



La Evaluación Docente en Chile

Editores

Jorge Manzi • Roberto González • Yulan Sun

Centro de Medición MIDE UC



FACULTAD DE
CIENCIAS SOCIALES
ESCUELA DE PSICOLOGÍA



La Evaluación Docente en Chile

Editores

Jorge Manzi • Roberto González • Yulan Sun



FACULTAD DE
CIENCIAS SOCIALES
ESCUELA DE PSICOLOGÍA



Acerca de los editores

Jorge Manzi es Profesor Titular de la Escuela de Psicología de la Pontificia Universidad Católica de Chile y Director de su Centro de Medición, MIDE UC. Es psicólogo y doctor en psicología; sus áreas de especialidad incluyen la psicología social y la psicometría. En este último campo ha realizado aportes principalmente en el ámbito de la medición educacional: participó en el proyecto de investigación que contribuyó a la reformulación de las pruebas de admisión a las universidades, ha conducido el desarrollo de un sistema de evaluación del progreso en el aprendizaje para alumnos de enseñanza básica y media aplicando modelos de valor agregado y ha dirigido diversos proyectos vinculados a la evaluación de profesores: el Sistema de Evaluación del Desempeño Profesional Docente, el programa de Asignación de Excelencia Pedagógica y el sistema de evaluación diagnóstica para egresados de pedagogía INICIA, entre otros. Pertenece a distintas sociedades científicas y cuenta con numerosas publicaciones en medios especializados nacionales e internacionales. Durante el 2006 participó en el Consejo Asesor Presidencial para la Calidad de la Educación, integra el Comité Técnico Asesor de la Prueba de Selección Universitaria (PSU) y ha sido consultor en proyectos vinculados al desarrollo de sistemas de evaluación de profesores en Portugal, México y Kosovo.

Roberto González es Vicerrector académico de la Pontificia Universidad Católica de Chile. Es psicólogo y Doctor en Psicología, y se ha especializado en el campo de la psicología social, la psicología política y la medición educacional. Su investigación y publicaciones en el área de la psicología social y política se han centrado en el estudio de las relaciones intergrupales, el prejuicio, las emociones intergrupales, los procesos de aculturación y conductas de ayuda. En el campo de la medición educacional, es cofundador del Centro de Medición MIDE UC, donde ha participado en el desarrollo de instrumentos y sistemas de medición a gran escala para evaluar el desempeño profesional docente y la selección de profesores de excelencia a nivel nacional. Es profesor asociado de la Escuela de Psicología y miembro de diversas organizaciones científicas norteamericanas y europeas; participa en comités editoriales de diversas revistas para evaluar artículos de divulgación científica a nivel nacional e internacional y forma parte del directorio de la Sociedad para Estudio Psicológico de Problemas Sociales (the Society for the Psychological Study of Social Issues).

Yulan Sun es Gerente del Proyecto Docentemás de MIDE UC, Centro de Medición de la Pontificia Universidad Católica de Chile. En esa condición, dirige el equipo que asesora, desde 2003 a la fecha, al Ministerio de Educación en la implementación de la Evaluación Docente. Es psicóloga y Magister en Psicología y sus áreas de especialidad incluyen la psicología educacional y la evaluación, habiendo desempeñado funciones de docencia, investigación y gestión de proyectos en ambas. Integró el equipo de Desarrollo Profesional Docente de la División de Educación General del Ministerio de Educación, correspondiéndole diseñar e implementar programas de capacitación para profesores y coordinar el plan piloto de la Campaña de Lectura, Escritura y Matemática (LEM). Ha sido docente en las carreras de Psicología y Pedagogía en distintas universidades, ha desarrollado talleres de formación para docentes de nivel escolar y de educación superior de distintas instituciones (P.U. Católica de Chile, U.de Chile, U. Católica del Norte y U. Diego Portales, entre otras), y ha sido consultora para proyectos educacionales dentro y fuera de Chile.

Índice

	Agradecimientos	7
	Introducción	9
Capítulo 1	Origen de la Evaluación Docente y su conexión con las políticas públicas en educación <i>Rodolfo Bonifaz</i>	13
Capítulo 2	El modelo de evaluación y los instrumentos <i>María Paulina Flotts · Andrea Abarzúa</i>	35
Capítulo 3	La implementación de la Evaluación Docente <i>Yulan Sun · Paulina Calderón · Nelson Valerio · Pablo Torres</i>	63
Capítulo 4	Resultados: qué dice la Evaluación Docente acerca de la enseñanza en Chile <i>Yulan Sun · Mónica Correa · Alvaro Zapata · Diego Carrasco</i>	91
Capítulo 5	Consecuencias de la Evaluación Docente <i>Flavio Cortés · María José Lagos</i>	137
Capítulo 6	Estudios de validez de la Evaluación Docente <i>Sandy Taut · Verónica Santelices · Jorge Manzi</i>	157
Capítulo 7	Desempeño docente: relaciones con antecedentes de los profesores y su contexto <i>Edgar Valencia · Jorge Manzi</i>	177
Capítulo 8	Investigación educacional a partir de la Evaluación Docente	195
	<i>Microgénesis de la enseñanza: zoom en el modo en que los profesores presentan contenidos disciplinarios</i> <i>Carlos Cornejo · David Silva · Himmbler Olivares</i>	197
	<i>¿Qué actividades realizan los docentes de NB1 para enseñar a leer en situación de evaluación docente? Enfoques tras las prácticas pedagógicas</i> <i>Viviana Galdames · Lorena Medina · Ernesto San Martín · Rosa Gaete · Andrea Valdivia</i>	200
	<i>Uso del tiempo e interacciones profesores - alumnos en la sala de clases</i> <i>Sergio Martinic</i>	204
	<i>Patrones instruccionales en Chile: la evidencia de la Evaluación Docente</i> <i>David Preiss</i>	209

Capítulo 9	Comentarios	213
	<i>La evaluación docente en Chile: perspectivas sobre validez</i>	215
	<i>José Felipe Martínez</i>	
	<i>Aportes para pensar las políticas de evaluación docente</i>	222
	<i>Pedro Ravela</i>	
	<i>La buena enseñanza y la agenda abierta de la evaluación docente ante un recurso escaso</i>	231
	<i>Gilbert Valverde</i>	
Capítulo 10	Conclusiones	239
	<i>Jorge Manzi · Roberto González · Yulan Sun</i>	
	Referencias	251

Agradecimientos

Este libro es fruto del arduo trabajo y la generosa contribución de muchas personas e instituciones; en primer lugar, de quienes hicieron posible la Evaluación Docente conciliando los legítimos intereses de sus representados, con la necesidad de evaluar la calidad de la docencia como un camino para mejorar la educación de los niños y jóvenes de Chile; también de quienes han coordinado su implementación, tanto a nivel nacional como en cada municipio desde 2003 hasta ahora, especialmente en el Ministerio de Educación, el CPEIP y los Departamentos y Corporaciones Municipales de Educación a lo largo de todo el país; y, desde luego, de los directores, jefes de UTP y, muy especialmente, los propios docentes, que han favorecido la instalación y apropiación de la evaluación allí donde se juega finalmente su impacto, en los establecimientos y las salas de clases de nuestro país.

También debemos hacer extensivo este agradecimiento a las autoridades, académicos y profesionales de la Universidad Católica que han colaborado con el desarrollo del *Proyecto Docentemás*, a través del cual MIDE UC ha asesorado al Ministerio de Educación, desde 2003 hasta la fecha en la implementación de la evaluación.

Finalmente, expresamos nuestro agradecimiento a todos los autores que contribuyeron con los distintos capítulos que conforman este libro; a los profesores consultores del *Proyecto Docentemás*, por su asistencia en la preparación de los ejemplos para el Capítulo 4, y:

- a Edgar Valencia, por su dedicado trabajo asistiendo a los editores en todo el proceso de producción del libro, especialmente en la preparación y análisis de datos y la revisión de referencias bibliográficas;
- a Karin Luna, por su colaboración en la preparación del Capítulo 3; a Lissette Sepúlveda, Beatriz Rodríguez, Loreto Tagle, Dina Cembrano, Valeska Grau, Carola Quinteros, Verónica Santelices, Sandy Taut y Mauricio Rojas, por su colaboración en la revisión de borradores de capítulos;
- y en particular hacia quienes han formado parte y conforman hoy el equipo del *Proyecto Docentemás*, por su compromiso profesional y personal con la responsabilidad que les ha cabido en la implementación de la Evaluación Docente, y su incansable participación en la revisión final de todos los capítulos de este libro, su impresión y lanzamiento:

Equipo–Proyecto Docentemás de MIDE UC (Abril 2011)

<i>Área Instrumentos de Evaluación</i>	<i>Área Implementación y Gestión Local</i>	<i>Área Operaciones</i>	<i>Área Análisis de datos</i>
M. Paz Cadena	M. Soledad Aravena	Juan Pablo Arancibia	Diego Carrasco
Mónica Correa	Marianela Calderón	Cristian Gallardo	Patricia Llona
Mónica Cortez	Paulina Calderón	Tatiana Góngora	Diego Rojas
Paloma Fernández	Andrea Fuentes	Magdalena Ocampo	Alvaro Zapata
M. Rosa García	Andrea Muñoz	Héctor Zamorano	
Alejandra Guzmán	Orianna Reyes		
Angela Haquin	Loreto Valderrama	<i>Área Capacitación y Selección</i>	<i>Apoyo Administrativo</i>
Fanisa Miño	Nelson Valerio	Paulina Caselli	M. José Abásolo
Carolina Requena		Daniela Maluenda	Sergio Alcázar
M. José Robles		M. Teresa Olivares	Claudia Díaz
Pablo Torres			

Introducción

En agosto del año 2003, tras un largo camino de negociaciones, Chile comenzó a implementar por primera vez un programa nacional de evaluación de profesores, el Sistema de Evaluación del Desempeño Profesional Docente, convirtiéndose así en uno de los países pioneros en la aplicación de este tipo de programas dentro de la región y en un caso de alto interés en el concierto internacional.

La Evaluación Docente marca un importante hito en las políticas de fortalecimiento docente durante la última década e involucra también un importante desafío desde el punto de vista técnico y logístico. Por una parte, impone la necesidad de desarrollar instrumentos de medición capaces de establecer el nivel real de desempeño de los profesores utilizando para ello múltiples fuentes: el juicio del propio evaluado, evidencia directa de su labor en aula y el reporte de pares y supervisores.

Por otra parte, dado que su aplicación descentralizada descansa en las autoridades educacionales y evaluadores pares de cada comuna, el programa exige un complejo sistema de coordinación con los 345 municipios que administran establecimientos educacionales. Y todo ello en el contexto de un programa de escala nacional que incluye cada año a un gran número de docentes y debe cumplir estándares de rigurosidad y confidencialidad en el manejo de la información, así como de equidad en las condiciones de evaluación para todos los docentes. Otro desafío no menos importante es el desarrollo de métodos que permitan informar los resultados de la evaluación de modo que ésta cumpla con su propósito formativo y de retroalimentación para los docentes, así como para sus directivos y el sistema educacional en su conjunto.

A la fecha se han completado ya ocho años de aplicación de la Evaluación Docente. El programa ha alcanzado cobertura nacional desde 2005, incluyendo a profesores de todas las comunas del país. Más de 67.000 docentes que ejercen en los principales niveles, modalidades y subsectores de enseñanza han participado en el proceso y alrededor de 14.000 de ellos han completado ya su segunda evaluación. Estos datos ilustran el desarrollo y consolidación creciente que experimenta el programa al enfrentar su noveno año de aplicación.

Desde el inicio de la evaluación, un equipo de especialistas de la Pontificia Universidad Católica de Chile fue convocado para asesorar técnicamente el proceso, función que lleva a cabo hasta la fecha el Centro de Medición MIDE UC, dependiente de la Escuela de Psicología.

La participación de una entidad técnica independiente en la Evaluación Docente fue una exigencia planteada por la normativa que rige el sistema desde su inicio, y el prestigio y capacidades técnicas de la Universidad Católica mostraron ser un factor que favoreció su desarrollo, en un contexto inicial complejo, de temor e incertidumbre por parte de muchos docentes y persistente resistencia de otros. Asimismo, la naturaleza académica de MIDE UC junto con la fructífera relación de colaboración establecida con la contraparte ministerial —el Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas (CPEIP)— permitieron, a la par de la ejecución misma, enriquecer el trabajo técnico involucrado en la evaluación, implementar una agenda de investigación especialmente centrada en estudiar la validez del programa y sus instrumentos, y promover la investigación educacional en un sentido más amplio, como resultado de la gran cantidad de datos y evidencia que el proceso permite recolectar. Cabe señalar que, además de los resultados mismos de la evaluación, cada año ésta recoge muestras de trabajo de miles de docentes, provenientes de todo el país, así como antecedentes adicionales, referidos por ejemplo, a su formación inicial y en servicio, sus condiciones laborales y su experiencia en el proceso de evaluación, entre otros.

Merced a su misión como centro académico comprometido con el mejoramiento de la calidad y equidad de la educación chilena, y por la responsabilidad que le ha cabido en la implementación de esta política educacional, MIDE UC emprendió la iniciativa de producir una publicación que permitiera sistematizar y dar a conocer la experiencia de estos ocho años de Evaluación Docente. Así surgió el presente libro; en él se busca dar cuenta del contexto y desarrollo histórico del programa, sus características y los resultados obtenidos hasta ahora; comunicar los datos de diversos estudios que se han realizado alrededor de la evaluación, por ejemplo, en torno a la validez de sus instrumentos, sus consecuencias y la relación de sus resultados con antecedentes de los profesores y del contexto en que ejercen. También fue parte del interés de los editores mostrar algunos ejemplos de la investigación sobre calidad docente y prácticas pedagógicas desarrollada empleando datos y/o evidencias aportadas por la evaluación, y situar la experiencia chilena en el contexto internacional a partir de la visión de expertos de otros países.

Al abordar esta gama de temas, el libro busca responder al interés de diversas audiencias motivadas y comprometidas con la educación y su desarrollo. En primer lugar está dirigido a los actores del sistema escolar, especialmente los docentes, directivos y sostenedores directamente responsables de proveer educación a los niños y jóvenes chilenos. También busca ser un aporte para las instituciones formadoras de profesores y para quienes desarrollan acciones de apoyo y capacitación orientadas a los docentes en servicio; para las autoridades y especialistas, dentro y fuera de Chile, que participan en el diseño de políticas educacionales; para los académicos e investigadores cuyo trabajo proporciona conocimiento científico para comprender y mejorar las prácticas educativas; y para el público general, crecientemente informado e interesado en la educación, por su importancia social y su valor para promover la equidad y el desarrollo del país.

A través de una estructura modular se ha tratado de favorecer que cada lector pueda seleccionar aquellos capítulos que responden mejor a su interés y, al inicio de los mismos, un breve resumen permite obtener una visión panorámica del contenido a abordar, el que se describe someramente en las líneas que siguen.

En el **Capítulo 1** se describe el contexto en que surgió la Evaluación Docente y el complejo proceso de negociaciones que condujo finalmente a la promulgación de la ley que consagró su obligatoriedad. También se señalan algunos factores adicionales que contribuyeron a dar factibilidad no solo política, sino también técnica a la implementación del sistema, incluyendo el conocimiento de experiencias internacionales y el inicio previo del programa de Asignación de Excelencia Pedagógica.

Enseguida, en el **Capítulo 2** se da cuenta del modelo de evaluación que subyace al sistema y los estándares que le sirven de base (el Marco para la Buena Enseñanza). Se describe la forma en que se construyen los instrumentos de evaluación, cómo se define el nivel de desempeño de los docentes y su resultado final, y se profundiza en la descripción del proceso de corrección del instrumento que mide directamente el desempeño de los profesores, el Portafolio.

La compleja operación que encierra la aplicación de la evaluación cada año es abordada en el **Capítulo 3**. Éste da cuenta de los actores y entidades participantes, así como las responsabilidades que les caben, y recorre cronológicamente los hitos principales que conforman un período de evaluación, mostrando cómo se coordina el proceso a nivel nacional y local. Mediante descripciones, datos y ejemplos, se expone la forma en que opera el sistema y su gradual consolidación y apropiación a nivel local.

El **Capítulo 4** permite conocer los principales resultados generados por la evaluación hasta ahora. Considerando los distintos períodos e instrumentos, se señala la distribución de los docentes en los cuatro niveles de desempeño, y se reportan datos adicionales, por ejemplo, sobre la relación entre resultado en la evaluación y situación laboral de los docentes. En su segunda sección, el capítulo profundiza en los resultados del Portafolio: describe los aspectos evaluados, reporta el desempeño de los docentes de acuerdo a la modalidad, nivel o subsector que imparten y, utilizando ejemplos concretos del tipo recolectado a través del Portafolio, ilustra las prácticas pedagógicas más y menos logradas, según indican los datos de la evaluación.

Por su parte, el **Capítulo 5** detalla las consecuencias asociadas a los resultados de Evaluación Docente: el programa de incentivos conocido como Asignación Variable por Desempeño Individual (AVDI), accesible a los docentes que han obtenido un nivel de desempeño Competente o Destacado y, especialmente, los Planes de Superación Profesional, acciones de apoyo implementadas a nivel local para los docentes con resultado Básico e Insatisfactorio.

El **Capítulo 6** presenta distintos estudios en torno a la validez de la Evaluación Docente, desarrollados por la unidad de Investigación de MIDE UC y por otros especialistas. En referencia al concepto de validez actualmente predominante, se revisan variados aspectos, tales como: la cobertura que los instrumentos de evaluación hacen del Marco para la Buena Enseñanza, la vinculación entre el resultado del docente y el rendimiento de los alumnos, y las consecuencias —positivas o negativas, esperadas y no esperadas—, que el programa demuestra tener para los profesores, los establecimientos y las municipalidades.

En el **Capítulo 7** se analiza la evidencia proveniente de distintos estudios acerca de la relación entre el desempeño de los profesores en la Evaluación Docente y antecedentes individuales de éstos (por ejemplo, formación inicial y condiciones laborales) y del contexto en que ejercen (características del establecimiento y la comuna en que se encuentra). El capítulo complementa datos provenientes de análisis descriptivos con otros obtenidos utilizando modelos lineales jerárquicos.

El **Capítulo 8** permite conocer estudios realizados por distintos académicos e investigadores acerca de la calidad docente y las prácticas de enseñanza presentes en nuestro país. Estas investigaciones se basan en datos y/o evidencia (especialmente videos de clases) recolectados por la Evaluación Docente. La forma en que los profesores de Matemática presentan contenidos propios de esta disciplina a estudiantes de 5° y 6° básico, el tipo de prácticas que emplean los docentes de Lenguaje para enseñar a leer, el uso del tiempo en aula y la identificación de patrones instruccionales recurrentes entre los docentes, son algunos de los tópicos abordados en este capítulo.

Tres reconocidos especialistas internacionales analizan y comentan el sistema de evaluación docente chileno en el **Capítulo 9**. Su análisis permite situar esta experiencia en el contexto mundial de políticas orientadas a evaluar y mejorar la calidad docente. Los comentaristas identifican algunos de los aspectos críticos que enfrenta, desde el punto de vista conceptual y metodológico, la evaluación de profesores y abordan temas de intenso debate en la actualidad como el uso de los resultados de los alumnos para este objetivo. También enuncian algunas preguntas, desafíos y recomendaciones para el programa chileno.

Finalmente, en las **Conclusiones** los editores retoman algunos de los temas centrales abordados en cada capítulo, enfatizando a partir de ellos el valor e implicancias que ha tenido la Evaluación Docente como política educacional en el caso de Chile. Asimismo, señalan algunos desafíos futuros, por ejemplo, respecto de los instrumentos de medición y el tipo de datos recolectados, el monitoreo de las consecuencias formativas asociadas al programa y el papel de la evaluación en una carrera profesional docente; e insinúan las oportunidades de expansión, profundización y mejoramiento que estos desafíos abren para el futuro de la Evaluación Docente.



1

Origen de la Evaluación Docente y su conexión con las políticas públicas en educación

Rodolfo Bonifaz

Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas, CPEIP.
Ministerio de Educación

Origen de la Evaluación Docente y su conexión con las políticas públicas en educación ⁽¹⁾

¿Cómo se origina la Evaluación Docente en Chile? ¿Quiénes participan en ese proceso y qué pasos o hitos claves es posible identificar para que llegara a concretarse el sistema de evaluación? ¿Cómo se armonizaron los intereses de las distintas partes involucradas: el gremio de profesores, las autoridades gubernamentales y los sostenedores municipales? ¿Qué elementos del contexto educacional chileno pueden haber favorecido la instalación del proceso? Estas son las preguntas centrales que busca abordar este capítulo. Se describe el origen del programa de evaluación docente, aplicado desde 2003 en Chile, y que está orientado a evaluar el desempeño profesional de los docentes de aula de establecimientos municipales del país. Se describe el contexto social e histórico en cuyo seno llegó a desarrollarse la evaluación, el extenso y cuidadoso proceso de negociación que condujo a su implementación, la estructura del sistema y su consolidación gradual tras ocho años de implementación.

El Sistema de Evaluación del Desempeño Profesional Docente representa uno de los esfuerzos en materia educacional más relevantes, actualmente en ejecución, orientados al fortalecimiento de la profesión docente. Desde su origen, esta política ha estado fuertemente conectada con un conjunto más amplio de programas e iniciativas promovidas por los gobiernos democráticos a partir de la década del 90 y que buscan dignificar y restituir el carácter profesional de la actividad pedagógica en Chile. El proceso a través del cual fue gestado este sistema de evaluación es reconocido como un ejemplo de construcción de acuerdos a través del diálogo y el entendimiento fructífero entre los principales actores que impulsaron esta política —docentes, municipalidades y gobierno— unidos por el interés compartido de mejorar la calidad y la equidad de la educación.

Dos son los caminos que permiten explicar el proceso de instalación del Sistema de Evaluación del Desempeño Profesional Docente (en adelante Evaluación Docente) como política educacional. El primero, desde una perspectiva histórica, lo ubica dentro del continuo de surgimiento y consolidación de las políticas educacionales orientadas al fortalecimiento de la profesión docente y la calidad de la educación, desde los inicios de la década de los 90. El segundo, desde una perspectiva procedimental, aborda las particularidades del proceso de negociación tripartita que definió las características de la Evaluación Docente y los antecedentes directos relacionados con su elaboración, entre ellos el Marco para la Buena Enseñanza y la creación del programa de Asignación de Excelencia Pedagógica.

1 Los editores expresan su especial agradecimiento por la preparación de este capítulo a Rodolfo Bonifaz, Coordinador del Área de Acreditación y Evaluación Docente del Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas (CPEIP) del Ministerio de Educación desde el año 2005 a la fecha.

Las políticas orientadas al mejoramiento de la calidad de la educación y el fortalecimiento docente

El comienzo de las políticas de fortalecimiento de la profesión docente

A comienzos de la década de los 90, la situación laboral de los profesores y el estatus de la profesión docente acusaban un marcado deterioro, a consecuencia de los cambios ocurridos a partir de 1973 y hasta fines de la década de los 80. Entre ellos, el más importante fue la pérdida por parte de los profesores de su condición de funcionarios públicos estatales (dependientes del Ministerio de Educación) para pasar a ser trabajadores regidos por el Código Laboral y dependientes de las municipalidades. Esta situación conllevó la pérdida de beneficios históricamente alcanzados por el gremio de profesores en términos de sus condiciones de ingreso, remuneraciones, mantención en servicio y sistema de jubilación por años de servicio, entre otros; a ello se sumaron despidos masivos como consecuencia de la flexibilidad laboral del nuevo contexto contractual y normativo (Assael, 2009).

Para enfrentar este complejo cuadro de deterioro en la identidad profesional y condiciones laborales de los docentes, tras el restablecimiento de la democracia a fines de los años 80, se introdujeron sucesivamente una serie de reformas y medidas, cuyo principio guía ha sido de manera clara y persistente el fortalecimiento de la profesión docente, con énfasis reparatorio en una primera etapa y posteriormente anclado al tema del mejoramiento de la calidad de la educación.

Con el propósito de devolver el carácter profesional a la enseñanza, los gobiernos de los años 90 optaron por seguir una estrategia multidimensional incluyendo los siguientes ámbitos: construcción de un discurso público sobre la importancia estratégica de la educación; nuevos y mayores incentivos para atraer candidatos a la profesión docente y retener a los docentes efectivos; un apoyo creciente a los profesores en ejercicio con mayores recursos para el desarrollo profesional; y un ordenamiento jurídico que ha mejorado ostensiblemente sus remuneraciones. Aun cuando estas medidas implicaron una ganancia significativa para la profesión docente, una vez que se fue avanzando también se hizo evidente para la sociedad que si lo que se pretendía era asegurar que todos los alumnos aprendieran en forma efectiva, entonces los profesores necesitaban ser apoyados en el desarrollo de sus competencias profesionales de manera más específica y pertinente a sus particulares requerimientos de formación.

Las políticas de fortalecimiento de la profesión docente pueden analizarse en perspectiva haciendo referencia a tres hitos determinantes. Estos son: la promulgación del Estatuto Docente, el impulso a la Reforma Educacional de mediados de los 90 y el acuerdo por la Calidad de la Educación a comienzos del año 2000. En lo que sigue se revisará brevemente cada uno de ellos.

Políticas de reparación y Estatuto Docente

Un hito prominente en el desarrollo de las políticas educacionales que buscaban fortalecer la profesión docente fue la publicación en el año 1991 de la Ley 19.070, que fija el Estatuto de los Profesionales de la Educación. Este cuerpo legal norma los requisitos, deberes, obligaciones y derechos de los profesionales de la educación. Su propósito principal fue incentivar la profesionalización de la labor docente y mejorar las condiciones de trabajo y remuneraciones asociadas a ella, haciendo reconocimiento explícito y legal del carácter profesional de la función docente. Entre otros aspectos, este estatuto definió las causales para poner término a la relación laboral entre docente y municipio. También estableció un mejoramiento del piso salarial a través de la Renta Básica Mínima Nacional (RBMN), y sobre dicho piso formuló una arquitectura de asignaciones para los docentes del sector municipal en función de su antigüedad y el perfeccionamiento acumulado, diferenciando además las remuneraciones en favor de los docentes con

responsabilidades directivas o técnicas, al modo de una carrera funcionaria tradicional. Respecto a las condiciones de trabajo, estipuló la duración máxima de la jornada laboral contratada con un mismo empleador (44 horas cronológicas) y estructuró dicha jornada en horas de docencia de aula (hasta 33 horas en contratos de 44 horas) y horas de actividades curriculares no lectivas.

Asimismo, el Estatuto Docente fijó normas para ingresar a la carrera para los profesionales de la educación del sector municipal, reglamentó el período de vacaciones de verano y el tiempo máximo de perfeccionamiento en dicho período, y estableció el derecho de participación en las actividades de la unidad educativa. Por último, y de especial importancia para el tema que ocupa este capítulo, el Estatuto reconoció el derecho al perfeccionamiento profesional, estimuló la formación en servicio y consagró la autonomía del docente en el ejercicio de su función así como su responsabilidad personal por su desempeño, el que sería evaluado a través de un sistema de calificaciones².

Reforma educacional y fortalecimiento de la profesión docente a mediados de los 90

Con posterioridad a la promulgación del Estatuto Docente, el fortalecimiento de la profesión docente continuó siendo una de las prioridades de los gobiernos siguientes.

