



Serie Evidencias:

Evaluación Docente: Su estado actual y pasos futuros

8 de noviembre de 2013

Año 2, N° 22

Existe consenso respecto del fundamental rol que cumplen los profesores en la mejora de la calidad de la educación. Para perfeccionar los procesos de enseñanza-aprendizaje se requiere conocer las fortalezas y debilidades de dicho proceso, para lo cual la evaluación docente es primordial.

En este número se revisan los resultados y recomendaciones de dos estudios respecto del sistema de evaluación docente chileno – uno de la Organización para la Cooperación y el Desarrollo Económico (OCDE) y otro del Ministerio de Educación (MINEDUC) en conjunto con el Programa de Naciones Unidas para el Desarrollo (PNUD)– los que, entre otras cosas, indican la necesidad de revisar y mejorar los instrumentos utilizados a la luz de sus objetivos.

Asimismo, y con el objeto de orientar el perfeccionamiento de este instrumento de medición, se realiza una revisión de la literatura respecto de los objetivos e instrumentos utilizados para este fin por diversos países e instituciones.

1. Introducción

Existe consenso respecto del importante rol que juegan los docentes en los sistemas escolares, siendo el conocido Reporte Coleman (1966)¹ uno de los primeros estudios en educación que dio cuenta de la magnitud de su influencia en los resultados de los estudiantes. Posteriormente, Goe L. (2007)², en una revisión de las investigaciones existentes, concluye que la calidad docente (medida de diversas formas³) afecta positivamente los resultados de los estudiantes. En esta misma línea, el Informe McKinsey (2007)⁴ releva la importancia de los docentes al establecer que la calidad de los sistemas educativos no puede exceder a la calidad de sus profesores.

A nivel nacional se han desarrollado iniciativas que apuntan a fortalecer y mejorar la calidad de la docencia en el país. Un primer grupo se ha enfocado en mejorar la formación inicial de los docentes, con medidas como la Beca Vocación de Profesor, que busca incentivar a los mejores alumnos a estudiar pedagogía (Barber y Mourshed 2008)⁵; el desarrollo de estándares

1 Coleman, J., Campbell, E., Hobson, C., McPartland, F., Mood, A., Weinfeld, F., York, R. (1966) Equality of educational opportunity. Washington, DC: U.S. Government Printing Office.

2 Goe, Laura. (2007) The link between teacher quality and student outcomes: A research synthesis. National Comprehensive Center for Teacher Quality.

3 La revisión señala que no hay consenso en la forma de medir (de manera observable) la calidad, por lo que considera cuatro: 1) calificaciones de los docentes, 2) características de los docentes, 3) prácticas docentes y 4) efectividad docente.

4 McKinsey & Company. (2007) How the world's best-performing school systems come out on top.

5 Barber, M. y Mourshed, M. (2008) Cómo hicieron los sistemas educativos con mejor desempeño del mundo para alcanzar sus objetivos. 1ª ed. Santiago, Chile: PREAL.

orientadores para carreras de pedagogía (que sugieren contenidos mínimos que deben manejar los egresados); la Prueba Inicia (que evalúa los conocimientos de los egresados de pedagogía en base a los mencionados estándares); y los convenios de desempeño con facultades de educación (recursos para mejorar áreas como currículum, docencia, entre otros, cuya entrega está sujeta a resultados de avance). Por otra parte, se han desarrollado iniciativas para mejorar el desempeño durante el ejercicio de la profesión docente, dentro de las cuales está el proyecto de ley que moderniza la Carrera Docente Inicial, cuyo fin es avanzar en la selección y retención de los mejores estudiantes para esta profesión, motivar la mejora y el buen desempeño y mejorar las condiciones de los profesores. Ello mediante mayores exigencias para los maestros, pero también mayores sueldos y más horas de libre disposición y estímulos al buen desempeño, particularmente para quienes enseñen a los alumnos de mayor vulnerabilidad⁶.

En relación también al ejercicio docente, un informe de la Organización para la Cooperación y el Desarrollo Económico (OCDE) del año 2009⁷, indica que más del 80% de los 25 países estudiados (todos pertenecientes a la OCDE) contaba a la fecha con algún tipo de evaluación para sus docentes (aun cuando el alcance de dichas evaluaciones, así como los instrumentos utilizados variaban entre países), señalando dentro de sus recomendaciones que “La evaluación regular debe considerarse como una parte rutinaria integrada de la vida profesional” (OCDE, 2009).

