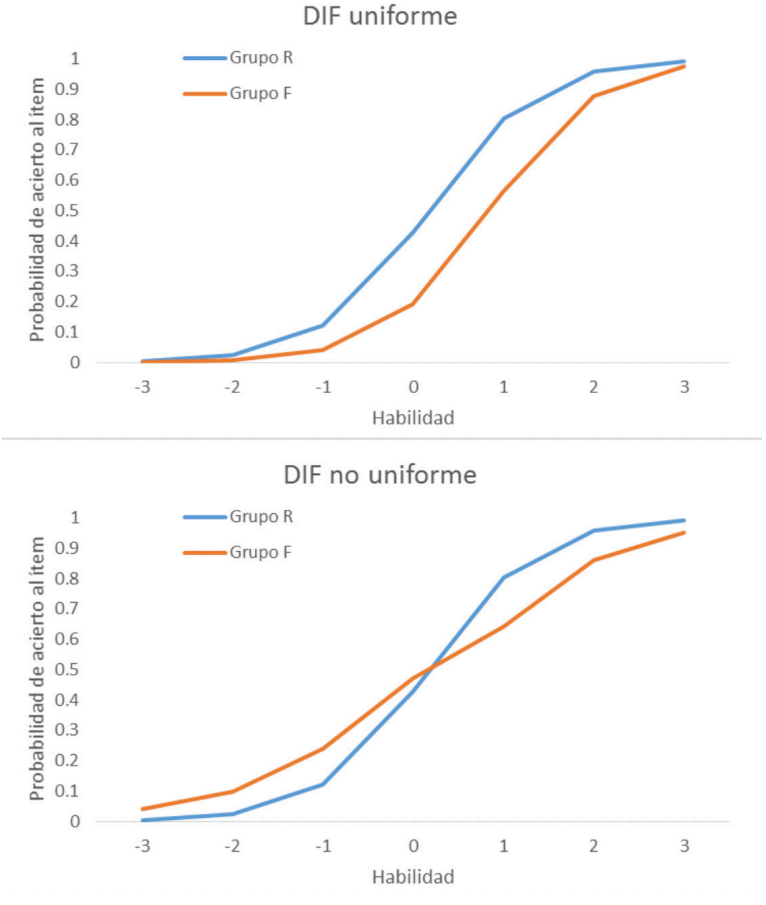

Figura 1. Ejemplo de DF uniforme y DIF no uniforme. *Fuente:* Elaboración propia.

La fase final del proceso de adaptación se refiere al establecimiento de normas para el instrumento adaptado. De acuerdo con Geinsinger (1994), las normas proporcionan información acerca del lugar que ocupa la puntuación de un individuo dentro de la población de sujetos que respondieron al mismo instrumento; al mismo tiempo que facilita la interpretación de las puntuaciones y ubicar, clasificar o diagnosticar a los sujetos bajo estudio.

Conclusiones

La calidad métrica de los instrumentos es una condición necesaria pero no suficiente para alcanzar resultados rigurosos, ya que el uso inapropiado de un instrumento puede comprometer la validez y confiabilidad del mismo (Evers et al., 2013). En este sentido, los instrumentos de medición deben contar con manuales de uso e interpretación que informen sobre el conjunto de evidencias que garantizan que se puedan realizar inferencias de manera válida y confiable (Muñiz y Fernández-Hermida, 2010).

Así, podría afirmarse que un test vale tanto como su manual (Muñiz y Fernandes-Harmida, 2010), pues en estos debe documentar el procedimiento llevado a cabo durante el desarrollo del test, las evidencias de validez y confiabilidad del instrumento, el procedimiento de aplicación del test, de interpretación (normativa o criterial) de las puntuaciones, así como cualquier tarea asociada con la transformación de puntuaciones y el establecimiento de criterios de selección (Byrne et al., 2009; Elosua, 2012; International Test Commission, 2010; 2017).

Por lo tanto, cuando se adapta un instrumento en una nueva población resulta inapropiado interpretar las puntuaciones a la luz de la información normativa asociada al instrumento original. Así, la adaptación de instrumentos de medición debe considerar la recolección de información necesaria que permita establecer las normas para valorar el desempeño de la población objetivo en el nuevo instrumento (APA, 2013; Geinsinger, 1994). Del mismo modo, el manual del test adaptado requiere contener más documentación que el original, pues deberá incluir: (a) toda la información relativa al proceso de traducción y adaptación que se llevó a cabo (Muñiz y Fernandez-Harmida, 2010); (b) información detallada de todos los cambios realizados acompañados de una justificación de tipo lingüística o psicométrica; (c) evidencias de equivalencia de constructo y de medición dentro de cada población; y (d) evidencias de equivalencia de puntuaciones entre las poblaciones (Van de Vijver y Tanzer, 2004).

De esta manera, se observa que la adaptación de los test no es un trabajo sencillo. Requiere de una labor técnica relevante que permita asegurar que las puntuaciones que se desarrollen, pero, más aún, las interpretaciones de las mismas, se fundamenten en evidencia de validez dentro de la población objetivo. Frente a esta necesidad, valdría no solo considerar la posibilidad de desarrollar métodos para las necesidades de adaptación en el país, sino incluso hacer uso de este tipo de ejercicios como medio de valoración de la validez de los instrumentos que actualmente se utilizan en la práctica educativa.

Referencias

Abad, F. J., Olea, J., Ponsoda, V., & García, C. (2011). *Medición de ciencias sociales y de la salud.* Madrid: Síntesis.

American Education Research Association, American Psychological Association & National Council on Measurement in Education (AERA/APA/NCME). (2014). *Standards for Educational and Psychological Testing 2014.* Washington, DC: American Education Research Association.

American Psychological Association (APA). (2013). *Going International: A Practical Guide for Psychologists. Using Translated/Adapted Measurement Scales Internationally.* Nueva York: APA.

Byrne, B.M., Oakland, T., Leong, F.T.L., van de Vijver, F., Hambleton, R.K., Cheung, F.M., & Bartram, D. (2009). A critical analysis of cross-cultural research and testing practices: Implications for improved education and training in psychology. *Training and Education in Professional Psychology, 3*(2), 94–105. http://doi.org/10.1037/a0014516

Casagrande, J. (1954). The Ends of Translation. *International Journal of American Linguistics, 20*(4), 335-340.

Elosua, P. (2003). Sobre la validez de los tests. *Psicothema, 15*(2), 315-321.

Elosua, P. (2012). Test publicados en España: usos, costumbres y asignaturas pendientes. *Papeles Del Psicólogo, 33*(1), 12-21.

Elosua, P., & López, A. (1999). Funcionamiento diferencial de los ítems y sesgo en la adaptación de dos pruebas verbales. *Psicológica, 20*(1), 23-40.

Evers, A., Muñiz, J., Hagemeister, C., HstmæLingen, A., Lindley, P., Sjöberg, A., & Bartram, D. (2013). Assessing the quality of tests: revision of the EFPA review model. *Psicothema, 25*(3), 283-291.

Geisinger, K. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment, 6*(4), 304-312.

Hambleton, R.K. (2001). The next generation of the ITC test translation and adaptation guidelines. *European Journal of Psychological Assessment, 17*(3), 164-172.

Hambleton, R.K. (2005). Issues, Designs, and Technical Guidelines for Adapting Tests Into Multiple Languages and Cultures. En R.K. Hambleton, P. Mereda y C. Spielberger (2005), *Adapting Educational and Psychological Tests for Cross-Cultural Assessment.* Mahwah, New Jersey: Lawrence Erlbaum Associates.

Hambleton, R.K., Mereda, P., & Spielberger, C. (2005). *Adapting Educational and Psychological Tests for Cross-Cultural Assessment.* New Jersey, London: Lawrence Erlbaum Associates.

He, J., & van de Vijver, F. (2012). Bias and Equivalence in Cross-Cultural Research. *Online Readings in Psychology and Culture, 2*(2). http://doi.org/10.9707/2307-0919.1111

Instituto Nacional para la Evaluación de la Educación (INEE). (2011). *La Teoría del Error de Traducción de Pruebas y las evaluaciones internacionales de TIMSS y PISA.* México, D.F.: INEE.

International Test Commission. (2010). *International Test Commission Guidelines for Translating and Adapting Tests.* Recuperado de http://www.intestcom.org

International Test Commission. (2017). *The ITC Guidelines for Translating and Adapting Tests.* Recuperado de http://www.intestcom.org

Klerk, G. (2008). Cross-cultural testing. En M. Born, C.D. Foxcroft & R. Butter (eds.), *Online Readings in Testing and Assessment.* International Test Commission.

Mereda, P. (2005). Cross-Cultural Adaptation of Educational and Psychological Testing. En R.K. Hambleton, P. Mereda & C. Spielberger (2005). *Adapting Educational and Psychological Tests for Cross-Cultural Assessment.* Mahwah, New Jersey: Lawrence Erlbaum Associates.

Miller, K., Mont, D., Maitland, A., Altman, B., & Madans, J. (2010). Results of a cross-national structured cognitive interviewing protocol to test measures of disability. *Quality & Quantity, 45*(4), 801–815. http://doi.org/10.1007/s11135-010-9370-4

Muñiz, J., Elosua, P., & Hambleton, R.K. (2013). Directrices para la traducción y adaptación de los tests: segunda edición. *Psicothema, 25*(2), 151–157.

Muñiz, J., & Fernández-Hermida, J.R. (2010). La opinión de los psicólogos españoles sobre el uso de los test. *Papeles Del Psicólogo, 31*(1), 108–121.

Muñiz, J., & Hambleton, R.K. (1996). Directrices para la traducción y adaptación de los test. *Papeles Del Psicólogo,* (66), 63–70.

Organization for Economic Co-operation and Development (OECD). (2010). *Translation and adaptation guidelines for PISA 2012.* Recuperado de http://www.oecd.org/pisa/pisaproducts/49273486.pdf

Organization for Economic Co-operation and Development (OECD). (2016). *PISA 2018. Draft analytical frameworks.* Recuperado de http://www.oecd.org/pisa/pisaproducts/PISA-2018-draft-frameworks.pdf

Solano-Flores, G., Backhoff, E., & Contreras-Niño, L.Á. (2009). Theory of Test Translation Error. *International Journal of Testing, 9*(2), 78–91. http://doi.org/10.1080/15305050902880835

Van de Vijver, F., & Hambleton, R. (1996). Translating tests: Some practical guidelines. *European Psychologist, 1*(2), 89–99.

