

Desarrollo de un enfoque sistémico para evaluar la docencia y fomentar una enseñanza eficaz

Evaluating and Supporting Effective Teaching: Developing a Systemic Approach

Linda Darling-Hammond

Universidad de Stanford, EE.UU.

Resumen

Este artículo argumenta a favor de un enfoque integral para crear un sistema que evalúe y apoye a los profesores. La autora sugiere que este sistema debería incluir cinco elementos clave: 1) estándares estatales comunes para la enseñanza que garanticen un aprendizaje real y que rijan toda la profesión; 2) evaluaciones de desempeño, basados en estos estándares, que orienten funciones estatales, tales como formar a los docentes, habilitarlos para el ejercicio de la profesión y certificar niveles avanzados de desempeño; 3) sistemas locales de evaluación inspirados en los mismos estándares para evaluar la calidad de la enseñanza in situ, basados en distintos criterios de medición de la práctica docente y del aprendizaje; 4) estructuras de apoyo que garanticen la disponibilidad de evaluadores capacitados y de mentores para los profesores que requieran ayuda adicional, así como la toma de decisiones justas en materia de recursos humanos; y 5) oportunidades de desarrollo profesional que permita a los profesores mejorar su labor y la calidad de la enseñanza. Sobre la base de investigaciones y ejemplos de buenas prácticas, este artículo expone a continuación cómo debería operar cada uno de estos elementos.

Palabras clave: evaluación de profesores, práctica docente, desarrollo profesional docente, modelo de valor agregado, enseñanza eficaz

Correspondencia a:

Linda Darling-Hammond
Universidad de Stanford, EE.UU.
485 Lasuen Mall, Stanford, EE.UU.
Correo electrónico: lindadh@suse.stanford.edu

El presente artículo proviene en gran parte de un informe más extenso publicado por el Stanford Center for Opportunity Policy in Education (SCOPE) en la Universidad de Stanford, titulado *Creating a comprehensive system for evaluating and supporting effective teaching* y disponible en <http://edpolicy.stanford.edu/publications/pubs/591>

© 2012 PEL, <http://www.pensamientoeducativo.org> - <http://www.pel.cl>

ISSN: 0719-0409 DDI: 203.262, Santiago, Chile
doi:10.7764/PEL.49.2.2012.1

Abstract

This article argues for a comprehensive approach to creating a system for evaluating and supporting teachers. The author suggests such a system should include five key elements: 1) Common statewide standards for teaching that are related to meaningful student learning and are shared across the profession; 2) Performance assessments, based on these standards, guiding state functions such as teacher preparation, licensure, and advanced certification; 3) Local evaluation systems aligned to the same standards, for evaluating on-the-job teaching based on multiple measures of teaching practice and student learning; 4) Support structures to ensure trained evaluators, mentoring for teachers who need additional assistance, and fair decisions about personnel actions; and 5) Aligned professional learning opportunities that support the improvement of teachers and teaching quality. Based on research and evidence of best practices, this article outlines how each of these five elements should operate.

Keywords: teacher evaluation, teaching practice, teacher development, value-added methods, effective teaching

Prácticamente todo el mundo concuerda en que es necesario revisar el sistema de evaluación docente de Estados Unidos. Los sistemas existentes no ayudan mucho a los profesores a mejorar o a distinguir claramente a los profesores exitosos de los que tienen dificultades. Las herramientas usadas no siempre reflejan los aspectos realmente importantes de una buena enseñanza. Es prácticamente imposible para los directores, sobre todo en los establecimientos más grandes, disponer de tiempo suficiente o tener el conocimiento necesario de las distintas disciplinas como para evaluar a todos los profesores que supervisan, y mucho menos para prestar todo el apoyo didáctico que puedan necesitar algunos docentes. Y muchos directores de establecimiento no han tenido acceso al desarrollo profesional y a todo el apoyo que necesitan para convertirse en buenos líderes y evaluadores de equipos docentes. De ahí que los procesos de evaluación actuales por lo general contribuyan muy poco, ya sea para el aprendizaje de los propios docentes o para la obtención de información oportuna y acertada que ayude a tomar decisiones en materia de recursos humanos.

Además, los criterios y métodos de evaluación de docentes varían considerablemente de un distrito y un establecimiento a otro y, por lo general, no tienen vínculo alguno con la forma en que se evalúa a los docentes en momentos clave de su carrera: esto es, cuando concluyen su período de formación preliminar, cuando se titulan y comienzan a ejercer, y cuando se convierten en docentes titulares de un establecimiento al amparo de una licencia para ejercer a largo plazo. Durante su carrera, la mayoría de los profesores se enfrentan a una plétora de estándares y orientaciones que determinan *qué* tienen que enseñar y *cómo* deben hacerlo. En otras palabras, muchos estados carecen de un sistema coherente para evaluar y mejorar la calidad de la enseñanza, lo que dificulta la búsqueda de soluciones eficaces a los problemas de la docencia.

Hoy en día la atención está puesta en identificar y apartar a los malos profesores. Pero lo que realmente se necesita es un tipo de evaluación docente que forme parte de un *sistema de enseñanza y aprendizaje* que fomente la mejora continua tanto para cada uno de los docentes como para la profesión en general. Tal sistema debiera incrementar el aprendizaje y las aptitudes de los docentes y, al mismo tiempo, garantizar que los profesores seleccionados y titulares de cursos puedan realmente contribuir al aprendizaje de los estudiantes durante toda su carrera.

Para tener un equipo docente altamente calificado se necesitan profesores bien formados, desde su contratación, siguiendo con la capacitación hasta su desarrollo profesional continuo. Se debe concebir el aprendizaje y la evaluación como parte integrante de un sistema completo que fomente la eficacia durante cada etapa de la carrera de un profesor. Dicho sistema debe garantizar que la evaluación docente esté relacionada con —y no aislada de— los programas de preparación e inducción, la práctica profesional cotidiana y un entorno pedagógico productivo.

Se requiere un sistema basado en estándares de enseñanza profesional vinculados con los estándares de aprendizaje, el currículo y la evaluación de los estudiantes que permitan forjar una relación fluida entre lo que los profesores hacen en la sala de clases y la forma en que se prepararan y son evaluados. Todo

sistema ideado para medir la producción de los profesores debería tomar en cuenta la práctica docente en el contexto de las metas curriculares y las necesidades de los estudiantes, así como las diferentes evidencias de la contribución de los profesores al aprendizaje de los estudiantes y al establecimiento en general.

Por último, se necesita un sistema integral que permita cumplir una serie de propósitos: la titulación, la contratación y la asignación de un cargo titular en un establecimiento; el apoyo a la supervisión y el desarrollo profesional; la identificación de los profesores que necesitan apoyo adicional —y, en algunos casos—, una reorientación profesional y el reconocimiento de docentes expertos que puedan contribuir al aprendizaje de sus pares, ya sea de manera informal o como tutores, capacitadores y líderes docentes. Algunos formuladores de políticas públicas también buscan asociar compensaciones al juicio sobre la eficacia docente, ya sea con sueldos diferenciados o relacionando dichas evaluaciones con responsabilidades adicionales que puedan redundar en mayores honorarios o sueldos.

El presente artículo describe un enfoque integrado que concilia estas metas con la carrera docente y con un sistema de desarrollo profesional que contribuye para la eficacia de todos los profesores en cada etapa de su carrera.

Comprender las cualidades del docente y la calidad de la enseñanza

En el marco del interés actual por medir la eficacia de la labor docente, cabe distinguir entre las *cualidades del docente* y la *calidad de la enseñanza*. Se puede entender por *cualidades del docente* el conjunto de rasgos personales, de habilidades y de comprensión que un individuo aporta a la enseñanza, incluida su predisposición a comportarse de ciertas formas. Las investigaciones sobre la eficacia docente, basadas en las calificaciones de los profesores y los logros de los estudiantes, pusieron en evidencia la importancia de las siguientes cualidades:

- un sólido conocimiento del contenido relacionado con lo que se debe enseñar; conocimientos sobre la forma de enseñar dicha materia (contenido pedagógico) y habilidad para adoptar prácticas pedagógicas y de evaluación productivas;
- la comprensión sobre el aprendizaje y del desarrollo de los estudiantes, incluidos aquellos que presentan diferencias o dificultades en el aprendizaje, y sobre cómo apoyar el aprendizaje del lenguaje y del contenido en quienes presentan deficiencias en el idioma en que se enseña;
- una habilidad general para organizar y explicar ideas, así como para establecer diagnósticos mediante la observación y la reflexión; y
- una capacidad de adaptación que permita a los profesores encontrar las soluciones adecuadas para atender las necesidades de los estudiantes en un entorno específico (Darling-Hammond, 2000; Darling-Hammond & Bransford, 2005; Wilson & Ferrini-Mundy, 2001).

La mayoría de los docentes, apoderados y formuladores de políticas también incluirían importantes actitudes en esta lista, tales como la disposición para:

- apoyar el aprendizaje de todos los estudiantes,
- enseñar de forma justa e imparcial,
- adaptar la enseñanza en pos del éxito de los estudiantes,
- continuar aprendiendo y mejorando, y
- colaborar con otros profesionales y con los apoderados para contribuir al desarrollo de los estudiantes en forma individual y al establecimiento en general.