Chile ha venido sistemáticamente incorporando este concepto en los últimos años. En efecto, si bien la legislación que rige a los docentes del sector municipal⁸, junto con otorgarles varios beneficios asociados a la estabilidad, incorporaba desde sus orígenes (a principios de los años 90) un sistema de evaluación, pero dicha norma no se concretó hasta la década siguiente. En 2003 se implementó el Sistema de Evaluación Docente para el sector municipal, administrado por el Ministerio de Educación (MINEDUC), que buscó fortalecer la profesión. Asimismo, con la ley que crea el Sistema de Aseguramiento de la Calidad escolar en Chile⁹, desde 2011 se considera la evaluación de los profesores –así como sus resultados– en diversas etapas del proceso de aseguramiento de la calidad, incluyendo dentro de las funciones de la recientemente creada Agencia de Calidad, el validar los sistemas de medición de los docentes que voluntariamente presenten los establecimientos educativos.

6 Boletín 8189-04.

7 Organización para la Cooperación y el Desarrollo Económicos - OCDE. (2009) Los docentes son importantes: atraer, formar y conservar a los docentes eficientes.

8 Decreto con Fuerza de Ley 1 de 1997 (Fija texto refundido, coordinado y sistematizado de la ley nº 19.070 que aprobó el estatuto de los profesionales de la educación, y de las leyes que la complementan y modifican).

9 Ley 20.129 (Sistema Nacional de Aseguramiento de la Calidad de la Educación Parvularia, Básica y Media y su fiscalización).

En vista del tiempo que ha pasado desde el inicio del Sistema de Evaluación chileno, así como conscientes del aporte que significa evaluar las políticas públicas¹⁰ para así corregirlas y mejorarlas, el MINEDUC decidió revisar dicho sistema, mediante dos trabajos complementarios: uno encargado a la OCDE y otro realizado en forma conjunta por el Centro de Estudios del MINEDUC y el Programa de Naciones Unidas para el Desarrollo (PNUD).

A continuación se reseñan algunos aspectos de contexto, tanto del sistema de evaluación docente chileno como de los estudios mencionados, presentándose posteriormente los resultados de ambos trabajos y las recomendaciones que de ellos se desprenden. Luego, se revisa la evidencia internacional respecto de las prácticas más efectivas para evaluar docentes en base a los objetivos buscados, en orden a orientar la evolución del sistema de medición docente chileno. El documento concluye con una breve sección que sintetiza y aporta reflexiones para el futuro.

2. Antecedentes de contexto

2.1. Sistema de Evaluación Docente Chileno

El Sistema de Evaluación Docente comenzó a funcionar en el año 2003 (primero voluntaria y después obligatoriamente). Está dirigido a todos los profesores que se desempeñan en establecimientos municipales, quienes son evaluados cada cuatro años y busca “fortalecer la profesión docente, favoreciendo el reconocimiento de las fortalezas y la superación de las debilidades de los docentes, con el fin de lograr mejores aprendizajes de sus alumnos y alumnas (...)”¹¹.

Está construido en base a los dominios, criterios y descriptores que provee el Marco para la Buena Enseñanza (MBE)¹² y se compone de cuatro instrumentos: Portafolio, Pauta de Autoevaluación, Entrevista por un Evaluador Par e Informes de Referencia de Terceros, que buscan reunir información y evidencia sobre el desempeño docente (ver Tabla 1 con breve descripción de cada uno, así como sus ponderaciones en la calificación final).

10 Las evaluaciones permiten detectar si los programas o políticas tienen los efectos esperados, fomentan la rendición de cuentas en el uso de los recursos y ayudan a distinguir aquello que no funciona de lo que efectivamente sí lo hace (Shahidur R. Khandker, Gayatri B. Koolwal y Hussain A. Samad. (2010) Handbook on Impact Evaluation. Quantitative Methods and Practices. The World Bank).

11 Reglamento sobre Evaluación Docente. Decreto N°192 de Educación del 30 de agosto de 2004.

12 El MBE establece aquello que se espera de los docentes en términos de lo que conocen y saben hacer, estableciendo cuatro dominios que se refieren a aspectos diferentes de la enseñanza: a) Preparación de clases; b) Creación de un ambiente propicio para el aprendizaje; c) Enseñanza para el aprendizaje de todos los estudiantes; y d) Responsabilidades profesionales. Más información en www.cpeip.cl

Tabla 1: Instrumentos utilizados en la evaluación docente.

Instrumento	Descripción	Ponderación
Portafolio	(i) Diseño e implementación de una unidad y respuesta a preguntas de su quehacer. (ii) grabación de clase y llenado de ficha sobre ella.	60%
Pauta de Auto-evaluación	Preguntas que invitan al docente a reflexionar sobre su práctica y valorar su propio desempeño profesional.	10%
Entrevista de Evaluador Par	Entrevista realizada por un profesor que ha sido capacitado para esta tarea.	20%
Informe de Referencia de Terceros	Evaluación de los superiores jerárquicos (Director y Jefe de la Unidad Técnico Pedagógica, UTP).	10%

Fuente: Elaboración propia.