Van de Vijver, F., & Tanzer, N.K. (2004). Bias and equivalence in cross-cultural assessment: an overview. *European Review of Applied Psychology, 54*(2), 119–135. http://doi.org/10.1016/j.erap.2003.12.004

Van de Vijver, F., & Watkins, D. (2006). Assessing Similarity of Meaning at the Individual and Country Level. *European Journal of Psychological Assessment, 22*(2), 69–77. http://doi.org/10.1027/1015-5759.22.2.69

Zumbo, B.D. (1999). *A handbook on the theory and methods of differential item functioning (DIF).* Ottawa: National Defense Headquarters. Recuperado de http://faculty.educ.ubc.ca/zumbo/DIF/handbook.pdf

Capítulo 6

La evaluación del egreso de la Licenciatura en Actividad Física y Deportes de la Universidad Autónoma de Baja California

José Alfonso Jiménez Moreno y Joaquín Caso Niebla

Introducción

La Universidad Autónoma de Baja California (UABC) es una de las 34 Universidad Públicas Estatales (UPES) de México. Dicha característica le impone ciertas condiciones para el desarrollo de su labor formativa en educación superior. Las universidades catalogadas como UPES son Instituciones de Educación Superior (IES) creadas por decreto de los congresos de cada Estado de la República, bajo la figura de organismos públicos descentralizados. Este tipo de instituciones tienen la función de docencia, generación y aplicación innovadora del conocimiento, así como de extensión y difusión de la cultura (Secretaría de Educación Pública [SEP], 2018). La UABC tiene la responsabilidad de ejercer estas funciones en el noroeste del país y de asegurar la formación competitiva de sus egresados. Para ello, al igual que el resto de las IES del país, adopta esquemas de evaluación orientados a la verificación de los resultados de la formación que promueve (Moreno, 2017).

Dentro de los diversos mecanismos de rendición de cuentas y responsabilidad social que utilizan las IES se encuentran diversas prácticas relacionadas con la evaluación educativa. En este sentido, la evaluación en la educación superior ha fungido como un mecanismo de valoración de la rendición de cuentas y competitividad (Cabrera, 2005).

En México, existen diversas formas de evaluar la educación superior. El antecedente más importante se encuentra en el Sistema Nacional de Evaluación de Educación Superior, estructura que en los años noventa soportaba decisiones de alto impacto en la vida de las IES, conformado

por el Consejo para la Acreditación de la Educación Superior, A.C. (COPAES), los Comités Interinstitucionales para la Evaluación de la Educación Superior (CIEES), la Oficina del Registro de Validez Oficial de Estudios (RVOE) de la Secretaría de Educación Pública (SEP), el Programa Integral de Fortalecimiento Institucional (PIFI[1]), el Programa para el Mejoramiento del Profesorado (PROMEP[2]), el Consejo Nacional de Ciencia y Tecnología (CONACYT) y el Centro Nacional para la Evaluación de la Educación Superior (CENEVAL) (ANUIES, 1995).

Si bien, en la actualidad, el Sistema Nacional de Evaluación de Educación Superior no funciona como un órgano, las instituciones y organismos antes mencionados mantienen su ejercicio, evaluando diversos componentes de la educación superior mexicana: el COPAES verificando a las instancias acreditadoras de los programas educativos; los CIEES evaluando los programas educativos; el PFCE y el PRODEP valorando condiciones institucionales y la productividad académica; el Sistema Nacional de Investigadores (SNI) vigilando la productividad de los investigadores; y el CENEVAL, concentrado en la valoración de los aprendizajes de los aspirantes por ingresar a la educación superior y la de sus egresados. En el caso particular de la UABC, ésta participa de todos los elementos de evaluación disponibles como medio de valoración de su nivel de competitividad en la formación de profesionistas.

Uno de los mecanismos de evaluación que mayormente se relaciona con la competitividad en este nivel de estudios es el Examen General para el Egreso de la Licenciatura (EGEL) del CENEVAL (Jiménez, 2017). El EGEL es un instrumento orientado a medir los conocimientos y habilidades mínimos esperados en diversas formaciones profesionales (CENEVAL, 2013). El EGEL comienza a aplicarse a mediados de los años noventa como un medio de valoración de los programas educativos de nivel superior (Gago, 1995). En su elaboración, estos instrumentos no son curriculares, es decir, no consideran ningún plan de estudios en particular para sustentar su desarrollo.

En este marco, la UABC ha sido una de las UPES con mayor participación en el EGEL. De los 63 programas educativos a nivel de licenciatura que oferta, 38 (60%) de estos participan en el EGEL. La intención de la evaluación del egreso de la UABC está establecida en su normatividad, reflejada en los artículos 80, 86 y 87 del Estatuto Escolar (UABC, 2014). Adicionalmente, su Plan de Desarrollo Institucional 2015-2019 (UABC, 2015) establece a la evaluación del egreso como un

1 *Se implementa a partir de 2001. Se le denomina a partir de 2017 como Programa de Fortalecimiento de la calidad Educativa (PFCE).*

2 *Actualmente Programa para el Desarrollo Profesional Docente (PRODEP).*

mecanismo que permite conocer la calidad de la formación de la institución. En este contexto, el EGEL forma parte importante de la vida institucional.

La UABC ha adoptado la política educativa actual que se fundamenta en la competitividad, manifestada, entre varias acciones, en la intención de valorar los conocimientos y habilidades de quien egresa de la institución. En este sentido, aprovecha la perspectiva de evaluación del EGEL que, al ser un instrumento no curricular, permite una comparación entre IES de todo el país para todas las disciplinas que cuentan con este instrumento. Cabe mencionar que 25 programas educativos de la institución, a falta de respectivo EGEL, no cuentan con herramientas para la evaluación de los aprendizajes logrados por sus egresados.

De manera particular, la Licenciatura en Actividad Física y Deportes (LAFyD) que se ofrece en la Facultad de Deportes (FD) en los tres campus de la UABC (Mexicali, Ensenada y Tijuana) es uno de los programas que no cuenta con mecanismos específicos de evaluación que permitan certificar el egreso de sus estudiantes. En el año 2017, la FD se preocupó por generar un insumo de retroalimentación de la formación de sus estudiantes, a través de una evaluación a gran escala de las competencias de sus egresados, lo que representa un ejercicio inédito en el país. Con esta intención, el Instituto de Investigación y Desarrollo Educativo (IIDE) de la UABC asumió el compromiso de desarrollar una propuesta metodológica que atendiera a las intenciones de la FD.

Actualmente, la UABC construye una evaluación del egreso que le permita valorar los conocimientos y habilidades de los egresados del programa de LAFyD, bajo una perspectiva comprensiva. Para dicho propósito se utilizarán dos instrumentos: el primero, enfocado en la medición de competencias de la LAFyD establecidas en el plan de estudios; el segundo, es un cuestionario dirigido a estudiantes para conocer las competencias intra e interpersonales que, de acuerdo con la literatura especializada, influyen en resultados de la formación profesional.

A continuación, se muestran las consideraciones metodológicas asociadas al diseño, desarrollo y validación inicial de estos instrumentos.

Selección del enfoque de evaluación

Como se apuntaba previamente, la evaluación de los conocimientos o habilidades que los egresados tienen al terminar su formación representa una de las formas actuales más utilizadas en México para identificar los resultados de la formación profesional (Jiménez, 2016). En este marco, la evaluación del egreso de la LAFyD en los diversos campus en las que se imparte (Mexicali, Ensenada

y Tijuana) representa la oportunidad de realizar un ejercicio inédito en México de la evaluación del egreso de una formación en esta disciplina.

Una de las perspectivas más relevantes en el enfoque de la propuesta de evaluación del egreso de la LAFyD es que se considera que la evaluación de la educación trasciende la aplicación de un instrumento o prueba (Mateo, 2006). En ese sentido, el primer eje de la propuesta trabajada entre el IIDE y la FD concibe a la evaluación del egreso no solo como la emisión de puntajes en una prueba, sino como un modo de conocer el grado en que los estudiantes de la LAFyD desarrollan las competencias comprometidas en dicho programa educativo, bajo un instrumento construido bajo la mirada de los actores educativos que intervienen en esa formación. Si bien, se pretende que los resultados funjan como un insumo para su retroalimentación, se parte de la idea que esta evaluación solo representa un fragmento de la complejidad que implica un programa educativo. En ese sentido, uno de los primeros ejes de la evaluación es asumir la limitante epistemológica que el método da para la interpretación de los resultados y la posterior retroalimentación de la formación de los estudiantes de la LAFyD.

Por otra parte, como se comentó previamente, el mayor antecedente de la evaluación del egreso de la LAFyD en México son los EGEL del CENEVAL. De acuerdo con el *Manual técnico del EGEL* (CENEVAL, 2013):

> El EGEL es una prueba nacional especializada por carrera profesional, cuyo propósito es identificar la medida en que los egresados de la licenciatura cuentan con los conocimientos y habilidades esenciales para el inicio del ejercicio profesional; de esta forma, se asume que la información resultante de su uso sistemático contribuye a establecer distintos aspectos de gran importancia, como: a) el grado de idoneidad de cada egresado con respecto a un estándar de formación nacional; b) el nivel de eficacia de los programas y modalidades de formación profesional que administran las IES y c) indicadores de rendimiento a partir de los cuales, los organismos evaluadores y acreditadores respectivos, pueden identificar los méritos de cada programa de licenciatura de las IES. (p. 13)

Frente a esta descripción, se puede observar que el EGEL está enfocado en la observación y la comparación de lo observado con un estándar nacional, independientemente de las caracterís-

ticas curriculares y contextuales de los programas educativos a los cuales se dirige; en ese sentido, no representa un ejercicio de evaluación que necesariamente dé respuesta a las características de los programas. Esto permite valorar la fuerte apuesta del EGEL por ser un elemento técnico enfocado en la medición, ya que el eje de su desarrollo es el instrumento en sí mismo, más no el objeto al que pretende dar respuesta. La intención de la presente propuesta de evaluación del egreso de la LAFyD pretende marcar una distancia respecto a esa perspectiva, trascendiendo la intención de la medición en sí misma, que da un lugar preponderante al instrumento y no a la estructura y orientaciones de los programas, o bien, a las perspectivas de los diferentes actores educativos. Esto representa un segundo eje de la evaluación del egreso de la LAFyD, que también resulta una diferencia relevante respecto al trabajo que el CENEVAL realiza con el EGEL.