En el contexto de los cambios al sistema educacional impulsados por el gobierno del Presidente Frei, el Informe del Comité Técnico Asesor para la Reforma Educacional de 1994 (Informe Brunner) concluyó que para asegurar la calidad de los aprendizajes resultaba imprescindible fortalecer la profesión docente, dotándola de las condiciones necesarias para su desarrollo profesional y para favorecer su desempeño al más alto nivel posible.

El discurso presidencial del 21 de mayo de 1996, junto con anunciar la consolidación de la Jornada Escolar Completa, planteó la voluntad gubernamental de promover la profesionalización de los docentes en el marco de un relanzamiento de la Reforma Educacional. Este re-impulso consideró elementos como el Programa de Fortalecimiento de la Formación Inicial Docente, el de Perfeccionamiento Fundamental, los programas de becas para estudiantes destacados que optaran por seguir la carrera de pedagogía, las pasantías y diplomados en el exterior y los premios de excelencia docente. Estas y otras iniciativas permitieron cristalizar los llamados cuatro pilares de la Reforma Educacional en curso en dicho momento: el Programa de Mejoramiento e Innovación Pedagógica, la Reforma Curricular, el Desarrollo Profesional Docente y la Jornada Escolar Completa.

Acuerdo por la Calidad de la Educación entre magisterio docente y gobierno

A fines de la década de los 90 y comienzos del nuevo siglo se inició un proceso de acercamiento entre los intereses de los docentes expresados a través de su organización gremial y los objetivos de la Reforma Educacional impulsada desde el gobierno.

Durante la segunda vuelta de las elecciones presidenciales del año 2000, el entonces candidato Ricardo Lagos suscribió en conjunto con el Colegio de Profesores el documento “Bases para un Compromiso para el Fortalecimiento de la Educación y la Profesión Docente”, firmado por las partes el 6 de enero de 2000. En dicho acuerdo, se expresaba la voluntad del futuro gobernante de fortalecer la educación de interés público, mejorar la participación social y magisterial en las decisiones educacionales y enfatizar el mejoramiento laboral y el desarrollo profesional de los docentes. También se manifestaba el compromiso a discutir el

.....

2 En la práctica, este sistema de calificaciones nunca logró ser implementado por la oposición gremial del magisterio, siendo luego reemplazado por un sistema de evaluación de los docentes de aula de carácter formativo, el Sistema de Evaluación del Desempeño Profesional Docente, según lo dispuesto por el literal d) del artículo 12° de la Ley 19.933, del 31 de enero de 2004.

reemplazo del vigente sistema de calificaciones incluido en el Estatuto Docente, que en la práctica no había sido aplicado. Tras concluirse la elección presidencial, en marzo de 2000 comenzaron las negociaciones entre el Ministerio de Educación y el Colegio de Profesores orientadas a concretar este acuerdo. Si bien el tema de los ajustes a las remuneraciones comunes y diferenciadas fue parte importante en ellas, la idea articuladora de las reuniones entre magisterio y gobierno no solo estuvo fundada en el tema salarial sino que se trabajó dentro del concepto más amplio de mejoramiento de la condición docente, en el marco de una voluntad común de contribuir a fortalecer la calidad y la equidad de la educación chilena.

Entre los acuerdos alcanzados en la negociación figuró el perfeccionamiento de la Jornada Escolar Completa, que contempló una evaluación y el compromiso de un debate para enfrentar sus desafíos. También se acordó fortalecer las capacidades de gestión y fiscalización del Ministerio de Educación para el efectivo cumplimiento del derecho a la educación y el buen uso de los recursos públicos, respetando al mismo tiempo los principios de la libertad de enseñanza y la autonomía de los sostenedores. De igual forma, se acordó incrementar las remuneraciones generales de los profesionales de la educación y mejorar las condiciones de trabajo de los docentes, considerando acciones como el aumento en una hora del tiempo no-lectivo, el perfeccionamiento de las normas de contratación, la implementación de programas para el análisis y prevención de las enfermedades profesionales propias de los docentes, el apoyo a profesores que se desempeñan en condiciones de vulnerabilidad y la reducción del número de alumnos por curso. Se acordó promover la creación de instancias de participación del profesorado creando los Consejos Regionales de Educación, estudiando perfeccionar otros espacios existentes, y promoviendo la incorporación de profesionales de la educación al Consejo Superior de Educación, medidas que tuvieron un valor como símbolo de reparación hacia el Magisterio.

Entre los temas de acuerdo, producto de la negociación, uno central fue el fortalecimiento de la profesión docente. En ese ámbito, los acuerdos incluyeron medidas como las modificaciones a la norma sobre perfeccionamiento y el financiamiento de pasantías dentro de Chile. También se aceptó la creación de un incentivo de carácter individual para reconocer y premiar la excelencia de los docentes de aula (la Asignación de Excelencia Pedagógica). Además, se estableció que los docentes así destacados serían convocados para poner sus saberes y experiencias al servicio de sus colegas por medio de la Red Maestros de Maestros. Finalmente, se acordó construir una propuesta conjunta entre el Colegio de Profesores, el Ministerio de Educación y la Asociación Chilena de Municipalidades para reemplazar el sistema de calificaciones establecido en el Estatuto Docente. Dicho sistema nunca logró implementarse debido a la resistencia de los profesores, quienes temían que se repitieran despidos injustificados al alero de un proceso evaluativo que era percibido como punitivo y mal fundado (Assael & Pavez, 2008). Entre las características centrales que debía tener el nuevo sistema se acordó que debía ser de carácter formativo y basarse en estándares de desempeño explícitos y previamente acordados.

Construcción y desarrollo del Sistema de Evaluación del Desempeño Profesional Docente

Bases para la construcción de la Evaluación Docente

El diálogo permanente entre las autoridades de los gobiernos de turno y del magisterio junto con la priorización del consenso por sobre el conflicto, fueron aspectos claves en el desarrollo de la política educacional que ha caracterizado y acompañado las reformas educacionales impulsadas en la década de los 90. En el caso de la Evaluación Docente, si bien un hito clave fue el acuerdo del año 2000, existe una serie de antecedentes adicionales que dieron cuenta de la necesidad de cambiar las calificaciones del Estatuto Docente y favorecieron un proceso de acercamiento entre docentes y gobierno sobre la materia.

El Primer Congreso Nacional de Educación organizado por el Colegio de Profesores en octubre de 1997, abordó en profundidad —entre otros importantes tópicos— el malestar del gremio con el sistema de calificaciones del Estatuto Docente. El rechazo a su carácter punitivo y el no orientarse al desarrollo profesional, mejoramiento económico o configuración de una carrera docente, además de la desconfianza que muchos profesores mantenían respecto a los directores, los jefes de UTP y las corporaciones encargadas de la calificación, propiciaron el surgimiento de una primera propuesta de evaluación docente por parte del Magisterio. En las conclusiones de este Congreso destacó el reconocimiento, por parte de los profesores, del derecho de la comunidad educativa a evaluar la gestión del establecimiento, y la necesidad y conveniencia de contar con una evaluación remedial y formativa de las falencias profesionales. Entre sus conclusiones, el Congreso recomendó disponer de un sistema de evaluación en cuya elaboración participara ampliamente el Colegio de Profesores, crear una comisión técnica de carácter nacional para diseñar instrumentos evaluativos realistas y objetivos, y convocar a una consulta nacional para la validación de la propuesta de un sistema nacional de evaluación del desempeño docente⁽³⁾. Sin lugar a dudas, el hecho que el Magisterio conceptualizara la evaluación de la calidad del ejercicio docente como una responsabilidad del gremio fue determinante para el éxito tanto del diseño como de la implementación de la Evaluación Docente.

Al año siguiente, en la negociación de 1998 con el gobierno, el gremio docente planteó la derogación del sistema de calificaciones y su reemplazo por un sistema de evaluación del desempeño profesional, lo que fue aceptado parcialmente por el gobierno, que entonces estuvo disponible solo para discutir acerca de medidas destinadas a perfeccionar el sistema de calificaciones del Estatuto Docente⁽⁴⁾.

Dada la alta convergencia entre las partes en la conveniencia de replantear el sistema de calificaciones, finalmente, en la negociación del año 2000, el Colegio de Profesores junto con el Ministerio de Educación acordaron las bases de un consenso acerca de los principios básicos para instalar un sistema de evaluación formativo orientado al desempeño profesional de los docentes. Dicho acuerdo, que enfatizó la función del Estado como responsable de su implementación en la perspectiva de potenciar un sistema de educación pública de calidad para todos, expresó que el proceso de evaluación debía desarrollarse de manera participativa y transparente como requisito para su función formativa y de desarrollo profesional⁽⁵⁾. En el acuerdo se estableció la creación de un comité técnico tripartito que tendría la responsabilidad de avanzar en la construcción del nuevo sistema, y en el que participarían el Colegio de Profesores, la Asociación Chilena de Municipalidades y el Ministerio de Educación.

En el trabajo del señalado comité técnico tripartito participaron diversos sectores que pusieron en juego sus propias posiciones ideológicas, políticas y técnicas. Esta diversidad enriqueció la discusión hasta alcanzarse finalmente un marco de acuerdo respecto a las características centrales que debía tener el sistema de evaluación. La comisión tripartita sesionó en dos subcomisiones, una relacionada con los estándares de desempeño que fundamentarían la evaluación y otra a cargo de desarrollar la propuesta global del sistema de evaluación. El trabajo del subcomité sobre estándares se basó en los estándares previamente definidos para la formación inicial docente (Ministerio de Educación, 2000) y culminó su trabajo con la aprobación del Marco para la Buena Enseñanza, tras dos consultas públicas realizadas en diciembre de 2001 y abril de 2002.

Una vez consensuado el Marco para la Buena Enseñanza y mientras la comisión tripartita continuaba avanzando en la elaboración de los lineamientos para implementar el sistema de evaluación, se resolvió realizar paralelamente una experiencia piloto de evaluación docente en cuatro comunas del país (Puerto Montt, Melipilla, Cerro Navia y Conchalí), que fuera adjudicada a un equipo de la Universidad Católica de Temuco. Esta experiencia tuvo como objetivo poner a prueba el tipo de evaluación que se

3 Colegio de Profesores de Chile A.G, Informe Final Primer Congreso Nacional de Educación, octubre de 1997.

4 Revista Docencia 20, agosto de 2003, p. 80. *Fin de las calificaciones: Nuevo Sistema Nacional sobre Evaluación del Desempeño Docente.*

5 Revista Docencia 10, junio de 2000, p. 59. *Criterios Fundantes de un Sistema de Evaluación de Profesionales de la Educación.*

estaba concordando, es decir, se basó en los estándares del Marco para la Buena Enseñanza, capacitó a los profesores para que se hicieran cargo de la evaluación de sus pares y utilizó instrumentos tales como un portafolio, observaciones o filmaciones de clases y entrevistas. El piloto se realizó durante el segundo semestre del año 2002 y, efectivamente, permitió apreciar cómo se comportaban los distintos instrumentos, resultando útil para continuar avanzando en la discusión y construcción del modelo que finalmente se acordaría.

El 24 de junio de 2003, el esfuerzo tripartito dio sus frutos y las partes suscribieron el documento “Informe Final Comisión Técnico Tripartita sobre la Evaluación del Desempeño Profesional Docente”⁽⁶⁾, que dio cuenta del consenso establecido y definió la estructura central del sistema de evaluación docente. Dicho informe técnico fue posteriormente ratificado por las máximas autoridades del Ministerio de Educación, la Asociación Chilena de Municipalidades y el Colegio de Profesores⁽⁷⁾. Finalmente, el 10 de julio de 2003, mediante una Consulta Nacional en la que participaron 65.846 profesores, que representaban un 80% de los docentes del ámbito municipal, fue aprobada la propuesta del nuevo sistema de evaluación por las bases del Colegio de Profesores con un 63,13% de los votantes a favor⁽⁸⁾.

La estructura del Sistema de Evaluación del Desempeño Profesional Docente consignada en el informe técnico emanado de la Comisión Técnica Tripartita⁽⁹⁾, definió así los elementos centrales que hasta hoy caracterizan la Evaluación Docente:

- **¿Para qué se evalúa?:** se evalúa con el propósito de contribuir al desarrollo y fortalecimiento sostenido de la profesión docente, específicamente al mejoramiento de la labor pedagógica de los educadores, con el fin de favorecer el aseguramiento de aprendizajes de calidad de los alumnos y aportar información valiosa a cada uno de los actores educativos, al sistema de formación inicial y continua de los docentes, y al sistema educativo en general. La evaluación tiene un carácter fundamentalmente formativo y constituye una oportunidad única para que el docente se conozca mejor desde el punto de vista profesional, identificando tanto sus fortalezas como los aspectos que puede mejorar.
- **¿Qué se evalúa?:** se evalúa lo propio de la misión del educador, es decir, la calidad de la enseñanza, a partir de los dominios y criterios establecidos en el Marco para la Buena Enseñanza. Se trata de una evaluación de carácter explícito, es decir, el docente conoce previamente los criterios a través de los cuales será evaluado. El sistema evalúa al docente exclusivamente por su ejercicio profesional, considerando el contexto y las condiciones de trabajo en que éste se desenvuelve. No se evalúa al docente por su desempeño funcionario–administrativo ni por el rendimiento escolar de sus alumnos.
- **¿Cuáles son las consecuencias de la evaluación para los profesores?:** para todos los profesores, en primer lugar, la evaluación significa un proceso de auto reflexión respecto de su desempeño profesional, que le permite identificar fortalezas y debilidades y proyectar un camino para potenciar las primeras y superar las segundas.

La evaluación final de cada profesor corresponde a una apreciación global relativa al desempeño profesional relacionado con el Marco para la Buena Enseñanza, sus dominios y criterios, que se establece en cuatro niveles de desempeño: Destacado, Competente, Básico e Insatisfactorio.

Los profesores evaluados como destacados y competentes tienen acceso prioritario a oportunidades de desarrollo profesional: ventajas en los concursos, pasantías en el extranjero, profesores guías de talleres, participación en seminarios académicos, entre otros; al tiempo que

.....

6 Este informe fue firmado por Carlos Eugenio Beca, Jaime Veas, Xavier Vanni y Rodolfo Bonifaz, en representación del Ministerio de Educación; Sadi Melo, Luisa Saldías y Alba Maldonado, en representación de la Asociación Chilena de Municipalidades; y Mario Aguilar, Jenny Assael, Isabel Guzmán, Hugo Miranda, en representación del Colegio de Profesores.

7 Suscribieron el documento: Sergio Bitar, Ministro de Educación; Pedro Sabat, Presidente de la Asociación Chilena de Municipalidades; y Jorge Pavez, Presidente del Colegio de Profesores.

8 Revista Docencia 20, agosto 2003, p. 81. *Fin de las calificaciones: Nuevo Sistema Nacional sobre Evaluación del Desempeño Docente.*

9 Ver Informe Final Comisión Técnico Tripartita sobre la Evaluación del Desempeño Profesional Docente, del 24 de junio de 2003.

se les abre la posibilidad de optar a percibir la Asignación Variable por Desempeño Individual, previa rendición de una prueba escrita de conocimientos disciplinarios y pedagógicos.

Los profesores que obtengan en su evaluación global el nivel de Básico o Insatisfactorio disponen de Planes de Superación Profesional gratuitos destinados a superar sus debilidades, los que serán financiados por el Ministerio de Educación, y diseñados y ejecutados por los respectivos sostenedores municipales⁽¹⁰⁾.

Los docentes que obtengan resultado Insatisfactorio deberán evaluarse nuevamente al año siguiente. Si en su segunda evaluación un docente resulta nuevamente Insatisfactorio, deberá dejar la responsabilidad de curso para trabajar durante el año su Plan de Superación Profesional con un docente tutor y será sometido a una tercera evaluación. En caso de mantener el desempeño Insatisfactorio deberá abandonar el sistema municipal.

- **¿Quiénes se evalúan?:** se evalúan cada cuatro años todos los docentes de aula del ámbito de la educación municipal que cuenten con, a lo menos, un año de ejercicio profesional, a excepción de aquellos que han obtenido resultado de nivel Insatisfactorio, quienes deben evaluarse al año inmediatamente siguiente con el propósito de asegurar que las falencias en el cumplimiento de estándares mínimos de desempeño no afecten significativamente el quehacer docente⁽¹¹⁾.
- **¿Cuáles son los instrumentos y fuentes de información para la evaluación?:** la evaluación se realiza sobre la base de las evidencias del desempeño profesional de cada docente recogidas a través de cuatro instrumentos: Autoevaluación, Entrevista por un Evaluador Par, Informes de Referencia de Terceros y Portafolio de desempeño pedagógico, todos los cuales deben guardar relación con los dominios, criterios y descriptores fijados en el Marco para la Buena Enseñanza y contar con la aprobación previa del CPEIP del Ministerio de Educación.
 - **Portafolio de desempeño pedagógico:** este instrumento tiene como función recoger, a través de productos estandarizados, evidencia verificable respecto de las mejores prácticas de desempeño del docente evaluado. Contempla la presentación de productos escritos, en los cuales el docente debe reportar distintos aspectos de su quehacer profesional y un registro audiovisual consistente en la grabación en video de una clase de cuarenta minutos de duración. La corrección de las evidencias del Portafolio debe ser realizada por profesores correctores del mismo nivel y subsector de aprendizaje del docente evaluado, especialmente entrenados para el desempeño de esta función.
 - **Autoevaluación:** tiene como función principal que el docente reflexione sobre su práctica pedagógica y valore su propio desempeño profesional a partir de una pauta previamente determinada.
 - **Entrevista por un Evaluador Par:** aplicada mediante una pauta de entrevista estructurada por un evaluador par previamente capacitado, esta entrevista consulta al docente acerca de su práctica por aproximadamente una hora. Una vez que la entrevista finaliza, el evaluador par aplica una rúbrica asignando, en base a ella, uno de los cuatro niveles de desempeño a cada respuesta recogida del docente, permitiendo contar con la apreciación de un docente de aula del mismo nivel y modalidad de aquel cuyo desempeño es evaluado.
 - **Informe de Referencia de Terceros:** pauta estructurada que debe ser completada por el Director y Jefe de la Unidad Técnico Profesional (UTP) del establecimiento y contiene distintas preguntas referidas al quehacer del docente evaluado. Para cada pregunta, el Director y el Jefe de UTP deben emitir, por separado, su evaluación del docente en uno de los cuatro niveles de desempeño, permitiendo recoger el juicio de los superiores jerárquicos respecto a la práctica del docente evaluado.



10 Para más información sobre las consecuencias de la Evaluación Docente y los Planes de Superación Profesional, ver Cortés & Lagos, en este volumen.

11 Posteriormente, por efecto de la ley 20.501, promulgada el 26 de febrero de 2011, se ha establecido que los docentes que obtienen nivel de desempeño Básico deben evaluarse al año subsiguiente, modificación que entrará en aplicación para los docentes que se evalúen a partir del año 2011.

Junto con lo anterior, los últimos tres instrumentos recogen información contextual a partir de la mirada del propio docente, el evaluador par y el director y el jefe técnico, sobre aquellos aspectos que pudiesen influir positiva o negativamente en el desempeño de cada docente evaluado. Estos antecedentes son puestos en conocimiento de la Comisión Comunal de Evaluación respectiva, la que debe considerarlos al momento de pronunciarse colegiadamente respecto al nivel de desempeño final de cada docente evaluado en la comuna.

Los cuatro instrumentos de evaluación permiten que la Comisión Comunal de Evaluación pueda contar con diversas miradas y juicios evaluativos acerca del nivel de desempeño profesional de cada docente con relación al Marco para la Buena Enseñanza y atendiendo al contexto en el cual cada uno de ellos se desempeña.

- **¿Quiénes determinan la evaluación final de cada docente?:** la evaluación final de cada docente es realizada por la Comisión Comunal de Evaluación respectiva, la que se encuentra integrada por docentes de aula previamente seleccionados y capacitados por el Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas del Ministerio de Educación (en adelante CPEIP) entre aquellos que anualmente postulan voluntariamente para desempeñar el cargo de evaluadores pares. Estos docentes han de contar con resultado previo de Destacado o Competente en la Evaluación Docente.
- **¿Cómo son corregidos los instrumentos?:** de los cuatro instrumentos involucrados en la evaluación, solo el Portafolio es sometido a un proceso de corrección propiamente tal, pues en los demás instrumentos los propios evaluadores han asignado el nivel de desempeño correspondiente, frente a cada pregunta o indicador evaluado.

Los Portafolios son corregidos en Centros de Corrección, los que han de funcionar en distintas universidades del país. Allí, los correctores (docentes de aula con experiencia en la modalidad, nivel y subsector que están evaluando) examinan la evidencia enviada por los docentes y asignan un nivel de desempeño a cada aspecto o indicador evaluado en el Portafolio.

Los resultados de la Entrevista por un Evaluador Par, son digitados por éste en un sistema computacional. Los resultados de la Autoevaluación y el Informe de Referencia de Terceros son digitados centralmente.

La información de todos los instrumentos se ingresa en un sistema computacional que recoge los puntajes y calcula el nivel de desempeño global de cada docente, generando un reporte de resultados a ser revisado por la respectiva Comisión Comunal de Evaluación.

Los distintos aspectos antes señalados quedaron expresados en la normativa que consagró la Evaluación Docente y en su forma de organización a partir del inicio de su implementación en el año 2003. En el apartado que sigue se describen algunos aspectos centrales de esa organización, en particular los distintos niveles (central, comunal y nivel escuela) y los principales actores responsables del proceso⁽¹²⁾.

Estructura del Sistema de Evaluación del Desempeño Profesional Docente

La estructura del sistema de evaluación docente está compuesta de tres niveles: el nivel central, el nivel comunal y el nivel escuela (ver Figura 1).

El **nivel central** que se encuentra integrado por el **CPEIP** del Ministerio de Educación, entidad que le corresponde la coordinación técnica para la adecuada aplicación de los procesos anuales de Evaluación Docente, la que para tales efectos dispone de las siguientes dos instancias técnicas asociadas: una asesoría técnica independiente y una comisión técnica asesora:

.....

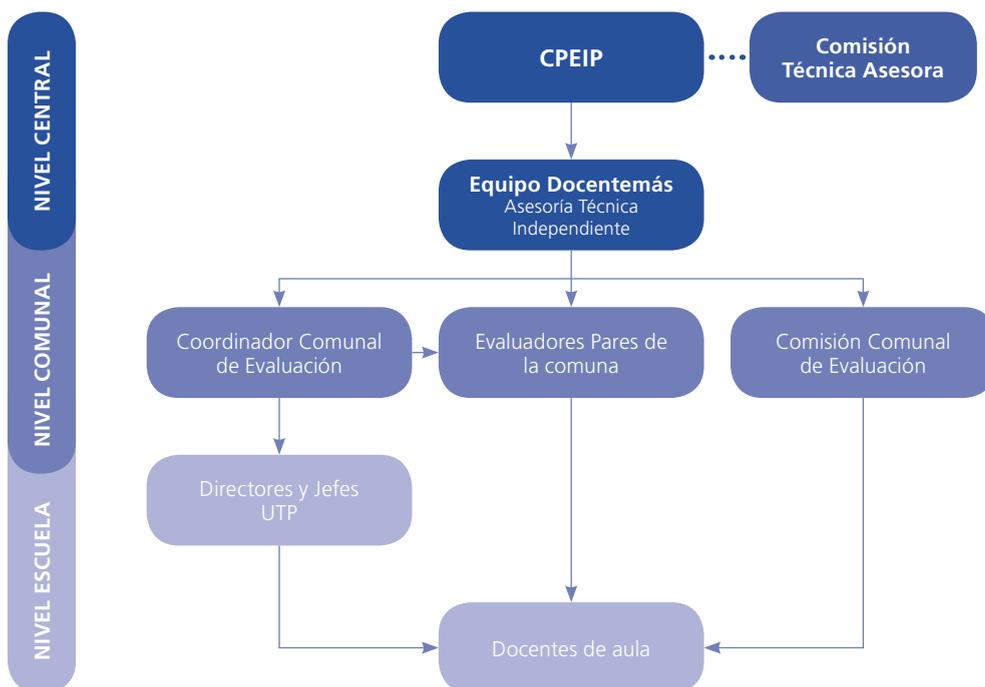
12 Para más información sobre la forma de implementación de la Evaluación Docente en cada período, ver Sun, Calderón, Valerio & Torres, en este volumen.

- **Asesoría Técnica Independiente:** la normativa establece que el CPEIP, para la realización de las funciones de coordinación técnica que debe desarrollar en el sistema de evaluación docente, debe necesariamente contar con una asesoría técnica independiente, impartida por entidades académicas de educación superior con experiencia comprobada en el área de la formación docente y la evaluación de desempeños profesionales a escala nacional.

Desde su creación a la fecha, esta asesoría ha estado a cargo de MIDE UC, Centro de Medición de la Pontificia Universidad Católica de Chile, específicamente a través de un equipo conformado por 36 profesionales, técnicos y administrativos, provenientes especialmente del campo de la educación y de la psicología, quienes bajo la supervisión del CPEIP colaboran en la implementación del conjunto de procesos y operaciones que requiere la Evaluación Docente.

- **Comisión Técnica Asesora:** la normativa también establece que para la realización de la coordinación técnica el CPEIP se hará asesorar por una Comisión Técnica Asesora, conformada por tres representantes de la asociación gremial de docentes con mayor representación en el país, tres académicos de educación superior designados por el CPEIP y tres representantes designados por la Asociación Nacional de Municipalidades con mayor representatividad del país.

• **Figura 1**
 • Estructura del Sistema de Evaluación del Desempeño Profesional Docente



A **nivel comunal** es necesario distinguir los roles que corresponden a la Comisión Comunal de Evaluación como entidad colegiada y a sus integrantes, el Coordinador Comunal de Evaluación y los Evaluadores Pares, como actores individuales:

- **Coordinador Comunal de Evaluación:** esta condición se encuentra asignada a quien desempeña el cargo de Jefe del Departamento de Administración Municipal de Educación o de Director de la Corporación de Educación Municipal, según corresponda; quien puede delegar la ejecución de las funciones administrativas inherentes a la Evaluación Docente en un Encargado de Evaluación, pero no así la responsabilidad del proceso, que le corresponde por reglamento.
- **Evaluadores Pares:** son docentes de aula del mismo nivel escolar, sector del currículo y modalidad del docente evaluado, a los que les corresponde la responsabilidad de aplicar la entrevista al docente, informarla y participar en la Comisión Comunal de Evaluación. Los evaluadores pares son anualmente seleccionados y preparados para realizar su función por el CPEIP, que establece las bases para esta selección, determinando el número de evaluadores pares requeridos para el proceso del año respectivo, mediante un proceso de postulación voluntario.
- **Comisión Comunal de Evaluación:** es la entidad colegiada que tiene la responsabilidad de aplicar localmente el sistema de evaluación y que se encuentra compuesta por el Coordinador Comunal de Evaluación y por el conjunto de los evaluadores pares de la comuna. Son funciones de la Comisión Comunal de Evaluación:
 - tomar conocimiento y aprobar los reportes de resultados, considerando la información del contexto del docente evaluado y pronunciarse respecto de su nivel de desempeño final, para lo cual podrá ratificar el nivel de desempeño indicado en el reporte de resultados o modificar la evaluación final de un docente mediante acuerdo especialmente fundado, con la aprobación de, a lo menos, dos tercios de sus integrantes.
 - conocer y resolver los recursos de reposición presentados por aquellos docentes que solicitan la reconsideración de su resultado, a la propia Comisión (única instancia reconocida por la normativa para resolver sobre ellos).