Los docentes que participan de la evaluación se clasifican de acuerdo a su nivel de desempeño en “Destacado”, “Competente”, “Básico” e “Insatisfactorio”. Cada uno de estos niveles está asociado a consecuencias como acceso a capacitación, reconocimientos, posibilidad de ser desvinculado, entre otras¹³.

A la fecha se han realizado más de 110 mil evaluaciones, en las cuales han participado 76.999 docentes, siendo la más reciente la del año 2012, cuyos resultados indican que más del 76% de los docentes tuvo un desempeño calificado como “Competente” o “Destacado”, mientras que menos del 1% tuvo uno en la categoría “Insatisfactorio”¹⁴.

13 Consecuencias: (i) uso de resultados como antecedentes para seleccionar profesionales (becas, puestos de trabajo, capacitación); (ii) reconocimientos monetarios para quienes obtienen calificación competente o destacado y aprueban una prueba de conocimientos disciplinarios y pedagógicos y; (iii) acciones de carácter formativo, llamados “Planes de Superación Profesional (PSP)” y causal de salida de la dotación docente para quienes salen evaluados en las 2 categorías inferiores, quienes, además, deben evaluarse con mayor frecuencia. Más información en www.docentemas.cl

14 En 2012 se realizaron 16.428 evaluaciones considerando los 5 niveles de enseñanza (Educación Parvularia, Educación Básica, Media, Especial y Adultos).

2.2. Revisiones de la Evaluación Docente

Durante el año 2011 y en el marco del Review on Evaluation and Assessment Frameworks for Improving School Outcomes¹⁵ de la OCDE, las autoridades chilenas de educación solicitaron a dicho organismo que revisara el marco general del sistema de Evaluación Docente en relación a sus procedimientos y a su vínculo con el desarrollo profesional de los profesores. Para ello, el equipo considerado por la OCDE¹⁶ realizó un trabajo cualitativo, que consideró el levantamiento de información respecto de las fortalezas y debilidades del sistema según la visión de los principales actores del mismo. Ello se tradujo en un intenso proceso de diálogo con diversos actores del sector escolar chileno que incluyó a las autoridades del Ministerio de Educación, al Consejo Nacional de Educación, al Colegio de Profesores, expertos en educación, diversos actores de municipalidades, visitas a escuelas en tres regiones del país, entre otros. Todo este trabajo fue complementado a través de la revisión de literatura y de un documento preparado por el Ministerio de Educación en base a una serie de preguntas elaboradas por el equipo a cargo de la evaluación.

Paralelamente, el MINEDUC junto al PNUD realizaron un estudio cuantitativo que buscaba responder: (i) ¿Cuánto discriminan y explican sus instrumentos en el resultado final de la evaluación docente?, y (ii) ¿Cómo se relaciona el desempeño de los docentes en la evaluación con el rendimiento de los estudiantes?¹⁷ Los resultados de ambos trabajos se presentan a continuación.

3. Principales hallazgos

3.1 Resultados generales

Uno de los resultados más importantes encontrados por el equipo evaluador de la OCDE es que Chile ya tiene instalada una cultura de evaluación entre los diversos actores del sistema, destacando que la mayoría de los docentes, si bien lo reconocen como un sistema mejorable, valoran como positivo y necesario su funcionamiento.

Asimismo, el informe reconoce que el sistema se basa en un marco adecuado (el MBE), puesto que provee información clara a los docentes respecto de qué se espera de ellos y también

15 Se trata de una iniciativa que busca aportar información respecto de las diferentes aproximaciones y objetivos de realizar evaluaciones en el ámbito educativo en los países participantes. Más información en <http://www.oecd.org/education/school/oecdreviewonevaluationandassessmentframeworksforimprovingchooloutcomes.htm>

16 El equipo estuvo compuesto por: Paulo Santiago, Francisco Benavides y Deborah Nusche (de la OCDE) y las expertas internacionales Charlotte Danielson y Laura Goe.

17 Alvarado, M., Canales, C., Falck, D., Ortega, M.E. (2012) La Evaluación Docente y sus instrumentos: discriminación del desempeño docente y asociaciones con los resultados de los estudiantes. MINEDUC-PNUD, Santiago, Chile.

permite delinear el trabajo de formación continua para aquellos docentes con calificaciones en las categorías inferiores.

Por otra parte, el informe señala que el objetivo formativo del sistema, por ejemplo, a través de los Planes de Superación Profesional (PSP) para los docentes evaluados en las dos categorías inferiores, debe ser desarrollado o fortalecido ya que ha quedado relegado respecto de aquél asociado a la rendición de cuentas. Por tanto, faltarían mecanismos para retroalimentar a los docentes respecto de sus evaluaciones, haciéndose necesario aumentar el diálogo sobre las prácticas docentes. En este sentido, se señala que dentro de los resultados de un estudio del Consortium for Chicago School Research (CCSR)¹⁸ se encontró que tanto los docentes como los directores de las escuelas valoraban positivamente las conversaciones posteriores al ejercicio evaluativo.