Cabe resaltar que el ejercicio de evaluación desarrollado en el EGEL no se realiza bajo una perspectiva curricular debido a la enorme diversidad de instituciones y programas educativos de todo el país, los cuales no necesariamente coinciden en contenidos ni perfiles de formación dentro de una disciplina en particular. Sin embargo, desarrollar una evaluación *ex profeso* para un programa educativo en particular representa una oportunidad de dar respuesta a necesidades formativas específicas.

Un tercer eje de este ejercicio de evaluación representa la perspectiva metodológica utilizada, la cual se desarrolló bajo un modelo comprensivo (Stake, 2006). La evaluación comprensiva implica la consideración de la diversidad de perspectivas que constituyen el fenómeno de interés, donde la pluralidad nutre la propuesta de evaluación. Esto es distante a la perspectiva que fundamentan los ejercicios actuales de la evaluación del egreso en México. De acuerdo con Stake (2006), la evaluación comprensiva debe tomar en cuenta los intereses y preocupaciones de todos aquellos que estén involucrados, debido a que la diversidad de perspectivas es lo que permite conformar una realidad implicada en los procesos educativos.

Es necesario que la evaluación a gran escala del egreso de la LAFyD incorpore a sí misma la diversidad de perspectivas, ya que, como comenta Santos (1988), la evaluación es un instrumento de poder tal que permite la construcción de una realidad en particular sobre una persona; siendo así, la construcción de la realidad sobre una persona no debe ser decisión del evaluador o del instrumento en aislado, sino de la diversidad de perspectivas y contextos que le rodean, que son los elementos epistemológicos que permiten su entendimiento y comprensión.

Para tener mayor claridad del modelo comprensivo, a continuación se abordan los elementos que constituyen "El reloj comprensivo" propuesto por Stake (2006), el cual contiene los acontecimientos que bajo este modelo se deben de considerar al momento del desarrollo de una evaluación. Su lectura no va en el sentido de las manecillas del reloj, sino que únicamente presenta los 12 pasos que una evaluación suele considerar, sin tener un orden establecido de los momentos en los cuales atender uno u otro, ya que ello dependerá del fenómeno por evaluar y de su contexto en particular:

1. Acordar con los clientes, personal del programa y posible público de la propuesta de evaluación.
2. Identificar los alcances del programa que se pretende evaluar.
3. Desarrollar una perspectiva general de los posibles alcances que el programa por evaluar puede tener.
4. Descubrir los propósitos inmersos en la evaluación del programa y los intereses relacionados con ello.
5. Conceptualizar los problemas asociados con el programa que se pretende valorar.
6. Identificar los datos necesarios que se necesitan recabar dados los problemas asociados con el programa.
7. Seleccionar a los observadores, jueces o instrumentos relacionados con los objetivos de la evaluación.
8. Observar antecedentes del programa y sus resultados.
9. Preparar estudios de caso relacionados con las temáticas del programa.
10. Validar y confirmar la información que se ha recabado.
11. Discernir y hacer corresponder las cuestiones identificadas en la evaluación del programa con los diversos públicos.
12. Dar un formato para el uso de la información por parte de las audiencias.

Como se puede observar, los elementos metodológicos que enmarcan la evaluación comprensiva que sustenta el ejercicio de la evaluación del egreso de la LAFyD no delimitan de forma determinista el diseño que debe guardar una evaluación. La metodología se centra básicamente en implementar los pasos que permiten el desarrollo de una evaluación en el momento que se requiera. Derivado de esta postura, la evaluación del egreso de esta licenciatura se centra más en

la identificación de los elementos lógicos que subyacen en la formación de los estudiantes como el elemento que soporta los instrumentos, considerando las opiniones e intereses de los involucrados.

La evaluación del egreso de la LAFyD es, en esencia, un ejercicio que rescata un modelo que centra su atención en la comprensión del fenómeno, más que en la definición metodológica de la búsqueda de objetividad para la emisión de un juicio de valor, que es la perspectiva tradicional de ejercicios a gran escala (Tristán y Pedraza, 2017). Un trabajo de evaluación, bajo la lógica propuesta, pretende recuperar la perspectiva de los involucrados, no solo con fines de validación de la información e instrumentos generados, con el fin de emitir juicios de valor más centrados en las particularidades de lo observado y, en menor medida, en la búsqueda de generalización y sistematización que caracteriza a las evaluaciones del egreso que se realizan a nivel nacional.

En este sentido, es el programa educativo de la LAFyD el centro de la evaluación, no los instrumentos que se desarrollan para generar resultados; de tal suerte que, son sus condiciones de trabajo cotidiano y las competencias enunciadas en el Plan de estudios lo que delimita su desarrollo metodológico.

De esta forma, el interés de esta evaluación se centra en emitir una retroalimentación de la formación de los próximos profesionistas en actividad física y deporte, considerando un contexto diverso (sobre todo debido a la característica de este programa educativo, que se oferta en tres campus diferentes), evitando mantenerse ajeno a una realidad intersubjetiva en donde el contexto de cada campus influye, sin duda, en los resultados de egreso. En síntesis, la evaluación del egreso de la LAFyD no se considera solo como un examen, sino como la construcción de una realidad sobre un fenómeno de interés a través de la mirada de los actores educativos.

De acuerdo con Stake (2006), la evaluación debe considerar que el objeto por evaluar es único, por lo que sus resultados ni conclusiones pueden ser generalizados. También, se entiende que la evaluación implica enfrentarse a cuestiones y prácticas humanas, por lo que es necesario entenderlas como acciones contextualizadas y mediante ello reconocer que se trabaja con experiencias vividas y no solo con aspectos conceptuales propios de una formación disciplinar.

Además de conocer los logros o competencias profesionales desarrolladas a lo largo de la formación en la FD, la metodología de la evaluación del egreso privilegia también la comprensión de las características del evaluando, implica entonces valorar aspectos adicionales a los resultados educativos, tal y como las habilidades intra e interpersonales que orillan a un egreso profesional exitoso (National Academies of Sciences, Engineering and Medicine, 2017). Lo anterior sugirió el desarrollo de un cuestionario dirigido a los estudiantes.

Metodología para la construcción de instrumentos

La elaboración de la propuesta de evaluación del egreso de la LAFyD bajo un modelo comprensivo consideró inicialmente la definición de participantes y sus roles dentro del proyecto. Se resalta la participación de enlaces entre las instancias desarrollados: IIDE, con la gestión de su dirección e infraestructura académica e institucional, así como la orientación metodológica en todo el proceso del proyecto, y la FD, con la intención de desarrollo de la evaluación y cuyos docentes fueron la base de la construcción de la evaluación. El IIDE fungió como el agente asesor metodológico, mientras que la FD como *stakeholder.*[3] De esta manera, se conjunta la visión epistemológica y metodológica respecto a la sensibilidad y experiencia de la formación en deportes, dando una función preponderante a quienes tienen mayor relación con los estudiantes de la LAFyD.

Para el desarrollo de los trabajos asociados a esta tarea evaluativa se requirió de la conformación de los siguientes comités:

- Comité para la revisión del marco de evaluación
- Comité para la revisión de las especificaciones de ítems
- Comité para la elaboración-validación de ítems

La planeación de las etapas de trabajo no se fundamenta en un desarrollo preordenado[4] (Stake, 2006), sino que fue flexible en función de las necesidades mismas de la FD, de los avances del trabajo y avances mostrados por los diversos docentes que, en esencia, representan el eje disciplinar del ejercicio evaluativo. La planeación de esta evaluación se apoyó en la lógica sugerida por "el reloj comprensivo" de Stake (2006).

---

3 *Término usualmente utilizado para referir a los interesados dentro de un proceso de evaluación educativa.*

4 *En el argot del enfoque de evaluación comprensivo, el término de evaluación preordenada refiere a aquellos ejercicios de evaluación que mantienen una estructura metodológica poco flexible, incluso desde el primer acercamiento del agente evaluador al objeto de evaluación. La principal crítica a esta perspectiva es que la evaluación se centra en la precisión técnica, independientemente del objeto por evaluar, lo que resulta paradójico ya que, al conducir una evaluación muy precisa, el evaluador olvida las características particulares del objeto de interés.*

Delimitación del marco de evaluación

La delimitación del marco de evaluación correspondiente requirió de la participación de docentes de la FD para el establecimiento de las competencias por evaluar, tomando como insumo lo expresado en el Plan de estudios de la LAFyD, el cual comparten los tres campus de la UABC en los que se oferta este programa (Mexicali, Ensenada y Tijuana). El marco de evaluación se desarrolló bajo una estrategia universal de definición de dimensiones de constructo, decantando las dimensiones en acciones y habilidades concretas que permitieran la construcción de ítems. El insumo principal de este ejercicio fueron las competencias que cada asignatura abona al perfil de egreso. Los docentes participantes identificaron, a partir del perfil de egreso, las competencias que se desarrollan en el aula y la importancia relativa de cada contenido, lo que permitió establecer los pesos específicos de cada área con base a su contribución a dicho perfil. Las competencias resultantes de este ejercicio, mismas que conforman las áreas del examen (tabla 1).

Tabla 1. *Áreas del examen para la evaluación del egreso de la LAFyD*

Competencia / Área	Habilidades por evaluar	Número de ítems en el examen
1. Evaluación y aplicación de procedimientos pedagógicos y metodológicos de la actividad física y deporte.	9	26
2. Aplicación de la currícula oficial de educación física en la educación básica mediante los programas educativos del área física.	14	38
3. Diseño de programas de actividad física, rehabilitación y de recreación.	11	31
4. Formulación de organizaciones de actividad física y deporte.	4	11
5. Evaluación del funcionamiento y adaptación del individuo a estímulos.	4	11
Total	**42**	**117**

Fuente: Elaboración propia.