Conforme a reglamento, la evaluación final de cada docente debe ser realizada por los evaluadores pares que integran la respectiva Comisión Comunal de Evaluación, con la excepción de aquel que realizó la entrevista al docente, quien no podrá concurrir con su voto a la decisión acerca del resultado final de ese profesor o profesora. La decisión final de la Comisión Comunal de Evaluación deberá contenerse en un Informe de Evaluación Individual para cada profesional de la educación y deberá fundarse en el reporte de resultados que integra ponderadamente los resultados del conjunto de los instrumentos de evaluación y en las variables contextuales que hubiesen dispuesto para tal efecto el docente evaluado, el evaluador par que lo entrevistó, el director y el jefe técnico del respectivo establecimiento.

A nivel escuela es posible distinguir los roles que le corresponde a los docentes evaluados y a los Directores y Jefes de Unidades Técnico Pedagógicas.

- **Docentes evaluados:** corresponden a todos los profesionales de la educación que, en conformidad a la letra a) del artículo 6° del Decreto con Fuerza de Ley 1, de 1996, del Ministerio de Educación, ejercen funciones en aula en establecimientos del sector municipal, administrados ya sea directamente por municipios o por corporaciones municipales y en los que habiendo sido municipales son administrados por corporaciones educacionales privadas, de acuerdo con las normas establecidas en el Decreto con Fuerza de Ley 1–3063, del Ministerio del Interior, de 1980. Conforme a la programación anual dispuesta para tales efectos por el CPEIP, corresponde que sean evaluados todos los docentes de aula del ámbito de la educación municipal, es decir, aquellos docentes que cumplen funciones en Educación Parvularia, Educación Básica, Educación Especial o Diferencial, Formación General de Enseñanza Media, Formación Diferenciada de Enseñanza Media Humanístico–Científica y especialidades de la Educación Media Técnico Profesional, y Educación de Adultos⁽¹³⁾.

13 Hasta el año 2010, se han incorporado ya los docentes de todos estos niveles, subsectores y modalidades de enseñanza, salvo aquellos que imparten Educación de Adultos y especialidades de la Educación Media Técnico Profesional.

- **Directores y Jefes de Unidades Técnico Pedagógicas:** a quienes corresponde emitir los informes de referencia de terceros, a través de una pauta estructurada especialmente diseñada por el CPEIP. En aquellos establecimientos en que no exista ninguno de dichos cargos, este informe deberá ser emitido por el superior jerárquico del profesor evaluado y por el Jefe Técnico Pedagógico comunal cuando exista.

Instalación y evolución del marco normativo

Mientras se elaboraba el proyecto de ley de evaluación docente y cuando aún no se iniciaba su tramitación parlamentaria, sobre la base de la fortaleza del consenso y del acuerdo técnico suscrito por las partes, se puso en marcha el proceso evaluativo en el año 2003. Su inicio fue gradual, con la evaluación en ese primer período solo de docentes generalistas de Primer Ciclo de Enseñanza Básica en 63 comunas del país y con la participación de 3.673 docentes.

El proyecto de ley que daba cuenta del acuerdo ingresó al Parlamento en noviembre de 2003 y nueve meses después fue aprobado, el 4 de agosto de 2004, con 59 votos a favor y 19 en contra. Tras su promulgación, la Ley 19.961 fue publicada en el diario Oficial el 14 de agosto de 2004. Con anterioridad, dando continuidad al proceso de evaluación implementado el año anterior, se inició la Evaluación Docente correspondiente al año 2004, en la que participaron 104 comunas y se evaluaron 1.719 docentes, incluyendo profesores generalistas de Primer Ciclo Básico y docentes de Segundo Ciclo que impartían los subsectores de Lenguaje y Comunicación, Matemáticas, Estudio y Comprensión de la Naturaleza, y Estudio y Comprensión de la Sociedad. No obstante lo anterior, la demora en la aprobación de la ley y el hecho de que el reglamento de evaluación docente aún no se encontrara aprobado hicieron que la cantidad de docentes evaluados durante este segundo año de aplicación se redujera ostensiblemente.

La Ley 19.961 dejó casi la totalidad de los aspectos operativos del proceso de evaluación docente a la determinación de un reglamento, que debería dictarse en el plazo de 120 días contados desde la publicación de la ley. Con fecha 30 de agosto de 2004, el Ministerio de Educación ingresó a la Contraloría General de la República la primera versión de dicho reglamento y, tras la introducción de variadas modificaciones a su texto y nueve meses de tramitación, la Contraloría tomó razón del Decreto Supremo de Educación 192, de 2004, que aprobó el reglamento sobre evaluación docente. Con ello se hizo posible que, al año siguiente, la tercera aplicación del proceso de evaluación docente pudiera normalizarse. Con ley y reglamento aprobados, en 2005 la Evaluación Docente tuvo por primera vez aplicación nacional (en todas las comunas) y fueron evaluados 10.665 docentes.

Para asegurar la validez legal de las evaluaciones realizadas con anterioridad a la aprobación de la Ley 19.961 y de su reglamento, a través de otra disposición legal⁽¹⁴⁾ se declaró válida, para todos los efectos, la evaluación docente realizada durante 2003 y 2004 en las comunas que voluntariamente se incorporaron a ella en dichos años, las que se encuentran consignadas en las Resoluciones Exentas N° 10.015 y N° 11.750, ambas del 2004, del Ministerio de Educación.

Con posterioridad, se han realizado tres modificaciones legales significativas para el sistema de evaluación. La primera de ellas⁽¹⁵⁾ tuvo por objeto establecer un bono compensatorio para los docentes que, tras haber obtenido un resultado Insatisfactorio en tres evaluaciones consecutivas, debieran a causa de ello hacer abandono de la dotación docente. Al mismo tiempo, se estableció una sanción legal consistente en

14 Artículo 5° transitorio de la Ley 19.933.

15 El inciso primero del artículo 36° de la Ley 20.079, estableció un bono pro calidad de la dotación docente para los profesionales de la educación que habiendo sido evaluados de conformidad a lo dispuesto en el artículo 70 del Decreto con fuerza de ley 1, de 1996, del Ministerio de Educación, hayan dejado de pertenecer a dicha dotación por encontrarse en la situación señalada en el inciso séptimo del señalado artículo. Este bono asciende a una suma entre 3 y 5,5 millones de pesos (equivalentes a US\$6.200 y US\$11.400, aproximadamente), dependiendo del promedio de las últimas doce remuneraciones del docente.

asignar un resultado de nivel Insatisfactorio a todos los docentes que, debiendo cumplir con la obligación legal de evaluarse, no lo hicieren sin existir para ello causa justificada⁽¹⁶⁾. Esta sanción ha permitido reducir ostensiblemente el incumplimiento de la obligación legal de evaluarse: en el año 2005, cuando la sanción todavía no existía, esta situación alcanzó a un 30% de los docentes inscritos para ser evaluados, proporción que se redujo a un 10,5% durante su primer año de vigencia (2006) y ha mostrado un descenso creciente hasta llegar a solo un 3,5% en el año 2010⁽¹⁷⁾.

Una segunda modificación legal⁽¹⁸⁾ facultó a los docentes a quienes restasen tres años o menos para cumplir la edad legal para jubilar, a eximirse voluntariamente de la Evaluación Docente, bajo la condición de que presentasen la renuncia anticipada e irrevocable a su cargo, la que ha de hacerse efectiva al cumplirse la edad legal de jubilación por el solo ministerio de la ley.

Por último, la recientemente promulgada Ley 20.501, del 26 de febrero de 2011, introdujo cuatro importantes modificaciones relacionadas con las consecuencias de la Evaluación Docente, a saber: reduce las oportunidades para superar un resultado Insatisfactorio de tres a dos evaluaciones consecutivas (tras las cuales el docente está forzado a abandonar la dotación municipal); reduce de cuatro a dos años el período entre evaluaciones, en el caso de los docentes que hubieren obtenido un resultado Básico y establece para éstos la exigencia de superar dicho resultado en un máximo de tres evaluaciones consecutivas, debiendo en caso contrario hacer abandono de la dotación docente; y, por último, faculta a los directores de establecimientos a despedir anualmente hasta un 5% de la dotación docente del respectivo establecimiento entre aquellos profesores cuyo resultado vigente en la Evaluación Docente sea Básico o Insatisfactorio⁽¹⁹⁾.

Iniciativas que favorecieron la construcción de la evaluación

Además de los elementos ya mencionados en la sección anterior de este capítulo, existen otros aspectos directamente relacionados con el proceso de construcción de la Evaluación Docente que contribuyeron a aportar al sistema una solidez que de otro modo no poseería. En los cuatro puntos siguientes se abordará brevemente cada uno de ellos: la realización del Seminario Internacional sobre Profesionalismo Docente y Aseguramiento de la Calidad, en mayo de 2001; las misiones de estudio realizadas a Irlanda y Escocia, en 2001 y a Cuba, en 2002; la construcción y aprobación del Marco para la Buena Enseñanza, durante el año 2002; y la implementación del programa de Asignación de Excelencia Pedagógica, en 2003.

Seminario internacional sobre profesionalismo docente

Con el propósito de sensibilizar a los actores educativos sobre la importancia de contar con un sistema de evaluación docente y de recoger la opinión de expertos internacionales sobre la materia, en mayo de 2001, el Ministerio de Educación organizó el "Seminario Internacional sobre Profesionalismo Docente y Aseguramiento de la Calidad". La actividad tuvo una concurrencia masiva reuniendo a más 600 personas, que incluyeron docentes, directivos, investigadores, sostenedores y autoridades ministeriales y gremiales.

.....

16 En conformidad al inciso tercero del artículo 36° de la Ley 20.079, "los profesionales de la educación que deban ser evaluados de conformidad al artículo 70 del Decreto con fuerza de Ley 1, de 1996, del Ministerio de Educación, y se negaren a ello sin causa justificada, se presumirán evaluados en el nivel de desempeño insatisfactorio, no tendrán derecho a los planes de superación profesional, mantendrán su responsabilidad de curso y la obligación de evaluarse al año siguiente".

17 En el año 2007 el incumplimiento alcanzó al 6,5% de los docentes; a un 5,0% el año 2008 y 4,5% el año 2009.

18 Véase el literal d) del artículo 10° de la Ley 20.158, del 29 de diciembre de 2006.

19 Las primeras tres modificaciones son aplicables a los docentes que se evalúen durante el 2011 en adelante. La facultad del director de despedir hasta un 5% de la dotación docente del establecimiento es aplicable transcurridos 90 días desde la publicación de la ley, pero solo por aquellos directores cuyos establecimientos hayan obtenido dentro de los últimos dos años la subvención por desempeño de excelencia, establecida en el artículo 15 de la ley 19.410.

Entre los aspectos destacados de este seminario, en lo que concierne a su impacto para la posterior instauración del sistema de evaluación docente, se encuentran los discursos de inauguración a cargo de la Ministra de Educación Sra. Mariana Aylwin y del presidente del Colegio de Profesores Sr. Jorge Pavez, quienes coincidieron en reconocer el rol del docente como factor clave para promover aprendizajes en los alumnos y, consistentemente, enfatizaron la necesidad de fortalecer la profesión docente. Otro aspecto relevante radicó en la experiencia y visión de los expertos invitados, que permitió dar a conocer: la importancia que la experiencia internacional atribuye al liderazgo directivo y al desarrollo de competencias en los docentes para mejorar la calidad de la educación⁽²⁰⁾, la experiencia cubana en evaluación docente⁽²¹⁾ y, especialmente, el rol de los estándares de desempeño en los procesos de evaluación docente y su aporte al desarrollo y responsabilización profesional⁽²²⁾.

Misiones de estudio

Entre mayo de 2001 y mayo de 2002 se llevaron a cabo tres misiones de estudio con el propósito de conocer sistemas de evaluación docente en países con sistemas educacionales diversos donde esta experiencia había sido exitosa. En estas misiones participaron representantes del Colegio de Profesores, la Asociación Chilena de Municipalidades, la Federación de Instituciones de Educación Particular (FIDE), la Asociación Gremial de Colegios Particulares de Chile (CONACEP) y el Ministerio de Educación. Los países visitados fueron Irlanda y Escocia, en el año 2001, y Cuba en el 2002.

En las misiones fue posible conocer distintos modelos de evaluación de profesores articulados con el sistema educacional y anclados en el contexto social, histórico, político y cultural, de los distintos países visitados. En Irlanda, por ejemplo, la evaluación de profesores se inició en 1999 como plan piloto y se trata de una evaluación externa, realizada completamente por inspectores, que realizan visitas a las escuelas examinando en ellas aspectos de la gestión, la adecuación entre currículum y las necesidades de los alumnos y los procesos de enseñanza y aprendizaje. Los inspectores utilizan múltiples fuentes de información, incluyendo entrevistas a los padres y alumnos, y después de varios días de trabajo participativo elaboran un informe y apoyan la implementación de sus recomendaciones. En Escocia los procesos de evaluación del sistema educacional en su conjunto están orientados completamente hacia lo formativo, dando gran autonomía a las escuelas para que gestionen mejoras en los aprendizajes, en el marco de una comunidad educativa que ha internalizado el espíritu del proceso de evaluación. Por último, en Cuba se observó que la evaluación de los profesores está anclada a un fuerte compromiso con el país, con las familias y con la revolución, y está vinculada con la evaluación de los aprendizajes y rendimiento de los alumnos en la escuela. En este sistema si bien el responsable último de la evaluación es el director del establecimiento, cooperan también con el proceso los subdirectores, supervisores, el sindicato de profesores y los representantes de la comunidad.

Estas misiones cumplieron con el objetivo de conocer en profundidad cómo se implementa y qué propósitos tiene la evaluación de profesores en distintos contextos. La experiencia permitió compartir y discutir ideas entre los representantes de los diversos sectores participantes, lo que contribuyó al enriquecimiento del trabajo a la hora de formular los lineamientos para la evaluación docente en Chile.

.....

20 Véase, por ejemplo, las ponencias de Hilary Emery "Lecciones de la Experiencia del Reino Unido en el desarrollo de estándares de calidad de la enseñanza", y de Françoise Delannoy "Profesionalismo Docente y Aseguramiento de la Calidad de la Enseñanza". *Seminario Internacional "Profesionalismo Docente y Calidad de la Educación"*. Santiago, 8 de mayo, 2001.

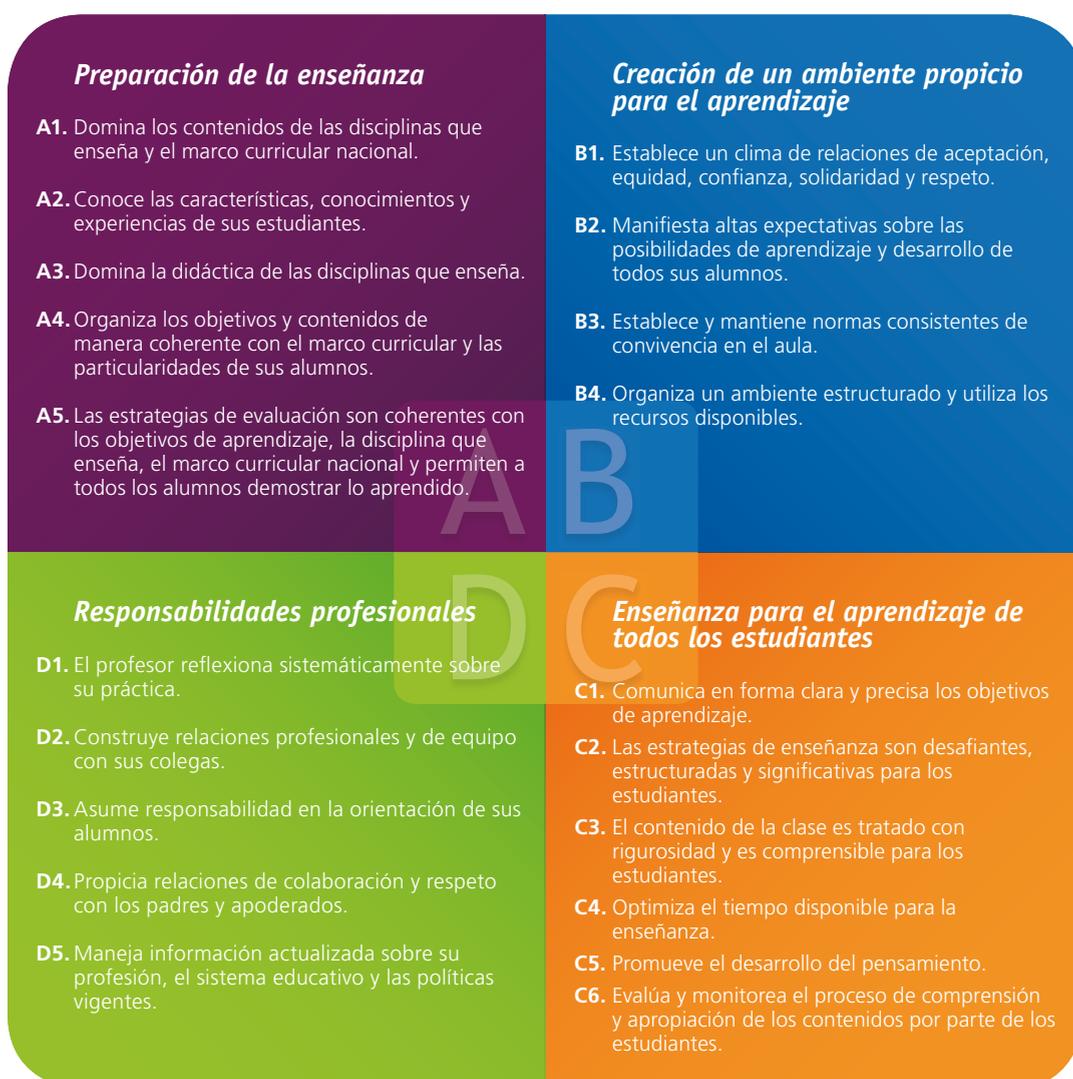
21 Véase ponencia de Héctor Valdés: "La Evaluación del Desempeño del Docente en Cuba". *Seminario Internacional "Profesionalismo Docente y Calidad de la Educación"*. Santiago, 8 de mayo, 2001.

22 Véase exposición de Charlotte Danielson: "Marco para el Mejoramiento de la Práctica Docente". *Seminario Internacional "Profesionalismo Docente y Calidad de la Educación"*. Santiago, 8 de mayo, 2001.

Elaboración del Marco para la Buena Enseñanza

La implementación de una evaluación del desempeño que fuese profesional y formativa demandaba el desarrollo de un modelo de evaluación basado en estándares, que permitiera a los docentes conocer las competencias que caracterizan un buen desempeño en aula, en relación a las cuales serían evaluados. Al mismo tiempo, un cuerpo de estándares debía servir de base para el desarrollo de instrumentos de evaluación válidos y que permitieran retroalimentar a los docentes (así como a otros actores del sistema escolar) a partir de sus resultados, contribuyendo así a cumplir el propósito formativo del proceso. La elaboración de tales estándares, y su validación y difusión entre los actores involucrados era un paso prominente para el desarrollo de la evaluación. A continuación se describirá brevemente este proceso que permitió contar con los estándares de desempeño que sirven de base al sistema, el Marco para la Buena Enseñanza (en adelante MBE):

● **Figura 2**
● *Dominios y Criterios del Marco para la Buena Enseñanza*



La elaboración del MBE estuvo a cargo de un subcomité sobre estándares creado a partir de la comisión tripartita sobre evaluación docente. El trabajo de la subcomisión se estructuró en relación a tres preguntas básicas acerca del desempeño del docente de aula: ¿qué es necesario saber?, ¿qué es necesario saber hacer? y ¿cuán bien se debe hacer? ó ¿cuán bien se está haciendo? Estas últimas dos preguntas dicen relación con los niveles de desempeño que deberían lograrse y cómo cada docente aprecia su propio desempeño en función de los estándares previamente definidos. Las respuestas a estas preguntas se tradujeron en el desarrollo de un marco conformado por cuatro dominios que agrupaban a su vez veinte criterios de ejercicio profesional, que ayudan a orientar el trabajo profesional del docente de aula. Los dominios y criterios del MBE, así como los descriptores más específicos en que éstos se descomponen, establecen los principales aspectos por los cuales los docentes deben ser evaluados y permiten identificar su nivel de desempeño.

Es importante mencionar que previamente se había desarrollado en el país una discusión y definición de estándares de desempeño para la formación docente, en el marco del Programa de Fortalecimiento de la Formación Inicial Docente, y dichos estándares llegaron a ser aplicados aunque solo voluntariamente por algunas universidades. El MBE, conjuntamente con recoger los lineamientos generales de esos estándares, los amplió y adecuó a la realidad de los docentes en servicio, de manera que respondieran al conjunto de responsabilidades de un profesor en el desarrollo de su trabajo diario, tanto en el aula como en la escuela y la comunidad. De esta forma, el MBE llegó a constituirse en referente técnico para el desempeño docente e instrumento oficial del Ministerio de Educación destinado a orientar la política de fortalecimiento de la profesión docente (Resolución Exenta de Educación N° 11.025, de 2004).

El MBE fue sometido a una consulta pública por establecimiento que se realizó, en una primera etapa, en diciembre de 2001, con la participación de 2.371 grupos de profesores, y con 4.325 en una segunda oportunidad, en abril de 2002. La Ministra de Educación de la época, Mariana Aylwin, señaló a propósito de la primera consulta, que “El Marco tiene como referente el compromiso de nuestra sociedad con el mejoramiento de los aprendizajes de todos los niños y jóvenes. Pero se centra en lo que los profesionales de la educación podemos hacer en la escuela y cara a cara con nuestros alumnos”⁽²³⁾.

En el proceso de consulta participaron en total 73.000 docentes, constituidos en 6.696 grupos de trabajo, que revisaron los contenidos fundamentales del MBE y, principalmente, propusieron precisiones en alguno de los conceptos utilizados. Al mismo tiempo, manifestaron diversas preocupaciones sobre el futuro proceso de evaluación, especialmente en relación con los evaluadores, los procedimientos y las consecuencias del proceso.

Entre los principales resultados de las consultas realizadas cabe mencionar que el 24,8% de los docentes participantes destacó el rol de la Evaluación Docente como aporte para indicar el tipo de desempeño que se espera que todos los profesores desarrollen. Además, un 21,5% señaló que la evaluación podría permitir que los profesores dialoguen y reflexionen sobre su práctica con temas comunes, y un 18,2% sostuvo que permitiría apoyar el mejoramiento de las prácticas en el aula (ver Tabla 1).

.....
23 Ministerio de Educación, Consulta sobre el Marco para la Buena Enseñanza, p. 2, Separata distribuida nacionalmente a través del Diario La Tercera el miércoles 21 de noviembre de 2002.

• **Tabla 1**

• *Datos de consulta nacional acerca del Marco para la Buena Enseñanza—Aporte de la evaluación a la profesión docente*

Aporte de la evaluación a la profesión docente	% que marca en primera opción
Indicar el tipo de desempeño que se espera que todos los profesores debieran desarrollar.	24,8%
Permitir que los profesores dialoguen y reflexionen sobre su práctica con temas comunes.	21,5%
Apoyar el mejoramiento de las prácticas en el aula.	18,2%
Mostrar a la sociedad que los profesores saben, saben hacer y hacen en su desempeño.	10,0%
Identificar lo que distingue a la profesión docente de las otras profesiones.	8,7%
Generar el cuerpo de estándares para el sistema de evaluación formativa del desempeño docente.	8,5%

En la consulta también se recogieron las preocupaciones de los docentes en torno a la implementación de la evaluación. A partir de respuestas a una pregunta abierta, se codificaron los temas de mayor recurrencia, los que incluyeron: la forma en que este cuerpo de estándares permitiría evaluar formativamente el desempeño, el tipo de desempeño que se espera que desarrollen, las instancias en que los profesores podrían dialogar y reflexionar sobre su práctica a partir del MBE, y la forma en que los estándares apoyarían el mejoramiento de las práctica en el aula (ver Tabla 2).

• **Tabla 2**

• *Datos de consulta nacional acerca del Marco para la Buena Enseñanza—Preocupaciones respecto del MBE y la profesión docente*

Temas	%	Ejemplos de respuestas
La forma en que este cuerpo de estándares permitirá evaluar formativamente el desempeño docente	25,3%	Necesidad de considerar la diversidad de realidades y contextos escolares en que se desempeñan los profesores.
		Mayor información sobre cómo se evaluará, quiénes los harán, la capacidad e imparcialidad de los evaluadores y los instrumentos que se utilizarán.
El tipo de desempeño que se espera que todos los profesores deban desarrollar	19,0%	Importancia del carácter formativo de la evaluación.
		Porque debe atender a la diversidad que existe entre los docentes.
		Porque se debe considerar la diversidad de los alumnos.
Las instancias en que los profesores puedan dialogar y reflexionar sobre su práctica a partir del Marco	16,5%	Por las distintas disponibilidades de equipamiento y recursos materiales de enseñanza–aprendizaje.
		Por la distinta formación inicial de los docentes.
		Necesidad de más espacio para el trabajo en equipo.
La manera en que los estándares apoyarán el mejoramiento de las prácticas en el aula	10,7%	Ausencia de espacios para la reflexión pedagógica que apoyen el trabajo docente en el aula.
		Insuficiente espacio de participación de los profesores en la toma de decisiones técnico–pedagógicas del establecimiento.
		Necesidad de contar con criterios y pautas que señalen qué y cómo mejorar para cumplir con los estándares propuestos.
		Se requiere de apoyo permanente con el fin de actualizarse en metodologías para la enseñanza.
		Preocupa que el uso de estándares solo sirva para estratificar el desempeño de los docentes y no para mejorar las prácticas del aula.

La Asignación de Excelencia Pedagógica

Un último antecedente importante para el desarrollo de la Evaluación Docente reside en la creación y puesta en marcha del programa de acreditación para percibir la Asignación de Excelencia Pedagógica (en adelante AEP). El protocolo de acuerdo del año 2000 suscrito por el Ministerio de Educación y el Colegio de Profesores⁽²⁴⁾, sentó las bases para la creación de esta asignación, expresada jurídicamente en la Ley 19.715 y en el DFL N°1, de Educación, del 15 de enero del 2002.

El programa AEP fue creado con la finalidad de fortalecer la calidad de la educación y, a la vez, reconocer y destacar el mérito de los docentes de aula que manifiestan habilidades y competencias de excelencia, favoreciendo su permanencia en el desempeño de esta función profesional.

En términos sustantivos, el programa AEP considera a un docente de excelencia como aquel que evidencia un desempeño profesional que de manera clara y consistente sobresale con respecto a lo que se espera para el conjunto de los indicadores evaluados que se encuentran contenidos en el MBE, el cual suele manifestarse por un amplio repertorio de conductas respecto a lo que se está evaluando, o bien, por la riqueza pedagógica que se agrega al cumplimiento de los indicadores. El proceso de acreditación se realiza sobre la base del desempeño evidenciado a través de la presentación de un portafolio y la rendición de una prueba escrita.