El estudio reconoce las ventajas de avanzar hacia una evaluación más descentralizada, destacando el potencial que tiene el nivel local (municipal) para reconocer las diferencias que tiene cada territorio para aplicar la evaluación docente, sin embargo reconoce que las capacidades de las municipalidades y sostenedores son muy disímiles, por lo que en algunas zonas se requeriría fortalecerlas.

El informe también recoge como positivo, los esfuerzos desplegados por potenciar la función del líder educacional dentro de las escuelas y lo reconoce como un actor clave del proceso de evaluación. En ese sentido, se destaca el avance al aumentar la autonomía y atribuciones de los directores para seleccionar y desvincular al equipo docente de la escuela.

Considerando lo anterior, el trabajo señala como un desafío para el sistema, avanzar en la mejora de las capacidades de los docentes y equipo directivo para evaluar y sacar el máximo provecho a los resultados de la evaluación al interior de las escuelas.

3.2 Respetto de los instrumentos

El documento también da cuenta de que el uso de más de un instrumento para evaluar el desempeño docente es algo importante, ya que incrementa la validez del sistema.

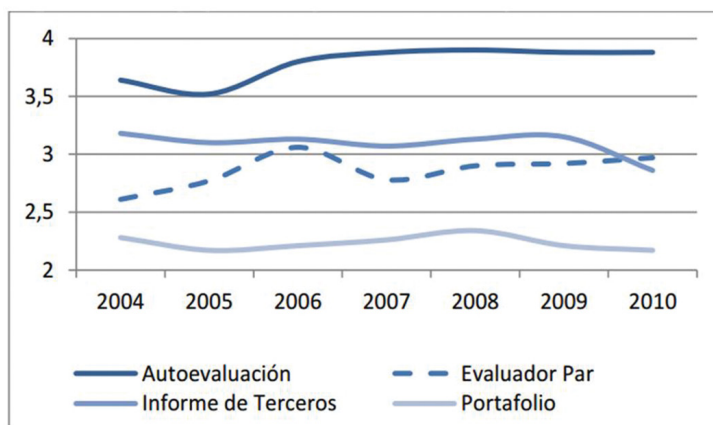
Sin embargo, reconoce que la pauta de autoevaluación es un instrumento que se debe reconsiderar puesto que tanto la forma en que está elaborado, como su aplicación, disminuyen su confiabilidad y validez como instrumento evaluativo. Ello coincide con las conclusiones del estudio del MINEDUC y PNUD respecto de la poca capacidad de entregar información de dicho instrumento, puesto que más del 95% de los docentes obtiene una puntuación de 3 o más puntos, de un máximo de 4 (Alvarado et al. 2012), lo que, en forma conjunta con la poca variabilidad de

18 Investigación sobre la implementación de un programa piloto de evaluación docente en las escuelas públicas en Chicago (<http://ccsr.uchicago.edu/publications/rethinking-teacher-evaluation-chicago-lessons-learned-classroom-observations-principal>)

los resultados, no permite identificar a los docentes que lo hacen bien de aquéllos que no.

En este sentido, el informe del MINEDUC y PNUD da cuenta de que, con excepción del portafolio, las calificaciones del resto de los instrumentos también se encuentran sesgadas hacia valores altos y tienen poca variabilidad¹⁹. La Figura 1 muestra cómo los resultados de la pauta de autoevaluación se han concentrado sobre los 3,5 puntos desde el año 2004, mientras que el promedio de resultados en el portafolio no supera el punto de corte para el nivel "Básico" (2,5 puntos). Resulta preocupante que un docente tenga niveles de competencia tan diferentes según sea el instrumento utilizado, lo que sugiere la necesidad de revisar los instrumentos y su combinación²⁰.

Figura 1: Evolución de los puntajes promedio por instrumento (2004 - 2010)



Fuente: Alvarado et al. (2012)

Dicho trabajo detecta también que la autoevaluación y la entrevista por un evaluador par no se relacionan con el desempeño de los estudiantes, a diferencia del portafolio y el informe de referencia de terceros. Así, tanto la pauta de autoevaluación como la entrevista del evaluador par no entregan información respecto del nivel de desempeño de los docentes, y por otra, tampoco son indicativas del impacto del desempeño del profesor en los aprendizajes de sus estudiantes.