La racionalidad detrás de la determinación del número de ítems fue a partir de estimar el tiempo que los estudiantes requerirán contestar a cada uno de ellos y el tiempo máximo posible de ejecución total. Ese estimado se distribuyó de forma proporcional al número de habilidades por evaluar en cada una de las áreas. La representación de la racionalidad para la estimación del número de ítems para cada área (*i*) se realizó de la siguiente manera:

$$i = \frac{\frac{h * 100}{\sum H} * 120}{100}$$

En esta expresión, *h* representa el número de habilidades por evaluar en el área, mientras que *H* es el total de habilidades de todo examen (42). El valor *h* permitió asegurar que el número de ítems estuviera determinado por la amplitud de las áreas determinadas por los académicos participantes. El valor 120 representa el total de ítems esperados de toda la prueba, considerando que cada ítem requeriría dos minutos por responderse en un tiempo máximo de duración de la prueba de cuatro horas[5]. El número total de ítems del examen fue de 117, ya que el máximo de 120 propuesto se vio afectado una vez que se definieron las habilidades definitivas que serán medidas por el instrumento.

Elaboración y validación de ítems

A diferencia de procesos de evaluación a gran escala en la cual los docentes fungen como reproductores de contenidos de especificaciones, el desarrollo de ítems en el proyecto de evaluación del egreso de la LAFyD representó la oportunidad de orientar formas de evaluación de los docentes de la FD hacia una metodología orientada en competencias, tal como lo establece el Modelo Educativo de la UABC (UABC, 2013), al mismo tiempo de aprovechar la experiencia docente en la definición de un instrumento que refleje la formación que ellos mismos promueven, ahora bajo un marco común.

El proceso de elaboración de ítems se apoyó de la impartición de talleres en los que participaron 45 docentes de los tres campus de la FD, pertenecientes a las diversas áreas de especialización del programa y con amplia experiencia en su práctica docente.

5 *La bibliografía en psicometría no es clara en lo que respecta al tiempo estimado de respuesta para un ítem de opción múltiple. A partir de la experiencia empírica de los involucrados en el proyecto, se optó por considerar un promedio de dos minutos por ítems. Se estableció un tiempo total de cuatro horas de ejecución total en la prueba, que representaba el tiempo máximo que la FD estaba dispuesta a dedicar a la evaluación y evitar el desgaste de los estudiantes.*

Construcción de cuestionarios

Usualmente, los ejercicios de evaluación a gran escala se fundamentan en el uso de una metodología que pone en el centro al aprendizaje, haciendo así uso de información adicional regularmente denominada *factores asociados*. En el caso de esta evaluación, se desarrolló un cuestionario para dichos fines el cual presenta el mismo nivel epistemológico que el instrumento que mide competencias del Plan de estudios. El diseño del cuestionario se realizó bajo el modelo de competencias intra e interpersonales desarrollado por la *National Academies of Sciences, Engineering and Medicine* (2017) y complementado con las definiciones teóricas de los indicadores por evaluar propuestas por diversos autores especialistas en esta materia (Renninger y Hidhi, 2006; Dugan, 2015; Paunesku et al., 2015; Blanco, Aguirre, Barrón y Blanco, 2016). Las dimensiones definidas en el cuestionario se muestran en la tabla 2.

Tabla 2. *Dimensiones del Cuestionario para estudiantes*

Dimensión	Indicadores por evaluar
Dedicación	Autocontrol Responsabilidad Trabajo duro Persistencia Establecimiento de metas
Sentido de pertenencia a la institución	Integración y adaptación a la UABC Integración y adaptación a la FD
Autoeficacia académica	Comunicación Atención Comprensión Orientación hacia la excelencia
Mentalidad de crecimiento	Percepción de competencia académica Percepción de competencia personal
Objetivos de utilidad y valores	Orientación al logro
Objetivos intrínsecos e intereses	Interés de formación en contenidos disciplinares específicos
Metas prosociales y valores	Orientación hacia la colaboración Orientación hacia propósitos comunes Actitudes cívicas
Perspectiva positiva sobre el futuro	Perspectiva positiva sobre el futuro profesional

Fuente: Elaboración propia.

Considerando los dos instrumentos desarrollados, es factible dar una retroalimentación precisa a todos los egresados de la LAFyD que los presenten, ya que los datos mostrarán de manera responsiva los resultados de una formación, particularmente, con una orientación a la resolución de problemas, en coincidencia con el Modelo Educativo de la UABC y las características principales del plan de estudios de la LAFyD. De manera coincidente con el enfoque de evaluación seleccionado, los resultados de los egresados facilitarán la retroalimentación de los elementos que el programa debe mejorar, ya que no solo considera el logro en función de las competencias enunciadas en el plan de estudios, sino que el procedimiento de desarrollo asegura que los resultados reflejen la diversidad formativa entre los campus que ofrecen el programa educativo, estableciendo un eje de comparación común para todos los estudiantes.

Esta evaluación se orienta, así, al conocimiento de las particularidades de formación de los egresados de la LAFyD, lo cual será un insumo relevante para procesos de mejora didáctica en la FD, así como en la reestructuración o modificación del programa de dicha licenciatura.

En el caso particular del cuestionario, permitirá conocer elementos que tradicionalmente no se conocen mediante este tipo de ejercicios. En este caso, la consideración de los resultados por campus resulta igual de valioso. Presentar los resultados de dedicación, sentido de pertenencia a la UABC, autoeficacia académica, intereses, valores sociales y visión positiva del aprendizaje alcanzado permitirá retroalimentar elementos que usualmente se relacionan con un egreso exitoso.

El ejercicio técnico de la construcción de los instrumentos resulta coincidente con la perspectiva del enfoque de evaluación seleccionado, ya que se generaron dos productos relevantes que reflejan de manera importante la perspectiva didáctica establecida en el Modelo Educativo de la UABC y en las pretensiones de formación de la LAFyD. En ese sentido, los contenidos por valorar en los egresados reflejan fielmente las intenciones formativas de la Universidad y las orientaciones didácticas de la comunidad académica que conforma la FD.

A manera de cierre

Las etapas brevemente descritas pretenden mostrar las generalidades de acción que permiten establecer una evaluación del egreso desarrollada bajo métodos aceptados en la evaluación educativa a nivel internacional, pero, particularmente, con la perspectiva de los docentes y autoridades de la FD de la UABC, favoreciendo así en la generación de instrumentos valiosos de retroalimentación bajo el marco propio de la institución.

En los avances de la evaluación del egreso de la LAFyD fue necesario asegurar la participación de docentes del programa educativo que aportaran sus conocimientos y experiencia en la constitución de las diversas etapas previamente descritas. Como resultado final de la experiencia, se espera una aproximación a una realidad en particular, delimitada no por la perspectiva del IIDE como evaluador, sino por una diversidad de posturas hacia un mismo fenómeno considerando a un grupo de *stakeholders* heterogéneos, al ser parte de tres campus en distintas partes del Estado de Baja California; el producto final de esta evaluación deberá ser entonces fiel reflejo del perfil del egresado que se pretende evaluar.

La evaluación del egreso de la licenciatura en México ha sido uno de los elementos estratégicos de implementación de una cultura de evaluación y de rendición de cuentas en la educación superior. La elaboración de proyectos de evaluación a gran escala se ha desarrollado también gracias a la posibilidad metodológica que la perspectiva tyleriana de la evaluación permitió (Tyler, 1998), centrando sus esfuerzos en la observación de resultados educativos esperados de un plan de estudios. Si bien, este tipo de evaluación ha llegado a un punto tal de desarrollo que sus mejoras se observan notablemente en los aspectos de medición, es necesario pensar la posibilidad de un ejercicio de evaluación de esta naturaleza bajo una perspectiva que resalte el papel del objeto de evaluación, más que el instrumento por sí mismo.

De entre las posibles opciones de enfoques de evaluación, la presente propuesta se enfocó en la evaluación comprensiva propuesta por Stake (2006). El enfoque comprensivo ha sido aplicado particularmente en la evaluación de programas educativos, dado su interés particular en las cuestiones de proceso y que el evaluador valore la experiencia de estar ahí. Esta perspectiva, si bien, representaba un reto operativo importante, es un enfoque mucho más acercado a una realidad de formación profesional. Las dinámicas educativas de la actualidad requieren dejar atrás el modelo tyleriano de enseñanza y mudar poco a poco a enfoques y metodologías con mayor apertura hacia la diversidad y la complejidad.

La presente propuesta ataca principalmente la necesidad de encuadrar una postura firme que dé sentido a todo el método utilizado, el cual resalte el valor del fenómeno por evaluar y proponer una postura enfocada únicamente en la medición del aprendizaje. Normalmente, cuando se piensa en evaluación a gran escala se hace referencia a elementos técnicos y operativos que, si bien son importantes, deben estar fundamentados enfoques que den un sentido y un encuadre a los resultados que una persona recibe sobre su formación.

De esta forma, la propuesta procura mantener una ilación lógica en función del encuadre del enfoque comprensivo de la evaluación y el modelo pedagógico de la UABC, que sustenta la formación profesional de la LAFyD; por tanto, la evaluación de las habilidades de los egresados deberá estar en concordancia con esta postura. Evidentemente, hay elementos que resultan tan operativos que, si bien, pueden repensarse en función de la perspectiva comprensiva, también pueden mantenerse prácticamente sin cambios como cuando se hace mediante las evaluaciones preordenadas (Stake, 2006). El ejemplo claro de ello es la consideración de una desagregación de los elementos por evaluar en el desarrollo de la prueba, acción necesaria para plantear un enunciado que permita la medición de los elementos las habilidades por evaluar.

Por otro lado, el presente trabajo destaca la utilidad que pueden tener los cuestionarios dirigidos a los estudiantes para la medición de aspectos no cognitivos. El cuestionario propuesto permitirá retroalimentar al programa educativo de la LAFyD en elementos relacionados con un egreso exitoso y no centrar la atención únicamente en las habilidades de carácter cognitivo.

Resalta la relevancia de incorporar la participación de diversos agentes dentro de todo el proceso del desarrollo de esta evaluación. La consideración de una cantidad importante de docentes de la FD favoreció a la construcción de la evaluación como el marco de una realidad; por tanto, el mayor soporte que los resultados deben tener no es solo una coherencia entre el enfoque y la metodología, sino un sustento de diversas perspectivas respecto al mismo fenómeno de interés.

Partiendo del hecho que la realidad de un fenómeno educativo no es en sí misma, sino en función de los actores y medios involucrados, la postura comprensiva que da sostén a la presente propuesta realza la diversidad de posturas frente al egreso de la LAFyD, de tal suerte que la conjunción de esta diversidad permita la construcción de un modelo unificador.