El portafolio del programa AEP reúne evidencias del trabajo del docente en una o más áreas de su quehacer profesional, en tanto que la prueba de conocimientos disciplinarios y pedagógicos implica la medición de los conocimientos teóricos declarativos que todo docente debe manejar, para asegurar que posee un profundo conocimiento y comprensión de la disciplina que enseña. Los profesionales que obtienen la AEP son reconocidos como profesores de excelencia por el Ministerio de Educación, a través de un certificado, obtienen un reconocimiento público por sus méritos profesionales, reciben una asignación económica, y tienen la posibilidad de postular a la Red Maestros de Maestros.

Uno de los mayores aportes del programa AEP para efectos de la Evaluación Docente fue cimentar una cultura de la evaluación entre los profesionales de la educación, permitiendo que éstos se familiarizaran con instrumentos orientados a establecer la calidad de su desempeño de manera individual, en el marco de un sistema de evaluación con cobertura nacional.

Entre las principales fortalezas obtenidas producto de la creación de la Asignación de Excelencia Pedagógica se encuentra el establecimiento de una nueva modalidad de asociación entre el Ministerio de Educación y las universidades del país para implementar políticas de alta complejidad a nivel nacional. Las universidades comprometidas en el programa AEP⁽²⁵⁾ debieron recibir asesoría técnica por parte del Educational Testing Service (ETS) y asumieron la responsabilidad de otorgar asistencia técnica especializada, bajo la orientación y supervisión del Ministerio de Educación, en el diseño de los instrumentos de acreditación. Asimismo, asumieron la responsabilidad de elaborar los procedimientos y rúbricas de corrección, capacitar a los profesores correctores, evaluar los portafolios en centros de corrección y elaborar los informes de resultados. Los positivos resultados obtenidos en la experiencia contribuyeron a establecer un modelo de trabajo, vinculando al Ministerio de Educación con entidades universitarias como contraparte técnica externa, que fue más tarde replicado para implementar la Evaluación Docente.

...

24 Véase Protocolo de Acuerdo entre el Ministerio de Educación y el Colegio de Profesores de Chile A.G., de fecha 14 de noviembre de 2000.

25 La asesoría para la implementación del programa AEP ha sido adjudicada desde su inicio en 2002 hasta la fecha de publicación del presente a la Universidad de Chile (Centro de Microdatos) y a la Pontificia Universidad Católica de Chile (Centro de Medición MIDE UC).

A lo largo de este primer capítulo se ha buscado dar cuenta del origen, desarrollo y gradual consolidación del Sistema de Evaluación del Desempeño Profesional Docente. Para ello, he intentado dar cuenta del contexto y los antecedentes que hicieron posible la construcción e implementación de este programa en nuestro país. Asimismo, he revisado someramente algunos factores que contribuyeron a la viabilidad política y técnica de un programa de esta magnitud y complejidad. A lo largo de este libro, el lector tendrá oportunidad de complementar y profundizar esta mirada, a través de distintas vertientes de análisis o aspectos del proceso: la forma en que se lleva a cabo de manera coordinada en las más de trescientas comunas del país en cada período de evaluación; el modelo de evaluación que subyace al sistema y la elaboración de sus instrumentos; los resultados hasta ahora obtenidos y la investigación educacional que se ha generado o favorecido al disponer tanto de datos como de evidencias del trabajo de decenas de miles de docentes de nuestro país.

Probablemente esta lectura contribuya a formarse una idea muy cabal de lo que involucra consensuar, instalar y consolidar un sistema de evaluación de profesores.

En algunas ocasiones he debido responder si el sistema de evaluación docente satisface plenamente las aspiraciones del Ministerio de Educación. Mi respuesta ha sido que no; porque no resulta esperable que un sistema de evaluación que ha sido concordado tripartitamente responda plenamente a los intereses de ninguno de los distintos actores que participaron de su construcción. No obstante lo anterior, el sistema de evaluación docente del cual trata este libro, fue en su génesis una respuesta aceptable para cada uno de los actores institucionales que lo propiciaron y concordaron, y a través del tiempo ha demostrado que era una alternativa factible de ser implementada.

Transcurridos ocho años desde su construcción y cuando su cobertura se encuentra casi totalmente desplegada, resulta posible analizar cuanto de su potencial se encuentra plenamente aprovechado y en qué medida podría resultar posible perfeccionarlo, sin poner en riesgo las bases que han permitido su despliegue y desarrollo. Esperamos que este libro contribuya a dichos propósitos.



2

El modelo de evaluación y los instrumentos

María Paulina Flotts · Andrea Abarzúa

Centro de Medición MIDE UC
Pontificia Universidad Católica de Chile

El modelo de evaluación y los instrumentos

¿Qué modelo de evaluación subyace al Sistema de Evaluación del Desempeño Profesional Docente aplicado en Chile desde 2003?, ¿cómo se construyen los instrumentos a través de los cuales se establece el nivel desempeño de los docentes?, ¿y cómo se define su resultado final en la evaluación? Estas son las principales interrogantes abordadas en el capítulo que sigue.

Después de hacer un breve análisis de los modelos de evaluación de profesores más conocidos y estudiados a nivel internacional, se describe el sistema de evaluación aplicado en Chile. Se da cuenta de los estándares que sirven de base a esta medición, y cómo a partir de ellos se elaboran los cuatro instrumentos utilizados. Se caracteriza el proceso de corrección del instrumento que mide directamente el desempeño de los profesores, el Portafolio, y se describe el procedimiento que conduce al resultado final de cada docente evaluado.



Este capítulo comienza con una breve revisión de los modelos de evaluación docente más estudiados a nivel internacional. Luego, y siguiendo la línea de una evaluación basada en estándares, se hace una descripción del Marco para la Buena Enseñanza y se describe cómo éste se relaciona con los instrumentos que forman parte del sistema de evaluación chileno. A continuación se hace una presentación detallada de cada uno de los instrumentos, describiendo sus principales características y cómo se lleva a cabo su construcción para garantizar que cumplan con las propiedades métricas esperadas para llevar adelante una evaluación confiable y válida; más adelante se aborda el tema del cálculo de puntaje en cada instrumento y la manera en que se informa al docente el resultado de su evaluación. Finalmente, se describe cómo el sistema cautela la incorporación de los elementos contextuales que caracterizan el quehacer docente, de modo de compatibilizar una evaluación estandarizada con la posibilidad de dar cuenta de las particularidades del trabajo que cada profesor lleva a cabo en los distintos contextos y realidades del país.

Evaluación basada en estándares

Los sistemas de evaluación de los docentes usualmente apuntan a dos objetivos parcialmente relacionados: la responsabilización profesional y el mejoramiento del desempeño (Stronge, 1997). El primero intenta asegurar que los servicios profesionales de los docentes cumplan con las expectativas sociales asociadas a esta actividad, identificando metas para los estudiantes, escuelas y sistema escolar que deben ser logradas por los docentes. Por su parte, el objetivo de mejoramiento profesional busca asegurar que la evaluación entregue a los docentes orientaciones para revisar sus prácticas y tomar decisiones de desarrollo profesional. Aunque estos dos objetivos están conceptualmente relacionados (en la medida que el mejoramiento de las prácticas debiera incidir en los logros de la actividad profesional), no existe un claro consenso en torno a las formas en que ambos propósitos deben ser recogidos en un sistema de evaluación ni respecto al peso o ponderación que se debe otorgar a los distintos aspectos que formen parte de éste. Un ejemplo ilustrativo de lo anterior es el amplio debate que se ha producido en el contexto norteamericano en torno al empleo

de métodos basados en mediciones de logros escolares para evaluar la efectividad de los profesores. Recientemente el informe de un connotado grupo de expertos en investigación educacional planteó importantes críticas al uso de este tipo de métodos, aun cuando se empleen los métodos estadísticos más adecuados para analizar los datos, como los de valor agregado. Estos expertos señalan que los resultados de los alumnos en mediciones estandarizadas no debieran tener una ponderación sustantiva en un sistema de evaluación docente, en que lo fundamental debe ser el uso de instrumentos que permitan medir el desempeño docente en forma directa, idealmente en un contexto en que se hayan establecido estándares para la docencia (Baker et. al., 2010). Otros expertos, en cambio, reconocen las limitaciones de los métodos de valor agregado, pero afirman que éstos no deben ser descartados, pues aunque aportan evidencia parcial y limitada acerca de la efectividad de los docentes, su confiabilidad no es inferior a la de metodologías alternativas (Glazerman, Loeb, Goldhaber, Staiger & Raudenbush, 2010). Estos especialistas recomiendan el empleo de métodos de valor agregado, aunque advierten que en el presente no es posible basar decisiones de contratación, desvinculación, incentivos o desarrollo profesional solo a partir de ellos, y en un sistema de evaluación docente debieran combinarse con otros antecedentes. Aunque este debate está aún en curso (y ambas partes reconocen la necesidad de perfeccionar las metodologías actualmente en uso), hay razonable acuerdo en que, si se busca conciliar los propósitos de responsabilización y mejoramiento, un sistema de evaluación de profesores debe incorporar múltiples instrumentos, perspectivas y fuentes de información. En este marco, no cabe desestimar a priori la incorporación en el sistema de evaluación docente chileno, de una medida de la efectividad del profesor a partir de los resultados de sus alumnos. Los desarrollos actuales en torno a los modelos de valor agregado debieran dar luz para analizar la conveniencia de incorporar este tipo de información en el futuro, y la mejor forma de hacerlo (por ejemplo, a nivel de escuela en lugar de individual por docente). Por ahora, es claro que en ese análisis se deben contemplar algunas condiciones que pueden limitar el uso de modelos de valor agregado en nuestro medio, atendiendo a la disponibilidad de mediciones de logro educativo, así como a factores contextuales de nuestro sistema educativo:

- a.** Los modelos de valor agregado solo pueden ser empleados en aquellos casos donde existen mediciones regulares de logro escolar, idealmente espaciadas anualmente. Como es sabido, en Chile las mediciones del SIMCE solo abarcan algunos ámbitos curriculares y no se aplican anualmente. Aunque está previsto un incremento de estas mediciones, parece improbable llegar a alcanzar la cobertura y frecuencia que se requeriría para analizar la efectividad de los docentes de todos los sectores de aprendizaje.
- b.** Los modelos de valor agregado tienden a entregar estimaciones inestables acerca de la efectividad de docentes individuales. Es decir, una proporción importante de docentes que aparecen como efectivos en un ciclo de evaluación, no son identificados como tales en el siguiente. Por ello, se recomienda acumular evidencia más sostenida en el tiempo antes de tomar decisiones importantes acerca de los docentes evaluados. Si no se cuenta con pruebas aplicadas anualmente esta espera puede implicar ciclos relativamente largos.
- c.** Los modelos de valor agregado no son útiles cuando los docentes atienden a grupos pequeños de estudiantes ni cuando existe alta movilidad de los docentes y sus alumnos. Esto conlleva una limitación seria para su aplicación en contextos donde se producen estas condiciones (como el sector rural o en zonas con alta movilidad demográfica).
- d.** Además de estas limitaciones especialmente pertinentes a nuestro contexto escolar, Baker et. al. (2010) identifican un conjunto de otras limitaciones de los modelos de valor agregado que es necesario considerar al momento de decidir la forma en que este tipo de antecedentes puede ser incluido dentro de un sistema amplio de evaluación docente. Estas limitaciones incluyen efectos como la composición socioeconómica desigual de las escuelas, la existencia de prácticas de selección, así como los efectos colaterales o previos de otros profesores (además de aquel que es evaluado).

En el caso chileno, la Evaluación Docente es fruto de un largo proceso en el que se combinaron negociaciones entre las partes interesadas, así como desarrollos paralelos en la formación docente y en el establecimiento

de un marco de estándares para la docencia (el Marco para la Buena Enseñanza)⁽¹⁾. El proceso incluyó un consenso inicial en torno a la necesidad de fundar la evaluación en esos estándares, construyendo instrumentos de medición pertinentes a ellos. El acuerdo, plasmado luego en una ley y reglamento de evaluación, enfatizó la necesidad de asegurar un sistema de evaluación que promueva el mejoramiento de las prácticas docentes, reconociendo y estimulando a quienes cumplen los estándares, y apoyando con acciones remediales a quienes están bajo el estándar. Pero, junto con este propósito formativo, se identificaron claras implicancias de responsabilización, al establecer el sistema altas consecuencias para los docentes bajo el estándar, especialmente aquellos calificados con el nivel de desempeño más bajo (Insatisfactorio). Estos profesores deben re–evaluarse en el año siguiente y abandonar la dotación docente si no logran mejorar esa calificación en tres evaluaciones consecutivas. La Ley 20.501 de Calidad y Equidad en la Educación, promulgada en febrero de 2011, endurece las consecuencias de la evaluación; por ejemplo, reduce a dos el número de evaluaciones consecutivas con resultado Insatisfactorio que conducen a la desvinculación del docente.

Aunque el modelo chileno de evaluación combina mejoramiento con responsabilización, su fundamentación en un marco de estándares de desempeño docente, lo constituye en un sistema que privilegia la entrega de retroalimentación para promover una reflexión personal y orientar decisiones de desarrollo profesional que permitan mejorar la práctica docente. Este enfoque evaluativo se funda en la convicción acerca del rol central que posee el mejoramiento continuo de las prácticas docentes como pilar de las políticas de mejoramiento en la calidad del proceso educativo y sus logros. En este contexto, se ha privilegiado un sistema de evaluación que identifique con precisión y en forma personalizada aquellos aspectos en que el desempeño profesional se encuentra por encima o por debajo de los estándares acordados. Para poder obtener información pertinente a estos propósitos, se ha recurrido a instrumentos de medición que permiten analizar con mayor riqueza y profundidad el desempeño, como son un portafolio (incluyendo la observación de una clase filmada), la entrevista de un par y el juicio de los superiores jerárquicos.

El modelo de evaluación de calidad docente en base a evidencia directa del desempeño que tienen los docentes ha ganado creciente interés. Un informe reciente que analiza las prácticas internacionales en materia de evaluación docente (Isoré, 2010), identifica varios ejemplos de sistemas de evaluación de profesores que se basan en estándares e incluyen medidas directas de desempeño profesional. El ejemplo más estudiado es el del National Board of Professional Teaching Standards en Estados Unidos, al que se han sumado recientes experiencias en algunos estados de ese país. En Latinoamérica, además del sistema chileno, hay avances en México, Colombia, Perú y Ecuador que apuntan a establecer sistemas de evaluación de docentes fundados en estándares o prescripciones acerca del desempeño.

Aunque los modelos de evaluación docente basados en estándares usualmente no incluyen medidas directas del logro de los estudiantes, la investigación existente indica que no están dissociadas del logro escolar. En el caso chileno, el informe nacional de resultados del SIMCE (Sistema de Medición de la Calidad de la Educación) ha incluido desde el año 2007 a la fecha evidencia acerca del logro de estudiantes según el número de profesores bien evaluados que han tenido. Dichos informes han documentado en forma consistente que, a medida que aumenta el número de profesores bien evaluados, también lo hace el rendimiento de los alumnos en el SIMCE. Algo similar arrojan otros estudios y análisis realizados a partir de los datos del SIMCE y la Evaluación Docente⁽²⁾.

La investigación internacional también aporta evidencia favorable acerca de la efectividad de los docentes que obtienen buenos resultados en evaluaciones de desempeño. Por ejemplo, un informe que resume la evidencia existente acerca de la efectividad de los profesores certificados con el National Board of Professional Teaching Standards, concluye que estos docentes también son más efectivos, en términos del aprendizaje logrado por sus estudiantes, que sus pares no certificados (Hakel, Koenig & Elliott, 2008).

.....

1 Para más detalles del proceso de construcción del sistema de evaluación chileno, ver Bonifaz, en este volumen.

2 Ver Taut, Santelices & Manzi, en este volumen.

Milanowski, Kimball y Odden (2005) también concluyen que profesores con buenos resultados en un sistema de evaluación basado en el marco de estándares docentes de Charlotte Danielson, demostraron ser más efectivos para promover el aprendizaje de sus estudiantes que sus pares con evaluaciones más bajas.

Como se señalaba al inicio de este capítulo, el modelo chileno se sustenta en una evaluación basada en estándares. Este tipo de evaluación implica contar con un marco explícito que señale lo que se espera de un buen docente y, junto con ello, disponer de instrumentos de medición que permitan recoger evidencia del desempeño de cada profesor. De esa forma es posible trazar la distancia que hay entre el desempeño esperado y el actual, visibilizando los aspectos mejor logrados y aquellos más débiles, es decir, estableciendo las bases para una evaluación formativa, que entregue información para progresar y mejorar y que incluso puede orientar los planes de formación en servicio que directores y sostenedores planifican para sus colaboradores.

En Chile, la referencia de estándares —aspecto esencial del modelo de evaluación adoptado— está dada en el Marco para la Buena Enseñanza, cuyos contenidos y procesos de validación se presentan a continuación.

El Marco para la Buena Enseñanza

Los estándares que definen una adecuada práctica profesional docente y que constituyen el fundamento para evaluar el desempeño de los profesores están sistematizados en el “Marco para la Buena Enseñanza” (en adelante, MBE).

Este documento tiene como un antecedente fundamental el marco para la formación inicial docente elaborado en el contexto del Programa de Fortalecimiento de la Formación Inicial Docente (FFID) (Ávalos, 2002). “Generalmente se considera que hay calidad docente cuando los maestros disponen de los conocimientos y de las capacidades suficientes para enfrentar las tareas y resolver los problemas propios de su esfera de trabajo. Pero, para que la calidad docente quede de manifiesto, los maestros deben ofrecer evidencia, en su desempeño profesional, de estar haciendo uso de esos conocimientos y capacidades (...) Al elaborar estándares docentes se busca establecer parámetros que indiquen el nivel de calidad necesario para lograr un desempeño adecuado de la profesión docente y, por otra parte, posibilitar la evaluación del desempeño docente, tanto en sus etapas iniciales como en las avanzadas” (Ministerio de Educación, 2000, p.4). Estos estándares de desempeño para la formación inicial de docentes se estructuraron en cuatro facetas que buscan dar cuenta de las principales áreas de acción profesional que son de responsabilidad del profesor (Faceta A: Preparación para la enseñanza: organización del contenido en función del aprendizaje del estudiante; Faceta B: Creación de un ambiente propicio para el aprendizaje de los alumnos; Faceta C: Enseñanza para el aprendizaje de todos los alumnos; y Faceta D: Profesionalismo docente).

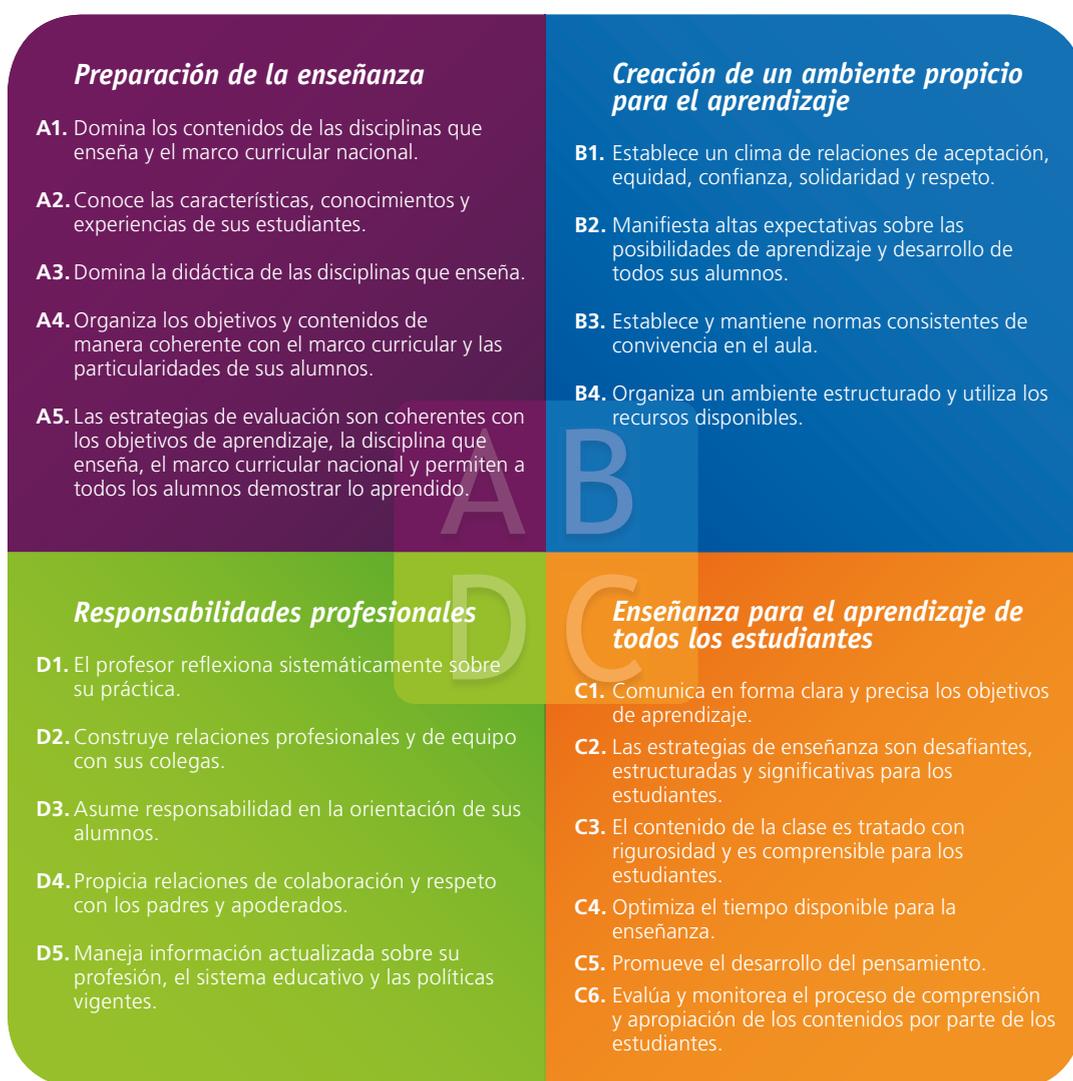
Sobre la base de este trabajo, una comisión técnica tripartita, compuesta por representantes del Ministerio de Educación, del Colegio de Profesores de Chile y de la Asociación Chilena de Municipalidades elaboró una primera versión del conjunto de estándares que constituyen una práctica efectiva y comprometida del quehacer docente, tomando como referencia la propuesta más conocida en este campo a nivel internacional: “Framework for Teaching” de Charlotte Danielson (publicado por primera vez en 1996). Esta primera versión de los estándares fue difundida y se sometió a procesos de consulta entre los profesores del país, producto de los cuales se consensuaron los criterios y descriptores que lo componen. En paralelo, se solicitó a reconocidas personalidades del mundo de la educación que, a modo consultivo, contribuyeran desde sus conocimientos y experiencia a la elaboración de este documento.

El MBE identifica el conjunto de responsabilidades que debe asumir un profesor en el desarrollo de su trabajo cotidiano con el propósito de contribuir significativamente al aprendizaje de sus alumnos. De esta forma, aborda tanto las responsabilidades que el docente asume en el aula, como aquellas que debe

cumplir a nivel de su escuela y también en la comunidad en que se inserta su trabajo. Se trata entonces de un marco normativo elaborado por expertos y validado en distintas instancias y con diversos actores; no es un modelo o teoría educativa, no explica cómo son los buenos docentes, sino que describe qué debe hacer un buen docente.

Para dar cuenta de todas estas responsabilidades, el MBE está organizado, según niveles crecientes de especificidad, en torno a aquellas destrezas, habilidades y conocimientos que son relevantes para el quehacer pedagógico y profesional docente. El nivel más inclusivo se denomina “dominios”, cada uno de los cuales agrupa una serie de “criterios”, y estos, a su vez, incluyen un conjunto de “descriptores”⁽³⁾ (ver Figura 1).

• **Figura 1**
• *Dominios y Criterios del Marco para la Buena Enseñanza*



3 El Marco para la Buena Enseñanza incluye 71 descriptores distribuidos en 20 criterios, por lo tanto, se recomienda consultarlo directamente. Este documento se encuentra disponible en diversos sitios web, entre ellos www.docentemas.cl. Al término de esta sección se ejemplifica la estructura dominio-criterios-descriptores.

- **Dominio A: Preparación de la enseñanza.** Manejo que tiene el profesor de la disciplina que enseña y de los principios y competencias pedagógicas necesarias para organizar el proceso de enseñanza. Considera el dominio del marco curricular, es decir, de los conocimientos, habilidades, competencias, actitudes y valores que los alumnos requieren alcanzar para desenvolverse en la sociedad actual. El profesor debe poseer un profundo conocimiento y comprensión de las disciplinas que enseña y de los conocimientos, competencias y herramientas pedagógicas que faciliten una adecuada mediación entre los contenidos, los estudiantes y el contexto de aprendizaje. Respecto de este dominio cabe precisar que el sistema de evaluación docente, tal como fue acordado por la Comisión Tripartita, no considera la medición del nivel de conocimiento disciplinario que tienen los profesores; aun cuando se evalúan competencias que implican poner en juego distintos tipos de conocimientos didácticos y pedagógicos, el manejo disciplinario específico del docente fue excluido del actual sistema de evaluación.
- **Dominio B: Creación de un ambiente propicio para el aprendizaje.** Ambiente y clima que genera el docente en los que tienen lugar los procesos de enseñanza y aprendizaje. Considera el carácter de las interacciones que ocurren en el aula, y la promoción, por parte del profesor, de la confianza, aceptación, equidad y respeto. Asimismo, alude a sus habilidades para generar un ambiente estimulante y al compromiso del profesor con los aprendizajes y el desarrollo de sus estudiantes.
- **Dominio C: Enseñanza para el aprendizaje de todos los estudiantes.** Habilidad del profesor para organizar situaciones interesantes y productivas que aprovechen efectivamente el tiempo para el aprendizaje, y que consideren los saberes e intereses de los estudiantes. Requiere la habilidad de organizar recursos de aprendizaje y monitorear en forma permanente los aprendizajes, con el fin de retroalimentar las propias prácticas y ajustarlas a las necesidades detectadas en los estudiantes.
- **Dominio D: Responsabilidades profesionales.** Capacidad de reflexionar consciente y sistemáticamente sobre la propia práctica y de reformularla para garantizar una educación de calidad para todos los estudiantes. La responsabilidad profesional también implica la capacidad de detectar las propias necesidades de aprendizaje, compromiso con el proyecto educativo del establecimiento y con las políticas nacionales de educación. Además, refiere a dimensiones del trabajo docente que van más allá del trabajo en aula y que involucran la interacción con los colegas, la comunidad y el sistema educativo en su conjunto.

A modo de ejemplo, el dominio D Responsabilidades profesionales se desagrega en cinco criterios:

- D1. El profesor reflexiona sistemáticamente sobre su práctica.
- D2. Construye relaciones profesionales y de equipo con sus colegas.
- D3. Asume responsabilidades en la orientación de sus alumnos.
- D4. Propicia relaciones de colaboración y respeto con los padres y apoderados.
- D5. Maneja información actualizada sobre su profesión, el sistema educativo y las políticas vigentes.