Por otra parte, el informe de la OCDE encuentra que pese a que en el proceso de evaluación participan variadas instituciones –el Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas (CPEIP) del MINEDUC, los municipios, diversas

19 Desde el año 2007, la varianza del instrumento ha sido menor a 0.15 puntos. (Alvarado et al 2012)

20 En el mejor de los casos, dichas diferencias se pueden explicar porque los instrumentos están midiendo aspectos diferentes de la práctica docente y en el peor de los escenarios, debido a problemas de los instrumentos.

universidades y un asesor técnico independiente²¹ (lo que debería otorgar credibilidad al proceso), los docentes dan poco crédito a las competencias de quienes corrigen el portafolio, lo que afecta la percepción de confianza hacia el sistema en su conjunto.

3.3 Recomendaciones

A continuación se presentan algunas de las recomendaciones realizadas por el equipo evaluador de la OCDE, quedando disponible su detalle en el informe completo²²:

a. Consolidar el Marco para la Buena Enseñanza:

(i) Mejorar los mecanismos de información a los docentes respecto de los estándares, indicadores y métrica con la que serán evaluados y difundir los estándares durante las diferentes etapas del desarrollo profesional. (ii) Revisar de forma permanente tanto los estándares como los indicadores y su alineación con los instrumentos utilizados para medirlos, como su nivel de actualización, de modo que reflejen adecuadamente la última evidencia respecto de buenas prácticas docentes.

b. Fortalecer la función formativa del proceso de evaluación:

(i) Fomentar el desarrollo de sistemas de evaluación docente internos en las escuelas con objetivo formativo, lo que requiere ajustarlos a la realidad y necesidad de los establecimientos y capacitar a los docentes y equipos directivos en cómo realizar y utilizar los resultados. Se recomienda que una institución externa realice un monitoreo constante de estos sistemas para armonizar los diferentes sistemas de evaluación formativa al interior de las escuelas. (ii) Potenciar las formas de retroalimentación a los docentes evaluados, asegurándose que independiente de sus resultados, se detecten las áreas que deben ser reforzadas y mejoradas.

c. Revisar los instrumentos de evaluación y procedimientos utilizados:

(i) Asegurar altos estándares de calidad en la preparación de correctores de portafolios, por ejemplo, mediante su certificación o acreditación. Asimismo, se aconseja utilizar más de un evaluador por portafolio con el fin de reducir posibles sesgos en la corrección²³. (ii) Centrar la evaluación en la observación de clases, por una parte incrementando las capacidades de los docentes y directivos para realizar observaciones internas, y por otra, aumentando el número de observaciones sobre las que se basa la evaluación. En ese sentido, se señala que la evidencia muestra que

21 El reglamento del sistema requiere que el CPEIP cuente con una asesoría técnica independiente, rol que ha estado a cargo del Centro de Medición de la Pontificia Universidad Católica (MIDE UC).

22 Para descargar el informe completo, visitar centroestudios.mineduc.cl

23 Actualmente, existe un proceso de doble corrección sólo para 20% de los portafolios, así como instancias de recalibración y corrección grupal.

“múltiples observaciones y múltiples observadores aumentan de manera importante la precisión de la puntuación”²⁴. (iii) Dado el doble objetivo del sistema (rendición de cuenta y formación), se aconseja revisar la mezcla de instrumentos para que éstos estén alineados de mejor manera con cada objetivo. Por ejemplo, se señala que si bien la pauta de autoevaluación presenta problemas (ver nota 17) para ser usado en la rendición de cuentas, sí puede ser utilizado para detectar áreas de formación requeridas por el docente. (iv) Mejorar el informe de referencia de terceros (el cual se reconoce como novedoso a nivel internacional), por ejemplo, a través de informar a los docentes respecto de qué aspectos del MBE están siendo abordado por cada pregunta, en orden a aumentar la reflexión y mejorar la retroalimentación al docente evaluado.

4. ¿Qué dice la evidencia internacional?

La literatura da cuenta de que no existe una única forma de medir efectividad docente, es decir, no existe una medida que permita describir de manera adecuada todo el aporte que realizan los docentes tanto a los estudiantes, como a la comunidad y la escuela en su conjunto (Goe 2007, Goe et al 2008²⁵). Por ello, es posible observar que los países usan diferentes instrumentos (combinados o de manera individual) para determinar qué docentes tienen o no un buen desempeño. Asimismo, los sistemas utilizados tienen objetivos diversos, lo que se traduce, entre otros, en que las consecuencias de dichas evaluaciones también varían entre sistemas²⁶.

Por lo tanto, la revisión que se realiza a continuación busca conocer qué información aporta la evidencia internacional respecto de las estrategias más efectivas para abordar los objetivos de los ejercicios evaluativos, así como también explorar las potenciales fortalezas y debilidades de los instrumentos en cuanto a su vínculo con la efectividad docente.