La evaluación del egreso de la LAFyD no se centra en juzgar, es decir, en emitir una conclusión de valor en función de un puntaje, sino en proporcionar información que facilite la toma de decisiones para mejorar al programa educativo. Para ello, se requiere forzosamente evitar a la evaluación como un elemento netamente técnico, cuyo interés se enfoque principalmente en lo metodológico, sino que se debe valorar en mayor medida un proceso dialógico con los actores educativos que favorezca el establecimiento de lo que se pretende evaluar. De esta manera, la diversidad de actores favorecerá la consideración de una multiplicidad de perspectivas sobre el egreso de los estudiantes de la Facultad de Deportes.

Referencias

Asociación Nacional de Universidades e Institutos de Educación Superior (ANUIES). (1995). *La evaluación y la acreditación de la Educación Superior en México.* Recuperado de http://publicaciones.anuies.mx/acervo/revsup/res101/txt6.htm

Blanco, H., Aguirre, J., Barrón, J., & Blanco, R. (2016). Composición factorial de la escala de autoeficacia académica en universitarios mexicanos. *Formación universitaria, 9*(2), 81-88.

Cabrera, V. (2005). El concepto calidad en la educación universitaria: clave para el logro de la competitividad institucional. *Revista Iberoamericana de Educación, 36*(12), 1-7.

Centro Nacional de Evaluación para la Educación Superior (CENEVAL). (2013). *Manual técnico del EGEL.* México: Ceneval.

Dugan, J. (2015). The measurement of socially responsible leadership: consderations in establishing psychometric rigor. *Educational, cultural and Psychological Studies,12*, 23-42.

Gago, A. (1995). Los Exámenes Nacionales y el Examen General de Calidad Profesional. *Revista de la Educación Superior, 24*(95), 89-93.

Jiménez, A. (2016). El papel de la evaluación a gran escala como política de rendición de cuentas en el Sistema Educativo Mexicano. *Revista Iberoamericana de Evaluación Educativa, 9*(1), 109-126. http://doi.org/10.15366/riee2016.9.1.007

Jiménez, A. (2017). La evaluación de los egresados de formación profesional en México: Reflejo de la implementación de la política de competitividad en la educación superior. *Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas,* (25), 1-18.

Mateo, J. (2006). Claves para el diseño de un nuevo marco conceptual para la medición y evaluación educativas. *Revista de Investigación Educativa, 24*(1), 165-186.

Moreno, C. (2017). Las reformas de educación superior pública en México: rupturas y continuidades. *Revista de la Educación Superior, 46*(182), 27-44.

National Academies of Sciences, Engineering and Medicine (2017). *Supporting students' college success. The role of assessment of intrapersonal and interpersonal competences.* Washington: The National Academies Press.

Paunesku, D., Walton, G., Romero, C., Smith, E., Yeager, D., & Dweck, C. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psichologycal science, 26*(6), 784-793.

Renninger, A., & Hidhi, S. (2006). Revisiting the conceptualization, measurement, and generation of interest. *Educational psychologist, 46*(3), 168-184.

Santos, M.A. (1988). Patología general de la evaluación educativa. *Infancia y aprendizaje,* 41, 143-158.

Secretaría de Educación Pública (SEP). (2018). *Instituciones de Educación Superior.* Recuperado de http://www.ses.sep.gob.mx/instituciones.html

Stake, R. (2006). *Evaluación comprensiva y evaluación basada en estándares.* Barcelona: Editorial Graó.

Tristán, A., & Pedraza, N. (2017). La objetividad en las pruebas estandarizadas. *Revista Iberoamericana de Evaluación Educativa, 10*(1), 11-31.

Tyler, R. (1998). *Principios básicos del currículo.* Buenos Aires: Troquel.

Universidad Autónoma de Baja California (UABC). (2013). *Modelo educativo de la UABC.* Mexicali: UABC.

Universidad Autónoma de Baja California (UABC). (2014). *Estatuto Escolar de la Universidad Autónoma de Baja California*. Mexicali: UABC.

Universidad Autónoma de Baja California (UABC). (2015). *Plan de Desarrollo Institucional 2015-2019.* Mexicali: UABC.

Capítulo 7

Apuntes sobre la perspectiva de calidad en la Educación Superior mexicana

Ana Patricia Gutiérrez Zavala y José Alfonso Jiménez Moreno

Introducción: dos posturas contrapuestas sobre la calidad educativa

Desde las últimas décadas del Siglo XX, la educación superior (ES) en México ha estado envuelta en una dinámica orientada a ofrecer formación profesional enfocada en la calidad. El uso del adjetivo de *calidad* en la práctica educativa no es algo baladí, ya que ha determinado el rumbo de las acciones institucionales de las universidades mexicanas. Si bien, el término de calidad en el contexto de la ES mexicana es algo del uso cotidiano, en este uso regular conviene una reflexión sobre sus implicaciones. Al respecto, diversas investigaciones en México y en el mundo han referido a la tendencia de analizar la calidad educativa en este nivel de formación (Campos, 2004; Márquez, 2004; Alvarado, Morales y Aguayo, 2016; Mungaray, Ocegueda, Moctezuma y Ocegueda, 2016). A pesar de la orientación de las líneas de acción institucionales hacia la búsqueda de la calidad, poco se ha cuestionado el sentido de dicho concepto en este nivel educativo, salvo casos de algunas investigaciones (Egido, 2005; Gautier, 2007; Olaskoaga, Marúm y Partida, 2015). El tema de calidad en ES puede abordarse desde distintos enfoques, y es necesario especificar la postura bajo la cual las acciones educativas se fundamentan, debido a que el concepto suele diferir considerablemente dependiendo de la perspectiva que la soporta; y, en consecuencia, ello impacta en las acciones educativas que le derivan. Frente a este problema, este capítulo pretende contextualizar el concepto de *calidad educativa* en ES como un discurso que puede ser ligado a una de dos posturas: una perspectiva de calidad en educación, que parte de una visión de derechos humanos y otra que parte desde una perspectiva con tendencias a lo empresarial y competitivo.

Por medio de una revisión de literatura acerca del tema de calidad, y concretamente en ES en México, es posible notar cómo dos posturas sobresalen en discursos que abordan temas como: la discusión conceptual de calidad, la mejora o aseguramiento de la calidad en los diversos procesos formativos, la evaluación de la calidad o la importancia de la calidad dentro de la formación universitaria. Epistemológicamente, no trata un tema sin partir de una perspectiva conceptual; en ese sentido, se puede decir que, independientemente de la postura (una orientada hacia los derechos humanos y otra asentada en la visión empresarial), ésta será la base del discurso y, si se quiere comprender ese discurso, se tendrá que, no sólo reconocer esa base ideológica, sino resaltar sus antecedentes e implicaciones.

Una postura inicial: la educación superior y los Derechos Humanos

La educación superior no es obligatoria para el Estado en ningún país del mundo; no obstante, cursar este nivel educativo es muy importante, pues es la culminación del trayecto educativo de una persona, e incluso va más allá de su relación con la obtención de un empleo y movilidad social (Huerta, 2012). El fundamento que sustenta lo anterior es que la educación es un derecho, mismo que no termina en la educación básica, sino que debe darse a lo largo de la vida (UNESCO, 2005). Asimismo, el derecho a la educación no sólo se manifiesta en la posibilidad de acceso a ella, sino a considerar derechos *dentro* o *en* la educación; es decir, a recibir una educación de calidad (Latapí, 2009).

Desde esta postura, la calidad es un elemento mismo de la educación, lo cual permite asegurar que los estudiantes sean formados bajo un mínimo irreductible que debe ser responsabilidad del Estado. La postura de derechos humanos considera al estudiante como el centro de la educación, pues la formación integral del ser humano debe ser su prioridad (Jongitud, 2017). En ese sentido, esta visión hace una crítica al uso e interpretación de resultados de las pruebas de medición educativa que regularmente se utilizan para determinar si los estudiantes han alcanzado, o no, los mínimos esperados en su formación, principalmente aquellas a gran escala y de corte cuantitativo, pues se cree que un resultado numérico no solo es incapaz de reflejar toda una realidad, sino que es posible que pueda ocultar cosas importantes. Por lo tanto, esta postura sostiene que "los promedios estadísticos camuflan todo tipo de falencias en cuestiones de género, raza, etnia o legua, las cuales son cruciales desde la perspectiva de los derechos humanos" (Tomasevski, 2003); más adelante se aborda la relación entre evaluación y calidad, por ahora basta resaltar la manera en que esta primera postura centra su atención no en formas de evaluación, sino en el aseguramiento de una formación plena.

Para conocer de mejor manera la perspectiva de calidad basada en los derechos humanos, debemos retomar a Tomasevski (2003), quien estableció un referente de la postura de derechos humanos, a partir de un sistema 4-A (por las iniciales de cada principio), el cual se conforma por: asequibilidad, acceso, aceptabilidad y adaptabilidad. Estos elementos son los que, de acuerdo con la autora, son definitorios para considerar que una educación es de calidad. Analicemos cada uno de ellos.

La *asequibilidad* concibe a la educación como un derecho social y, por lo tanto, el Estado debe asegurarse de garantizar su obligatoriedad y gratuidad, asimismo, es un derecho cultural que no debe excluir a las minorías, pues debe respetar la diversidad. Bajo este principio, el Estado debe garantizar que la educación sea alcanzable por toda su población. El *acceso* refiere que verdaderamente todos tengan posibilidades de ingreso a la educación, busca eliminar la exclusión y discriminación para poder acceder a la educación; que todos puedan concluir su formación obligatoria, y del mismo modo, tener la posibilidad de ingresar a la ES. Por su parte, la *aceptabilidad* refiere un conjunto de criterios que la educación debe cumplir, como los relativos a la seguridad y salud en la escuela, o bien, algunas cualidades profesionales de los profesores para fomentar una adecuada formación de los estudiantes; la aceptabilidad obliga al Estado a establecer, controlar y exigir el cumplimiento de determinados mínimos, tanto en instituciones públicas como privadas. Finalmente, la *adaptabilidad* se fundamenta en la manera en que las escuelas se transforman para poder acomodarse a las necesidades sociales y de los estudiantes.

Cada uno de los principios del sistema 4-A se considera un indicador obligatorio, en otras palabras, si no son atendidos todos, entonces no se puede hablar de una educación de calidad. Hay que recordar que esta mirada está en contra de ideas que dictan que "la educación se compra y se vende, se adquiere por un precio, constituye un servicio regulado por el derecho comercial, no por el derecho público de los derechos humanos" (Tomasevski, 2003, p. 7). Vaya, la postura de la educación de calidad fundamentada en los Derechos pareciera no establecer como prioritario ofrecer una formación normalizada por medio de intereses puramente económicos. En este sentido, la educación, entre ellas la orientada al nivel superior, no debe considerar como eje de su quehacer los objetivos comerciales de las instituciones o del Estado.