Estas cualidades destacadas en las investigaciones sobre la enseñanza quedaron plasmadas en los estándares adoptados por el National Board for Professional Teaching Standards de Estados Unidos (NBPTS) y, en el caso de los docentes principiantes, por los estados que integran el Interstate New Teacher Assessment and Support Consortium (INTASC). El INTASC ha asumido un papel preponderante en el desarrollo de estándares y evaluaciones para docentes principiantes, las que constituyen una base de conocimientos para la enseñanza fundada en investigaciones sobre desarrollo, aprendizaje, malla curricular y enseñanza.

Estos estándares se han usado como criterios para la formación de los profesores y para su licenciamiento en más de 40 estados durante el último decenio, con lo cual han permitido desarrollar una base más

sólida para fomentar una enseñanza eficaz. Varios estudios sugieren que la habilitación profesional de los docentes se ha convertido en un factor de peso para predecir la eficacia de la práctica docente en estados que han elevado el nivel de exigencia de sus estándares (Clotfelter, Ladd, & Vigdor, 2007; Goldhaber & Brewer, 2000).

La *calidad de la enseñanza* se refiere a una enseñanza sólida que permita aprender a una amplia gama de estudiantes. Dicha enseñanza cumple con las demandas de la disciplina, con las metas de enseñanza y con las necesidades de los estudiantes en un entorno dado. La calidad de la enseñanza depende en parte de las cualidades del docente —sus conocimientos, sus habilidades y su disposición—, pero también se ve fuertemente influenciada por el entorno en que se imparte. El entorno, en tanto, se define por aspectos clave como los sistemas curriculares, la formación y especialización de los docentes y lo que deben enseñar, y las condiciones en las que se enseña. Un excelente profesor puede verse impedido de ofrecer enseñanza de alta calidad cuando el currículo es deficiente y no dispone de material o herramientas de evaluación adecuados. Asimismo, un docente bien formado puede tener un mal desempeño si se ve obligado a enseñar fuera de su campo de especialidad o en malas condiciones; por ejemplo, sin material didáctico adecuado, en un espacio que no cumple con los estándares de calidad, si no dispone de tiempo suficiente, o en clases con demasiados alumnos.

Así, un docente con buenas cualidades puede incrementar las probabilidades de lograr una enseñanza eficaz, pero no es una garantía de éxito. Toda iniciativa que apunte a desarrollar la eficacia y la calidad de la enseñanza debe tomar en cuenta no solo la forma de identificar, recompensar y aprovechar las cualidades y habilidades de los docentes, sino también de propiciar entornos de enseñanza que permitan la aplicación de buenas prácticas. Por lo tanto, para lograr una enseñanza eficaz se necesitan, además de las cualidades individuales de los docentes, políticas que propicien un entorno adecuado de aprendizaje y enseñanza.

Un enfoque sistémico de la evaluación de los docentes y del apoyo a una enseñanza eficaz

Hay una conciencia creciente de que se necesita un enfoque sistémico que busque alcanzar la eficacia en la labor docente tomando en cuenta las etapas de contratación, formación, inducción y desarrollo continuo. Además de estándares claros para el aprendizaje, sumados a un material curricular y de evaluación de alta calidad, dicho enfoque debería incluir cinco elementos clave:

1. *estándares estatales comunes* para la enseñanza relacionados con un aprendizaje significativo y compartidos por toda la profesión,
2. *evaluaciones de desempeño, basadas en estos estándares, que sirvan de pauta para las funciones estatales*, tales como formar a los docentes y otorgarles licencias y certificados de desempeño avanzados,
3. *sistemas locales de evaluación inspirados en los mismos estándares* para evaluar la calidad de la enseñanza in situ, basados en distintos criterios de medición de las prácticas de enseñanza y del aprendizaje,
4. *estructuras de apoyo* que garanticen la disponibilidad de evaluadores capacitados y de mentores para los profesores que requieran ayuda adicional, así como decisiones justas en materia de recursos humanos, y
5. *oportunidades de desarrollo profesional homogeneizadas* que permitan a los profesores mejorar su labor y la calidad de la enseñanza.

Sobre la base de investigaciones y ejemplos de buenas prácticas, se detalla a continuación cómo debería operar cada uno de estos elementos.

Empezar por los estándares

«Si no sabes adónde vas, cualquier camino te sirve», afirma el gato de *Alicia en el país de las maravillas*. Lo mismo vale para la educación. Sin una comprensión clara de lo que deberían aprender los estudiantes y de cómo la enseñanza puede ayudarlos, se avanza a ciegas. Es fundamental desarrollar una visión común de las metas educativas y de las características de una formación de apoyo si se quiere construir un sistema que contribuya a una enseñanza eficaz.

Los estándares estatales o nacionales pueden ayudar a establecer una concepción clara de las metas de aprendizaje y de los tipos de formación que fomentan una comprensión disciplinaria (e interdisciplinaria)

de cada asignatura y entre ellas. Estas concepciones deben ser respaldadas por marcos curriculares y materiales cuidadosamente ideados, así como por sistemas de evaluación que permitan medir rigurosamente los conocimientos y las capacidades de los estudiantes con respecto a todos los objetivos fijados en los estándares.

Las Common Core State Standards, recientemente adoptadas, y demás estándares estatales semejantes de Estados Unidos son muy similares a las metas curriculares comunes planteadas en países como Finlandia, Japón, Singapur y Corea del Sur. Estos países han preparado una serie de estándares, materiales de respaldo curricular y sistemas de evaluación asociados que constituyen un entorno fundamental para el desarrollo y la evaluación de los docentes. Gran parte de la formación profesional continua de los docentes ocurre mientras desarrollan, en colaboración con sus pares, las lecciones y herramientas de evaluación que usan en la sala de clases. Estas instancias de aprendizaje colectivo cobran un carácter muy analítico e intensivo si se adoptan estrategias como el *lesson study* ('estudio de lecciones'), la investigación-acción sobre las prácticas, o los *learning circles* ('círculos de aprendizaje').

El consenso sobre las metas de aprendizaje permite lograr el alineamiento de los estándares de enseñanza, tal como han hecho el NBPTS y el INTASC. Estos esfuerzos por lograr cohesión entre los estándares de aprendizaje y los estándares de enseñanza se expresan en términos de desempeño, es decir, qué deben saber y ser capaces de hacer los docentes para ayudar al estudiante a aprender, en vez de cuántas horas de clases y talleres deben impartir para ganar puntos. De esa forma, se ha comenzado a basar la formación y el desarrollo de los docentes en la práctica.

No obstante, poco se ha hecho para vincular esos estándares a las evaluaciones in situ de los profesores en cada distrito. En efecto, las evaluaciones locales varían considerablemente de un distrito a otro, y muchas veces se basan en listas de comportamientos de los profesores que no guardan relación con la eficacia. De ahí que los docentes se enfrenten a numerosas señales inconexas en el transcurso de su carrera y pierdan sistemáticamente la oportunidad de desarrollar prácticas de calidad. Con un enfoque exhaustivo de la evaluación docente se crearían sistemas de evaluación mucho más útiles para el otorgamiento de licencias estatales y de certificados avanzados, y se usarían dichos sistemas como un marco para la organización de evaluaciones in situ más pertinentes.

Crear sistemas de evaluación basados en el desempeño

Se sabe que los sistemas de evaluación basados en el desempeño, cuando están bien diseñados, permiten medir aspectos de la eficacia de la docencia reflejados en los logros de los estudiantes. Entre ellos figuran los *sistemas de evaluación de desempeño de docentes* estandarizados como los que se usan en la certificación nacional otorgada por el NBPTS y en las licencias que se otorgan a profesores principiantes en estados como Connecticut y California, así como los *sistemas de evaluación de profesores basados en estándares* usados en algunos distritos (analizados más adelante). La ventaja de estos sistemas es que ambos permiten *documentar* y ayudar a los profesores a alcanzar un mayor nivel de eficacia, pues la participación en estos procesos contribuye al aprendizaje tanto de los docentes que se someten a la evaluación como de los formadores a quienes se capacita para que hagan las evaluaciones.

Idealmente, los estados deberían crear un sistema de licencias escalonado que permita habilitar a los profesores nuevos y reconocer a los docentes experimentados en función de su desempeño; un sistema que sirva de marco para el aprendizaje y el desarrollo profesional que se busca medir con las evaluaciones locales. Se usaría una evaluación estatal de desempeño de la enseñanza en el Nivel 1 para la licencia inicial que da derecho a ejercer la profesión (generalmente llamada *licencia preliminar* o *provisoria*), y luego nuevamente en el Nivel 2 para demostrar que se concluyó el período de inducción con éxito y que el docente alcanzó el nivel de competencia necesario para la obtención de la licencia profesional. Por lo general esto sucede antes de que los distritos locales evalúen la asignación de cargos titulares, y puede influir en dicha decisión.

Algunos estados tienen un tercer nivel de reconocimiento que certifica la experiencia y permite a los docentes asumir mayores responsabilidades, como las de tutor o profesor guía (*master teacher*), que pueden implicar compensaciones adicionales. La designación a estos cargos a veces depende de la certificación nacional del NBPTS y otras veces de las evaluaciones estatales. Entre estas dos instancias, se administrarían

evaluaciones locales de desempeño en el trabajo basadas en los mismos estándares, con el fin de garantizar un desarrollo coherente y constante durante toda la carrera del docente.

Nuevo México es un ejemplo de estado que creó un sistema que combina una evaluación estatal de desempeño para el otorgamiento de licencias y evaluaciones locales basadas en los estándares para efectos de gestión de recursos humanos; esto, con la ayuda de asociaciones de docentes, capacitadores e investigadores.

Otorgar licencias en función del desempeño.