4.1 Utilizar diferentes instrumentos y combinarlos apropiadamente

En relación a la dificultad de identificar a los buenos docentes mediante un instrumento, el trabajo de Kane et al (2013)²⁷ realiza una importante contribución. Los autores, a través de un ejercicio experimental, encuentran que es posible identificar a los profesores que tienen mayor impacto en los resultados de

24 “(...) research has shown that multiple observations and multiple observers greatly increases the accuracy of the scores” OCDE 2013, pag 116. Traducción no oficial.

25 Goe, L., Bell, C. y Little, O. (2008) Approaches to Evaluating Teacher Effectiveness: A Research Synthesis. National Comprehensive Center for Teacher Quality.

26 Información más detallada respecto de las características de diversos sistemas de evaluación puede encontrarse en OCDE 2009 (ver nota 7).

27 Kane, T., McCaffrey, D., Miller, T. y Staiger, D. (2013) Have we identified effective teachers? Validating measures of effective teaching using random assignment. Bill & Melinda Gates foundation.

los estudiantes (docentes más efectivos), utilizando una medida compuesta de efectividad docente, es decir, una estrategia que combine diversos instrumentos.

En esta misma línea de trabajo, el proyecto Measures of Effective Teaching (en adelante MET)²⁸, definió nueve principios para el uso de medidas de efectividad docente, recomendando que se utilicen varios instrumentos con el objetivo de capturar la naturaleza compleja de la enseñanza (MET, 2013a, MET, 2013b, Milhaly et al 2013)²⁹. Esta conclusión es compartida por otros autores como Gordon et al (2006)³⁰ y Goe (2007).

Sin embargo, hay que tomar en cuenta que la forma en que se combinan los instrumentos no es una decisión trivial. Se debe considerar que, por una parte, la elección debe estar guiada por el objetivo evaluativo y, por otra parte, la validez de las medidas no recae sólo en la calidad del instrumento, sino también en su capacidad específica de medir la dimensión a observar (ver nota 25). En este sentido Milhaly et al (2013) desarrollan un modelo para determinar una forma de combinar las distintas mediciones e instrumentos que se utilizan en diferentes estados de los Estados Unidos, señalando que la decisión respecto a cómo ponderarlos no responde sólo a una decisión técnica, sino a una decisión respecto de cuál es el énfasis deseado para cada indicador. Por ejemplo, como se menciona en la siguiente sección, si bien la literatura indica que las medidas de valor agregado permiten identificar a los profesores efectivos, no muestran las prácticas de dichos docentes, por lo que para este último objetivo se requiere de otro instrumento, como la observación de clases.

4.2 Escoger instrumentos efectivos

Por otra parte, la literatura recomienda utilizar instrumentos que entreguen medidas consistentes de "efectividad docente", que son aquéllos que evalúan directamente los procesos y actividades que acontecen al interior de la sala de clase (ver nota 25), entre los que destacan:

a. Observación de aula

Las observaciones de aula se han posicionado como un instrumento confiable para evaluar la práctica pedagógica (Goe

et al 2008; Kane et al 2011³¹) dado que son instrumentos flexibles que pueden diseñarse en función de su objetivo.

Se debe considerar que su calidad depende de dos condiciones: i) la confiabilidad de sus resultados, lo que a su vez requiere que sus dimensiones y subdimensiones capturen la esencia de la "efectividad", y ii) los observadores deben estar entrenados para evaluar correctamente y generar puntuaciones consistentes. Un ejemplo es la pauta de observación Classroom Assessment Scoring System (CLASS) puesto que su protocolo de uso exige instancias de calibración grupal y doble codificación para asegurar consistencia³² y cuenta con un programa de entrenamiento para observadores y entrenadores, el cual vela por un uso consistente del instrumento. Por su parte, proyectos como el MET (ver nota 29), han demostrado que sus resultados se relacionan con el desempeño de los estudiantes.

Estos instrumentos permiten mapear las prácticas de los docentes de buen desempeño y generar valiosa información para la mejora, aunque no existe todavía suficiente evidencia respecto de los cambios que se observan en las prácticas pedagógicas una vez que se empiezan a realizar observación y retroalimentación con esta pauta, ni de la validez que los docentes entregan al instrumento.

b. Encuestas de percepción a estudiantes

La evidencia señala que las encuestas de percepción de los estudiantes son válidas y confiables para la evaluación de los docentes en los distintos niveles educativos (Peterson et al 2000)³³. Incluso han demostrado ser más confiables que la pauta de autoevaluación de docentes y la evaluación de los directores, ya que se relacionan de forma significativa y consistente con medidas de valor agregado (Blach R. 2012)³⁴ y con la mejora en el aprendizaje que obtienen estudiantes de otros cursos que reciben instrucción del mismo profesor (MET, 2010)³⁵. Asimismo, han demostrado ser muy útiles para retroalimentar el trabajo docente (ver nota 35) y orientar las prioridades de capacitación de los establecimientos.