Dentro de las posturas que defienden a la educación como un derecho humano, también destaca la idea de la importancia de la responsabilidad social que tienen las instituciones para con su entorno, cuestión que se acentúa en la ES. De ahí que surgiera el concepto de *responsabilidad*

social universitaria (RSU) (Vallaeys, 2008; 2014). Esta perspectiva se refiere a que las IES, a través de sus programas educativos, deben hacer posible que cada persona desarrolle sus capacidades en plenitud para que sea un ciudadano activo en sociedad y que se comporte de manera profesional bajo determinados valores, como la igualdad y la justicia. En este sentido, se puede afirmar que la RSU dentro de las IES es un enfoque que orienta las acciones institucionales hacia la formación integral de los individuos bajo un contexto y una región con determinadas necesidades, no solo dándole prioridad a los conocimientos, habilidades y competencias del ámbito académico que el estudiante pueda llegar a desarrollar. Incluso, pensar que la educación superior solo se orienta al desarrollo de conocimientos, habilidades y competencias, sería una importante reductibilidad de su función (Vallaeys, 2014).

La perspectiva de calidad basada en los Derechos Humanos no es una postura que deba tomarse a la ligera o que baste solo con incrementar el ingreso a las aulas universitarias[1], sino que el Estado, a través de su Sistema Educativo, debe promover que la educación sea también asequible, aceptable y adaptable. Esto no es un asunto menor, ya que requiere la implementación de políticas y acciones que se orienten al cumplimiento de estos principios; pero, además, de valorar formas de evaluación que facilitan el cumplimiento de esta responsabilidad social de las instancias universitarias.

La otra cara de la moneda: la empresarialización de la educación superior

Dentro de las diversas políticas y acciones en el ámbito de la educación también existe, como ya se mencionó, otra postura desde la cual se puede abordar el tema de calidad: la empresarial o mercantil. Varios autores han mencionado este aspecto, el cual se observa claramente en la educación superior (Ruíz, 2013; Bernasconi y Celis, 2017). Es interesante analizar cómo las universidades no solo han incorporado poco a poco un discurso basado en tendencias mercantiles, sino han adoptado un proceso de empresarialización dentro de sus políticas y acciones cotidianas (Ibarra, 2005).

---

1 *Que tampoco es un tema que deba tomarse a la ligera. El Estado Mexicano ha hecho un esfuerzo notable en esta materia, considerando el incremento que la matrícula en educación superior en México ha tenido. Por ejemplo, tan solo para nivel Licenciatura, en el ciclo 2010-2011 había un total de 2,773,088 estudiantes matriculados, mientras que en el ciclo 2016-2017 esa cifra fue de 4,096,139 estudiantes (Asociación Nacional de Universidades e Institutos de Educación Superior [ANUIES], 2011; 2018).*

Algunos de los primeros en profundizar en la relación entre universidad, empresa y Estado fueron Etzkowitz y Leydesdorff (1995), quienes, en la perspectiva de la *triple hélice*, resaltaron las influencias de la vida empresarial en las relaciones entre el Estado y las instituciones educativas; desde entonces el tema es vigente y cada vez más notorio.

El discurso sobre cómo dirigir una empresa ha llegado al de las IES, para permear en sus objetivos y formas de gestión. Se habla de formación en competencias (Cano, 2008), se establecen formas de acción y evaluación educativa con una clara alusión a la competitividad (Smith, 2014), la cual es una palabra adaptada del discurso económico. Es a través de esos enfoques de competencia bajo los cuales se desarrollan modelos educativos institucionales y se establecen los logros educativos esperados por quienes estudian en el nivel educativo superior, que, a su vez, son orientados para insertarse en el mercado laboral.

El sentido de competencia no solo se deja ver en los resultados de los estudiantes, sino también se ve entre instituciones educativas, lo cual se puede observar en *rankings* (y en la importancia que se le da a los mismos), los cuales no hacen más que posicionar y comparar con fines de un reconocimiento oficial que permita decir cuál institución es "la mejor" dentro de un listado. Al respecto, existen diversas críticas al uso de los rankings como medio de mejora de la educación (Enders, 2015). En esencia, las posturas críticas a esta temática refieren al escaso o nulo valor de retroalimentación respecto a las mejoras que las instituciones y programas educativos requieren para hacer frente de mejor manera a las condiciones en las cuales se desenvuelven. Adicionalmente, su valor metodológico está en duda.

La empleabilidad también es un factor importante dentro de este enfoque. En afán de querer responder a una demanda laboral, la cual es exigida tanto por empleadores como estudiantes que persiguen la obtención de un empleo, las IES pueden ser vistas como simplemente formadoras de recursos humanos; es decir, como un recurso, más que como profesionistas. Incluso, dentro de los indicadores que fundamentan las evaluaciones y acreditaciones de programas educativos en México[2], suelen considerar a la perspectiva de los empleadores como elemento primordial en la retroalimentación de programas educativos (Salas, 2013; Cortés y Hernández, 2018). Desde esta

2 *La evaluación de programas educativos se realiza por medio de los Comités Interinstitucionales para la Evaluación Superior (CIEES) y las acreditaciones a través de organismos reconocidos por el Consejo para la Acreditación de la Educación Superior (COPAES). Ambos son Asociaciones Civiles y no órganos del Gobierno.*

perspectiva, la RSU se vería relegada, pues la prioridad de la formación superior no se orientaría hacia responder a las necesidades de desarrollo social, sino que los objetivos prioritarios se enfocarían para brindarle al estudiante universitario las herramientas necesarias para insertarse en el mercado laboral y ser competitivo para las necesidades empresariales.

Orientaciones sobre ambas perspectivas de calidad en el caso mexicano

Antes de contextualizar el enfoque que ha permeado en México, para entender el desarrollo de ambas perspectivas es necesario apuntar que esta idea de "dos posturas de calidad" no es planteada por primera vez en este escrito; hay sustento teórico que le precede. Tal es el caso de Gautier (2007), quien indicó que la calidad podía ser vista desde una postura utilitarista o de desarrollo humano, las cuales coinciden con las mencionadas hasta este momento. A su vez, Martínez (2010) habló de dos posturas epistemológicas, una de enfoque moderno, la cual busca orden, estandarización y certeza a través de objetividad y cuantificación, considerando al mundo como estático; y otra postura posmoderna que hace énfasis en la incertidumbre, verdades múltiples y coexistencia de muchos puntos de vista sobre una misma realidad, en un mundo que se encuentra en cambio constante. Si se quisieran relacionar los enfoques, la postura de derechos humanos sería afín a la postura de desarrollo humano de Gautier y a la posmoderna de Martínez; mientras que la postura empresarial se asemejaría a la utilitarista de Gautier y a la moderna expuesta por Martínez.

Si bien, este manuscrito solo presenta una mirada superficial de lo que representa cada una de las posturas sobre calidad educativa en educación superior, es necesario contextualizar lo que ha ocurrido en México respecto al tema. Se puede iniciar diciendo que, a la fecha, pareciera que prevalece una de las posturas, y ésta es la de enfoque empresarial. Para una mejor explicación, vale referir a una serie de factores que han fungido como los antecedentes que han permitido sostener esta perspectiva.

Aunque se pretende resaltar lo ocurrido en México, tampoco se pretende negar que hay cambios que afectan de manera global, existen sucesos que ocurren en uno o varios países y afectan otros. En cada país ha habido cambios en cuanto a ES se refiere, y a pesar de que han sido de manera distinta, la globalización, como bien lo señala su concepto, pretende impactar de manera mundial. Concretamente se puede hablar del gran cambio que ocurrió después de los años ochenta, al pasar del Estado Benefactor al Estado Evaluador (Bleiklie, 1998).

Bajo la consideración de políticas económicas neoliberales y, por tanto, de apertura económica, la relación entre universidades-Estado-sector productivo (Etzkowitz y Leydesdorff, 1995) comenzó a sufrir transformaciones. El Estado comenzó a valorar a la educación superior como un medio importante para el desarrollo de la competitividad económica. En ese sentido, en los años ochenta, en Europa, se desarrolló una fuerte modificación de la política y gestión de la educación superior. Esa "nueva gestión" (Dill, 1998) se enfocaba en hacer uso de las universidades como medio de regulación de la competitividad esperada por parte del Estado. La trasformación de la relación universidades-Estado-sector productivo se enfocó entonces en la valoración de la educación superior como insumo de desarrollo económico. En esa transición, el Estado pasó de ser un agente benefactor o financiador de la educación superior como bien social, a ser un gestor de los recursos que se requieren para que las universidades realicen su función formativa.

El problema principal radica en la consideración de medios de evaluación constantes como mecanismo de implementación de políticas de competitividad del Estado, incentivando así la competitividad de la formación profesional. De esta forma, el Estado genera la cultura deseada para la competitividad económica mediante un *poder amable* (Han, 2014), en la cual se gestionan recursos para la función sustantiva de las universidades mediante la evaluación. La evaluación, así, funge como mecanismo de implementación de competitividad a través del condicionamiento de los resultados a incentivos.

En los años noventa, esto se manifestó claramente en la política educativa mexicana (Jiménez, 2016b), específicamente en el nivel superior, como consecuencia directa del proceso de Modernización Educativa de educación básica. En los años noventa, comenzó el interés de evaluación de la calidad de la educación superior mexicana. Como consecuencia, comenzó el desarrollo e implementación del Sistema Nacional de Evaluación de la Educación Superior, conformado por el Consejo para la Acreditación de la Educación Superior, A.C. (COPAES), el otorgamiento del Registro de Validez Oficial de Estudios (RVOE) de la Secretaría de Educación Pública (SEP), el Programa Integral de Fortalecimiento Institucional (PIFI[3]), el Programa para el Mejoramiento del Profesorado (PROMEP[4]), el Sistema Nacional de Investigadores (SNI) del Consejo Nacional de Ciencia y Tecnología (CONACYT), así como la creación del Centro Nacional de Evaluación para la Educación Superior (CENEVAL) (Jiménez, 2016a). Si bien, al año 2018, el Sistema no existe como

3 *Actualmente Programa de Fortalecimiento de la Calidad Educativa (PFCE).*

4 *Actualmente Programa para el Desarrollo Profesional Docente, para el Tipo Superior (PRODEP).*

tal[5], la mayoría de sus agentes mantienen funciones en activo. En todos los casos, se desarrollan actividades que dan cuenta de la implementación del enfoque evaluador del Estado Mexicano como medio de regulación de la competitividad.