Para mejorar la calidad de la formación y las cualidades de los docentes, los estados deberían basar sus decisiones iniciales sobre el otorgamiento de licencias en evidencias más fehacientes del nivel de competencia de los profesores en vez de limitarse a exigir que completen una serie de cursos o que hayan sobrevivido ciertos períodos de tiempo en las salas de clases. Desde la década de los ochenta, la necesidad de otorgar licencias con más confianza se ha traducido en la adopción de exámenes específicos para profesores en todos los estados. Sin embargo, estos exámenes, consistentes generalmente en pruebas de selección múltiple que miden conocimientos y aptitudes básicos, no permiten medir con mucha exactitud la capacidad de los profesores para enseñar con eficacia a los niños. Además, en muchos casos estos exámenes miden los conocimientos de los docentes *antes* de que empiecen o terminen su formación profesional, con lo cual no son una herramienta adecuada para reflejar el nivel de formación de los docentes.

Varios estados han incorporado evaluaciones de desempeño a los procesos de otorgamiento de licencias. Estas evaluaciones del desempeño, que pueden proporcionar información útil para el proceso de acreditación, han resultado ser buenas herramientas para mejorar la formación y la capacitación de los docentes, así como para determinar sus competencias. Por ejemplo, en la Performance Assessment for California Teachers (PACT) los profesores deben documentar su planificación para una unidad de enseñanza en relación con los estándares estatales, adaptarlos a los alumnos sujetos a condiciones especiales y a aquellos que todavía estén aprendiendo inglés, grabar en video y criticar las clases, y recabar y evaluar evidencias del aprendizaje de los alumnos. Luego el material es calificado por profesores de secundaria y de nivel superior especialmente capacitados para ello. La evaluación Connecticut Beginning Educator Support and Training (BEST) usa un sistema similar de portafolios para otorgar la licencia profesional a los profesores principiantes (que llevan dos o tres años ejerciendo).

Al igual que la evaluación nacional del NBPTS, las calificaciones de los profesores principiantes sometidos a las evaluaciones PACT y BEST han resultado ser útiles para predecir la mejora del desempeño de sus alumnos en los exámenes estatales (Newton, 2010; Wilson & Hallum, 2006). En los estados donde ya se han usado, las evaluaciones de desempeño de profesores principiantes no solo han permitido medir aspectos de la eficacia en la enseñanza, sino que también han ayudado a mejorar la eficacia tanto de la labor de los participantes como de los programas de formación a los que asisten.

Una vez concluida la evaluación, los candidatos afirman haber aprendido acerca de la profesión. Las facultades de formación de profesores que participan en la calificación de los participantes describen cómo este proceso genera una comprensión común de la enseñanza de calidad, ayuda a los candidatos a buscar formas de mejorar su formación, y sienta las bases para la planificación de la inducción de los docentes y del desarrollo profesional. Las carreras de pedagogía recopilan información sobre todos sus candidatos por área de estudio y dimensiones de la enseñanza, y la usan para mejorar su currículo, la calidad de la formación y el diseño de los programas (Darling-Hammond, 2010a). Estos datos agrupados, aplicados a la acreditación, pueden constituir una sólida base para decidir qué modelos de programa se debería aprobar y ampliar, y cuáles se debería cerrar en caso de que no mejoren lo suficientemente como para permitir a sus candidatos demostrar que son capaces de enseñar.

Actualmente, más de 25 estados se han unido bajo el alero del Teacher Performance Assessment Consortium para crear un modelo común de evaluación para el otorgamiento de una licencia inicial, inspirada en la labor realizada en esos estados, que pueda usarse a nivel nacional para basar la formación y las acreditaciones en el desempeño y en la capacidad de los docentes para contribuir al aprendizaje de los alumnos.

Usar la certificación nacional del NBPTS para reconocer la experiencia adquirida

En un principio se desarrolló un enfoque basado en estándares para la evaluación de los docentes, a través de la labor del NBPTS, que desarrolló estándares para una enseñanza de calidad en más de 30 áreas de enseñanza definidas por asignatura y nivel de desarrollo de los alumnos. A continuación, la entidad desarrolló evaluaciones para reunir evidencias de la práctica y el desempeño de los docentes en un portafolio que incluye videos de clases con comentarios, planes de lecciones y evidencias del aprendizaje de los alumnos. Estas evidencias son clasificadas por evaluadores capacitados y expertos en el mismo campo pedagógico usando rúbricas que definen dimensiones pedagógicas fundamentales, tales como la planificación basada en el conocimiento del aprendizaje de los alumnos; clases que usan estrategias eficaces para atender las necesidades de los alumnos; y evaluaciones y retroalimentaciones que inciden en los planes futuros y permiten a los alumnos mejorar su trabajo. Varios estados han adoptado la certificación nacional del NBPTS, diseñada para identificar a los profesores experimentados, como criterio para el otorgamiento de bonos salariales y otras formas de reconocimiento, como la designación de mentores o profesores titulares.

Algunos estudios arrojaron que el proceso de evaluación para obtener la certificación nacional del NBPTS permite distinguir a aquellos profesores que posteriormente son más eficaces a la hora de aumentar el rendimiento de los estudiantes que los que no logran obtener la certificación (Bond, Smith, Baker, & Hattie, 2000; Cavaluzzo, 2004; Goldhaber & Anthony, 2005; Smith, Gordon, Colby, & Wang, 2005; Vandevort, Amrein-Beardsley, & Berliner, 2004).

Igualmente importantes son los numerosos estudios que han demostrado que la participación de los docentes en el proceso del NBPTS contribuye a su formación profesional y fomenta cambios en su práctica. En efecto, los docentes se dan cuenta de que el proceso de análisis de su propio trabajo y del de sus alumnos en función de los estándares les permite evaluar mejor el aprendizaje de los alumnos, así como a medir los efectos de sus propias acciones y a modificarlas, de ser necesario. Por último, los docentes muchas veces desarrollan nuevas prácticas recomendadas en los estándares y evaluaciones (Athanasos, 1994). Los profesores relatan mejoras considerables en su desempeño en cada área evaluada, a saber, planificación, diseño e impartición de clases; gestión de cursos; diagnóstico y evaluación del aprendizaje de los alumnos; aplicación de los conocimientos pedagógicos; y participación en una comunidad de aprendizaje, y los estudios observacionales han demostrado que estos cambios efectivamente ocurren (Sato, Wei, & Darling-Hammond, 2008; Tracz, Sienty, Todorov, Snyder, Takashima, Pensabene, Olsen, Pauls, & Sork, 1995; Tracz, Sienty, & Mata, 1994).

Además, la participación de un establecimiento entero en el proceso de certificación nacional del NBPTS puede ayudar a los profesores a consolidar su eficacia colectiva. Algunos establecimientos han hecho participar a todo su cuerpo docente en el proceso para que incrementen su experiencia, y han registrado mejoras considerables en el nivel de retención del personal y en el rendimiento de los alumnos (Dean & Darling-Hammond, en preparación).

Idear un sistema de evaluación local basado en estándares

Si los sistemas de evaluación docente de los distritos se basaran en los mismos estándares que los de los sistemas estatales de certificación y otorgamiento de licencias, ambos contribuirían al fortalecimiento de la formación y desarrollo de los docentes. Estos sistemas locales de evaluación deberían tener tres componentes:

1. *Evaluaciones de las prácticas basadas en estándares.* Estas evaluaciones deberían basarse no solo en observaciones o grabaciones en video de la práctica en clase, sino también en una comprensión de los planes curriculares de los docentes, de sus tareas y de sus apreciaciones en relación con el contenido que imparten y los estudiantes a quienes enseñan.
2. *Evidencias de la contribución de los profesores a la labor de sus pares y del establecimiento como un todo.* La educación es una tarea colectiva y el aprendizaje de los alumnos es el resultado de los esfuerzos concertados de los profesores, por lo que estas contribuciones son valiosas y deberían incluirse en el proceso de evaluación.
3. *Evidencias de la contribución de los profesores al aprendizaje de los alumnos.* Estas evidencias deberían

evaluarse usando múltiples fuentes de información procedentes tanto de la documentación originada en la sala de clases como de otras evaluaciones, que sean apropiadas y válidas para el currículo y para los alumnos.

Estos datos deben cotejarse unos con otros, de manera integrada, a la luz del contexto pedagógico (Skinner, 2010).

Evaluaciones de la práctica docente basadas en estándares

Las evaluaciones de la práctica basadas en estándares que se practican en algunos distritos están significativamente relacionadas con el progreso de los alumnos, y ayudan a los profesores a mejorar su práctica y su eficacia (Milanowski, Kimball, & White, 2004). Al igual que los sistemas de evaluación de desempeño de docentes descritos anteriormente, estos sistemas de observación de la práctica de los profesores en la sala de clases se basan en estándares de enseñanza profesional fundados en investigaciones sobre docencia y aprendizaje. En ellos se utilizan protocolos sistemáticos de observación, administrados por evaluadores capacitados, con el fin de examinar la práctica docente en una serie de dimensiones, tales como la organización y la gestión de la clase, la planificación bien organizada del currículo, la capacidad de transmitir el contenido a los alumnos de manera accesible, la evaluación del aprendizaje de los estudiantes, y la diferenciación de la enseñanza para atender las distintas necesidades de los alumnos.