Destaca entre las ventajas su costo efectividad, al no requerir entrenamiento ni el uso de grandes cantidades de tiempo; al permitir resguardar el anonimato y aplicarse con alta frecuencia,

28 Más información sobre este proyecto se puede obtener en www.metproject.org

29 Measures of Effective Teaching (MET) (2013a) Feedback for better teaching. Nine principles for using measures of effective teaching. Bill & Melinda Gates foundation; Measures of Effective Teaching (MET) (2013b) Ensuring fair and reliable measures of effective teaching. Culminating findings from the MET Project's three-year study. Bill & Melinda Gates foundation; Milhaly, K., McCaffrey, D., Staiger, D. y Lockwood, J. (2013) A composite estimator for effective teaching.

30 Gordon, R., Kane, T. y Staiger, D. (2006) Identifying effective teachers using performance on the job. The Brookings Institution. Discussion paper 2006-01.

31 Kane, T., Taylor, E., Tyler, E. y Wooten, A. (2011) "Evaluating Teacher Effectiveness". Education Next, Summer

32 Para más información ver <http://www.teachstone.org/about-the-class/>

33 Peterson, K., Wahlquist, C., y Bone, K. (2000) Student Surveys for School Teacher Evaluation. Journal of Personnel Evaluation in Education, 14:2, pp. 135-153

34 Blach, R. (2012) The validation of a student survey on teacher practice. Dissertation for the degree of Doctor of Philosophy. Vanderbilt University

35 Measures of Effective Teaching (MET) (2010) Learning About Teaching: Initial Findings from the Measures of Effective Teaching Project (Research Paper). Bill & Melinda Gates foundation.

lo que beneficia la retroalimentación y perfeccionamiento de la práctica por parte de los docentes. Sin embargo, se debe tener en cuenta que este instrumento no es adecuado para evaluar aspectos como: cobertura curricular, desarrollo de contenidos, conocimiento, organización de la clase, etc.

c. Medidas de valor agregado del logro académico de los estudiantes

Las medidas de valor agregado se han posicionado como una forma confiable de cuantificar la calidad de los docentes, dado que se asocia directamente el progreso académico de los alumnos durante el período de tiempo instruido por el docente a la efectividad de dicho maestro. El ejemplo de uso más citado en los textos es el Tennessee Value-Added Assessment System (TVAAS), conocido como uno de los sistemas más sofisticados de valor agregado (Eckert, J. y Babrowski, J., 2010)³⁶. Se debe tener en cuenta que su uso requiere contar con pruebas anuales (o bianuales) de la asignatura asociada al docente, y que dichas mediciones deben contar con ciertas propiedades que permitan medir el avance en los aprendizajes de los menores, así como también deben poder aislar los factores que influyen en los aprendizajes de los alumnos que no tienen que ver con el profesor evaluado. Ello, junto con otros problemas detectados por la literatura en algunas mediciones de valor agregado (como por ejemplo, que es una medida inconsistente por presentar volatilidad -Darling-Hammond et al 2012³⁷) deben ser considerado en su construcción para minimizar sus efectos.

Otras medidas

Los instrumentos presentados no agotan las posibilidades. La exhaustiva revisión de maneras de medir efectividad docente que realiza Goe et al (2008) destaca también el uso de información de autoevaluaciones por parte de los docentes, evaluación de los directores, entre otras. Sobre el primero, y en línea con lo que destaca el estudio del MINEDUC y el informe de la OCDE, existen problemas de confiabilidad y validez, pero pueden ser útiles para señalar aspectos no observados por otras mediciones, como identificar el tipo de capacitación requerida.

5. Comentarios finales

En vista de los resultados presentados en este documento, tanto en relación a los estudios sobre la evaluación docente, como en la revisión de la evidencia respecto de diversos sistemas e instrumentos utilizados para evaluar profesores, se concluye, en primer lugar, que los sistemas de evaluación de docentes son valiosos, puesto que permiten contar con información importante para, entre otras cosas, identificar aquellos docentes que requieren apoyo y las áreas específicas en que lo necesitan, así

36 Eckert, J. y Babrowski, J. (2010) Should Value-Added Measure Be Used for Performance Pay?. Phi Delta Kappan.

37 Darling-Hammond, L., Amrei-Beardsley, A., Haertel, E. y Rothstein, J. (2012) Evaluating Teacher Evaluation. Phi Delta Kappan.

como también a aquellos maestros que están haciendo un buen trabajo y, por tanto, observar sus prácticas; retroalimentar el trabajo de los profesores; tomar decisiones respecto de la carrera docente (promoción, cambio de función, desvinculación, etc.).