Cada criterio contiene un número variable de descriptores. Por ejemplo, el criterio D1. *El profesor reflexiona sistemáticamente sobre su práctica*, se operacionaliza en tres descriptores, y para cada uno de ellos el MBE incluye una explicación o descripción. A continuación se ejemplifican estos descriptores:

- *Descriptor D.1.1: Evalúa el grado en que los alumnos alcanzaron los aprendizajes esperados.* Explicación: Como parte de sus prácticas regulares, el profesor recoge evidencias de los aprendizajes de sus estudiantes, a partir de los productos de sus trabajos y de procedimientos específicos de evaluación, que le permiten establecer el nivel de logro de los aprendizajes esperados.

- *Descriptor D.1.2: Analiza críticamente su práctica de enseñanza y la reformula a partir de los resultados de aprendizaje de sus alumnos.*
Explicación: El profesor analiza críticamente las fortalezas y debilidades de sus prácticas, con relación a sus efectos sobre los aprendizajes de sus alumnos, lo que le permite reformularlas para hacerlas más efectivas y pertinentes.
- *Descriptor D.1.3: Identifica sus necesidades de aprendizaje y procura satisfacerlas.*
Explicación: A partir de la evaluación de sus prácticas, el profesor busca de manera sistemática identificar sus necesidades de actualización y de desarrollo profesional. Revisa material bibliográfico atinente a su profesión y a las disciplinas que enseña. Aprovecha las oportunidades de desarrollo profesional ofrecidas por el establecimiento y el sistema educativo.

El MBE aborda entonces las competencias necesarias para una labor docente de calidad. Este documento, de uso y conocimiento de cada uno de los docentes que participan en el proceso de evaluación, es el referente único y obligatorio para la evaluación de su desempeño profesional. Debe señalarse, sin embargo, que no se trata de un “modelo de evaluación”, es decir, no indica qué específicamente debe conocer un profesor, en qué grado debe mostrar una determinada competencia y sus descriptores no están planteados como indicadores operacionales y observables; por su parte, los dominios y criterios del Marco suelen estar formulados de manera tal que combinan distintos elementos y que, para efectos de evaluación, requieren ser separados o descompuestos.

Por las razones antes expuestas, el proceso de definición de indicadores de evaluación en base a este referente ha implicado desagregar los diversos descriptores en elementos observables de la práctica docente, trabajo que ha sido llevado a cabo por expertos en evaluación de MIDE UC. Tal como se mostrará a continuación, todos los instrumentos de evaluación del Sistema de Evaluación del Desempeño Profesional Docente son creados y se evalúan en concordancia con lo planteado en este Marco, considerando que se han generado indicadores de evaluación como producto intermedio.

La existencia del MBE y la solidez del proceso de construcción y validación del mismo constituyen un antecedente clave para sostener un sistema de evaluación del desempeño profesional docente como el que existe en Chile⁽⁴⁾, y es en base a este Marco que es posible plantear una evaluación que considera distintos ámbitos del desempeño profesional docente, incluido el trabajo que hace el profesor en el aula.

Mapeo del Marco para la Buena Enseñanza e instrumentos

Por tratarse de una evaluación referida a criterios, cuyos estándares están contenidos en el MBE, resulta importante especificar la forma en que se construyen los instrumentos de la Evaluación Docente, para asegurar su consistencia con dichos estándares. Más adelante en este capítulo se hará una descripción detallada de esos instrumentos. Sin embargo, para dar contexto a la información que se presentará a continuación, es conveniente mencionar aquí que la evaluación utiliza cuatro instrumentos: Autoevaluación, Informe de Referencia de Terceros (que completan el director y Jefe de UTP del establecimiento en que trabaja el profesor), Entrevista por un Evaluador Par y Portafolio (compuesto de dos módulos: el Módulo 1, referido al diseño, implementación y evaluación de una unidad pedagógica y el Módulo 2, que considera la filmación de una clase).

Tal como se desprende de la presentación del Marco, la naturaleza y alcance de los distintos dominios, criterios y descriptores es muy diversa; incluye temáticas tan variadas como por ejemplo, establecer un clima de relaciones de aceptación, equidad, confianza, solidaridad y respeto (criterio B.1), y manejar

.....

4 Ver Bonifaz, en este volumen.

información actualizada sobre la profesión, el sistema educativo y las políticas vigentes (criterio D.5). Por su parte, los instrumentos que forman parte del sistema de evaluación son también muy diversos, en la medida que recogen distintas perspectivas del desempeño docente: desde la opinión que entrega el supervisor directo, hasta la evidencia directa del trabajo del profesor en el aula que queda registrada en la grabación de una clase (video).

Dado lo anterior, el primer paso en el proceso de elaboración de los instrumentos es definir un mapeo con los descriptores del MBE en función de la forma esperada de medirlos; este mapeo es una sistematización de la pregunta, “dada la naturaleza de este descriptor, ¿con qué instrumento es susceptible de ser medido?”. De esta manera, se genera una matriz de doble entrada en la que se cruzan los descriptores del Marco y los distintos instrumentos en función de la adecuación de estos últimos para abordar los primeros.

Para ejemplificar lo anterior, y aún sin profundizar en la especificidad de los descriptores, considérese la naturaleza del Dominio B “Creación de un ambiente propicio para el aprendizaje” y el D “Responsabilidades profesionales”... ¿resulta razonable pensar que ambos son susceptibles de ser evaluados con igual pertinencia a través del video? Probablemente no: mientras que en la clase filmada se puede observar directamente la manera en que el profesor interactúa con los estudiantes y cómo genera un ambiente de confianza, aceptación, equidad y respeto, es poco probable que en ella se registre evidencia acerca de asuntos ligados a las responsabilidades profesionales, como la reflexión consciente y sistemática sobre la propia práctica, la capacidad de detectar las propias necesidades de aprendizaje o el compromiso con el proyecto educativo del establecimiento y con las políticas nacionales de educación. Complementariamente, es más factible que el superior jerárquico del profesor tenga más evidencia sobre este último tipo de materias que sobre la dinámica específica de la interacción profesor alumno que surge al interior de una sala de clases, por lo que resulta más razonable que los descriptores del dominio D tengan una mayor representación en el Informe de Referencia de Terceros que los descriptores del dominio B (y éstos, a su vez, estén más representados en la evaluación del video).

A nivel de descriptores, hay algunos que solo se evalúan con un instrumento y otros en dos o más. Por ejemplo, considérense los descriptores “Conoce las particularidades familiares y culturales de sus alumnos” (A.2.2.) y “Conoce las diferentes maneras de aprender de sus estudiantes” (A.2.4.). Ambos son susceptibles de ser evaluados en los Módulos 1 y 2, pero solo el primero parece además posible de ser evaluado por el Director y Jefe de UTP, quienes, a través de las planificaciones que hace el profesor pueden tener evidencia para hacer su evaluación; complementariamente, para saber si el docente conoce las diferentes maneras de aprender de sus estudiantes se requiere un contacto mucho más permanente y constante con la realidad concreta del aula, por lo que podría ser más difícil de evaluar para los superiores jerárquicos del profesor. Por esto, en el mapeo ambos descriptores aparecen como pertinentes en el caso del Portafolio, pero solo el primero lo es para el Informe de Referencia de Terceros.

Solo con el propósito de ilustrar lo anteriormente expuesto, a continuación se presenta un extracto del mapeo, que permite apreciar la relación entre descriptores e instrumentos según la naturaleza de ambos, y también el que algunos descriptores puedan ser evaluados con un instrumento en particular, o con dos o más (ver Figura 2).

• **Figura 2**
 • *Ejemplo de mapeo de instrumentos y MBE*

Dominio A: Preparación de la enseñanza					
Criterio A2: Conoce las características, conocimientos y experiencias de sus estudiantes.					
Descriptor	Módulo 1 Portafolio	Módulo 2 Portafolio	Informe de Referencia de Terceros	Entrevista Ev. Par	Auto-evaluación
A.2.1. Conoce las características de desarrollo correspondientes a las edades de sus estudiantes.	•			•	•
A.2.2. Conoce las particularidades familiares y culturales de sus alumnos.	•	•	•	•	•
A.2.3. Conoce las fortalezas y debilidades de sus estudiantes respecto de los contenidos que enseña.	•	•		•	•
A.2.4. Conoce las diferentes maneras de aprender de los estudiantes.	•	•		•	•
Dominio B: Creación de un ambiente propicio para el aprendizaje					
Criterio B1: Establece un clima de relaciones de aceptación, equidad, confianza, solidaridad y respeto.					
Descriptor	Módulo 1 Portafolio	Módulo 2 Portafolio	Informe de Referencia de Terceros	Entrevista Ev. Par	Auto-evaluación
B.1.1. Establece un clima de relaciones interpersonales respetuosas y empáticas con sus alumnos.		•	•	•	•
B.1.2. Proporciona a todos sus alumnos oportunidades de participación.		•	•	•	•
B.1.3. Promueve actitudes de compromiso y solidaridad entre los alumnos.	•	•	•	•	•
B.1.4. Crea un clima de respeto por las diferencias de género, culturales, étnicas y socio-económicas.		•	•	•	•

Este mapeo de descriptores del MBE e instrumentos de evaluación tiene asociada también otra dimensión: la variabilidad en el tiempo. Es razonable pensar que la cobertura total del Marco (con sus 71 descriptores) resulta inabordable en un mismo momento, o al menos, resulta desaconsejable desde la perspectiva de la extensión de cada instrumento y del tiempo que los docentes que participan en el proceso de evaluación deben dedicarle. Ante esto, el sistema de evaluación docente diseñó un mecanismo de rotación de los descriptores que son medidos a través del tiempo con los distintos instrumentos, cautelando que exista un número de descriptores comunes entre años. De este modo, parte de los instrumentos varía año a año, lo que permite ir cubriendo la amplia diversidad de descriptores del Marco para Buena Enseñanza a través del tiempo.

En suma, la cobertura del MBE está garantizada desde dos perspectivas: la distinta naturaleza de los instrumentos que posibilita abordar cada descriptor desde una forma pertinente de medirlo, y la posibilidad de ir alternando descriptores en los distintos años de implementación del sistema. Por otra

parte, y desde la perspectiva de la construcción de los instrumentos, el mapeo de los descriptores se ha constituido en una herramienta clave para guiar el proceso: constituye una “carta de navegación” que orienta la elaboración de preguntas y pautas de evaluación o rúbricas.

Instrumentos de evaluación

a. Descripción de los instrumentos de evaluación

Para evaluar el cumplimiento de los estándares expresados en el Marco para la Buena Enseñanza, el Sistema de Evaluación del Desempeño Profesional Docente utiliza cuatro instrumentos: Autoevaluación, Informes de Referencia de Terceros (evaluación de superiores jerárquicos), Entrevista por un Evaluador Par (evaluación de un colega) y Portafolio (evaluación a partir de evidencia directa del trabajo en aula).

Estos instrumentos recogen distintas perspectivas del desempeño profesional del docente, para así obtener una imagen lo más completa posible de las competencias de cada docente que participa en la evaluación.

El resultado final del docente consiste en una integración de los resultados obtenidos en estos cuatro instrumentos, cada uno de los cuales recibe una ponderación establecida legalmente en el Reglamento sobre Evaluación Docente. El Portafolio pesa un 60% de la evaluación final; la Entrevista por un Evaluador Par, un 20%; y tanto el Informe de Referencia de Terceros como la Autoevaluación, 10%. Cuando los docentes obtienen un resultado Insatisfactorio y se evalúan por segunda o tercera vez, el Portafolio se pondera en un 80%, la Autoevaluación en un 5%, la Entrevista en un 10% y los Informes de Referencia de Terceros en un 5%.

A continuación se presenta una descripción de cada instrumento.

Pauta de Autoevaluación

Se trata de un instrumento que aborda el desempeño docente desde la perspectiva del propio evaluado. Este instrumento se estructura en una serie de ítems (idénticos para todos los docentes) derivados del MBE y que están orientados a que cada profesor reflexione acerca de su propia práctica y, a partir de dicha reflexión, evalúe su desempeño.

Cada profesor recibe un ejemplar de la Pauta de Autoevaluación y cuenta con cinco semanas para completarla. El instrumento contiene, para cada uno de los cuatro dominios del Marco, una pauta estructurada en base a una rúbrica con indicadores cuyo cumplimiento define a qué nivel de desempeño —Destacado, Competente, Básico e Insatisfactorio— corresponde su quehacer profesional. El docente debe calificar así su propio desempeño usando una de esas categorías.

Dado que la Evaluación Docente considera la importancia que tienen los factores contextuales en el trabajo de cada profesor, y asimismo, las vivencias particulares que puedan estar afectando positiva o negativamente su trabajo durante el período de evaluación, en una última sección de este documento se solicita al docente detallar situaciones de esta naturaleza —si las hubiera—, de manera tal que puedan ser consideradas en su evaluación final⁵.

.....

5 Tal como se verá en la parte final de este capítulo, esta información relativa a consideraciones contextuales es considerada por las Comisiones Comunales de Evaluación, instancia local que cierra el proceso de evaluación de los docentes.

Informes de Referencia de Terceros

Este instrumento⁽⁶⁾ recoge la apreciación que tienen los superiores jerárquicos acerca del desempeño profesional del docente. Aquí se espera que directores y jefes de UTP, quienes tienen la oportunidad de observar y conocer el trabajo del profesional en la escuela, puedan calificar su desempeño. Esta evaluación también consiste en una pauta estructurada en base a una rúbrica, con indicadores cuyo cumplimiento define el nivel de desempeño del docente a partir de lo que señala el MBE; el superior jerárquico debe, basándose en su conocimiento del trabajo del profesor, calificar su desempeño en cada uno de los ámbitos que se abordan empleando los cuatro niveles de desempeño del sistema (Destacado, Competente, Básico e Insatisfactorio).

En forma homóloga a lo que ocurre en la Autoevaluación, existe una sección específica que permite que directores y jefes de UTP puedan registrar si, a su juicio, existen situaciones de carácter contextual que puedan haber afectado el desempeño del docente, para que puedan ser consideradas por la Comisión Comunal de Evaluación al momento de analizar el resultado final del profesor. Todos los docentes deben contar al menos con un Informe de Referencia de Terceros, por tanto, en los casos de establecimientos en que no existe ninguno de estos cargos⁽⁷⁾, el informe debe ser emitido por el superior jerárquico del docente evaluado y por el jefe técnico pedagógico comunal si existe. Quienes emiten este informe cuentan con aproximadamente un mes para completarlo.

Entrevista por un Evaluador Par

Este instrumento permite recoger información acerca del desempeño del docente desde la perspectiva de un colega, un profesor que desempeña su trabajo en condiciones similares a las del docente evaluado, y que se denomina "Evaluador Par". Para cumplir este rol, los docentes postulan voluntariamente y participan de un proceso de selección, capacitación y entrenamiento que los habilita para aplicar correctamente el instrumento. Los Evaluadores Pares son docentes del sector municipal con al menos cinco años de experiencia profesional docente, preferentemente trabajan en la misma comuna del docente evaluado, y deben desempeñarse en el mismo nivel, sector y modalidad de enseñanza de éste.

Para llevar a cabo la entrevista, cada Evaluador Par se traslada al establecimiento del docente evaluado y aplica la entrevista en una fecha y horario que ha sido previamente concordado con el docente que participa en el proceso de evaluación⁽⁸⁾. Las preguntas que forman parte de este instrumento se desprenden directamente del MBE y en cada una de ellas se espera que el docente dé cuenta de su experiencia y conocimientos, así como del análisis que hace de su propia práctica. Al igual que en los dos instrumentos antes descritos, en una última sección el Evaluador Par solicita al docente en evaluación que describa posibles situaciones contextuales que puedan haber afectado su desempeño profesional en el período que se evalúa y que, a su juicio, deben tenerse en cuenta al determinar la evaluación final. El Evaluador Par tiene la obligación de tomar registro de cada una de las respuestas del docente e, inmediatamente después de aplicada la entrevista, debe evaluar cada una de ellas como destacadas, competentes, básicas o insatisfactorias, en base a una pauta de corrección específica para cada pregunta. La aplicación de la entrevista dura aproximadamente 50 a 60 minutos y los Evaluadores Pares disponen de dos meses para completar el total de entrevistas asignadas.

En la capacitación a los Evaluadores Pares que realiza MIDE UC cada año, éstos son preparados tanto para la correcta aplicación de la entrevista, como para su posterior codificación en base a la rúbrica o

.....

6 Este instrumento, tal como se verá más adelante en el capítulo, tuvo en el año 2010 cambios importantes en su diseño, lo que trajo consigo cambios a nivel de resultados. A continuación se describe la versión más actualizada del instrumento.

7 Es el caso, por ejemplo de docentes que se desempeñan en escuelas rurales uni, bi o tridocentes.

8 Los Evaluadores Pares no pueden entrevistar a docentes o parientes de su mismo establecimiento educacional, entre otras inhabilidades descritas en el Reglamento sobre Evaluación Docente.

pauta de corrección. Una rúbrica es un conjunto de pautas que permiten distinguir niveles de desempeño en las dimensiones que son objeto de evaluación. La rúbrica está compuesta por uno o más rasgos o dimensiones que sirven de base para juzgar el desempeño frente a la tarea —en este caso, la respuesta del docente a cada pregunta de la Entrevista—; definiciones y ejemplos que clarifican el significado de cada rasgo o dimensión; y una escala de valores con la que se puntúa cada dimensión (Goodrich, 1997; Hart, 1994; Hermann, Aschbacher & Winters, 1992; Tombari & Borich, 1999).

Así, cada pregunta se acompaña de una rúbrica que identifica cuatro niveles de desempeño para evaluar la respuesta del docente; todas las rúbricas siguen una misma lógica en su proceso de construcción: se comienza por especificar el desempeño del nivel Competente, tomando para ello lo que el MBE señala como estándar de calidad en el o los descriptores asociados, y luego se describen los niveles Básico e Insatisfactorio; en último lugar se define el desempeño Destacado. Para hacer estas distinciones, se trabaja sobre la definición general de cada uno de estos niveles de desempeño:

- Insatisfactorio: indica un desempeño que presenta claras debilidades en el indicador evaluado y éstas afectan significativamente el quehacer docente.
- Básico: indica un desempeño profesional que cumple con lo esperado en el indicador evaluado, pero con cierta irregularidad (ocasionalmente). Esta categoría también puede usarse cuando existen algunas debilidades que afectan el desempeño, pero su efecto no es severo ni permanente.
- Competente: indica un desempeño profesional adecuado en el indicador evaluado. Cumple con lo requerido para ejercer profesionalmente el rol docente. Aun cuando no es excepcional, se trata de un buen desempeño.
- Destacado: indica un desempeño profesional que clara y consistentemente sobresale con respecto a lo que se espera en el indicador evaluado. Suele manifestarse por un amplio repertorio de conductas respecto a lo que se está evaluando, o bien, por la riqueza pedagógica que se agrega al cumplimiento del indicador.

Desde el año 2009 cada Evaluador Par debe completar un registro online de las entrevistas aplicadas y los puntajes obtenidos por cada docente en cada una de las preguntas que componen este instrumento y así son consideradas en el resultado final de cada evaluado⁹. Para este registro online, se utiliza una herramienta ad hoc del software diseñado por el equipo técnico de MIDE UC.

Portafolio

Este instrumento permite recoger muestras directas o evidencias concretas del quehacer de cada docente, posibilitando que éste dé cuenta de su desempeño en las acciones profesionales centrales del trabajo en el aula. Se estructura en dos módulos: uno en que se presenta información referida al diseño, implementación y evaluación de una unidad pedagógica, y otro que consiste en la filmación de una clase. En cada módulo se incluyen determinados productos¹⁰ (ver Figura 3).

.....

9 Hasta el período de evaluación 2008, los Evaluadores Pares completaban su evaluación de la entrevista en un formulario que era luego enviado al equipo técnico de MIDE UC para su digitación e ingreso a las bases de datos.

10 Esta estructura de módulos y productos se ha mantenido estable a lo largo del tiempo, aun cuando los subproductos específicos que componen cada producto sí han experimentado variaciones a través de los años.

• **Figura 3**
 • *Estructura del Portafolio*

MODULO 1	MODULO 2
Producto 1 Unidad Pedagógica	Producto 4 Clase grabada
Descripción de la unidad pedagógica	Video de la clase grabada
.....
Análisis de las dificultades de los alumnos	Ficha de la clase grabada
.....
Análisis de la unidad realizada	(*) Fotocopias de los recursos de aprendizaje utilizados en la clase grabada
Producto 2 Evaluación de la unidad pedagógica
Evaluación de la unidad y Pauta de corrección	(*) El docente puede adjuntar los recursos de aprendizaje utilizados cuando estime que éstos podrían no verse claramente en el video (por ejemplo, guías de trabajo individual, páginas del texto escolar, imágenes o dibujos empleados por el docente, evaluaciones individuales, etc.).
.....	
Reflexión a partir de los resultados obtenidos en la evaluación	
.....	
Retroalimentación de la evaluación	
Producto 3 Reflexión Pedagógica	

El Módulo 1 incluye una serie de productos escritos relativos a la implementación del trabajo pedagógico, en los que principalmente se solicita al docente describir su trabajo y reflexionar acerca de él. El primer producto consiste en la descripción de la implementación de una unidad pedagógica de ocho horas, y su análisis en base a ciertas directrices. Esta unidad pedagógica debe ser desarrollada con alguno de los cursos que el docente tiene a su cargo, y debe enmarcarse en una de dos opciones de Objetivo Fundamental Vertical y Contenido Mínimo Obligatorio que cada año se proponen para la Evaluación Docente para cada curso o nivel. Los docentes deben elegir una opción y en base a ella desarrollar objetivos para la unidad en su conjunto y para cada una de las clases que la componen. También deben describir cada una de las clases que implementaron como parte de esta unidad pedagógica, indicando fecha, duración, actividades realizadas, recursos utilizados, entre otros aspectos. Luego deben responder una serie de preguntas orientadas a analizar críticamente la experiencia de implementación de la unidad y su propia práctica pedagógica.

El segundo producto del Módulo 1 consiste en la presentación de una evaluación de aprendizaje de los alumnos en esa unidad pedagógica. El requisito principal de esta evaluación es que permita obtener información sobre lo que han aprendido los estudiantes en esa unidad en particular, lo cual debe ser descrito en detalle por el profesor. Si se trató de una prueba escrita, el docente debe enviar una copia de este instrumento y detallar las respuestas correctas o los criterios utilizados para evaluar cada respuesta como correcta. Si el docente evaluó una ejecución (de un instrumento musical, de una disertación, u otros) debe describir en qué consistió la instancia de evaluación y transcribir las instrucciones entregadas a los estudiantes (orales o escritas). Además, el docente debe presentar la pauta de corrección o lista de cotejo que utilizó para evaluar a sus estudiantes. Al igual que en el producto anterior, el docente debe también responder a preguntas relacionadas con la experiencia de aplicación de esta evaluación, su análisis y el uso que dio a esa información (retroalimentación a los alumnos, retroalimentación para mejorar la propia práctica, entre otros aspectos).

Un tercer y último producto del Módulo 1 consiste en la reflexión sobre el quehacer docente del último año. En este producto se solicita al docente que describa experiencias pedagógicas específicas y luego las analice, o bien que responda a ciertas preguntas que apuntan a situaciones típicas del quehacer docente. Por ejemplo, pueden estar referidas a la promoción de valores en los estudiantes, trabajo colaborativo con docentes de otros subsectores, detección de las propias necesidades de aprendizaje, dificultades para aprender más comunes en los estudiantes, entre otros.

Por otra parte, el Módulo 2 del Portafolio contiene solo un producto, que es la filmación de una clase. En este producto el docente debe coordinar, a través del director de su establecimiento, una fecha y hora de grabación de una clase de 40 minutos. En esta clase, se espera que el docente dé muestras de su trabajo pedagógico y para ello se le solicita que durante al menos diez minutos explique o enseñe un contenido, procedimiento o habilidad a sus estudiantes. Además de presentar su clase filmada, el docente debe completar una breve ficha de descripción de ésta y adjuntar los recursos de aprendizaje fotocopiables que utilizó (si los hubiera), ya que éstos podrían no apreciarse con claridad en el video.

Para la presentación del Portafolio, cada docente recibe un manual que contiene instrucciones muy específicas para elaborar los productos asociados a cada módulo, de manera tal que todos los docentes entreguen lo solicitado siguiendo un formato estandarizado. Este manual contiene instrucciones generales para la elaboración del Portafolio y las normas que rigen su elaboración. También contiene una especificación de los descriptores del MBE que serán considerados al evaluar cada producto solicitado, de manera de hacer explícito al docente los criterios con que será evaluado en cada uno de ellos. Por su parte, la filmación de la clase del Módulo 2 se encuentra a cargo de un camarógrafo acreditado que provee el Sistema de Evaluación del Desempeño Profesional Docente y no tiene costo alguno para el profesor; los camarógrafos son capacitados y existen una serie de medidas para asegurar que la filmación sea audible, con una adecuada calidad de la imagen, que muestre tanto al profesor como a los alumnos y que refleje en forma fidedigna lo que sucede en la clase; es decir, se garantiza que la clase filmada sirva para los propósitos de la evaluación.

Los módulos del Portafolio pueden ser elaborados de manera completamente independiente; incluso, el docente puede optar por desarrollar el Módulo 1 con uno de los cursos con que trabaja, y el Módulo 2 con otro curso, siempre y cuando desarrolle ambos en el ciclo y subsector en que le corresponde evaluarse. Los docentes tienen aproximadamente dos meses para completar los dos módulos del Portafolio.

La evidencia que entrega el docente en su Portafolio es evaluada por profesores correctores en un complejo proceso que abarca alrededor de dos meses. La base de esta corrección son las rúbricas que cada año MIDE UC construye para evaluar esta evidencia. Tal como se señaló previamente al describir la Entrevista, estas rúbricas se basan en el Marco para la Buena Enseñanza y describen cuatro niveles de desempeño posibles (Destacado, Competente, Básico e Insatisfactorio). A modo de ejemplo se presenta a continuación la manera de operacionalizar la evaluación de uno de los descriptores del Marco que se aborda en el Módulo 2 del Portafolio (filmación de una clase):

Descriptor B.1.1: Establece un clima de relaciones interpersonales respetuosas y empáticas con sus alumnos.

Explicación: El profesor genera relaciones de respeto con sus alumnos, los escucha, comenta sus aportes y los enriquece con sus comentarios y con los del resto del curso, así como a partir de otros recursos que los ilustren. Favorece sistemáticamente las interacciones relacionadas con los aprendizajes, abriendo espacios para preguntas, cuestionamientos, aportes, opiniones, comentarios de todos los alumnos y alumnas, no solo de los más aventajados.