Estos protocolos generalmente proporcionan indicadores de eficacia pedagógica relacionados con una serie de estándares profesionales lo suficientemente concretos para ayudar a formular observaciones y proporcionar retroalimentación a los docentes; dichos estándares describen prácticas cuyos resultados en el aprendizaje de los alumnos se han demostrado en investigaciones. Estos protocolos difieren de los cuestionarios abiertos que permiten a los evaluadores determinar prioridades en la sala de clases desde el punto de vista idiosincrásico y también de los enfoques conductistas tradicionales que describen conductas aisladas de enseñanza que pueden o no fomentar el aprendizaje (como tener una pizarra ordenada, mantener un ritmo acelerado de enseñanza o tener metas claramente definidas).

En un estudio llevado a cabo en tres distritos que usan sistemas de evaluación basados en estándares, los investigadores encontraron importantes correlaciones positivas entre los puntajes de los profesores y el progreso de sus alumnos en pruebas estandarizadas (Milanowski, Kimball, & White, 2004). En estos distritos, se basaron las evaluaciones formativas y sumativas de los docentes en estándares de práctica bien formulados a partir de observaciones de la enseñanza, entrevistas pre- y postobservación y, en algunas ocasiones, de elementos como planes de lecciones, tareas y muestras de los trabajos de los alumnos.

Un aspecto crucial para que estos sistemas funcionen es que incluyan observaciones hechas en distintas salas de clases por parte de evaluadores expertos a lo largo del año. Estos sistemas buscan múltiples fuentes de datos que reflejen la práctica docente del profesor, y proporcionan retroalimentación oportuna y valiosa a este último. En los establecimientos más grandes, donde puede haber hasta cien profesores, es imposible para un director llevar a cabo evaluaciones intensivas como estas para cada profesor todos los años. (En el ámbito empresarial, por ejemplo, por lo general se considera que para tener un nivel de control adecuado se necesita un supervisor para cada siete empleados). En los distritos donde se logró solucionar la brecha entre la necesidad de evaluaciones de alta calidad y la disponibilidad de los directores se recurrió a subdirectores, coordinadores de departamentos y profesores guía, quienes asumieron como evaluadores, o bien se creó un ciclo de evaluaciones más intensivas por períodos (por ejemplo, cada dos o tres años) para los docentes bien calificados.

Los estudios sobre los sistemas de evaluación docente basados en estándares sugieren que, cuanto más el comportamiento y la labor de los docentes en las salas de clases reflejan los estándares profesionales de práctica, más contribuyen al aprendizaje de los alumnos (Milanowski, Kimball, & White, 2004). Estos resultados llevaron a un investigador a concluir que el hecho de supeditar el ascenso y las compensaciones de los docentes a sus conocimientos y habilidades —y luego usar sistemas de evaluación que ayuden a desarrollar dichas habilidades— a la larga puede producir más cambios positivos en la práctica que la mera evaluación basada en los puntajes de los alumnos (Hassel, 2002). Sus conclusiones son acertadas: varios estudios importantes han demostrado recientemente que el desempeño de los alumnos no mejora cuando se evalúa y se recompensa a los docentes únicamente por mejorar los puntajes de sus alumnos (Fryer,

2010; Martins, 2009; Springer, Ballou, Hamilton, Le, Lockwood, McCaffrey, Pepper, & Stecher, 2010).

Las evaluaciones docentes basadas en estándares pueden complementarse con otras medidas, que también toman en cuenta factores contextuales como el contenido enseñado y los alumnos, a saber:

- aportes a las metas del establecimiento mediante el trabajo con los pares, los apoderados y los alumnos; contribuciones al desarrollo y alineamiento curricular; iniciativas para mejorar el establecimiento en sí; involucramiento en instancias de desarrollo profesional más amplio; e
- iniciativas para incrementar el aprendizaje de los alumnos (basadas en evaluaciones y documentación originadas en clase, en muestras de los trabajos de los alumnos y en exámenes estandarizados, ya sean exámenes propios del colegio, distritales o estatales, siempre que sean válidos y pertinentes).

Aportes a las metas del establecimiento

Es importante recordar que la educación es un trabajo colectivo y que los establecimientos exitosos son aquellos que logran una combinación adecuada de destrezas y habilidades y que permiten a la gente colaborar entre sí.

Más del 90% de los profesores del país afirma que sus colegas aportan a la eficacia de su enseñanza (Metlife Foundation, 2009). En un estudio realizado recientemente por economistas, se cuantificó el incremento provocado en el aumento del aprendizaje de los alumnos por la experiencia colectiva de equipos de profesores. El estudio arrojó que los avances más importantes se atribuyen a profesores más experimentados y mejor calificados capaces de trabajar en equipo. Asimismo, se determinó que el aprendizaje entre pares en pequeños grupos de profesores parece ser el factor que más permite predecir mejoras en el desempeño de los alumnos a largo plazo (Jackson & Bruegmann, 2009). Otro estudio reciente concluyó que los alumnos obtienen mejores resultados en matemática y lectura cuando van a un establecimiento donde hay mayor nivel de colaboración entre profesores para mejorar el nivel del establecimiento (Goddard & Goddard, 2007).

Por consiguiente, cualquier sistema de evaluación debe documentar y reconocer no solo los conocimientos y habilidades de los docentes, sino también su contribución al trabajo del establecimiento en general. Dicha contribución puede incluir *conocimientos y habilidades* específicos, su participación en *prácticas pedagógicas compartidas* o iniciativas específicas de apoyo a los alumnos, y su aporte al *aprendizaje colectivo y a la mejora del establecimiento*.

Conocimientos y habilidades

Los establecimientos escolares necesitan una mezcla de conocimientos y habilidades, entre otras cualidades, para informar las decisiones curriculares y atender las necesidades de los alumnos. Además de las materias y la pedagogía que los profesores suelen aprender en su área de especialidad, en muchos establecimientos también puede ser necesario contar con conocimientos especializados para apoyar a los alumnos que todavía estén aprendiendo inglés o que requieran una educación diferencial. Asimismo, saber las lenguas que los alumnos hablan en casa es fundamental para poder comunicar con sus apoderados y especialmente útil cuando permiten a los profesores hacer participar a los niños y sus familias y a ayudar a sus pares a hacer lo mismo. En algunos contextos puede ser importante tener experiencia en técnicas pedagógicas específicas, tales como las de apoyo a la alfabetización, pues ayudan a los estudiantes y a los colegas a mejorar sus resultados.

Este enfoque contrasta con los esquemas de evaluación y compensación basados directamente en los puntajes obtenidos por los alumnos en los exámenes, los que pueden traer una serie de consecuencias imprevistas y disfuncionales, que se describen a continuación. Odden, Kelley, Heneman y Milanowski (2001) señalan:

Los sistemas de compensación basados en los conocimientos y habilidades proporcionan un mecanismo para vincular las remuneraciones a los conocimientos y habilidades (y, por extensión, al desempeño) deseados en los profesores [...]. El concepto de *compensación basado en conocimientos y habilidades* proviene del sector privado, donde se desarrolló para

incitar a los empleados a adquirir habilidades nuevas y más complejas, o requeridas por un empleador específico. La remuneración basada en conocimientos y habilidades también tenía por objetivo fortalecer una cultura organizacional que valorara el crecimiento y el desarrollo personal de los empleados, y definir un camino claro para aumentar las competencias profesionales (pp. 1-2).

Odden et al. (2001) dan varios ejemplos de planes de evaluación y compensación basados en conocimientos y habilidades. Por ejemplo, en Coventry (Rhode Island) se ofrecen remuneraciones por contar con la certificación nacional del NBPTS y por desarrollar habilidades en *pedagogía auténtica*, autorreflexión, educación diferencial, y participación familiar y comunitaria; todas las estrategias vinculadas con los logros de los alumnos, según las investigaciones. El condado de Douglas (Colorado) ofrece compensaciones por asistir a cursos sobre metas del distrito relacionadas, por ejemplo, con la evaluación de los alumnos o la enseñanza orientada a grupos heterogéneos. Vaughan Learning Center, un establecimiento particular subvencionado de Los Ángeles (California), ofrece incentivos a quienes tengan títulos y certificaciones pertinentes, así como a aquellos docentes que aporten conocimientos y habilidades relevantes para la misión del establecimiento, por ejemplo, docentes formados en alfabetización, en la enseñanza del inglés como segunda lengua, en la inclusión de alumnos de educación diferencial y en tecnología.

Prácticas pedagógicas compartidas

Los establecimientos que busquen un enfoque más coherente de la pedagogía tienen que instar a los profesores a compartir sus prácticas, principalmente aquellos que tienen un impacto positivo en el rendimiento de los estudiantes. Por ejemplo, el uso de evaluaciones formativas para proporcionar retroalimentación a los alumnos y darles la oportunidad de revisar su trabajo ha resultado, en varias docenas de estudios, tener un efecto muy positivo en el aprendizaje de los alumnos (Black & William, 1998). Los profesores que enseñan a sus alumnos métodos metacognitivos específicos para desarrollar su lectura, su redacción y su capacidad de resolución de problemas matemáticos han ayudado a aumentar su capacidad para adquirir habilidades complejas (Darling-Hammond & Bransford, 2005). Los ejemplos abundan.

En algunos sistemas, los profesores obtienen reconocimiento por demostrar que han adoptado nuevas prácticas relacionadas con las metas de su establecimiento o distrito, tales como ejercicios de alfabetización comunes en distintas clases, el uso de evaluaciones formativas en la planificación y modificación de la enseñanza, o la adopción de un nuevo sistema de enseñanza de la escritura. Siempre que se puede se documentan dichas prácticas y se recaban evidencias de cómo estas afectaron a la participación y al aprendizaje de los alumnos. Además de las prácticas pedagógicas específicas, un profesor también puede documentar la forma en que logra aumentar la asistencia o el cumplimiento de las tareas por parte de los alumnos gracias a la concertación periódica de reuniones con los apoderados y llamadas telefónicas. Así, el profesor puede presentar evidencias que demuestren mejoras en los resultados de un alumno, tales como mejores notas, tasas de graduación y el acceso a la educación superior. La lógica que sustenta el uso de estas medidas de eficacia pedagógica yace en el hecho de que contribuyen al desarrollo de los docentes y mejoran las condiciones de aprendizaje de los estudiantes.