En segundo lugar, se desprende que para que un sistema de evaluación docente sea efectivo, es importante alinear los objetivos del sistema de evaluación a los instrumentos que se utilizan. Por tanto, corresponde revisar si la mezcla de instrumentos que hoy en día se utilizan para detectar a los docentes "efectivos", está alineada, tanto con el objetivo formativo de la medición, como con el objetivo de rendición de cuentas. Ello implica asegurar que los instrumentos que se usan para cada uno de los objetivos sean apropiados. En el caso del objetivo de rendición de cuentas, se deben incorporar las medidas que realmente permitan identificar de mejor manera el esfuerzo y la efectividad de los docentes, puesto que con ellos se tomarán decisiones importantes respecto del desarrollo de su carrera profesional. Ello sugiere revisar el uso de la pauta de autoevaluación, ya que no aporta en cuanto a identificar a los docentes más o menos efectivos. Asimismo, se debe avanzar en el fortalecimiento y uso de observación de aulas (con las mejoras mencionadas en la sección anterior), como un componente principal para determinar las categorías en que los docentes son calificados. Junto con ello se debe avanzar en generar las condiciones para el uso de pruebas de valor agregado que, una vez aislados los factores que no tienen que ver con los profesores, permiten (además de monitorear el aprendizaje de los estudiantes) identificar a los docentes más efectivos de manera objetiva y, por ejemplo, recopilar información respecto de sus prácticas.

Lo mismo ocurre en el caso del objetivo formativo de la evaluación, donde la autoevaluación sí podría jugar un rol relevante, potenciando su uso como instrumento de retroalimentación e información, por ejemplo, para el equipo directivo de la escuela. Asimismo, la observación de clases puede ser un aporte dentro de las escuelas si se fomenta el desarrollo de las capacidades para que tanto el equipo directivo como docente puedan realizar observación de aula y aprovechar la riqueza de la información que estos ejercicios entregan.

La evidencia también apunta hacia la información que surge de los estudiantes que, para los niveles educativos más altos, puede aportar datos muy valiosos tanto para el proceso de mejora continuo del profesor (objetivo formativo de la evaluación docente) como para el equipo directivo respecto del desempeño docente (objetivo de rendición de cuentas), revelando aspectos que no son posibles de capturar con el resto de las mediciones. En este caso, y al igual que para el resto de los instrumentos, se debe cuidar de no generar incentivos inapropiados que podrían mermar la efectividad del mismo o afectar las prácticas pedagógicas y, por tanto, la calidad de la enseñanza. Por ejemplo, no se aconseja basar la decisión de desvinculación de un docente solamente en la encuesta a los estudiantes.

Por su parte, respecto de la evaluación de los directores, la literatura señala que este instrumento tiene un alto potencial,

lo que sugiere que es un componente muy adecuado de incluir en un sistema de evaluación, siempre que cuenten con la preparación para realizarlo, para que su trabajo sea efectivo. En este sentido, el informe de la OCDE señala que se debe asignar un rol clave a los directivos en la evaluación docente, ya que estarían preparados porque conocen el contexto educacional en que se desenvuelve el profesor y, por tanto, podrían utilizar su observación de las prácticas para informar la evaluación docente, retroalimentando de manera directa a los profesores. En esta línea el plan de formación de directores de excelencia, que busca otorgar y mejorar estas competencias en los directivos, es un apoyo esencial para avanzar hacia una aplicación local de la evaluación docente.

Es necesario no perder de vista que varios instrumentos analizan “la práctica docente”, la cual no es estática, sino que se caracteriza por ser perfectible y moldeable. Por ello, es crucial que los docentes conozcan el detalle de los instrumentos con los que se les evaluará y los estándares asociados, de manera que haya claridad respecto de qué se mide y cómo dichos instrumentos influyen en los resultados que obtienen. Asimismo, deben proveer a los docentes de retroalimentación abierta y de calidad. Esta información será significativa para los docentes siempre que ellos reconozcan en el evaluador experiencia y objetividad. Ello genera un círculo continuo de mejora, puesto que sirve no sólo a los docentes, primeros interesados, sino también al resto de la comunidad, por ejemplo, detectando prácticas docentes novedosas y, entregando información que permita mejorar la calidad de la educación que se imparte en el país.

Así, después de una década de aplicación, parece apropiado perfeccionar el sistema de evaluación docente para hacer un mejor uso de dicho instrumento, con el foco de mejorar el desempeño de los profesores y, finalmente, la calidad de la educación de millones de niños.

Este documento incluye los principales resultados del estudio “OECD Reviews of Evaluation and Assessment in Education: Teacher Evaluation in Chile 2012” realizado por la OCDE y del estudio “La Evaluación Docente y sus instrumentos: discriminación del desempeño docente y asociación con los resultados de los estudiantes” realizado por el Centro de Estudios MINEDUC y el PNUD.

Para descargar el informe completo de la OCDE pinche [AQUÍ](#) y el del Centro de Estudios MINEDUC y el PNUD pinche [AQUÍ](#).