••• Rúbrica⁽¹¹⁾

Destacado	<p>Cumple con las condiciones de “Competente” y además: El profesor motiva la participación de los alumnos explicitando el valor de la diversidad de opiniones y respuestas como elementos enriquecedores del proceso de aprendizaje.</p> <p style="text-align: center;">O</p> <p>Si hay alumnos que no participan en las actividades de aprendizaje de la clase, el profesor realiza acciones para que participen.</p>
Competente	<p>El profesor ofrece espacios de participación a todos los alumnos, sin privilegiar a uno o a un grupo de ellos.</p> <p style="text-align: center;">Y</p> <p>Todos los espacios de participación que ofrece el profesor están relacionados con los aprendizajes de la clase.</p>
Básico	<p>El profesor ofrece espacios de participación a todos los alumnos, sin privilegiar a uno o a un grupo de ellos.</p> <p style="text-align: center;">Y</p> <p>La mayoría de los espacios de participación que ofrece el profesor están relacionados con los aprendizajes de la clase.</p>
Insatisfactorio	<p>No cumple alguna de las condiciones de “Básico”.</p> <p style="text-align: center;">O</p> <p>Casi siempre el profesor está desatento a las intervenciones y preguntas de los alumnos durante la clase.</p>

Tal como se aprecia en el ejemplo anterior, el nivel Competente de la rúbrica recoge la descripción que el MBE hace del descriptor en cuestión; el nivel Básico se basa en el Competente, pero es menos completo, y el Insatisfactorio representa aquel desempeño que no cumple en modo alguno con lo esperado. Por su parte, el nivel Destacado va más allá de lo esperable, se trata de un desempeño de alta riqueza pedagógica.

Corrección del Portafolio

Tanto el Módulo 1 como el Módulo 2 del Portafolio son evaluados por docentes de aula, denominados profesores correctores, mediante rúbricas o pautas de evaluación. Estas rúbricas son creadas en forma conjunta por especialistas en medición y docentes con reconocida trayectoria en cada subsector de enseñanza. Los profesores correctores reciben un entrenamiento en el contenido y forma de aplicación de estas pautas previo al proceso de corrección, además, deben tener un perfil profesional específico, que se detalla a continuación.

Los profesores correctores deben ser docentes que han acreditado una experiencia de al menos cinco años haciendo clases en el nivel y subsector de los docentes cuyos módulos revisarán. Además debe

11 Este es un ejemplo de rúbrica para ilustrar aquellas efectivamente utilizadas; dado el carácter confidencial de estas últimas el ejemplo aquí presentado tiene solo un propósito ilustrativo.

tratarse de docentes en ejercicio que cumplan con el requisito de poseer título profesional o habilitación de MINEDUC para ejercer. Estos docentes, previo a participar en el proceso de corrección de alguno de los dos módulos, deben aprobar una serie de pruebas de selección y luego demostrar que tienen las competencias necesarias para un proceso de este tipo, que implica un buen manejo de rúbricas de evaluación.

De acuerdo a la normativa que reglamenta la Evaluación Docente, la corrección de cada módulo debe realizarse de manera tal que se asegure imparcialidad en la aplicación de la rúbrica. Para esto, cada profesor corrector evalúa individualmente los módulos que se le asignan, asignación que se hace al azar. Los profesores correctores no tienen acceso a la identidad de los docentes cuyos módulos evalúan, pero si eventualmente un corrector se percatara de que está evaluando la evidencia enviada por un docente que conoce (por ejemplo, al verlo en la filmación de la clase), debe abstenerse de evaluarlo y ese módulo es reasignado a otro corrector. Además, en el proceso de corrección se implementan una serie de mecanismos para asegurar y monitorear que la evaluación que realiza cada profesor corrector esté rigurosamente apegada a la pauta de corrección, puesto que la evaluación de los docentes debe ser completamente equitativa. Estos procedimientos se detallan más adelante en este mismo capítulo.

Otro aspecto importante de destacar es que el proceso de corrección es confidencial, lo que implica que los profesores correctores deben guardar estricta reserva tanto sobre las pautas de evaluación como respecto a la evidencia que corrigen, tal como establece el Reglamento sobre Evaluación Docente. Para dar cumplimiento a este aspecto, el proceso de corrección de Portafolio se realiza bajo estrictas normas de seguridad, protegiendo de esta forma tanto la confidencialidad como la integridad del material enviado por cada docente que participa en el proceso de evaluación; todos quienes participan en el proceso de corrección firman un Convenio de Confidencialidad, documento legal en que comprometen resguardar este aspecto hasta por cinco años luego de terminado su trabajo. Para realizar el proceso de corrección de cada año, entre los meses de noviembre y enero distintas facultades universitarias forman Centros de Corrección de Portafolio. Estos Centros de Corrección se organizan de manera tal que se encargan de proveer las instalaciones necesarias para el desarrollo del proceso de corrección y se responsabilizan por el reclutamiento de personas idóneas para cada uno de los cargos necesarios. Todos los Centros de Corrección cuentan con protocolos específicos de trabajo, de manera tal que en cada uno de ellos se realiza la evaluación de los módulos de los docentes siguiendo los mismos parámetros de rigurosidad, calidad técnica y eficiencia. Este proceso es dirigido en forma directa y permanente por el equipo técnico de MIDE UC, cuyos profesionales supervisan y garantizan la estandarización de los procedimientos en los Centros. Para hacer este trabajo, existe una serie de herramientas y aplicaciones del software desarrollado por MIDE UC que permiten monitorear el manejo de material, la asignación aleatoria de módulos a los correctores, la calibración o acuerdo entre ellos y la productividad, por mencionar los aspectos más relevantes.

Las etapas de trabajo del Centro de Corrección pueden ser resumidas de la siguiente manera:

- a. *Instalación del Centro:*** consiste en el reclutamiento de los profesionales necesarios y acondicionamiento de los espacios de trabajo de los profesores correctores y supervisores de corrección. También se prepara una sección destinada a la digitación de las evaluaciones realizadas y una bodega con estrictas normas de seguridad.
- b. *Entrenamiento de supervisores de corrección y profesores correctores:*** consiste en capacitaciones impartidas por expertos en las pautas de evaluación, de carácter teórico y práctico para asegurar una comprensión cabal de cada una de las dimensiones a evaluar.
- c. *Marcha Blanca del proceso de corrección:*** una vez que todos los integrantes del Centro de Corrección se encuentran preparados para cumplir sus funciones, el Centro completo funciona a modo de prueba durante una semana completa, de manera de poner a prueba los conocimientos adquiridos. En esta etapa se llevan a cabo todos los procedimientos de la corrección real, con la salvedad de que todo lo que se corrige en esta etapa es corregido nuevamente durante el período de corrección definitiva. Si el Centro de Corrección aprueba este proceso, demostrando

el manejo técnico y rigurosidad necesaria para completar sus tareas, se da paso al proceso de corrección real y definitiva. En esta etapa se determina si los profesionales que se ha seleccionado para desempeñarse como profesores correctores y supervisores de corrección son capaces de realizar correctamente su trabajo, y por tanto, también se decide su continuidad en el proceso.

- d. *Corrección definitiva*: consiste en el proceso de corrección real de los módulos del Portafolio. Se encuentra programada de manera tal que todos los módulos sean corregidos durante el mes de enero, y de esta forma, asegurar que los resultados de esta evaluación siga las siguientes etapas del proceso, y así todos los docentes reciban sus resultados (que incluyen los de los otros tres instrumentos, más la consideración de sus antecedentes contextuales a nivel comunal) a fines de marzo.

Tal como ya se mencionó, un aspecto crucial de la evaluación de los módulos del Portafolio es que se trate de un proceso objetivo y rigurosamente apegado a las pautas de evaluación que se generan en base al MBE. Todo el trabajo que hace un Centro de Corrección busca garantizar que los módulos sean corregidos bajo estándares de calidad técnica, orientados a asegurar una corrección imparcial y confiable. En el período de marcha blanca y en el de corrección definitiva se llevan a cabo los siguientes procesos orientados a asegurar la imparcialidad y equidad del proceso de corrección:

- *Doble corrección*: un 20% al azar de los módulos de cada subsector es evaluado por dos correctores distintos. Para asegurar la validez de este proceso, los correctores no saben cuál de los módulos que se les asigna será corregido dos veces, y tampoco tienen conocimiento de quién será el otro corrector que corregirá ese módulo. Cada día, todos los correctores participan en este proceso de doble corrección para al menos dos de los módulos que se les asigna. Si los correctores aplican rigurosamente la pauta de evaluación, los puntajes que asignan al módulo no debieran diferir entre sí, dentro de un cierto rango de tolerancia. Si esta diferencia existe se da paso al proceso de recalibración.
- *Recalibración*: Cuando se detectan diferencias en la evaluación de dos correctores en el proceso de doble corrección, este módulo es corregido por un tercero, el supervisor de la sala de corrección. Este supervisor evalúa nuevamente el módulo y en base a su juicio se determina cuál de los correctores evaluó alejándose de la pauta de corrección. Con esta información, al día siguiente y antes de comenzar la corrección de los módulos asignados, el supervisor se reúne con el corrector “descalibrado” y revisa con él la pauta de corrección y el módulo específico en cuestión, de manera de hacerle ver la razón de su equivocación y asegurarse que comprende bien la pauta de evaluación (proceso de recalibración). Si un corrector debe ser citado en reiteradas oportunidades a recalibración, es excluido del proceso.
- *Correcciones grupales*: al inicio de cada semana, todos los correctores de un ciclo y subsector corrigen un mismo módulo, con el fin de asegurar que todos estén aplicando las pautas de corrección con los mismos criterios. Esta corrección grupal es dirigida por los profesores especialistas que elaboraron la pauta de corrección y los supervisores de cada sala de corrección. Una vez corregido el módulo grupal, se da paso a una discusión en que los correctores argumentan la forma en que evaluaron ese módulo.

b. Construcción de los instrumentos

Uno de los aspectos esenciales a tener en cuenta en el proceso de construcción de los instrumentos de evaluación es el equilibrio entre estabilidad y cambio o, en otras palabras, la necesidad de mantener elementos comunes a través de los años y la necesidad de introducir cada año ciertas modificaciones a los instrumentos. Lo primero (estabilidad) obedece al requerimiento de contar con un sistema de evaluación comparable en el tiempo, mientras que lo segundo (cambio) permite una cobertura razonable del MBE, posibilita introducir mejoras al sistema cada año y, a nivel práctico, ayuda a evitar la ocurrencia de malas prácticas (como por ejemplo, la copia o falseamiento de la información). Por

ejemplo, en el caso del Portafolio, y tal como ya fue indicado previamente en este capítulo, tanto los productos solicitados al profesor como las rúbricas de corrección mantienen cierta estabilidad entre un año y otro; además de la comparabilidad, esto obedece a que hay aspectos muy sustantivos o relevantes del trabajo pedagógico que siempre deben abordarse (como la capacidad de planificar una unidad pedagógica, evaluar el aprendizaje de los estudiantes y reflexionar sobre la propia práctica). Pero junto con ello, los Portafolios y sus rúbricas no son idénticos entre años, se hacen cambios en base a lo que indican los análisis de datos de los períodos previos, para dar cuenta de las particularidades de los distintos subsectores y modalidades que van entrando a la evaluación y para prevenir la copia o presentación falseada de Portafolios de años anteriores.

En el caso de los otros instrumentos, aquel en que se han implementado más cambios es el Informe de Referencia de Terceros. En el año 2009 se usó una escala de 8 puntos en lugar de 4, como un modo de favorecer mayor variabilidad en las respuestas de los directores y jefes de UTP, dándoles la posibilidad de hacer distinciones más finas entre los docentes que evalúan; de acuerdo a los resultados obtenidos, esta medida no fue en la dirección esperada (en general, se tendieron a mantener puntajes en el rango superior de la escala y con una baja variabilidad). En el año 2010 se introdujo en este instrumento una nueva variación: las preguntas se acompañaron de una rúbrica y se pidió una justificación escrita de las evaluaciones “Destacado”; esto implicó algunos cambios en los resultados, concordantes con lo esperado⁽¹²⁾. A continuación, para ilustrar lo anteriormente expuesto, se presenta un extracto de este instrumento en sus versiones 2009 y 2010 (ver Figuras 4 y 5).

• **Figura 4**
 • *Extracto del Informe de Referencia de Terceros - Versión 2009*

Dominio A: Preparación de la enseñanza									
Criterio A.1. Domina los contenidos de las disciplinas que enseña y el marco curricular nacional.									
Criterio A.4. Organiza los objetivos y contenidos de manera coherente con el marco curricular y las particularidades de sus alumnos.									
		Insatisfactorio	Básico	Competente	Destacado				
1. Diseña unidades y planes de trabajo coherentes con el currículum oficial.	S/I	1	2	3	4	5	6	7	8
2. Emite opiniones, argumentos o explicaciones que muestran dominio de las materias o temáticas que enseña.	S/I	1	2	3	4	5	6	7	8
3. Diseña actividades coherentes con las características de los alumnos (conocimientos previos, intereses, dificultades, etc.).	S/I	1	2	3	4	5	6	7	8

.....
 12 Ver Sun, Correa, Zapata & Carrasco, en este volumen.

• **Figura 5**
 • *Extracto del Informe de Referencia de Terceros - Versión 2010*

Dominio A: Preparación de la enseñanza	
Criterio A.2	Conoce las características, conocimientos y experiencias de sus estudiantes.
Criterio A.4	Organiza los objetivos y contenidos de manera coherente con el marco curricular y las particularidades de sus alumnos.
Criterio A.5	Las estrategias de evaluación son coherentes con los objetivos de aprendizaje, la disciplina que enseña, el marco curricular nacional y permiten a todos los alumnos demostrar lo aprendido.

Responda las rúbricas 1, 2 y 3, considerando los criterios del MBE recién presentados, los que pueden estar recogidos en forma completa o parcial por las conductas o indicadores asociados a estas rúbricas.

1	S/I	Insatisfactorio	Básico	Competente	Destacado
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
		No cumple la condición establecida para el Nivel Básico.	Demuestra conocimiento de aquellos contextos personales y sociales de los alumnos que influyen en el aprendizaje. O Sabe cómo integrar en la preparación de sus clases las experiencias de sus alumnos.	Demuestra conocimiento de aquellos contextos personales y sociales de los alumnos que influyen en el aprendizaje. Y Sabe cómo integrar en la preparación de sus clases las experiencias de sus alumnos.	Demuestra conocimiento de aquellos contextos personales y sociales de los alumnos que influyen en el aprendizaje. Y Sabe cómo integrar en la preparación de sus clases las experiencias de sus alumnos. Y <i>Reúne otras condiciones que lo hacen sobresalir en los indicadores (justifique en el recuadro siguiente).</i>
<i>Solo si marcó Destacado, fundamente aquí describiendo la(s) conducta(s) sobresaliente(s)</i>					

En suma, los instrumentos van experimentando cambios en forma regulada y planificada. Se mantiene una estructura estable a pesar de que algunas de sus tareas cambian, y esto da cuenta de un sistema flexible, en constante revisión y que se va perfeccionando a partir de la experiencia y la acumulación de conocimiento.

Un proceso de construcción de instrumentos como el antes descrito exige la realización de una serie de etapas orientadas a asegurar la validez y pertinencia de las modificaciones que se realizan año a año. Para esto, en cada ciclo de evaluación se cuenta con la participación de docentes especialistas en los subsectores evaluados, quienes conforman comisiones de trabajo donde se revisan las propuestas de cambios a los instrumentos asegurando que cada modificación resulte pertinente para la variedad de docentes que participan en el proceso de evaluación de ese año. En forma paralela, cuando un subsector nuevo se incorpora a la evaluación, se forman equipos de investigación en terreno que realizan procesos de observación de clases, entrevistas a docentes de aula y revisión de documentos, para estudiar en detalle las modificaciones que podrían ser necesarias para que los instrumentos de evaluación sean pertinentes a estos nuevos grupos.

Todos estos procesos de revisión y estudio se traducen en una primera propuesta de cada instrumento. Estas propuestas son sometidas a estudios pilotos con docentes de aula de los distintos subsectores evaluados, quienes responden las preguntas o instrucciones nuevas. Esto permite determinar en qué medida las instrucciones o preguntas se orientan en la dirección esperada, asegurando su claridad para diversos tipos de docentes, y además permite asegurar que docentes de diversos ámbitos del aprendizaje puedan estar en condiciones de responder adecuadamente a los productos solicitados. Estos docentes, al concluir esta etapa de testeo, participan en discusiones grupales donde se proponen nuevas modificaciones y sugerencias a los instrumentos en desarrollo. Producto de esta etapa, los instrumentos son escritos en su versión definitiva.

En el caso del Portafolio y la Entrevista por un Evaluador Par, no solo se deben realizar estudios piloto sobre las preguntas e instrucciones, sino también sobre las pautas de evaluación. Para esto, se realiza una aplicación de estos instrumentos a una muestra de docentes de perfil similar al de los docentes evaluados, y a partir de los productos que éstos entregan, se realiza un primer testeo del comportamiento psicométrico de las pautas de evaluación, orientado a asegurar su confiabilidad y su capacidad de distinguir entre niveles de desempeño. En el caso del Portafolio, una vez que se alcanzan estándares de calidad técnica satisfactorios, desde el punto de vista de acuerdo interjueces, consistencia interna y distinción de niveles de desempeño, se formulan de forma definitiva las rúbricas, cuyo comportamiento psicométrico sigue siendo monitoreado permanentemente en el proceso de corrección definitiva. Para esto, un equipo de especialistas realiza análisis en forma simultánea durante todo el proceso de corrección, lo que permite detectar señales de riesgo en cuanto a la confiabilidad en la evaluación de los correctores.

c. Cálculo del resultado del docente en cada instrumento

Autoevaluación, Informe de Referencia de Terceros y Entrevista

En cada uno de estos instrumentos se calcula un resultado promediando el conjunto de los indicadores asociados a cada uno de los dominios del MBE. Este promedio se informa al docente en términos de una categoría de desempeño (que engloba un rango de puntajes). A continuación se presenta un ejemplo de cómo se reportan los resultados de la Pauta de Autoevaluación:

• **Figura 6**
 • *Ejemplo de reporte de resultados Pauta de Autoevaluación – Versión 2010*

	Resultados			
	I	B	C	D
Dominio A: Preparación de la enseñanza				●
Dominio B: Creación de un ambiente propicio para el aprendizaje			●	
Dominio C: Enseñanza para el aprendizaje de todos los estudiantes				●
Dominio D: Responsabilidades profesionales				●
Niveles de desempeño I: Insatisfactorio B: Básico C: Competente D: Destacado				

Según se desprende de la gráfica anterior, este docente considera que su desempeño profesional es muy bueno, y que éste se destaca especialmente en su preparación para enseñar y promover el aprendizaje de todos sus estudiantes, como asimismo, en el cumplimiento de sus responsabilidades profesionales. Comparativamente, considera que su capacidad de crear un ambiente propicio para el aprendizaje es de un buen nivel, más no sobresaliente.

Para calcular el puntaje final de la Autoevaluación, la Entrevista y el Informe de Referencia de Terceros, se calcula un promedio de los puntajes de cada dominio del MBE. Todos los dominios del Marco tienen, por tanto, el mismo peso en el puntaje final del docente en cada uno de estos instrumentos de evaluación.

Portafolio

En el caso del Portafolio, los distintos dominios del MBE se operacionalizan en distintas dimensiones de evaluación. En total, son ocho dimensiones, cada una de las cuales está asociada a un producto del Módulo 1 o al Módulo 2, según se indica en la siguiente tabla:

● **Tabla 1**
● *Productos del Portafolio y dimensiones de evaluación asociadas*

Módulo 1	
Producto 1: Unidad Pedagógica	Dimensión A: Organización de los elementos de la unidad
	Dimensión B: Análisis de las actividades de las clases
Producto 2: Evaluación de la unidad pedagógica	Dimensión C: Calidad de la evaluación de la unidad
	Dimensión D: Reflexión a partir de los resultados de la evaluación
Producto 3: Reflexión pedagógica	Dimensión E: Reflexión pedagógica
Módulo 2	
Producto 4: Clase grabada	Dimensión F: Ambiente de la clase para el aprendizaje
	Dimensión G: Estructura de la clase
	Dimensión H: Interacción pedagógica

A continuación se presentan con mayor detalle las ocho dimensiones evaluadas en el Portafolio:

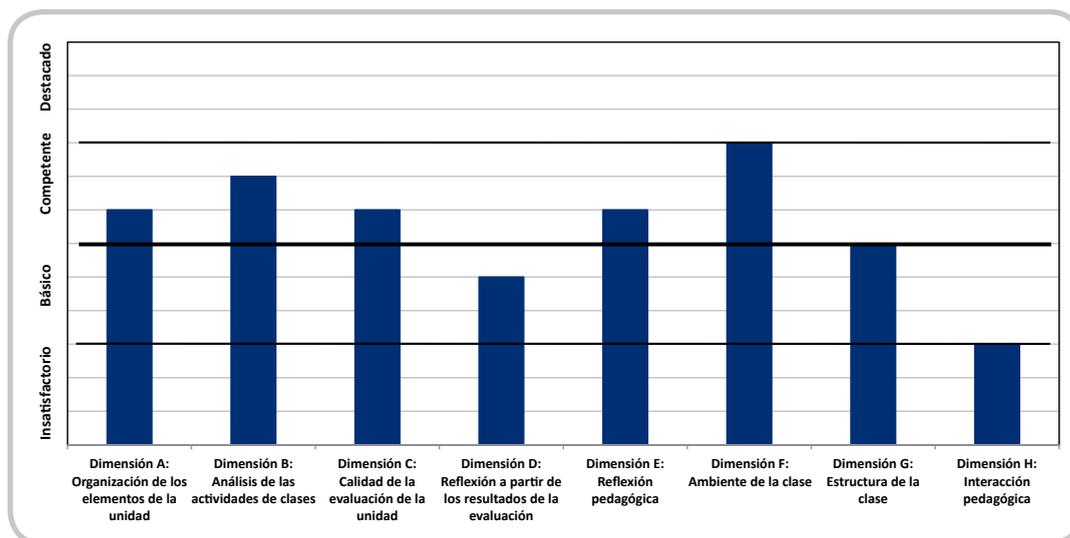
- Dimensión A: Organización de los elementos de la unidad. En esta dimensión se evalúa la capacidad del docente para establecer objetivos de aprendizajes que estén correctamente formulados y que sean coherentes entre sí, como asimismo, la capacidad de abordar pedagógicamente los objetivos de aprendizaje propuestos y realizar actividades coherentes con ellos. También se evalúa la capacidad de organizar las clases de la unidad en una secuencia pedagógica coherente.
- Dimensión B: Análisis de las actividades de las clases. En esta dimensión se evalúa la capacidad del docente para diseñar estrategias que le permitan abordar las dificultades que presentan los alumnos durante su proceso de aprendizaje. También se considera la capacidad de reconocer las fortalezas de una unidad pedagógica y cómo éstas favorecen el aprendizaje de los alumnos, y

junto con ello, los aspectos que dificultan dicho aprendizaje.

- Dimensión C: Calidad de la evaluación de la unidad. Se evalúa la capacidad del docente de diseñar evaluaciones que permitan determinar en sus estudiantes el logro de los aprendizajes de una unidad pedagógica. Más específicamente, se considera la capacidad de formular instrucciones e ítems claros y de diseñar pautas de corrección que identifiquen adecuadamente las respuestas o desempeños esperados frente a dichas instrucciones o ítems.
- Dimensión D: Reflexión a partir de los resultados de la evaluación. En esta dimensión se evalúa la capacidad del profesor para orientar a los alumnos para mejorar sus aprendizajes a partir del análisis de los aspectos logrados y no logrados en la evaluación y, junto con ello, la capacidad de utilizar los resultados de ella como un insumo para analizar su propio trabajo pedagógico y planificar acciones futuras.
- Dimensión E: Reflexión pedagógica. Evalúa la capacidad que tiene el docente para reflexionar sobre distintos aspectos de su práctica, tales como su capacidad para preparar diversas actividades de enseñanza que permiten a alumnos de distintas características lograr un mismo objetivo de aprendizaje; identificar factores que desmotivan a los alumnos y diseñar nuevas actividades para superarlos; e identificar los factores que dificultan la existencia de un clima propicio para el aprendizaje y plantear acciones para restablecerlo.
- Dimensión F: Ambiente de la clase para el aprendizaje. En esta dimensión se evalúa la capacidad del docente para lograr que los alumnos se focalicen en las actividades propuestas, manteniendo normas de convivencia que favorezcan el desarrollo de la clase, entregando instrucciones claras y supervisando el desarrollo de las actividades. Además se evalúa la capacidad para generar instancias equitativas de participación para los alumnos, promoviendo la interacción entre ellos.
- Dimensión G: Estructura de la clase. Se evalúa la capacidad del profesor de proponer actividades de aprendizaje que contribuyan a lograr los objetivos planteados en el tiempo disponible; además, se considera en qué medida promueve que los alumnos tengan un primer acercamiento a los contenidos que se abordarán y cuál es su capacidad para finalizar la clase con instancias en que los alumnos puedan consolidar los aprendizajes trabajados.
- Dimensión H: Interacción pedagógica. En esta dimensión se evalúa la capacidad del docente de explicar contenidos o procedimientos usando estrategias que promuevan una comprensión acabada por parte de los alumnos. Junto con ello, se atiende a la capacidad de formular preguntas de calidad que promuevan una interacción propicia para el aprendizaje y, asimismo, de aprovechar las preguntas e intervenciones de los alumnos por medio de estrategias que les permitan profundizar sus conocimientos. Finalmente, en esta dimensión se evalúa la capacidad del profesor de implementar situaciones de aprendizaje acordes a los contenidos y metodologías propias del subsector que enseña.

Las ocho dimensiones evaluadas en el Portafolio tienen el mismo peso relativo en el resultado del docente en este instrumento. A cada docente se informa el puntaje obtenido en cada dimensión, en términos de una categoría de desempeño, de forma similar a los otros instrumentos. A continuación se presenta un ejemplo:

••• **Figura 7**
 ••• *Imagen del gráfico con resultados en dimensiones del Portafolio–Versión 2010*



Este perfil de resultados muestra que el mejor desempeño de este docente es en una de las dimensiones del Módulo 2, capacidad de generar un buen ambiente de clase, seguida por la dimensión del Módulo 1 relativa al análisis de las actividades de las clases; también la organización de los elementos de la unidad, la calidad de la evaluación implementada y la reflexión pedagógica, son áreas con un desempeño relativo más alto. Complementariamente, el área con más debilidades es la interacción pedagógica. En otras palabras, no hay un desempeño homogéneo en las distintas áreas evaluadas, ni tampoco se concentran fortalezas y debilidades en uno de los módulos, sino que se observan aspectos mejor logrados que otros tomando en cuenta las distintas acciones que lleva a cabo un profesor en el ejercicio de su labor.

d. Los instrumentos de evaluación y la consideración de elementos de contexto

Un tema que ha sido preocupación permanente entre los actores involucrados en el sistema es cómo compatibilizar una evaluación estandarizada, con iguales condiciones de aplicación y análisis para todos, con una evaluación que considere las particularidades de los contextos sociales, económicos y culturales en que cada profesor lleva a cabo su trabajo. La respuesta a esta inquietud radica en dos ámbitos principales; por una parte, en los rigurosos procedimientos que se siguen durante el proceso de construcción de los instrumentos y, por otra, en el trabajo que realiza a nivel local la Comisión Comunal de Evaluación cuando analiza los resultados del docente en cada uno de los instrumentos a la luz de las consideraciones contextuales que se recogen en tres de ellos (Autoevaluación, Entrevista por un Evaluador Par e Informe de Referencia de Terceros, según fue descrito en la sección anterior).