Aportes a la labor de los colegas y del establecimiento educativo

Por último, los sistemas de evaluación también deberían tomar en cuenta la forma en que los docentes contribuyen a la labor de sus colegas y del establecimiento en general, tal como se hace en países como Singapur y en muchos distritos exitosos de Estados Unidos. Hay distintas formas de contribuir: desarrollando y compartiendo currículos; emitiendo observaciones y orientando a los colegas para ayudarlos; asumiendo papeles de liderazgo en iniciativas orientadas a mejorar las condiciones de un establecimiento; manteniendo una comunicación permanente con los apoderados; y muchas otras actividades que permitan al equipo docente desarrollar una labor más eficaz. En Singapur, estas actividades pesan considerablemente en las evaluaciones, y constituyen señales de liderazgo que pueden situar a un profesor en distintos niveles de su ascenso profesional (Darling-Hammond & Rothman, 2011).

Evidencias del aprendizaje de los alumnos

Si el aprendizaje de los alumnos es la meta principal de la enseñanza, no cabe duda de que debería ser un factor de peso a la hora de evaluar las competencias de un docente. Una propuesta conocida es la de usar los puntajes de los exámenes estandarizados para la *evaluación de valor agregado* en los estados o distritos como indicador clave del nivel de eficacia de los profesores.

Problemas con las mediciones de valor agregado

Los investigadores han levantado muchos reparos a la opción de basar las evaluaciones individuales de los docentes en resultados de exámenes anuales rendidos por los estudiantes. La simplicidad de los exámenes actuales, que se centran en habilidades básicas y usan principalmente preguntas de selección múltiple, suscita ciertas interrogantes acerca de la tendencia a enseñar específicamente para preparar los exámenes en detrimento de otros tipos de aprendizaje como la escritura, la investigación y la resolución de problemas complejos. En efecto, hay estudios que demuestran que la eficacia medida de los docentes difiere considerablemente en función de los exámenes administrados (Bill & Melinda Gates Foundation, 2010; Corcoran, Jennings, & Beveridge, 2011; Lockwood, McCaffrey, Hamilton, Stetcher, Le, & Martinez, 2007; Rothstein, 2011). Los docentes que obtienen puntajes elevados en las mediciones de valor agregado muchas veces obtienen resultados no tan buenos en exámenes más conceptuales.

También hay otras interrogantes relacionadas con las realidades de la enseñanza y de la educación en general: así como es difícil atribuir el progreso de un estudiante a profesores específicos, también es difícil distinguir la influencia del docente de otros factores como las condiciones del establecimiento y la situación en el hogar, entre otros factores que dependen del alumno. Así, inciden en el aprendizaje los distintos profesores, los apoderados, los tutores y los apoyos pedagógicos extraescolares, además de una serie de factores de la escuela (Braun, 2005; McCaffrey, Koretz, Lockwood, & Hamilton, 2005). En el resumen de su investigación sobre los modelos de valor agregado, Henry Braun (2005) (Educational Testing Service [ETS]), concluye:

Siempre se puede producir estimaciones de lo que el modelo identifica como efectos del docente. Sin embargo, estas estimaciones abarcan las contribuciones de una serie de factores, siendo las que dependen de los docentes solo una de ellas. Por consiguiente, termina siendo problemático medir los efectos de la labor del docente y presentarlos como indicadores fidedignos de su eficacia (p. 10).

De hecho, las estimaciones más optimistas concluyen que apenas entre un 7% y un 10% de la variación total en el rendimiento estudiantil puede atribuirse a un profesor específico del alumno (Rivkin, Hanushek, & Kain, 2000). Los factores más importantes, que suelen representar cerca del 60% de dicha variación, son factores socioeconómicos relacionados con la situación de cada alumno y con la composición colectiva de cada sala de clases o establecimiento. El resto o bien corresponde a factores escolares ajenos a los docentes (por ejemplo, las influencias de currículos específicos, la disponibilidad de materiales pedagógicos útiles, el tiempo que los docentes tienen para dedicar a los alumnos, el tamaño de las clases y la disponibilidad de especialistas en enseñanza), o bien no se explica. Además, en las evaluaciones hechas todos los años en la mayoría de los estados, los rezagos propios del verano que suelen afectar considerablemente los puntajes de los alumnos de menores ingresos también reducen su nivel de progreso, cosa que suele atribuirse al profesor del año siguiente (Alexander, Entwisle, & Olson, 2007).

Los estudios han concluido que, debido a los numerosos factores que pesan en el aprendizaje de los alumnos, la calificación de la eficacia de los docentes es muy variable. Un estudio realizado en cinco grandes distritos urbanos reveló, por ejemplo, que de los profesores mejor evaluados en un año (el 20% mejor calificado), tan solo entre un 25% y un 35% ocupaba un lugar similar en la clasificación del año siguiente, mientras que una proporción comparable caía a los últimos lugares. Y entre los peor evaluados, solo entre un 20% y un 30% quedaba en posición similar el año siguiente, mientras que entre un 25% y un 45% ascendía a los primeros lugares en un período de un año (Sass, 2008).

Si las clasificaciones por valor agregado realmente midieran las competencias básicas o la eficacia de un docente, no debería haber movimientos tan bruscos. De hecho, hay investigaciones que apuntan a que los aumentos de puntaje en los exámenes asociados con la labor de los docentes se ven afectados por

diferencias entre los estudiantes que se les asignan, aun cuando dichas diferencias quedan supuestamente «controladas» en el modelo estadístico. Un mismo profesor resulta ser más eficaz cuando imparte clases a alumnos más aventajados que cuando enseña a alumnos con distintos problemas educativos, principalmente aquellos con necesidades especiales o que están aprendiendo el idioma inglés (Newton, Darling-Hammond, Haertel, & Thomas, 2010).

Estas distorsiones se traducen en mediciones equivocadas de la eficacia pedagógica y, por lo tanto, terminan desincentivando a los profesores a hacerse cargo de los estudiantes que más lo necesitan. Este desincentivo puede llegar a consolidar la tradicional práctica de asignar profesores inexpertos a los estudiantes más necesitados. También puede exacerbar la tendencia de los establecimientos a rechazar estos estudiantes.

Ocurre algo similar con los profesores que tienen muchos alumnos con necesidades especiales y aquellos que trabajan en programas para alumnos talentosos, y que ya obtuvieron los mejores resultados posibles en el examen (Amrein-Beardsley & Collins, en prensa).

Estos problemas, entre otros, han llevado a las principales entidades de investigación a advertir que el modelo de valor agregado es inadecuado para tomar decisiones sobre cada profesor. Un importante informe de RAND Corporation concluye: «Los hallazgos actualmente son insuficientes para respaldar el uso del modelo de valor agregado para tomar decisiones importantes sobre un docente o un establecimiento» (McCaffrey, Koretz, Lockwood, & Hamilton, 2005, p. 20). Y el Board on Testing and Assessment del National Research Council concluye lo siguiente: «No debería usarse el sistema de evaluación de la eficacia de los docentes basada en el modelo de valor agregado para tomar decisiones de operación, pues los resultados son demasiado variables como para considerarse justos o fiables» (National Research Council, Board on Testing and Assessment, 2009, p. 10).

Hay que idear buenos sistemas que no penalicen a los docentes por impartir clases a estudiantes en situación escolar menos aventajada. Sistemas que, al contrario, incentiven directamente a los docentes a trabajar con estos alumnos (mediante reconocimiento y recompensas).

Para comprender cómo los profesores inciden en el aprendizaje de los estudiantes, se necesita más información sobre sus prácticas y el contexto en que se desenvuelven de la que se obtiene con el modelo de valor agregado. Se debe proporcionar evidencias multirrepresentativas del aprendizaje de los alumnos, acompañadas de un análisis del contexto pedagógico y de los alumnos del profesor. Este material debe estar integrado con evidencias adicionales de la práctica docente, y debe estar orientado a mejorar la enseñanza.

Usar otras evidencias del aprendizaje de los alumnos

Las limitaciones del análisis basado en el valor agregado no significan que los distritos no pueden reconocer y recompensar a los profesores por lograr un elevado nivel de aprendizaje entre sus alumnos, o crear incentivos para que ayuden a otros profesores y se preocupen de los alumnos más necesitados. Se pueden usar otras herramientas de medición del aprendizaje, tales como pre- y postests realizados por los distritos o los establecimientos, o incluso material recabado por los mismos profesores.

Dicha evidencia puede proceder de evaluaciones de aula y documentos desarrollados en clase usando instrumentos estandarizados, como por ejemplo el Developmental Reading Assessment, o mediciones pre- y postests en asignaturas o campos curriculares específicos (desarrollados por un profesor, por departamentos, o por el personal de los colegios y distritos). También se puede incluir evidencias de los logros de los alumnos relacionados con determinadas actividades pedagógicas, tales como investigaciones científicas, artículos o proyectos de arte. Algunos distritos se basan en documentación cuidadosamente recopilada por los profesores sobre el aprendizaje de un grupo heterogéneo de alumnos a largo plazo, como la contemplada en los portafolios para postular a la certificación nacional del NBPTS.