Este cambio fue y es considerado como importante en la historia del país, ya que ha tenido un impacto directo en la política actual de la ES, así como en lo cotidiano de la vida institucional, incluyendo los procesos de docencia e investigación. Desde entonces, se ha podido identificar una fuerte influencia política y económica a través de la evaluación, rendición de cuentas, acreditación y mecanismos para otorgar financiamiento en la concepción de lo que es una *educación superior de calidad*.

La *calidad* como un concepto derivado del Estado evaluador

Para describir más a fondo este cambio de Estado benefactor a Estado evaluador en la vida universitaria es oportuno retomar el tema de la globalización. Ésta es un concepto que surgió en la segunda mitad del siglo XIX (Saxe-Fernández, 2000), en sus inicios fue utilizado en el ámbito económico, pero se refiere a un proceso que abarca además aspectos políticos, sociales, culturales y tecnológicos; se relaciona con una conexión e intercomunicación entre todos los países del mundo, como una sociedad que pareciera no tener fronteras gracias a los medios de comunicación y de transporte y a las nuevas tecnologías. Por ello, fue posible la ampliación de la dependencia política, económica y cultural entre los países; por una apertura a relacionarse con otros y competir por el desarrollo económico. La globalización en ES cobró fuerza después de la Segunda Guerra Mundial, de manera particular en México se pudo percibir un cambio importante una vez que el Estado comenzó a desarrollar funciones evaluadoras, ejemplo de ello es la implementación de agencias centradas en el desarrollo de evaluaciones de los diversos elementos de la ES.

En México, previo a los años noventa, la ES guardaba una relación paternalista con el Estado. Antes de que la evaluación educativa se convirtiera en un tema protagónico para la educación, en el país permeaba el concepto de Estado benefactor, surgido después de la Segunda Guerra Mundial para tratar de hacer frente al malestar y descontento social ocasionado por la misma; se quería atender a las necesidades de la sociedad. El antecedente del Estado benefactor se mani-

5 *En gran parte debido al otorgamiento de autonomía al Instituto Nacional para la Evaluación de la Educación (INEE) en 2013. Este nombramiento dotaría al INEE de facultades para la organización de un Sistema Nacional de Evaluación Educativa.*

festó en Alemania, pues el primer modelo de ordenamiento que implementó el primer gobierno después de la guerra, conocido como modelo social de mercado, se adoptó en la mayoría de los países industrializados occidentales, nombrado entonces "*Welfare State*" o "Estado de bienestar" (Vázquez, 2005).

Esta etapa benefactora tuvo un punto de quiebre, el cual es el antecedente de la etapa evaluadora. Se empezó a incrementar el número de las universidades y se les proporcionaba financiamiento sin que tuvieran que rendir cuentas de cómo se usaba el dinero que se les otorgaba, no existía la rendición de cuentas como tal. Al Estado benefactor se le comenzaron a señalar características negativas que hicieron que surgiera una necesidad de cambio. Una de esas características fue que "se había creado una industria costosa y de baja calidad" (Revueltas, 1993, p. 217) y ese discurso pronto llegó al ámbito educativo, pues fue cuando comenzó a cuestionarse si en las instituciones se ofrecía educación de calidad, principalmente porque hasta entonces, no se pedían resultados que avalaran el uso eficiente de los recursos que se otorgaban a las instituciones de educación superior (IES), y como consecuencia, la evaluación y la búsqueda de la calidad permitió establecer criterios y requisitos para la regulación presupuestal.

Así, el Estado evaluador comenzó a coordinar y regular a la ES, así como comprobar si en las instituciones se respondía a exigencias y desafíos científicos, económicos, sociales y tecnológicos orientados a la competitividad del Estado, en aras de dar respuesta a la inversión que este rubro representaba para el país. Se empezó a prestar atención a lo que ocurría dentro de las IES, se cuestionaba si el financiamiento que se le otorgaba a las universidades de alguna forma se traducía en resultados de impacto en aprendizaje; por lo que fue necesario implementar políticas de evaluación que se tradujeran en el aumento de la calidad, es decir, en el aumento de competitividad. Esta idea se ve claramente señalada en la postura de De la Garza (2003), primer coordinador de los CIEES, justificando el origen de la evaluación de programas educativos en ES como medio de búsqueda de calidad. Para De la Garza, la transición de la concepción del Estado benefactor hacia la adopción del Estado evaluador en la ES se debió a la restricción del financiamiento público para este nivel educativo, debido a una aparente percepción de su baja calidad y poca pertinencia[6]. En este sentido, bajo la postura del Estado, la evaluación lleva la responsabilidad de combatir el pro-

6 *Esta afirmación resulta paradójica, ya que previo a la implementación de un sistema de evaluación, no era posible conocer parámetros de calidad; sin embargo, sirvió como un argumento falaz para la implementación de políticas de competitividad y financiamiento.*

blema de baja calidad y pertinencia de los programas educativos, verificando para ello los logros obtenidos por parte de las IES.

Para autores como López, Lagunes y Recio (2009), que México adoptara la perspectiva europea del Estado Evaluador, ha sido el cambio más significativo en materia de política educativa, pues fue donde el sistema de ES mexicano fue objeto de evaluaciones externas por distintos organismos internacionales como el Consejo Internacional para el Desarrollo de la Educación (CIDE), El Banco Mundial (BM) y la Organización para la Cooperación y el Desarrollo Económico (OCDE). Además, como señalaron los autores mencionados, también México se hizo miembro de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) y de la Comisión Económica para América Latina (CEPAL), lo cual llevó a que el país recibiera recomendaciones respecto al rumbo que debía tomar la ES.

Hasta este punto vale resaltar que, en México, desde los años noventa, se comienza a hacer énfasis en evaluaciones con fines de condicionar financiamiento y también de garantizar que se ofrezca educación de calidad en las IES (Jiménez, 2017). Esta transición en México fue relevante, ya que ha sido la base para la implementación de políticas y acciones orientadas en la evaluación de la ES. En las últimas dos décadas, esa perspectiva ha transitado de la competitividad a la meritocracia, en la cual se asume como válido el hecho que sean las IES mismas (y sus agentes, como los docentes y estudiantes) quienes sean responsables de su propio desarrollo y productividad, delegando la responsabilidad productiva desde el Estado hacia los individuos.

Una precisión necesaria: la política educativa y la evaluación como medios de calidad

Las políticas públicas son las maneras en las que el Estado propone y se involucra en la educación del país; para Amador (2008), referenciando a la Ley General de Educación, una política educativa se entiende como la intervención del Estado en la educación, que va desde la acción sistemática que el Estado lleva a cabo a través de las instituciones docentes hasta todas aquellas acciones que realice para asegurar, orientar o modificar la vida cultural de una nación. Estas acciones de intervención han ido transformándose según el momento histórico, político y social del país y de la influencia de contextos de países externos, por lo que no se podría hablar de una consistencia a nivel nacional (recordando nuevamente el peso de la globalización).

Históricamente, pueden ser resaltados tres momentos que han influido en el cambio de las políticas educativas en México; el periodo entre 1920 y hasta 1946 se caracterizó por el desarrollo social y cultural; después de esa fecha, en el desarrollo de la economía moderna y la educación como promotora de la movilidad social y, en los años 80, cuando México atravesaba una crisis económica, las políticas se centraron en reorientar el rol que desempeñaba la ES en el desarrollo económico y social (López et al., 2009). Como una descripción a este último periodo, los autores apuntaron que la educación pasó a ser un elemento de comercio del capital cultural. A través de la libre competencia se pretendió atraer a los "clientes" que aspiraban a poseer capital cultural para trasladarlo como competencias productivas en el mercado de laboral.

Un aspecto al que hay que prestar atención dentro de este tema es al "control". Hay varios autores que lo llaman "regulación", pero finalmente, ambos conceptos van encaminados a la acción de dirigir de manera dominante las acciones de otros. El control tiene lugar en esta discusión, ya que la evaluación, para muchos autores, es una forma en la que el Estado tiene control sobre las IES. Luengo (2003) explicó lo siguiente al respecto:

> La transformación más significativa en estos años, fue un nuevo dispositivo de regulación gubernamental hacia las universidades, consistente en los siguientes instrumentos básicos: la articulación de procedimientos de evaluación, los programas de financiamiento extraordinario y las exigencias específicas de cambio a las instituciones. Este dispositivo ha permitido conducir a las instituciones educativas en el rumbo de las estrategias y programas gubernamentales y en sintonía con las necesidades de la economía y la sociedad. (p. 10)

Los procedimientos de evaluación de los que habla el autor se fueron haciendo cada vez más específicos; las políticas de evaluación, por lo tanto, comenzaron a responder a ese rumbo nuevo de estrategias y programas para atender justamente a las necesidades económicas y sociales. De esta forma, la evaluación de la educación superior en el país empezó a abarcar a instituciones, programas, proyectos, personal académico y a alumnos. El control comenzó a implementarse a través de todos esos aspectos.

Reflexiones finales

Las posturas de *calidad educativa* expuestas en este manuscrito representan, aparentemente, dos enfoques antagónicos. Por una parte, la tendencia de Derechos Humanos nos orienta a considerar que la educación debe responder hacia la población, a través de garantizar el acceso a la misma, su adaptabilidad, asequibilidad, así como su adaptación, aun cuando la educación superior no sea obligatoria para el Estado; por su parte, la postura empresarial se enfoca fuertemente en dar respuesta a la necesidad del Estado Mexicano de atender a las demandas competitivas de una apertura económica y a su pertenencia a organismos internacionales.

El debate sobre ambas posturas no resulta solo un ejercicio de reflexión, sino una necesidad de valorar la práctica misma de la educación superior en México. Las consecuencias de cada postura en los procesos de evaluación educativa que en la actualidad son parte de la dinámica cotidiana en todo el país, representa solo una de las ventanas bajo las cuales se reflejan las consecuencias de la adopción de una postura sobre la *calidad de la educación*.