Como ya fue indicado previamente, durante el proceso de construcción de los instrumentos y las rúbricas o pautas de corrección participan diversos profesionales, la mayoría de ellos profesores de aula de los distintos ciclos y subsectores, quienes analizan las preguntas, directrices e instrucciones de los instrumentos a partir de su experiencia profesional, de modo de evaluar cuán “sintonizadas” están con las distintas realidades o contextos que enfrenta un docente en su ejercicio profesional. Lo anterior se complementa con los estudios piloto y procesos de análisis también descritos, llevados a cabo cada

año para poner a prueba las preguntas, directrices e instrucciones que componen los instrumentos y estudiando sus propiedades psicométricas, analizando si algunas de ellas podrían beneficiar o perjudicar a uno o más grupos de docentes que se desempeñen en áreas específicas. Finalmente, cabe mencionar que es una preocupación constante recoger retroalimentación de los propios docentes evaluados a través del cuestionario complementario y diversos focus groups con profesores y directivos que se llevan a cabo durante distintos momentos del año.

Tal como fue señalado previamente, para la inclusión de consideraciones del contexto de trabajo específico de cada profesor, el docente que participa en el proceso de evaluación puede explicitar tanto en la Pauta de Autoevaluación como en la Entrevista por un Evaluador Par si considera que durante el año en que se lleva a cabo la evaluación han ocurrido eventos puntuales que podrían afectar su desempeño profesional, o bien, si éste se lleva a cabo bajo ciertas condiciones particulares que es necesario atender. Lo mismo puede ser indicado por el director del establecimiento y el jefe de la unidad técnico pedagógica en el Informe de Referencia de Terceros. La información recogida en estos tres instrumentos constituye un registro que es analizado cuidadosamente por la Comisión Comunal de Evaluación, instancia colegiada formada por todos los Evaluadores Pares de la comuna más el coordinador comunal.

El trabajo de esta comisión consiste, en primer lugar, en un análisis de los resultados aportados por los cuatro instrumentos de evaluación para cada docente de la comuna. El Reglamento sobre Evaluación Docente establece que esta comisión tiene la atribución de ratificar el resultado así obtenido (dado por la ponderación de los puntajes de los instrumentos), o bien, modificar esa categoría, pudiendo hacerlo por un nivel de desempeño superior o inferior, de acuerdo a las consideraciones contextuales sometidas a su consideración. Para hacer modificaciones, la normativa exige un quórum y acuerdo específico entre los integrantes de la Comisión. Además, en el caso de modificar el resultado del profesor, ésta debe consignar los fundamentos de su decisión, en el informe de resultados que recibirá el docente, de manera de asegurar un proceso de evaluación transparente. Además, las Comisiones Comunales tienen la oportunidad de hacer recomendaciones a los profesores de su comuna (por ejemplo, invitarlos a cursos de perfeccionamiento que se dictarán ese año, o incentivarlos a postular a asignaciones por su buen desempeño).

Consideraciones finales

La Evaluación Docente se basa en los estándares señalados en el MBE y, desde esa perspectiva, se fundamenta en lo que de manera consensuada se considera una buena práctica profesional. El que se trate de una evaluación referida a criterios posibilita el que los resultados que cada docente obtiene pueden ser entendidos a la luz de lo que se concibe como un desempeño esperado, ilustrando así de manera clara cuáles son los aspectos mejor logrados y cuáles son aquellos que requieren mejora. Como se señalaba al comienzo del capítulo, esto posibilita que la evaluación tenga un carácter formativo.

Estos estándares son varios y diversos, tal como el desempeño profesional docente es complejo y compuesto por múltiples variables y dimensiones. Por eso, es razonable que se plantee un sistema de evaluación que aborde el fenómeno de interés desde múltiples perspectivas y a través de distintos instrumentos. Cualquier mirada única que además se captura por un instrumento particular es, por definición en este contexto, parcial e incompleta.

Desde el punto de vista técnico de la calidad de los instrumentos que forman parte de la Evaluación Docente, se ha puesto especial cuidado en cautelar que los procedimientos de construcción y corrección sigan rigurosamente las recomendaciones y estándares internacionales en procesos de este tipo. Entre ellas, cabe destacar la validación con expertos, la realización de estudios piloto, la existencia de pautas de corrección únicas y los procesos de monitoreo de la calibración de los correctores.

En suma, el Sistema de Evaluación del Desempeño Profesional Docente en Chile es un proceso sólido, que ha sido diseñado de forma coherente con el principio formativo que lo guía y que se ha implementado bajo rigurosos procedimientos que dan garantía de que la información que entrega es confiable. Como sistema de evaluación, cada uno de sus componentes se encuentra orientado a proveer oportunidades de aprendizaje a los docentes, retroalimentar y enriquecer su práctica pedagógica, favorecer la reflexión didáctica y, por último, contribuir al reconocimiento de aquellos docentes que ponen sus conocimientos y capacidades al servicio del mejoramiento de la calidad de los aprendizajes de los estudiantes.



3

La implementación de la Evaluación Docente

Yulan Sun · Paulina Calderón
Nelson Valerio · Pablo Torres

Centro de Medición MIDE UC
Pontificia Universidad Católica de Chile

La implementación de la Evaluación Docente

¿Cómo se aplica la Evaluación Docente a escala nacional?, ¿quiénes la implementan y qué rol cumplen los sostenedores y directivos?, ¿cómo se procesan los instrumentos de evaluación?, ¿cómo se define el resultado final de cada docente evaluado y qué tipo de retroalimentación recibe?

Para responder a estas y otras interrogantes, en este capítulo se expone el marco normativo de la Evaluación Docente, las distintas entidades y actores del sistema educacional que participan en su implementación y las etapas con que se lleva a cabo. Mediante descripciones, datos y ejemplos, se da cuenta de la forma en que opera el sistema y su consolidación gradual como una herramienta de evaluación confiable y objetiva. La presentación tiene siempre como perspectiva el sentido y propósitos finales de la evaluación, e ilustra distintas iniciativas comunales que muestran cómo una gestión local informada y activa es vital para alcanzarlos.

Características generales de la Evaluación Docente

El Sistema de Evaluación del Desempeño Profesional Docente (en adelante Evaluación Docente), se implementa en Chile desde el año 2003 y constituye un proceso obligatorio para los más de 70.000 docentes de aula que realizan su trabajo en establecimientos municipales, cuya matrícula corresponde al 42,1% del total de alumnos del país (Ministerio de Educación, 2010).

La Ley 19.933, promulgada en 2004, estableció el reemplazo del sistema de calificaciones del Estatuto Docente por un nuevo proceso de evaluación con carácter formativo, cuya implementación se inició con la incorporación voluntaria de 63 y 104 comunas en los años 2003 y 2004 respectivamente, tras un acuerdo firmado entre el Ministerio de Educación, la Asociación Chilena de Municipalidades y el Colegio de Profesores.

La obligatoriedad del nuevo sistema de evaluación quedó consagrada al promulgarse la Ley 19.961 sobre Evaluación Docente (de agosto de 2004) y el Reglamento de la misma (Decreto 192 de Educación, de agosto de 2004). Al año siguiente, el artículo 36 de la Ley 20.079 estableció la creación del "bono pro calidad de la dotación docente", destinado a los profesores que deben abandonar la dotación a causa de su mal resultado en el proceso, y señaló que aquellos que, debiendo evaluarse, se negaren a ello, se presumirán evaluados con nivel Insatisfactorio, no tendrán derecho a los Planes de Superación Profesional y tendrán la obligación de evaluarse al año siguiente. Por último, en febrero de 2011, se promulgó la Ley 20.501 que introdujo cambios en las consecuencias de la evaluación, señalando que los docentes con nivel Insatisfactorio deben abandonar la dotación tras haber obtenido ese resultado en dos ocasiones consecutivas (no tres, como hasta entonces). Además, los docentes que resulten calificados con desempeño Básico en tres evaluaciones consecutivas, o en forma alternada con desempeño Básico o Insatisfactorio en tres evaluaciones consecutivas, también deberán dejar la dotación.

La Evaluación Docente está basada en estándares que miden el desempeño de los docentes en función de los criterios establecidos en el Marco para la Buena Enseñanza (Ministerio de Educación, 2003). En base a él se construyen los cuatro instrumentos de evaluación⁽¹⁾:

- la **Autoevaluación**, que completa el propio evaluado (10% del puntaje total);
- el **Portafolio**, a través del cual cada docente entrega evidencias directas de su trabajo, incluyendo la grabación de una clase (60%);
- la **Entrevista por un Evaluador Par**, realizada por un profesor de características similares al docente evaluado, empleando una pauta estructurada y una rúbrica de corrección (20%);
- y los **Informes de Referencia de Terceros**, emitidos por el Director y el Jefe de la Unidad Técnico Pedagógica (UTP) del establecimiento (10%).

Para los docentes que vuelven a evaluarse tras obtener un nivel Insatisfactorio en el período anterior, la ponderación del Portafolio sube al 80% y los demás instrumentos la reducen a la mitad.

Cuatro son los niveles de desempeño que identifica la evaluación: *Insatisfactorio*, *Básico*, *Competente* (el nivel esperado) y *Destacado*. Los docentes deben evaluarse cada cuatro años, salvo cuando han obtenido un resultado Insatisfactorio (en cuyo caso deben repetir el proceso al año siguiente) o Básico (debiendo entonces evaluarse tras dos años, de acuerdo a la Ley 20.501).

Como muestra la Tabla 1, en ocho años de implementación la Evaluación Docente ha alcanzado una amplia cobertura, abarcando la mayor parte de los 71.797 docentes de aula de la dotación municipal actual⁽²⁾, incluyendo los tres niveles educacionales que concentran la mayor parte de la matrícula escolar: Educación Parvularia, Enseñanza Básica y Educación Media. En el futuro el sistema deberá incorporar nuevas modalidades educativas, como la Educación de Adultos y la Educación Técnico Profesional.

• **Tabla 1**
• *Datos por período de evaluación 2003–2010*

	2003	2004	2005	2006	2007	2008	2009	2010
Niveles y modalidades educacionales	PC	PC SC	PC SC EM	PC SC EM	PC SC EM	PC SC EM EP	PC SC EM EP EE	PC SC EM EP EE
Comunas (*)	63	104	330	342	337	339	342	340
Docentes evaluados	3.673	1.719	10.665	14.190	10.413	16.015	15.699	11.061
Docentes re-evaluados (**)	-	-	-	-	1.906	956	5.306	5.624
Docentes que se niegan a ser evaluados sin causa justificada	-	-	-	1.664	731	846	743	402
Docentes con evaluación objetada (***)	68	2	2	39	18	10	12	9

PC= Primer Ciclo; SC= Segundo Ciclo; EM= Educación Media; EP= Educación Parvularia; EE= Educación Especial

(*) A partir de 2005, el sistema cubre todas las comunas del país. En la tabla se consideran solo aquellas en las que al menos un docente concluyó la evaluación en el período.

(**) Corresponde a un subconjunto de la fila anterior: Docentes que se evalúan por segunda vez tras cumplirse 4 años de su resultado Básico, Competente o Destacado anterior.

(***) La evaluación de estos docentes se declara "objetada" o inválida por haberse constatado una violación a las condiciones de veracidad exigidas por el proceso. Por tanto, estos docentes no concluyen el proceso con un resultado válido y deben evaluarse nuevamente al año siguiente.

.....

1 Para mayor información sobre el Marco y la construcción de los instrumentos, ver Flotts & Abarzúa, en este volumen.

2 Información extraída de la Base de Datos del Bono de Reconocimiento Profesional (BRP), MINEDUC, Enero 2011.

Actores y entidades responsables

La legislación vigente estipula cuatro entidades participantes:

- a. el Ministerio de Educación (MINEDUC) a través del Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas (CPEIP);
- b. una comisión técnica asesora conformada por académicos de educación superior, representantes del Colegio de Profesores y de la Asociación Chilena de Municipalidades;
- c. un equipo técnico asesor independiente que corresponde a una institución de educación superior contratada por el MINEDUC (hasta ahora este rol lo ha desempeñado la Pontificia Universidad Católica de Chile a través de su Centro de Medición MIDE UC); y
- d. las Comisiones Comunales de Evaluación, a través del Jefe del Departamento de Administración Educacional Municipal (DAEM o DEM) o el Director de la Corporación Municipal de Educación de cada comuna³.

A continuación se describe el rol que cumple cada una de estas entidades en torno a la Evaluación Docente.

a. *Ministerio de Educación–CPEIP*

A través del CPEIP, el MINEDUC es la entidad oficialmente responsable de la Evaluación Docente. Dentro de sus roles se encuentra coordinar la implementación de ésta y velar por el cumplimiento de la normativa legal que regula el proceso y la articulación entre la evaluación y las distintas políticas orientadas al desarrollo profesional docente. Además, es responsable de coordinar y supervisar los Planes de Superación Profesional, destinados a los docentes que no alcanzan el nivel mínimo esperado en la evaluación⁴.

El CPEIP también se encarga de seleccionar la institución que asume el rol de asesor técnico para implementar la evaluación y dicta las bases que definirán sus tareas, correspondiéndole orientar y supervisar su trabajo. De igual forma, es su responsabilidad definir los niveles, subsectores y modalidades educacionales que se incorporan al proceso cada año, determinar la calendarización anual del mismo y realizar la selección de los docentes que ejercerán el rol de Evaluadores Pares en cada período. Asimismo, deben emitir informes técnicos para que las Comisiones Comunales de Evaluación puedan resolver los recursos de reposición presentados por aquellos docentes que, disconformes con el resultado de su evaluación, interponen una solicitud para su revisión.

b. *Comisión Técnica*

En el Reglamento sobre Evaluación Docente se establece que, para la coordinación técnica del proceso, el CPEIP contará con la asesoría de una comisión, conformada por tres representantes de la asociación gremial de docentes con mayor representación en el país, tres académicos de educación superior y tres representantes designados por la asociación nacional de municipalidades con mayor representatividad del país. A través de reuniones coordinadas por el CPEIP, esta comisión toma conocimiento de los distintos procedimientos implementados para llevar a cabo la evaluación, especialmente en lo que concierne a los lineamientos seguidos para la construcción de los instrumentos de evaluación, formulando recomendaciones para su elaboración.

.....

3 Si bien cada Comisión Comunal está integrada por el Jefe DAEM o Director de la Corporación Municipal de Educación y los Evaluadores Pares, como el rol de estos últimos tiene un carácter temporal, la función reside permanentemente en el único integrante estable, quien asume la responsabilidad de dicha comisión: el Jefe DAEM o Director de la Corporación Municipal.

4 Estos planes son diseñados e implementados por cada comuna, con fondos transferidos por CPEIP. Para mayor información sobre ellos, ver Cortés & Lagos, en este volumen.

c. Equipo técnico asesor–MIDE UC

Para la ejecución de la Evaluación Docente, el MINEDUC contrata la asesoría técnica de una entidad especializada. Desde el año 2003 a la fecha, ese rol ha sido desempeñado por la Pontificia Universidad Católica de Chile a través de su Centro de Medición, MIDE UC. En este centro, dependiente de la Escuela de Psicología y especializado en el ámbito de la medición educacional, se conforma el equipo responsable de poner en marcha y desarrollar los diversos procesos, acciones y productos que demanda la implementación de la evaluación, tanto en sus aspectos técnicos como operativos.

Entre las tareas del equipo técnico asesor de MIDE UC se encuentran: la elaboración y validación de los instrumentos de evaluación junto con sus manuales y rúbricas de corrección; el apoyo a la gestión de las autoridades comunales para la implementación de la evaluación a nivel local; la producción y distribución nacional de los instrumentos; la conformación de una red de productores audiovisuales para llevar a cabo las grabaciones de clases del Portafolio; el diseño y supervisión del proceso de corrección de Portafolios realizado en distintas universidades; el procesamiento de datos de los instrumentos; y la generación de los informes de resultados a nivel de docentes, establecimientos y municipios.

El equipo de MIDE UC también es responsable de mantener sistemas de información y apoyo para los docentes evaluados, y capacitar a los distintos actores del proceso, tales como Evaluadores Pares y correctores de Portafolio.

Además, dada su naturaleza académica, MIDE UC realiza investigaciones para generar conocimiento a partir de la evidencia y los datos recogidos por la Evaluación Docente.

d. Coordinación Comunal de Evaluación

En Chile los establecimientos educacionales públicos son administrados de manera descentralizada por municipalidades, a través de Departamentos de Administración de Educación Municipal (DAEM o DEM) y Corporaciones Municipales de Educación.

En la Evaluación Docente participan todas las comunas que cuentan con al menos un docente en ejercicio en establecimientos de dependencia municipal, incluyéndose bajo este criterio a un total de 346 unidades⁵. Cada comuna cuenta con una autoridad educativa representante para el proceso: el **Coordinador Comunal de Evaluación**. Este rol lo desempeña el Jefe DAEM o el Director de la Corporación respectiva, quien vela por la coordinación de la evaluación a nivel comunal.

La incorporación de los Coordinadores Comunales refleja el carácter descentralizado de la evaluación y su participación es fundamental pues, en su calidad de máximas autoridades educativas a nivel local, son responsables de la administración y seguimiento del proceso en cada comuna. Entre sus principales facultades y funciones están: elaborar un registro de los docentes que deben evaluarse cada año; resolver las solicitudes que los docentes pueden presentar para suspender su evaluación hasta el período siguiente; recibir los instrumentos de evaluación y distribuirlos a los destinatarios correspondientes, recuperarlos y entregarlos al equipo de MIDE UC; y convocar a la Comisión Comunal de Evaluación y asegurar su buen funcionamiento, entre otras. Además, les corresponde resolver situaciones particulares o excepcionales que afectan a docentes de su comuna y, en general, cumplen un importante rol al proporcionar información oficial y dar respuesta a las inquietudes de los actores a nivel local.

.....
5 No se considera a la comuna Antártica, por no contar con establecimientos educacionales municipales, al tiempo que Quinta Normal se considera como dos unidades locales para efectos de la Evaluación Docente, pues coexisten un DEM y una Corporación que tienen a cargo distintos establecimientos.

En la mayoría de las comunas, el Coordinador Comunal es apoyado por un **Encargado de Evaluación**, funcionario del DAEM o Corporación que colabora en la ejecución de las tareas contempladas en el proceso.

Los Coordinadores Comunales y Encargados de Evaluación son apoyados por el equipo técnico de MIDE UC a través de variadas estrategias presenciales y a distancia, de manera que puedan llevar a cabo de forma exitosa sus tareas. Entre las primeras se cuentan talleres y seminarios de capacitación, así como reuniones de trabajo en terreno. Entre las segundas están el contacto (telefónico y por correo electrónico) permanente con todos los municipios, además de la elaboración y entrega de manuales, documentos y materiales multimedia orientados a apoyar la labor del Coordinador Comunal.

Implementación anual de la evaluación

El proceso anual de evaluación involucra distintos hitos o etapas (ver Figura 1). En lo que sigue se describirán los más relevantes.

a. Inscripción de docentes a evaluar

El primer hito en la implementación anual de la evaluación es la generación del registro de docentes que deben evaluarse cada año. Para ello, tras iniciarse el año escolar en el mes de marzo, los Coordinadores Comunales deben inscribir a los docentes que cumplen con los requisitos de evaluación, de acuerdo a la normativa legal y a los lineamientos definidos por el CPEIP en cuanto a los niveles, subsectores y/o modalidades vigentes para el período correspondiente.

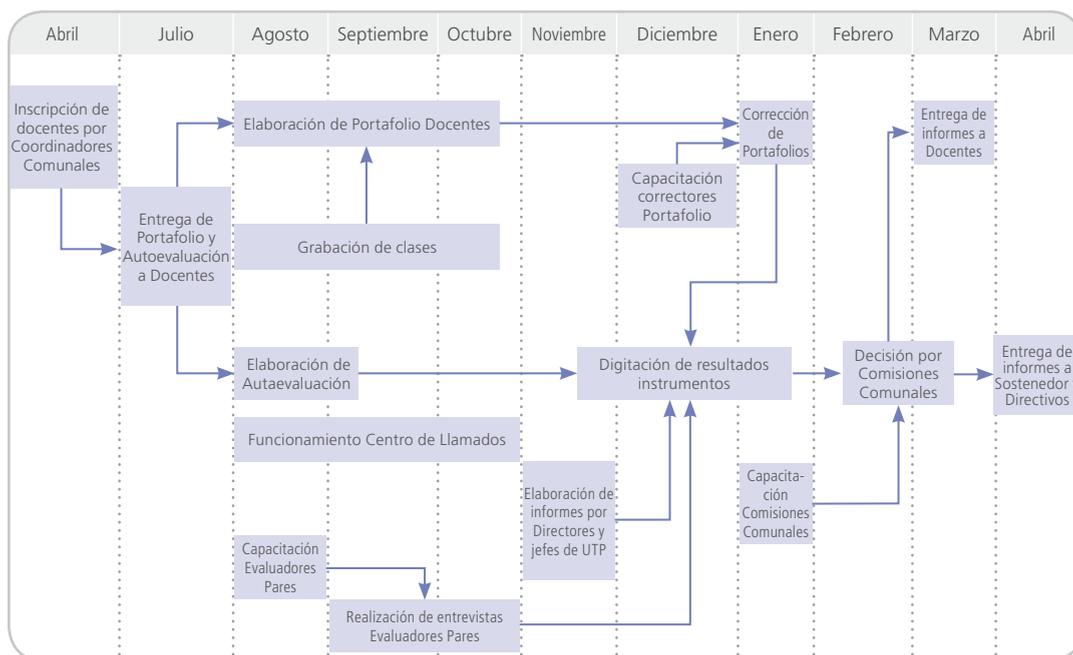
El registro de docentes así constituido experimenta modificaciones a lo largo del período de evaluación, de acuerdo a los distintos “estados” que pueden atravesar los profesores inicialmente inscritos:

- la mayoría concluye el proceso como “**evaluado**”, completando su evaluación en forma normal.
- algunos docentes cambian al estado “**eximido**” si se verifica que se encuentran solo en su primer año de ejercicio profesional en un establecimiento educacional, han sido seleccionados como Evaluadores Pares para el período, o bien, han presentado la renuncia anticipada e irrevocable a su cargo.
- un docente puede ser “**suspendido**” cuando el Coordinador Comunal ha acogido una solicitud de éste para postergar su evaluación por razones debidamente acreditadas. El Reglamento sobre Evaluación Docente establece como causales admisibles para esa resolución: razones de fuerza mayor; permiso sin goce de remuneraciones superior a tres meses (que afecte al proceso de elaboración de las evidencias de evaluación); traslado de establecimiento o de curso con fecha posterior al 30 de abril del año en evaluación y la realización de actividades de formación profesional fuera del territorio nacional, autorizadas por el sostenedor, durante el período de elaboración de evidencias de evaluación.
- cuando se constata que un docente ha faltado a las condiciones de veracidad exigidas para la evidencia que debe presentar, entregando para el proceso información falsa, copiada y/o elaborada por terceros, su evaluación es “**objetada**”, debiendo repetir el proceso al año siguiente, junto con quedar expuesto a sanciones administrativas⁶.
- por último existen docentes que, estando inscritos, **se niegan a completar su evaluación sin causa justificada**, aplicándose en ese caso lo señalado por el artículo 36 de la Ley 20.079, esto es, la asignación de un resultado Insatisfactorio y la obligación de evaluarse nuevamente en el período siguiente⁷.

6 En cada período existe un porcentaje muy marginal (0,1% del total de evaluados) de docentes que incurrir en esta situación (ver Tabla 1 en este capítulo).

7 La cantidad de docentes en esta situación ha disminuido en los períodos sucesivos de evaluación (ver Tabla f1 en este capítulo).

• **Figura 1**
 • Principales hitos de la implementación anual de la Evaluación Docente



Nota: En la figura se presentan los hitos de acuerdo al orden y momento **aproximado** en que ocurren. Las fechas específicas pueden variar.

b. Producción y distribución de los instrumentos de evaluación

Además de la construcción misma de los instrumentos de evaluación, las pautas y procedimientos para su corrección, abordados en detalle en otro capítulo de este volumen⁽⁸⁾, la Evaluación Docente requiere el diseño, impresión y distribución de todos los materiales necesarios para llevar a cabo el proceso. Como parte de su rol, los Coordinadores Comunales son responsables de recibir los instrumentos y entregarlos oportunamente a los destinatarios correspondientes (docentes evaluados, evaluadores pares, directivos de escuelas y liceos).

Una preocupación fundamental en esta etapa es proporcionar a todos los participantes la información necesaria para llevar a cabo su evaluación de manera adecuada. Para ello, cada docente recibe una **Carpeta Portafolio** que contiene: un documento con información general del proceso (por ejemplo fechas claves y procedimientos a seguir para entregar su evidencia de evaluación), la *Pauta de Autoevaluación*, en la que deberá reportar la forma en que califica su propio desempeño y el *Manual Portafolio*, con todas las instrucciones necesarias para completar este instrumento⁽⁹⁾.

También los docentes reciben un *Cuestionario Complementario*, que pueden completar voluntariamente. En él se recopilan antecedentes acerca de su formación profesional, carga de trabajo, actividades de perfeccionamiento y expectativas laborales, entre otros. También se formulan preguntas acerca de su experiencia al evaluarse (por ej. dificultades enfrentadas, beneficios obtenidos y opinión sobre los

8 Ver Flotts & Abarzúa.

9 En forma complementaria, los docentes deben descargar del sitio web de la Evaluación Docente (www.docentemas.cl) un archivo electrónico con las Hojas de Respuesta, en las que deben consignar la información solicitada para el Portafolio.

instrumentos). Los datos de este cuestionario, que es respondido y entregado por la gran mayoría de los docentes (99%), son un insumo importante para retroalimentar el proceso de evaluación, además de permitir el desarrollo de investigación⁽¹⁰⁾ y un análisis más completo de los resultados arrojados por los instrumentos.



En la Carpeta Portafolio cada docente recibe los materiales necesarios para completar su evaluación.

Otro importante destinatario de los materiales de evaluación son los Evaluadores Pares, quienes reciben un manual con todas las indicaciones para llevar a cabo la entrevista y las rúbricas para calificar las respuestas recogidas.

Finalmente, los Directores y Jefes de UTP reciben, para cada docente que deben evaluar, un ejemplar del informe a completar (pauta estructurada de preguntas y rúbricas para la calificación del desempeño en cada aspecto evaluado).

El envío del material se programa en conjunto con los Coordinadores Comunales y Encargados de Evaluación, sin costo para éstos y atendiendo a variables como la ubicación geográfica, para asegurar el estricto cumplimiento de plazos y la equidad en el acceso a información y los tiempos disponibles para todos los participantes. El despacho y recepción de materiales es monitoreado a través de un sistema computacional y el equipo técnico de MIDE UC mantiene sistemas complementarios de comunicación con las comunas para garantizar un exhaustivo seguimiento de la distribución de materiales.

Cabe señalar que, junto con su distribución física, la versión electrónica de los instrumentos y materiales de evaluación se pone a disposición de los docentes y del público general en el sitio web www.docentemas.cl, desde donde pueden ser consultados y descargados durante el período en que los docentes completan su evaluación.