Se podría incluir un análisis de resultados de exámenes estandarizados, cuando sea pertinente, resguardando la relación entre los exámenes y el currículo, y su idoneidad para los alumnos en cuestión. Así, se podría necesitar distintos tipos de evaluación para distintos tipos de estudiantes, por ejemplo,

pruebas de inglés para quienes estén aprendiendo el idioma y exámenes alternativos para algunos alumnos con necesidades educativas especiales.

Con esta evidencia el profesor puede constituir un portafolio que demuestre y explique los progresos de los alumnos en un amplio rango de resultados de aprendizaje, de tal forma que se tomen en cuenta los antecedentes y características de cada uno. En algunos establecimientos, los profesores usan sus propias evaluaciones semestrales para medir el progreso de los estudiantes incluyendo indicadores ideados específicamente para las metas de aprendizaje de determinados alumnos (por ejemplo, estudiantes que requieren una educación diferencial o los que estén aprendiendo inglés). Algunos establecimientos y distritos tienen exámenes comunes que usan en determinados niveles y clases. Estas mediciones, en el marco de un portafolio de evidencias, pueden ayudar a documentar la eficacia de la enseñanza para el cumplimiento de metas curriculares específicas.

Las evaluaciones del aprendizaje en asignaturas específicas pueden incluir muestras de trabajos de escritura calificados (con todos los borradores) o de lectura, ejercicios de matemática, pruebas de conocimientos en ciencias o historia, o incluso representaciones de obras musicales. Estas mediciones suelen proporcionar mejores evidencias del aprendizaje de los alumnos en una asignatura o materia determinada porque están ideadas en función del currículo. También es más probable que midan los efectos de la labor de un profesor en particular y esté a la disposición de la mayoría o la totalidad de los alumnos. Un profesor puede incluso documentar los premios que ayudó a sus alumnos a obtener en la competencia científica de Westinghouse u otros logros específicos, siempre poniendo en evidencia su papel en dichos logros.

En Long Beach (California), un distrito reconocido por sus logros y premios, se evalúa a los profesores haciendo un seguimiento de su desempeño en función de los California Standards for the Teaching Profession. Asimismo, los profesores y administradores trabajan juntos en la definición de metas para los estudiantes y la mejora de la práctica docente a nivel del establecimiento, así como metas colectivas a nivel de asignaturas o departamentos y metas individuales. El progreso hacia esas metas se toma en cuenta tanto en las autoevaluaciones como en las evaluaciones administradas por supervisores. Así, la persona evaluada propone cómo se evalúa el logro de sus objetivos, usando la siguiente evidencia:

- Seguimiento y observación de su labor.
- Análisis de antecedentes y casos concretos.
- Logros y progreso en un programa o asignatura definidos.
- Pre- y postests preparados por el profesor, el departamento o el establecimiento.
- Pruebas relacionadas con el currículo.
- Uso de documentación audiovisual, de ser pertinente y estar disponible.
- Autoevaluaciones de los alumnos.
- Reuniones de evaluación con los alumnos y apoderados.
- Antecedentes del desempeño de los alumnos en el pasado.
- Ejemplares de trabajos de los alumnos para ilustrar su evolución.
- Investigación-acción (Long Beach Unified School District, 2003).

El distrito crea oportunidades explícitas y continuas para que los establecimientos, departamentos y grupos a cargo de una asignatura revisen distintos tipos de trabajos y resultados de los alumnos con el fin de medir y comparar la evolución de los cursos, reflexionar acerca de las estrategias pedagógicas y de los currículos, solucionar los problemas de alumnos o grupos de alumnos, y establecer planes de mejora. De esa forma, el sistema fomenta el aprendizaje y el trabajo colectivo.

En el programa de ascenso profesional de Arizona, donde hace más de dos decenios se requiere el uso de evidencias sobre el aprendizaje de los alumnos para evaluar la eficacia de la labor docente en los distritos que participan en el programa, uno de los aspectos clave es el establecimiento de metas por parte del profesor como pauta para recabar evidencias del aprendizaje de los alumnos. Un estudio sobre los programas de ascenso profesional de Arizona reveló que, a la larga, los profesores que participaban en estos programas incrementaban su capacidad de idear instrumentos para evaluar el aprendizaje de los alumnos en las salas de clases, así como de desarrollar y evaluar pre- y postests para definir resultados medibles en áreas difíciles de cuantificar como el arte, la música y la educación física, y de monitorear el aumento del aprendizaje de los alumnos en función de sus planes de acción. También han demostrado ser más concientes de la importancia de desarrollar un currículo sólido, así como de alinearlo con las metas

del distrito y poner énfasis en materias, habilidades y estrategias pedagógicas de alta calidad (Packard & Dereshiwsky, 1991). De esa forma, el desarrollo y uso de evidencia del aprendizaje de los alumnos estuvo asociado con mejoras en la práctica docente.

Crear estructuras que respalden una evaluación eficaz, justa y de alta calidad

Una de las limitaciones más graves de las reformas a los sistemas de evaluación de docentes es que muchas se han centrado en el diseño de herramientas para hacer un seguimiento de la labor de los profesores sin necesariamente prever los elementos estructurales de un sistema de evaluación de calidad. Estos elementos deberían incluir, como mínimo:

- evaluadores bien capacitados,
- asesoría a los profesores que la requieran,
- estructuras de gobierno que permitan tomar decisiones correctas en materia de recursos humanos,
- y recursos para mantener y monitorear el sistema.

Los sistemas de evaluación de alta calidad requieren directores y evaluadores con conocimientos profundos sobre enseñanza y aprendizaje, que sepan cómo evaluar la labor docente, proporcionar retroalimentación útil y planificar el desarrollo profesional con miras a fomentar el aprendizaje de los docentes. La carencia de estos conocimientos y de la preparación necesaria ha sido un importante problema para la validez, la precisión y la utilidad de muchos sistemas de evaluación de profesores. Las respuestas a estas necesidades son múltiples:

- Una mejor *preparación de los directores*, sumada a la aplicación de una licencia profesional para directores basada en el desempeño. El enfoque de Connecticut es un buen ejemplo: la preparación de los directores se centra en el liderazgo pedagógico, la supervisión del equipo docente y el desarrollo profesional. En los procesos de otorgamiento de licencias basados en el desempeño se pide a los directores que demuestren que son capaces de evaluar clases grabadas en video y preparar una retroalimentación apropiada y planes de apoyo profesional.
- *Una capacitación específica e intensiva en evaluación y supervisión*. Hay capacitaciones para evaluadores en estados como Connecticut y programas como el Teacher Advancement Program, que ofrece cursos de varios días para capacitar a los directores en el uso de instrumentos de evaluación estandarizados y estrategias de retroalimentación y seguimiento.
- *Participación de profesores titulares experimentados* en algunos aspectos de las actividades de supervisión y evaluación. Por ejemplo, en algunos distritos se recurre a la participación de coordinadores de departamentos y profesores titulares en función de la materia impartida.

También es importante proporcionar asesoría a los profesores que la necesiten —algo que pocos directores tienen tiempo de prever— e idear procesos y estructuras de toma de decisiones que permitan tomar decisiones justas y eficientes en materia de recursos humanos.

En Rochester (Nueva York) se han adoptado iniciativas innovadoras para proporcionar un mejor conocimiento de las materias en el proceso de evaluación y motivar a los profesores a trabajar y ser evaluados en equipo. Los profesores tienen la opción de ser observados y evaluados por un profesor titular o evaluador capacitado, en caso de que así lo deseen, además de su supervisor administrativo. Los profesores titulares son elegidos por el panel administrativo del programa Rochester Career In Teaching (CIT), consistente en seis administradores designados por el distrito y seis docentes elegidos por el sindicato de profesores.

Los profesores también pueden optar por el Performance Appraisal Review for Teachers (PART). El sistema PART, acordado en Rochester en 1987, conlleva un enfoque colectivo de la evaluación de la labor docente. Se trata de grupos autoseleccionados de profesores de un mismo establecimiento o de distintos colegios que idean un proyecto anual o de varios años para mejorar la calidad de la enseñanza y del aprendizaje, establecen metas para ellos mismos y luego son evaluados, no individualmente sino como miembros de un equipo docente. De esa forma, se fomenta la colaboración y se contextualiza la evaluación en el marco de la labor cotidiana de los profesores. El sistema PART ha permitido fomentar la innovación, los enfoques interdisciplinarios, el aprendizaje por proyectos y las evaluaciones basadas en el desempeño, lo que ha ayudado a avanzar en la construcción de una práctica de enseñanza colectiva.

Desde luego, un sistema de evaluación basado en estándares de práctica profesional también debe ser capaz de remover a quienes no logran cumplir dichos estándares tras recibir asesoría. Los sistemas de evaluación más estables que han tenido éxito en tomar medidas adecuadas en materia de evaluación y recursos humanos tanto para profesores principiantes como experimentados, son aquellos que han usado programas de asesoría y evaluación de pares en los que se recurre a mentores expertos encargados de algunos aspectos de la evaluación y de proporcionar asesoría a los profesores que la necesiten. Estos sistemas de colaboración entre sindicatos y directorios de establecimientos, que comprenden procesos estandarizados y asesorías para los profesores puestos bajo supervisión, son más eficaces que los sistemas de evaluación tradicionales a la hora de mejorar la labor de los docentes o bien de apartarlos de la profesión sin provocar roces con los sindicatos.