Este ejercicio nos ofrece la oportunidad de reiterar lo expuesto por Olaskoaga et al. (2015), respecto a que la *calidad* no es un término definido. Al contrario, debe caracterizarse por su indefinición y comenzar a entenderse como una perspectiva política que las IES asumen como válida para su contexto, necesidades y actores académicos.

Es claro que en este manuscrito se ahondó en la descripción de la perspectiva mercantilista o empresarial de la calidad, fundamentada por las intenciones de competitividad del Estado Mexicano. Vale considerar la amplitud de investigación sobre la perspectiva de Derechos Humanos, que parece ser adoptada por la Educación Básica de México sin mayor problema (como puede valorarse claramente en el texto del INEE, 2014); sin embargo, ¿la incorporación del enfoque de Derechos es posible en el contexto de la educación superior? ¿Un esquema basado en Derechos (que, como se abordó, resulta antagónico a la necesidad de mercantilizar el objeto), es compatible en una dinámica educativa en la cual el recorte presupuestal ha marcado la pauta? Sin duda, la evaluación de los diversos aspectos de la educación superior ha sido un mecanismo relevante en esta dinámica; sin embargo, ¿resulta factible conjugar la evaluación de las 4-A en un contexto en el cual lo relevante es la identificación de los logros (como indicio de competitividad)?

Adicional a este debate, ¿las IES son conscientes de las diversas formas de entendimiento de la *calidad educativa* cuando se plantean los Planes de Desarrollo y las políticas institucionales que de ellos emanan? ¿Resulta necesario considerar la posibilidad de que las IES cuestionen su

modelo, su contexto y objetivos? Esta reflexión se enfatiza en las universidades de carácter público, que regularmente suelen ser autónomas; en ese sentido, ¿estas universidades determinan de forma soberana su política institucional orientada hacia la calidad o es resultante solo de las presiones mercantiles y financieras?

Referencias

Alvarado, E., Morales, D., & Aguayo, E. (2016). Percepción de la calidad educativa: caso aplicado a estudiantes de la Universidad Autónoma de Nuevo León y del Instituto Tecnológico de Estudios Superiores de Monterrey. *Revista de la Educación Superior, 45*(180), 55-74.

Amador, J. (2008). *La evaluación y el diseño de políticas educativas en México.* (Documento de trabajo número 35). Centro de Estudios Sociales y de Opinión Pública.

Asociación Nacional de Universidades e Institutos de Educación Superior (ANUIES). (2011). *Anuario Estadístico de Educación Superior. 2010.* México: ANUIES.

Asociación Nacional de Universidades e Institutos de Educación Superior (ANUIES). (2018). *Anuario Estadístico de Educación Superior. 2017.* México: ANUIES.

Bernasconi, A., & Celis, S. (2017). Higher Education Reforms: Latin America in Comparative Perspective. *Education Policy Analysis Archives, 25*(67), 1-11.

Bleiklie, I. (1998). Justifying the Evaluative State: New public management ideals in Higher Education. *Journal of Public Affairs Education, 4*(2), 87-100.

Campos, B. (2004). The balance between higher education autonomy and public quality assurance: Development of the Portuguese system for teacher education accreditation. *Education Policy Analysis Archives, 12*(73), 1-33.

Cano, M. (2008). La evaluación por competencias en la educación superior. Profesorado. *Revista de Currículum y Formación de Profesorado, 13*(3), 7-16.

Cortés, E. y Hernández, N. (2018). Evaluación y acreditación de las Escuelas Normales. Perspectiva histórica y realidades. *Debates en evaluación y currículum. Congreso Internacional de Educación, 3*(3).

De la Garza, E. (2003). Las universidades politécnicas. Un nuevo modelo en el sistema de educación superior en México. *Revista de la Educación Superior, 32*(126).

Dill, D. (1998). Evaluating the 'Evaluative State': Implications for research in Higher Education. *European Journal of Education, 33*(3), 361-377.

Egido, I. (2005), Reflexiones en torno a la evaluación de la calidad educativa. *Tendencias Pedagógicas, 10*, 18-28.

Enders, J. (2015). Una "carrera armamentista" en la academia: los rankings internacionales y la competencia global por crear universidades de clase mundial. *Revista de la Educación Superior, 4*(176), 83-109.

Etzkowitz, H., & Leydesdorff, L. (1995). The Triple Helix. University-Industry-Government Relations: A laboratory for knowledge based economic development. *EASST Review, 14*(1), 14-19.

Gautier, E. (2007). Educación de calidad. Comentarios a la Nueva propuesta de OREALC/Unesco. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación, 5*(3), 29-35.

Han, B-C. (2014). *Psicopolítica.* Barcelona: Herder.

Huerta, J. (2012). El rol de la educación en la movilidad social de México y Chile. *Revista Mexicana de Investigación Educativa, 17*(52), 65-88.

Ibarra, E. (2005). Origen de la empresarialización de la Universidad: El pasado de la gestión de los negocios en el presente manejo de la Universidad. *Revista de la Educación Superior, 2*(134), 13-37.

Instituto Nacional para la Evaluación de la Educación (INEE). (2014). *El derecho a una educación de calidad. Informe 2014*. México: INEE.

Jiménez, A. (2016a). *Evaluación a gran escala de la licenciatura: hacia un modelo comprensivo.* Tesis de doctorado. Universidad Nacional Autónoma de México, México.

Jiménez, A. (2016b). El papel de la evaluación a gran escala como política de rendición de cuentas en el sistema educativo mexicano. *Revista Iberoamericana de Evaluación Educativa, 9*(1).

Jiménez, A. (2017). La evaluación de los egresados de formación profesional en México: Reflejo de la implementación de la política de competitividad en la educación superior. *Archivos Analíticos de Políticas Educativas, 25*(48).

Jongitud, J. (2017). El derecho humano a la educación superior en México. *Revista de la Educación Superior, 46*(182), 45-56.

Latapí, P. (2009). El derecho a la educación. Su alcance, exigibilidad y relevancia para la política educativa. *Revista Mexicana de Investigación Educativa, 14*(40), 255-287.

López, M., Lagunes, C., & Recio, C. (2009). *Políticas públicas y Educación Superior en México.* Ponencia. X Congreso Nacional de Investigación Educativa, México.

Luengo, E. (2003). *Tendencias de la Educación Superior en México: Una lectura desde la perspectiva de la complejidad.* Trabajo presentado en Seminario sobre Reformas de la Educación Superior en América Latina y el Caribe, Colombia.

Márquez, A. (2004). Calidad de la educación superior en México. ¿Es posible un sistema eficaz, eficiente y equitativo? Las políticas de financiamiento de la educación superior en los noventa. *Revista Mexicana de Investigación Educativa, 9*(21), 477-500.

Martínez, J. (2010). La Construcción de Indicadores y Evaluación de la Calidad en Centros Educativos. Seis experiencias en México. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación, 8*(5), 133-153.

Mungaray, A., Ocegueda, M., Moctezuma, P., & Ocegueda, J. (2016). La calidad de las Universidades Públicas Estatales de México después de 13 años de subsidios extraordinarios. *Revista de la Educación Superior, 45*(177), 67-93.

Olaskoaga, J., Marúm, E., & Partida, M. (2015). La diversidad semántica y el carácter político de las nociones de calidad en la Educación Superior en México. *Revista de la Educación Superior, 44*(173), 85-102.

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO). (2005). *Hacia las sociedades del conocimiento.* París: UNESCO.

Revueltas, A. (1993). Las reformas del Estado en México: del Estado benefactor al Estado neoliberal. *Política y Cultura*, (3), 215-229.

Ruíz, M. (2013). La empresarialización de la educación superior en la Unión Europea. Neocolonialismo y Segregación. *Horizontes Sociológicos, 1*(2), 47-55.

Salas, I. (2013). La acreditación de la calidad educativa y la percepción de su impacto en la gestión académica: el caso de una institución del sector no universitario en México. *Calidad en la educación, 38*, 305-333.

Saxe-Fernández, J. (2000). Globalización, poder y educación pública. En: D. Cazés, E. Ibarra y L. Porter (coords.), *Estado, Universidad y Sociedad: entre la globalización y la democratización* (pp. 51-80). México: UNAM.

Smith, W. (2014). The global transformation toward testing accountability. *Education Policy Analysis Archives, 22*(116). http://doi.org/10.14507/epaa.v22.1571.

Tomasevski, K. (2003). *Derechos económicos, sociales y culturales. Informe de la Relatora Especial sobre el Derecho a la Educación.* Presentado de conformidad con la resolución 2002/23 de la Comisión de Derechos Humanos. Naciones Unidas. Nueva York: Naciones Unidas.

Vallaeys, F. (2008). "Responsabilidad social universitaria": una nueva filosofía de gestión ética e inteligente para las universidades. *Educación Superior y Sociedad, 13*(2), 191-220.

Vallaeys, F. (2014). La responsabilidad social universitaria: un nuevo modelo universitario contra la mercantilización. *Revista Iberoamericana de Educación Superior, 5*(12), 105-117.

Vázquez, J. (2005). Neoliberalismo y Estado benefactor. El caso mexicano. *Aportes, Revista de la Facultad de Economía, 30*, 51-76.

Evaluación Educativa.
Experiencias de investigación en Posgrado

ISBN 978-607-98270-3-8
DOI 10.29410/QTP.18.10

Esta obra se terminó de producir en diciembre de 2018.
Su edición y diseño estuvieron a cargo de:

Qartuppi

Qartuppi, S. de R.L. de C.V.
http://www.qartuppi.com

Esta obra se publica bajo una Licencia Creative Commons Atribución-NoComercial 4.0 Internacional.

https://creativecommons.org/licenses/by-nc/4.0/deed.es

Evaluación Educativa. Experiencias de investigación en Posgrado.

En esta obra se incluyen diversas investigaciones en torno a la evaluación educativa en México; incursiona en ámbitos teóricos, metodológicos y aplicados de la evaluación abonando a la reflexión, debate y avance de este tema.

Se integran siete trabajos sobre evaluación docente y del aprendizaje, derivados de los trabajos de tesis de Maestría y Doctorado en Ciencias Educativas que ofrece el Instituto de Investigación y Desarrollo Educativo, de la Universidad Autónoma de Baja California. Los capítulos que conforman la obra evidencian la originalidad, rigurosidad y diversidad que caracteriza a ambos programas educativos.

ISBN 978-607-98270-3-8

9 786079 827038

Qartuppi