Entre los elementos clave de estos sistemas figuran no solo las herramientas de evaluación, sino también la experiencia de los mentores o profesores evaluadores (docentes experimentados en las mismas materias y niveles escolares que los evaluados y que han reservado tiempo para desempeñarse como mentores y así ayudar a sus pares) y el uso de un sistema estandarizado de evaluación en el que participan profesores y administradores que emiten recomendaciones para la toma de decisiones en materia de recursos humanos basándose en evidencias de las evaluaciones. Esta comisión mixta supervisa el trabajo de los mentores que prestan asesoría tanto a los docentes principiantes como a los profesores veteranos que enfrentan dificultades. En función de los informes entregados por los mentores y directores, la comisión decide cuáles profesores acceden a un cargo titular, cuáles deben perfeccionarse durante un año más y cuáles son apartados de la carrera. Asimismo, la comisión decide, antes del término del año escolar, qué profesores puestos bajo supervisión han mejorado lo suficientemente para continuar ejerciendo en el distrito.

En estos sistemas, hay una mayor tasa de docentes principiantes que se mantienen en la profesión gracias a la asesoría recibida, mientras que aquellos que abandonan el distrito generalmente lo hacen por voluntad del distrito más que por renunciar. Entre los profesores veteranos puestos bajo supervisión, muchos logran mejoras suficientes para salir adelante, y entre un tercio y la mitad desiste por voluntad propia o a pedido del distrito. Como las asociaciones de profesores participan en la creación y gestión de los programas, diseñados para garantizar un proceso justo, se evitan injusticias a la hora de despedir a un profesor.

Generar oportunidades de aprendizaje profesional acordes a las necesidades

Por último, es importante vincular el desarrollo profesional formal y las oportunidades de aprendizaje en el trabajo al sistema de evaluación. Las evaluaciones deberían fomentar continuamente el establecimiento de metas en áreas en las que los profesores deseen mejorar, así como aportar herramientas específicas de apoyo y formación profesional y generar oportunidades para compartir experiencia en el marco del reconocimiento de las fortalezas y necesidades de los profesores.

Además, se debe proporcionar oportunidades de aprendizaje profesional de alta calidad. En vez de los talleres extraescolares llamativos, excesivamente ambiciosos y mal articulados que predominan en Estados Unidos, los profesores deberían poder contar con el tipo de programa de aprendizaje continuo y bien enfocado que ha demostrado ser eficaz para mejorar la enseñanza. En efecto, una revisión de estudios experimentales arrojó que los programas de desarrollo profesional de menos de 14 horas anuales sobre un tema específico no tenían incidencia alguna en el aprendizaje de los alumnos, mientras que los programas de alta calidad de aproximadamente 50 horas en períodos de 6 a 12 meses permitían incrementar el rendimiento estudiantil en 21 puntos porcentuales en promedio (Yoon, Duncan, Lee, Scarloss, & Shapley, 2007).

Estas oportunidades de alta calidad se caracterizan por:

- centrarse en el aprendizaje y la enseñanza de contenidos específicos del currículo,
- organizarse en torno a problemas reales de la docencia,
- estar relacionados con el trabajo que desarrollan los docentes con los niños,
- incluir análisis de la enseñanza y del aprendizaje de los alumnos,
- ser intensivos, continuos y a largo plazo,

- estar respaldados por sesiones de capacitación, modelamiento de prácticas, observación y retroalimentación,
- contemplar una labor colaborativa de los profesores en el marco de comunidades de aprendizaje profesional,
- y estar incorporados en el establecimiento y en las clases a la planificación del currículo, de la docencia y de la evaluación.

Dichas oportunidades pueden incluir centros de formación intensiva centrados en estrategias específicas o en la enseñanza de determinados currículos, intercalados con oportunidades para que los docentes hagan ensayos en la sala de clases, obtengan asesoría, reflexionen juntos sobre sus experiencias, revisen y mejoren sus métodos, y afinen sus habilidades con ciclos iterativos de práctica, reflexión y perfeccionamiento. Asimismo, pueden comprender instancias para analizar colectivamente las observaciones, los videos de clases o muestras de trabajos de los alumnos; grupos de estudio; proyectos de investigación-acción; observaciones de colegas; y la planificación y evaluación colectiva en equipos por nivel escolar o asignatura (Darling-Hammond & Richardson, 2009).

Las oportunidades para estudiar y aprender estrategias específicas fundamentales para la práctica individual son importantes y pueden desarrollarse en cursos específicos o participando en redes o conferencias sobre una materia determinada. También es importante organizar instancias de aprendizaje profesional que generen capacidad colectiva y coherencia curricular entre los miembros del equipo docente de un establecimiento. Los mejores sistemas prevén períodos para que los profesores trabajen y aprendan juntos durante la jornada escolar, como suele hacerse en países de Europa y Asia con elevados niveles de calidad, donde los docentes suelen tener entre 15 y 25 horas semanales para planificar su labor y trabajar juntos (Darling-Hammond, 2010b). Con el tiempo, estas iniciativas colectivas producen mejores resultados para el aprendizaje de los alumnos que los esfuerzos producidos en forma individual por cualquier profesor.

Estas oportunidades de aprendizaje profesional también deberían ser concebidas en el marco de una carrera profesional que incite a los profesores a adquirir y compartir experiencia y cree los canales necesarios para ello, de manera que los profesores asuman papeles de mentor y profesor guía, de especialista en currículo y evaluación, y de coordinadores de actividades orientadas a mejorar el establecimiento.

Por otro lado, los establecimientos y distritos deben desarrollar las condiciones que permitan a los profesores y directores tener el apoyo organizacional y pedagógico suficiente para desplegar un sistema de evaluación de docentes que propicie el aprendizaje continuo. Por ejemplo, los profesores y directores deben contar con el tiempo y la orientación necesarios para desarrollar una concepción común de enseñanza eficaz, examinar su práctica en busca de evidencias de aprendizaje, debatir sobre sus formas de concebir el aprendizaje y sobre lo que consideran ser una evidencia de aprendizaje, promover la reflexión y aprender a proporcionar retroalimentación útil. Sin estas condiciones, que dependen del entorno escolar, cualquier programa de evaluación docente resulta insuficiente para incentivar la formación continua.

El reto para la implementación generalizada de una buena práctica pedagógica (Elmore, 1996) yace en el desarrollo de un amplio liderazgo educativo y en la formación de expertos, por un lado, y en la capacidad de fomentar el diseño de organizaciones eficaces que sustenten todo el sistema, por el otro. De ahí que se deba concebir el perfeccionamiento de la práctica como un esfuerzo colectivo antes que individual.

Para transformar los sistemas, se debe estructurar los incentivos de manera que fomenten la colaboración y la divulgación de conocimiento entre las organizaciones, en vez de la competencia. Es necesario compartir los conocimientos para desarrollar no solo organizaciones de aprendizaje sino un sistema educativo orientado al aprendizaje. Esta ha sido la estrategia principal que ha permitido las mejoras logradas en Finlandia, por ejemplo, donde se estimula la evaluación y el estudio de la práctica en y entre las salas de clases, entre establecimientos asociados en las distintas regiones y en el sistema en general (Darling-Hammond, 2010b).

Discusión

En resumen, se debe concebir los sistemas de evaluación docente no solo como instrumentos o procesos de evaluación, sino también en el marco de los sistemas de políticas en los cuales se desarrollan y de las condiciones escolares necesarias para fomentar el aprendizaje y mejora continuos. Para ello, se debe:

- hacer participar a los docentes en el desarrollo del sistema y en los procesos de toma de decisiones;
- reconocer y fomentar las contribuciones colectivas al desempeño general del establecimiento, así como desarrollar criterios claros de éxito que todos los profesores interesados puedan cumplir, en vez de sistemas de cuotas que hagan competir a los profesores entre ellos;
- reunir evidencias válidas de la eficacia pedagógica basadas en variadas mediciones y fuentes de información, tales como:
 - evaluaciones de la práctica fundadas en observaciones de distintas salas de clases y en el análisis de materiales de otras clases (por ejemplo, planes de estudio, tareas para los alumnos y muestras de trabajos) mediante herramientas de evaluación basadas en estándares que proporcionen muestras de la planificación, la enseñanza, el entorno de aprendizaje y la evaluación de los alumnos;
 - evidencias de la contribución de los profesores al aprendizaje de sus alumnos mediante una serie de herramientas válidas que evalúen adecuadamente el currículo impartido;
 - evidencias de la contribución de los profesores a la labor de sus pares y del establecimiento;
- tomar en cuenta las necesidades de los alumnos y efectuar una evaluación válida y apropiada de la enseñanza y del aprendizaje de todos los alumnos, incluidos aquellos con necesidades especiales y los que estén aprendiendo inglés;
- propiciar oportunidades de aprendizaje profesional continuo y de alta calidad que permitan a los profesores cumplir los estándares requeridos.

A medida que se ha hecho más apremiante la necesidad de mejorar los resultados de los alumnos, han surgido muchas iniciativas para medir y mejorar la eficacia de la enseñanza. Estas iniciativas rendirán más frutos cuanto más fomenten prácticas que se sabe que contribuyen al aprendizaje de los alumnos y se incorporen a sistemas que también *desarrollen* una mayor competencia pedagógica. Tales sistemas se basarán en estándares de enseñanza profesional y en la enseñanza de materias enmarcadas en currículos bien definidos, sacarán el máximo provecho de las asesorías y generarán numerosas oportunidades para que los profesores ayuden a sus pares y a los establecimientos a mejorar. En última instancia, son las políticas que generan medidas cada vez más válidas para garantizar la eficacia de la enseñanza —y que crean sistemas innovadores para reconocer, desarrollar y aprovechar la labor de los docentes expertos— las que ayudan a crear una profesión docente más eficaz.

El artículo original fue recibido el 17 de agosto de 2012
El artículo revisado fue recibido el 3 de septiembre de 2012
El artículo fue aceptado el 5 de octubre de 2012

Referencias

- Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2007). Lasting consequences of the summer learning gap. *American Sociological Review*, 72, 167-180.
- Amrein-Beardsley, A., & Collins, C. (en prensa). *The SAS® Education Value-Added Assessment System (EVAAS®): Its intended and unintended effects in a major urban school system*. Tempe, AZ: Arizona State University.
- Athanases, S. (1994). Teachers' reports of the effects of preparing portfolios of literacy instruction. *Elementary School Journal*, 94(4), 421-439.
- Bill & Melinda Gates Foundation (2010). *Learning about teaching: Initial findings from the Measures of Effective Teaching Project*. Seattle: Autor.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-75.
- Bond, L., Smith, T., Baker, W., & Hattie, J. (2000). *The certification system of the National Board for Professional Teaching Standards: A construct and consequential validity study*. Greensboro, NC: Center for Educational Research and Evaluation.
- Braun, H. (2005). *Using student progress to evaluate teachers: A primer on value-added models*. Princeton, NJ: ETS Policy Information Center.
- Cavaluzzo, L. (2004). *Is National Board Certification an effective signal of teacher quality?* (National Science Foundation No. REC-0107014). Alexandria, VA: The CNA Corporation.
- Clotfelter, C., Ladd, H., & Vigdor, J. (2007). *How and why do teacher credentials matter for student achievement?* (NBER Documento de trabajo N°12828). Cambridge, MA: National Bureau of Economic Research.
- Corcoran, S. P., Jennings, J. L., & Beveridge, A. A. (2011). *Teacher effectiveness on high- and low-stakes tests*. Documento de trabajo. Nueva York: New York University.
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Educational Policy Analysis Archives*, 8(1), 1-44.
- Darling-Hammond, L., & Bransford, J. (2005). *Preparing teachers for a changing world: What teachers should learn and be able to do*. San Francisco: Jossey-Bass.
- Darling-Hammond, L., & Richardson, N. (2009). Teacher learning: What matters? *Educational Leadership*, 5(66), 46-53.
- Darling-Hammond, L. (2010a). *Evaluating Teacher Effectiveness: How Teacher Performance Assessments can Measure and Improve Teaching*. Washington DC: Center for American Progress. Para más información ver <http://tpafieldtest.nesinc.com/>
- Darling-Hammond, L. (2010b). *The flat world and education: How America's commitment to equity will determine our future*. Nueva York, NY: Teachers College Press.
- Darling-Hammond, L., & Rothman, R. (2011). *Teacher and leader effectiveness in high-performing education systems*. Washington, DC: Alliance for Excellent Education and Stanford, CA: Stanford Center for Opportunity Policy in Education.
- Dean, S., & Darling-Hammond, L. (en preparación). *Using National Board Certification to Improve Schools*. Stanford, CA: National Board Resource Center and the Stanford Center for Opportunity Policy in Education.
- Elmore, R. (1996). Getting to scale with good educational practice. *Harvard Educational Review*, 66(1), 1-26.
- Fryer, R. G. (2011). *Teacher incentives and student achievement: Evidence from New York City Public Schools* (NBER Documento de trabajo N° 16850). Cambridge, MA: National Bureau of Economic Research.
- Goddard, Y., & Goddard, R. D. (2007). A theoretical and empirical investigation of teacher collaboration for school improvement and student achievement in public elementary schools. *Teachers College Record*, 109(4), 877-896.
- Goldhaber, D., & Anthony, E. (2005). *Can teacher quality be effectively assessed?* Seattle, WA: University of Washington and the Urban Institute.
- Goldhaber, D., & Brewer, D. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis*, 22(2), 129-145.
- Hassell, B. C. (2002). *Better pay for better teaching: Making teacher compensation pay off in the age of accountability*. Progressive Policy Institute 21st Century Schools Project. Recuperado el 18 de noviembre de 2004 de <http://www.broadfoundation.org/investments/education-net.shtml>

- Jackson C. K., & Bruegmann, E. (2009, agosto). *Teaching students and teaching each other: The importance of peer learning for teachers*. Washington, DC: National Bureau of Economic Research.
- Lockwood, J. R., McCaffrey, D. F., Hamilton, L.S., Stetcher, B., Le, V. N., & Martinez, J. F. (2007). The sensitivity of value-added teacher effect estimates to different mathematics achievement measures. *Journal of Educational Measurement*, 44 (1), 47-67.
- Long Beach Unified School District (2003). *Teacher evaluation handbook*. Long Beach, CA: Autor.
- Martins, P. (2009). *Individual teacher incentives, student achievement and grade inflation* (Documento de discusión N° 4051). Londres, Reino Unido: Queen Mary, University of London, CEG-IST and IZA.
- McCaffrey, D. F., Koretz, D., Lockwood, J. R., & Hamilton, L. S. (2005). *Evaluating value-added models for teacher accountability*. Santa Mónica, CA: RAND Corporation.
- McCaffrey, D. F., Lockwood, J., Koretz, D., & Hamilton, L. (2005). *Evaluating value-added models for teacher accountability*. Santa Mónica, CA: RAND Corporation.
- MetLife Foundation (2009). *The MetLife survey of the American teacher: Collaborating for student success*. Nueva York: Autor.
- Milanowski, A., Kimball, S. M., & White, B. (2004). *The relationship between standards-based teacher evaluation scores and student achievement*. Wisconsin: University of Wisconsin-Madison, Consortium for Policy Research in Education.
- National Research Council, Board on Testing and Assessment (2009). *Letter report to the U.S. Department of Education*. Washington, DC: Autor.
- Newton, S. P. (2010). *Predictive validity of the performance assessment for California teachers*. Stanford, CA: Stanford Center for Opportunity Policy in Education. Recuperado de <http://scale.stanford.edu/>
- Newton, X., Darling-Hammond, L., Haertel, E., & Thomas, E. (2010). Value-added modeling of teacher effectiveness: An exploration of stability across models and contexts. *Educational Policy Analysis Archives*, 18 (23). Recuperado de <http://epaa.asu.edu/ojs/article/view/810>
- Odden, A., Kelley, C., Heneman, H., & Milanowski, A. (2001, noviembre). *Enhancing teacher quality through knowledge- and skills-based pay, CPRE Policy Briefs*. Philadelphia: Consortium for Policy Research in Education, University of Pennsylvania.
- Packard, R., & Dereshiwsky, M. (1991). *Final quantitative assessment of the Arizona career ladder pilot test project*. Flagstaff, AZ: Northern Arizona University.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2000). *Teachers, schools, and academic achievement* [Revised] (Documento de trabajo N° 6691). Cambridge, MA: National Bureau of Economic Research.
- Rothstein, J. (2011). *Review of learning about teaching: Initial findings from the measures of effective teaching project*. Boulder, CO: National Education Policy Center.
- Sass, T. (2008). *The stability of value-added measures of teacher quality and implications for teacher compensation policy*. Washington, DC: National Center for Analysis of Longitudinal Data in Education Research.
- Sato, M., Wei, R. C., & Darling-Hammond, L. (2008). Improving teachers' assessment practices through professional development: The case of National Board Certification. *American Educational Research Journal*, 45, 669-700.
- Skinner, K. J. (2010). *Reinventing evaluation: Connecting professional practice with student learning*. Boston, MA: Massachusetts Teachers Association.
- Smith, T., Gordon, B., Colby, S., & Wang, J. (2005). *An examination of the relationship of the depth of student learning and National Board certification status*. Boone, NC: Office for Research on Teaching, Appalachian State University.
- Springer, M. G., Ballou, D., Hamilton, L., Le, V., Lockwood, J. R., McCaffrey, D. F., Pepper, M., & Stecher, B. M. (2010). *Teacher pay for performance: Experimental evidence from the Project on Incentives in Teaching*. Nashville, TN: National Center on Performance Incentives, Vanderbilt University.
- Tracz, S. M., Sienty, S., Todorov, K., Snyder, J., Takashima, B., Pensabene, R., Olsen, B., Pauls, L., & Sork, J. (1995, abril). *Improvement in teaching skills: Perspectives from National Board for Professional Teaching Standards field test network candidates*. Trabajo presentado en la Annual Meeting of the American Educational Research Association, San Francisco, CA.
- Tracz, S.M., Sienty, S., & Mata, S. (1994, febrero). *The self-reflection of teachers compiling portfolios for National Certification: Work in progress*. Trabajo presentado en la Annual Meeting of the American Association of Colleges for Teacher Education, Chicago, IL.
- Vandevort, L. G., Amrein-Beardsley, A., & Berliner, D. C. (2004). National Board certified teachers and their students' achievement. *Education Policy Analysis Archives*, 12(46), 117.

- Wilson, M., & Hallum, P. J. (2006). *Using student achievement test scores as evidence of external validity for indicators of teacher quality: Connecticut's Beginning Educator Support and Training program*. Berkeley, CA: University of California at Berkeley.
- Wilson, S. M., Floden, R., & Ferrini-Mundy, J. (2001). *Teacher preparation research: Current knowledge, gaps, and recommendations*. Seattle: Center for the Study of Teaching and Policy, University of Washington.
- Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement*. Issues & Answers Report, REL 2007–No. 033. Recuperado el 15 de octubre de 2008 de http://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/REL_2007033.pdf