10 Los datos generados por la evaluación, incluyendo los de este Cuestionario, son de propiedad del Ministerio de Educación y pueden ser obtenidos por distintos investigadores siguiendo un protocolo determinado por el CPEIP y respetando las normas de reserva en el manejo de la información establecidas en el marco jurídico que regula la Evaluación Docente. En el caso del Cuestionario, además, éste incluye un formulario de consentimiento informado en que el docente puede expresar si autoriza o no que la información que ha proporcionado pueda ser utilizada con fines de investigación y estudios.

c. Aplicación de instrumentos

Los docentes reciben la Carpeta Portafolio con los materiales necesarios para su evaluación antes de iniciarse las vacaciones escolares de invierno. Para completar la Autoevaluación disponen de cinco semanas y para elaborar el Portafolio, de doce. Durante ese período, pueden realizar consultas a un Centro de Llamados especialmente habilitado para estos efectos. Además, asociado al sitio web www.docentemas.cl funciona en forma permanente un sistema de consultas por correo electrónico.

A continuación se describen en mayor detalle algunos aspectos específicos de la aplicación de los instrumentos de la evaluación⁽¹¹⁾.

Portafolio

El Portafolio se compone de dos módulos. En el Módulo 1 el docente debe dar cuenta de la implementación de una unidad pedagógica (de 8 horas de clase) y una evaluación utilizada para establecer el aprendizaje de los alumnos en ella, así como responder distintas preguntas que exigen la reflexión sobre su propio trabajo. En el caso del Módulo 2, debe presentar el video de una clase (de 40 minutos) y completar una ficha con información sobre ésta (objetivos de aprendizaje y contenidos abordados, entre otros).

Para realizar el Módulo 1, cada docente puede elegir libremente la unidad pedagógica a presentar, siempre que ésta aborde una de las dos opciones entregadas de objetivos y contenidos curriculares correspondientes al nivel, subsector o modalidad en que se evalúa.

La grabación de la clase para el Módulo 2 se lleva a cabo en fecha y hora previamente informadas al docente. La clase puede o no formar parte de la unidad presentada en el Módulo 1, pero debe abordar objetivos y contenidos del currículum correspondientes al nivel, subsector o modalidad por las que se evalúa.

La grabación de clases es realizada por camarógrafos capacitados por el equipo de MIDE UC y no tiene costo para los docentes. Su ejecución implica una compleja operación logística que se extiende por 11 semanas y cubre todo el país, incluyendo grandes ciudades, pueblos pequeños, islas y localidades fronterizas. Para ello, el equipo de MIDE UC selecciona a una red de productores especializados, responsables de planificar, conducir y supervisar el trabajo de 400 camarógrafos a nivel nacional. Los productores son capacitados y se les proporciona materiales impresos y audiovisuales para entrenar a sus equipos de camarógrafos en protocolos formales y técnicos indispensables para asegurar que las grabaciones se realicen bajo condiciones estandarizadas que permitan su corrección.

Los productores encargados de las grabaciones coordinan con el Director y/o Jefe de UTP de cada establecimiento el día y hora de filmación de cada docente. Así se asegura la factibilidad de la grabación, evitando que sea afectada por factores exógenos como la realización de actividades extraordinarias, como actos o celebraciones dentro del establecimiento.

La realización de las grabaciones es monitoreada permanentemente por el equipo de MIDE UC a través de los datos que los encargados de grabaciones ingresan a un sistema computacional y del contacto con los Coordinadores Comunales y Encargados de Evaluación. Esto permite identificar y resolver prontamente los casos en que se presentan dificultades y reprogramar grabaciones cuando es necesario.

Cabe señalar que, como parte de los protocolos de evaluación, al concluir la grabación de su clase, cada docente debe verificar los primeros y últimos minutos de la cinta junto con el camarógrafo, para

.....

11 Para mayor información sobre las características de cada instrumento y el proceso de su construcción, ver Flotts & Abarzúa, en este volumen.

asegurarse de que ésta se vea y escuche adecuadamente. Solo entonces se cierra el proceso de filmación, el camarógrafo deposita la cinta en un sobre que cierra con un sello de seguridad y lo entrega al docente, quien debe entregarlo al Coordinador Comunal, junto con el resto de su evidencia de evaluación, al cumplirse el plazo de entrega del Portafolio.



Como parte de su portafolio, cada docente debe presentar la filmación de una clase de 40 minutos.

Entrevista por un Evaluador Par

Este instrumento consiste en una pauta de entrevista estructurada aplicada por un Evaluador Par, quien toma nota de las respuestas del docente, las analiza aplicando una rúbrica (pauta de corrección), también proporcionada por la Evaluación Docente, y les asigna un nivel de desempeño (Destacado, Competente, Básico o Insatisfactorio).

En cada período de evaluación, aproximadamente 1.200 docentes desempeñan este rol, por el cual reciben una remuneración⁽¹²⁾. El Reglamento sobre Evaluación Docente establece que los Evaluadores Pares deben cumplir los siguientes requisitos:

- desempeñarse en al menos un establecimiento educacional municipal;
- ser docente de aula en ejercicio con al menos cinco años de experiencia en el sistema escolar formal;
- desempeñarse en el mismo nivel, sector curricular y modalidad del colega que evalúa;
- no haber sido sancionado en un sumario administrativo en los últimos cuatro años;
- tener evaluación docente con un resultado final Destacado o Competente, o ser docente con derecho a percibir la Asignación de Excelencia Pedagógica o integrante de la Red Maestros de Maestros; y
- haber cumplido y aprobado el programa de capacitación para Evaluadores Pares.

12 La Ley 19.961 fija un monto a transferir a cada municipio por docente evaluado en el período, recursos que deben ser íntegramente empleados para la remuneración de los Evaluadores Pares.

Aquellos docentes que cumplen con estos requisitos y tienen interés en desempeñarse como Evaluadores Pares en un período dado, deben realizar un proceso de postulación conducido por el CPEIP. Éste revisa los antecedentes de los postulantes y lleva a cabo una pre-selección atendiendo a los criterios señalados por el Reglamento, así como a otros antecedentes relevantes, tales como haber obtenido la Asignación Variable por Desempeño Individual.

Los docentes pre-seleccionados deben asistir a una capacitación intensiva de dos días —a cargo del equipo técnico de MIDE UC— para aprender a aplicar tanto la entrevista como la pauta para su corrección. Esto es crucial para asegurar que las entrevistas se apliquen de manera estandarizada y todos los docentes sean evaluados bajo las mismas condiciones.

Cada año, durante un mes, se llevan a cabo alrededor de 70 seminarios de capacitación a lo largo de todo el país, en los que participa un total aproximado de 1.300 docentes. Una vez concluidos los seminarios, se define la nómina de docentes que podrán ejercer el rol de Evaluador Par en el período, excluyéndose una porción menor de participantes (alrededor del 10%) que han demostrado dificultades serias ya sea para la realización adecuada de la entrevista o, más comúnmente, para la correcta aplicación de las rúbricas de evaluación.

Las entrevistas se llevan a cabo durante dos meses y en promedio cada Evaluador Par realiza doce. La asignación de los docentes a entrevistar es realizada por el Coordinador Comunal y debe atender a diversas consideraciones como la cantidad de docentes inscritos en la comuna, los niveles y modalidades por los que se evalúan y las inhabilidades e incompatibilidades establecidas por el Reglamento sobre Evaluación Docente⁽¹³⁾. Por lo mismo, es frecuente que los docentes sean entrevistados por pares de comunas vecinas, pues en algunos casos su comuna cuenta con un número reducido de establecimientos educacionales, lo que dificulta la selección y/o asignación de Evaluadores Pares no afectados a condiciones de inhabilidad.

Una vez realizada la asignación, el Evaluador Par debe contactarse con cada docente para concordar la fecha y hora de la entrevista. Ésta normalmente dura entre 50 y 60 minutos y es realizada fuera de la jornada de clases para no alterar el trabajo normal de ambos profesores en el aula.

Hasta el año 2008 los resultados de las entrevistas eran digitados por el equipo de MIDE UC a partir de las entrevistas entregadas por los Evaluadores Pares. Desde el año 2009, en cambio, son los propios evaluadores quienes ingresan los resultados de cada docente entrevistado en un sistema computacional en plataforma Internet. Además de agilizar el registro de los datos, esto permite que cada Evaluador Par sea plenamente responsable de las evaluaciones que ha realizado.

Informes de Referencia de Terceros

Los Directores y Jefes de UTP deben emitir un Informe para cada docente evaluado en su establecimiento educacional⁽¹⁴⁾. Este informe consiste en una pauta estructurada de preguntas o indicadores, basados en el Marco para la Buena Enseñanza, frente a los cuales el evaluador debe señalar el nivel de desempeño (Destacado, Competente, Básico o Insatisfactorio) que refleja la práctica habitual del profesor.

Anualmente, el CPEIP define un período de cuatro semanas para que los Directores y Jefes de UTP completen los informes, cuyas pautas impresas reciben del Coordinador Comunal. Una versión electrónica de esta pauta es publicada con tres meses de antelación en el sitio web www.docentemas.cl. Esto permite

.....

13 El Reglamento establece que es *incompatible* con el rol de Evaluador Par: cumplir funciones directivas en un establecimiento educacional y realizar funciones de coordinación del sistema de evaluación durante el mismo período de evaluación. También establece como condiciones de *inhabilidad*: evaluar a un docente del mismo establecimiento y evaluar a un pariente político o biológico (hasta el tercer grado de consanguinidad y segundo grado de afinidad).

14 En aquellos establecimientos en que no existe ninguno de dichos cargos, este informe es emitido por el superior jerárquico del profesor evaluado y por el Jefe Técnico Pedagógico comunal cuando existe.

que los directivos conozcan anticipadamente los aspectos que deberán evaluar en los docentes y así puedan generar mejores estrategias y acciones de recolección de la información necesaria para emitir un juicio fundado sobre el desempeño de cada docente.

En la misma línea, a partir del año 2009 se incorporó en el sitio web una sección específicamente orientada a los Directores y Jefes de UTP, con sugerencias de fuentes y métodos de recolección de información útiles para evaluar el desempeño de los docentes. Asimismo, a través de actividades de capacitación para los Coordinadores Comunales y también para los propios directivos a partir de 2010, se ha enfatizado crecientemente el rol de éstos tanto en la evaluación misma de los docentes, como en la gestión orientada a analizar y utilizar la información generada por el proceso para apoyar un mejor desempeño en los profesores a su cargo. Debe mencionarse que a esto se suman iniciativas locales que las comunas desarrollan en forma autónoma, generando, por ejemplo, instancias de reflexión entre directivos para fomentar la realización de informes de evaluación más completos y fidedignos.

Potenciando la evaluación de directivos: la experiencia de Talca.

Por iniciativa local en la ciudad de Talca, capital de la Región del Maule, a partir del 2005 se realizan talleres con Directores y Jefes de UTP para abordar la elaboración de sus informes de evaluación. Equipos directivos de distintos establecimientos, entre los casi 50 con que cuenta la comuna, unifican criterios y trabajan en conjunto para realizar evaluaciones objetivas y fundadas.

Los talleres abordan distintos temas a lo largo del año: en abril se revisa el Reglamento sobre Evaluación Docente; en mayo–junio trabajan analizando los descriptores del Marco para la Buena Enseñanza y su expresión en la práctica docente, y concuerdan en base a ello criterios para observar clases de los docentes y luego retroalimentarlos. Posteriormente, en octubre, antes del período de completación de los informes de directivos, las autoridades locales organizan una jornada de análisis de las dimensiones e indicadores de evaluación que contiene la pauta del informe. Los participantes fijan criterios comunes para la interpretación de esos indicadores y su aplicación a la hora de calificar el desempeño de los docentes a su cargo.

Los responsables de esta iniciativa destacan la motivación de los directivos por participar en esta instancia de trabajo y su buena respuesta ante esta forma de aproximación a la evaluación mediante el trabajo en equipo.

d. Procesamiento de datos

Después de recibir los instrumentos completados por los docentes, Evaluadores Pares y directivos, los Coordinadores Comunales deben enviarlos al equipo de MIDE UC, proceso que se coordina y monitorea según lo ya descrito para la entrega inicial de materiales en este mismo capítulo.

Una vez recibido en MIDE UC, cada instrumento es individualizado mediante la asignación de un código de barras que se registra en el sistema computacional y lo vincula a un docente en particular. Esto permite verificar la recepción de todos los instrumentos correspondientes a cada docente y hacer un riguroso seguimiento de ellos hasta que se completa su digitación y/o corrección.

Los datos de Autoevaluaciones e Informes de directivos son digitados directamente en un sistema computacional especialmente diseñado para la Evaluación Docente, en tanto que aquellos de las Entrevistas por Evaluadores Pares, tal como se ha indicado antes, son ingresados directamente por éstos

(en el mismo sistema). Con el objeto de garantizar la total fiabilidad en el registro de los datos, el sistema computacional incluye diversas formas de validación para evitar que errores accidentales puedan afectar los datos registrados y, como consecuencia de ello, el resultado de la evaluación. Además, todos los instrumentos digitados son doblemente ingresados por dos digitadores independientes.

Por su parte, la digitación de los datos de evaluación de los portafolios es antecedida por un complejo proceso de corrección, que se describe a continuación.

e. Corrección de Portafolios

La evaluación de Portafolios se realiza en Centros de Corrección, constituidos en distintas universidades del país y cuya selección requiere la aprobación del Ministerio de Educación a través del CPEIP. Entre los años 2003 y 2010, siete universidades han desempeñado este rol en al menos una oportunidad:

• **Tabla 2**
• *Centros de Corrección de Portafolios 2003–2010*

Centros de corrección	2003	2004	2005	2006	2007	2008	2009	2010
Pontificia Universidad Católica de Chile (*)	•	•	•	•	•	•	•	•
Pontificia Universidad Católica de Valparaíso	•		•	•	•	•	•	•
Universidad de Concepción	•		•	•	•	•	•	•
Universidad de Chile	•	•	•					
Universidad de La Serena			•	•	•	•	•	•
Universidad de Santiago de Chile						•	•	•
Universidad de Talca						•		

(*) Este centro de corrección no es implementado por MIDE UC, sino por la Facultad de Educación de la universidad.

El equipo de MIDE UC es responsable del diseño y planificación del proceso de corrección, así como de la supervisión (presencial y a distancia) de cada Centro de Corrección. Además, se encarga de definir la distribución de los portafolios a los centros, considerando el nivel, subsector y modalidad que será revisado por cada uno de ellos.

La marcha de cada Centro de Corrección requiere la implementación de la infraestructura y el equipamiento necesario así como la selección del personal que desempeñará la tarea, incluyendo especialmente los correctores y supervisores de corrección.

Las normas que rigen la Evaluación Docente exigen que los portafolios sean corregidos por docentes de aula, que son especialmente seleccionados y capacitados para llevar a cabo esta tarea. Tanto los correctores como sus supervisores deben contar con formación y experiencia en el mismo nivel, subsector y modalidad de los docentes cuya evidencia corrigen. Así por ejemplo, los portafolios del subsector Matemática de Educación Media solo son corregidos por docentes de Educación Media que ejercen en ese subsector, los de Educación Parvularia solo son evaluados por educadoras de párvulos, etc.

Aunque la preparación de cada Centro de Corrección se inicia varias semanas antes, la corrección misma de los portafolios se lleva a cabo fundamentalmente en el mes de enero (período de vacaciones escolares de verano) y es una actividad remunerada para los docentes participantes. Dependiendo del número de

docentes en evaluación (y, por tanto, de portafolios a corregir), en cada año se requiere la participación de más de 300 profesores (ver Tabla 3).

• **Tabla 3**
• *Dotación de Correctores y Supervisores de Corrección por Año*

Período	Portafolios corregidos	Centros de Corrección	Correctores	Supervisores
2003	3.741	4	60	8
2004	1.721	2	43	7
2005	10.667	5	226	30
2006	14.230	4	325	43
2007	10.431	5	268	39
2008	16.029	6	375	64
2009	15.712	5	423	86
2010	11.070	5	347	63

Una serie de procedimientos buscan garantizar que la corrección de Portafolios de todos los docentes evaluados se desarrolle cumpliendo estrictos criterios y estándares de calidad. Aunque esos procedimientos son abordados en detalle en otro capítulo de este volumen⁽¹⁵⁾, brevemente se pueden mencionar:

- **marcha blanca:** consiste en un período de funcionamiento de todo el centro en “modo de prueba”, es decir, si bien se siguen todos los procedimientos regulares de corrección y control, al concluir la etapa, se eliminan todos los puntajes generados en ella. Posteriormente esos portafolios son nuevamente asignados para ser revisados al iniciarse la corrección definitiva.
- **corrección grupal:** cada grupo de correctores (es decir, uno o más conjuntos de correctores que evalúan portafolios de docentes de una modalidad, nivel y subsector), revisan y discuten la evaluación de un mismo módulo de Portafolio. Los puntajes asignados por cada evaluador son contrastados con la corrección de referencia (entregada por el equipo de MIDE UC), lo que permite aclarar dudas y reforzar las indicaciones para la correcta aplicación de las rúbricas.
- **doble corrección:** un porcentaje de los portafolios es corregido por dos docentes de manera independiente y “ciega” (cada uno desconoce las puntuaciones asignadas por el otro). Ambas correcciones son cotejadas y, cuando no alcanzan un nivel de coincidencia previamente establecido, el portafolio debe ser evaluado por un supervisor, cuya corrección permite dilucidar aquella corrección que se ha “descalibrado”.
- **recalibración:** procedimiento individual realizado con cada corrector respecto del cual se detectan (a través de la doble corrección) imprecisiones en la aplicación de la rúbrica. El procedimiento es realizado por un supervisor, tras haber revisado la evidencia corregida por el evaluador.

Como se mencionara antes, tanto los aspectos técnicos como operativos de la corrección siguen estrictos protocolos fijados por el equipo de MIDE UC. Éstos son transferidos a cada centro a través de actividades de capacitación, reuniones informativas, manuales e instructivos, entre otros, y su propósito es resguardar la igualdad de condiciones para la corrección de los portafolios de todos los docentes. De igual forma, se definen requerimientos en cuanto a las instalaciones y el equipamiento necesario para la operación de un centro de corrección, así como medidas de seguridad para garantizar el resguardo del material.

.....
15 Ver Flotts & Abarzúa.

f. Toma de decisiones por las Comisiones Comunales de Evaluación

El marco regulatorio de la Evaluación Docente deposita en el nivel local la atribución de definir el nivel de desempeño final que obtiene cada profesor evaluado: el Reglamento sobre Evaluación Docente establece que, en cada comuna, esa decisión es responsabilidad de la Comisión Comunal de Evaluación, entidad colegiada integrada por el Coordinador Comunal y por aquellos docentes que se han desempeñado como Evaluadores Pares en ella, durante el período¹⁶.

Para cumplir su tarea, por medio de un sistema computacional las comisiones acceden a un informe con los resultados obtenidos por cada docente de la comuna en los distintos instrumentos y su nivel de desempeño global en base a la ponderación de éstos. Además, las comisiones toman conocimiento de la información sobre el contexto en que se desempeña el profesor evaluado, que es consignada por el propio docente en su Autoevaluación, por el Director y Jefe de UTP en los Informes de Referencia de Terceros, y por el Evaluador Par en la entrevista.

La inclusión de la información de contexto junto con los resultados de los instrumentos, hacen posible que el sistema de evaluación sea sensible a la complejidad de los procesos de enseñanza–aprendizaje y a la variedad de contextos en que éstos se desarrollan.

En este marco y considerando las facultades que le confiere el Reglamento sobre Evaluación Docente, cada comisión puede *ratificar*, o bien, *modificar* por acuerdo fundado, el nivel de desempeño global de cada docente arrojado por los puntajes ponderados de los instrumentos. Para *ratificar* el nivel de desempeño se requiere la mayoría simple de los Evaluadores Pares presentes en la respectiva sesión de la comisión; en cambio, para *modificarlo* el Reglamento exige que tal decisión sea aprobada por al menos dos tercios de los Evaluadores Pares y obedezca a razones fundadas que deben quedar consignadas en el informe que recibe el docente evaluado. Este informe también se genera por medio de un sistema computacional que reúne tanto los puntajes de los instrumentos como la decisión de la Comisión Comunal, que ésta ingresa a través de una plataforma en Internet. Junto con la decisión misma, la Comisión debe consignar las razones que la fundamentan. Asimismo, ya sea que haya modificado o no su resultado final, la Comisión puede incorporar en el informe otras observaciones y sugerencias dirigidas al docente, por ejemplo, para resaltar aquellos aspectos en que su desempeño fue mejor evaluado y felicitarlo, formular recomendaciones para mejorar las debilidades identificadas en su práctica o sugerir su participación en instancias comunales de perfeccionamiento, entre otros. En la Figura 2 se presenta un ejemplo real con la información consignada por una Comisión Comunal de Evaluación de la Región del Maule al modificar el resultado de un docente de Básico a Competente.

● **Figura 2**
● *Observación registrada por Comisión Comunal de Evaluación*
●

Observaciones de la Comisión Comunal de Evaluación:

La Comisión Comunal de Evaluación de acuerdo al Reporte de Resultados le reconoce y valora a la docente la capacidad de análisis de las actividades de las clases y generar un ambiente propicio para el aprendizaje. Según las Consideraciones Contextuales se valora su capacidad creativa para realizar las clases, su gran disposición al cambio y al trabajo en equipo. Genera instancias de participación motivadoras para sus alumnos y comunidad educativa. Se le recomienda que se perfeccione en el área de evaluación e interacción pedagógica.

.....
16 El Reglamento sobre Evaluación Docente establece que los Evaluadores Pares tienen derecho a voz y voto para pronunciarse sobre el resultado final de los docentes, mientras que los Coordinadores Comunales solo tienen el primero.

Cabe mencionar que la mayoría de las decisiones tomadas por las comisiones ratifican el nivel de desempeño de los docentes y solo alrededor del 5% de los evaluados ve modificado su resultado final. En las dos terceras partes de esos casos, la decisión implica subir el resultado al nivel inmediatamente superior (por ejemplo, de Básico a Competente).

Las Comisiones de todo el país deben sesionar durante una semana (justo antes del inicio del año escolar) y completar en ese período la toma de decisiones para todos los docentes evaluados en cada comuna. El equipo de MIDE UC monitorea el avance del proceso para garantizar la generación oportuna de todos los informes de resultados destinados a los docentes evaluados.

Posteriormente, las comisiones deben sesionar nuevamente si les corresponde resolver los **recursos de reposición** interpuestos por los docentes de la comuna disconformes con el resultado de su evaluación. El Reglamento sobre Evaluación Docente define las causales que pueden dar origen a estos recursos y señala que, para resolverlos, las comisiones deben contar con un informe técnico emitido por el CPEIP.

Para apoyar el trabajo de las Comisiones Comunales de Evaluación, el equipo de MIDE UC es responsable de implementar seminarios de capacitación a los que son convocados todos sus integrantes. Anualmente se llevan a cabo aproximadamente 70 seminarios, de un día de duración, en distintas ciudades del país. En ellos se abordan aspectos normativos y operativos relevantes para el proceso de toma de decisiones de la comisión, de acuerdo a los lineamientos que establece el Reglamento sobre Evaluación Docente.

g. Informes de Resultados

Dado que el principal objetivo formativo de la Evaluación Docente es apoyar el mejoramiento de las prácticas pedagógicas de los docentes, los informes de resultados cumplen un rol estratégico dentro del proceso. Bajo la premisa de que, mientras más específica y cercana a la práctica profesional sea la retroalimentación entregada, mayor será su utilidad para orientar el mejoramiento profesional de los docentes, los informes de resultados ofrecen retroalimentación especialmente detallada en cuanto a las fortalezas y debilidades identificadas en el Portafolio de cada docente.

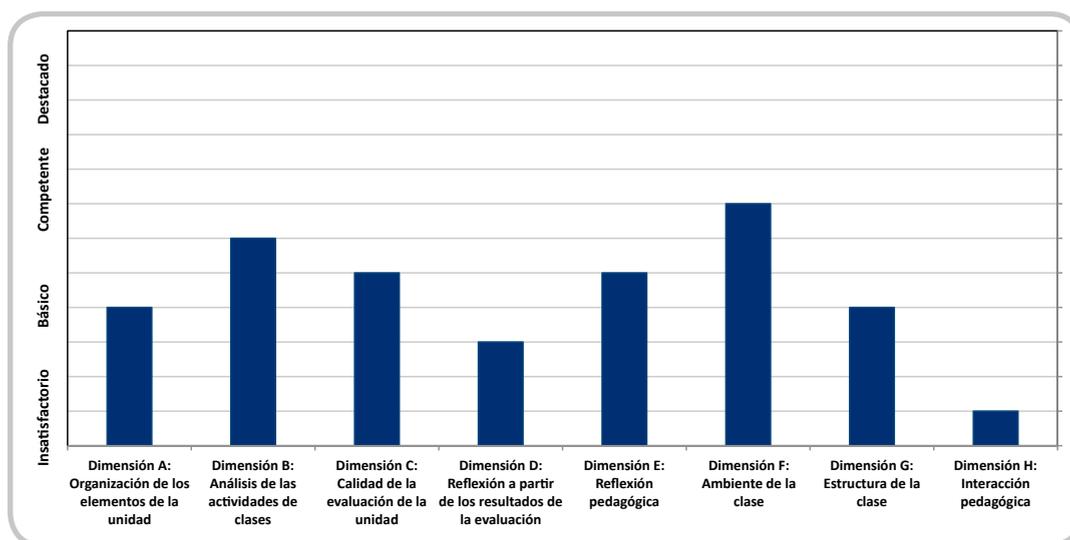
La Evaluación Docente emite informes para tres destinatarios: el propio evaluado, el equipo de gestión del establecimiento educacional y el sostenedor municipal. A continuación se describe brevemente el contenido de estos tres informes.

Informes de Evaluación Individual

Este informe es entregado por el Coordinador Comunal a cada docente evaluado y reporta su desempeño en los distintos instrumentos y su resultado final de acuerdo a la decisión tomada por la Comisión Comunal de Evaluación.

En el caso del Portafolio, se reportan los resultados obtenidos en ocho dimensiones de evaluación, cinco basadas en la corrección del Módulo 1 (evidencia escrita), y tres en el Módulo 2 (clase filmada). Para cada dimensión se especifica el nivel de desempeño alcanzado por el docente, en relación con los cuatro niveles identificados por el sistema: Insatisfactorio o Básico, si el desempeño se ha ubicado por debajo del estándar esperado, y Competente o Destacado, si lo ha alcanzado o sobresale respecto de él.

• **Figura 3**
 • *Imagen del gráfico de resultados del Portafolio en el Informe de Evaluación Individual⁽¹⁷⁾*



Además, para cada dimensión, el docente recibe información detallada en términos de: las prácticas características de un desempeño consistente con el estándar (nivel Competente), los aspectos logrados que se identificaron en su evidencia (aquellos indicadores de evaluación en que su desempeño fue calificado como Competente o Destacado) y aquellos que requieren ser desarrollados o mejorados (nivel Básico o Insatisfactorio). La Figura 4 muestra un extracto del informe individual que ilustra lo dicho.

Los resultados de la Entrevista por un Evaluador Par y los Informes de Referencia de Terceros se reportan de manera conjunta en una tabla de doble entrada que indica para cada dominio del Marco para la Buena Enseñanza, el nivel de desempeño obtenido por el docente expresado en la escala de cuatro niveles del sistema. Una segunda tabla, similar a ésta, da cuenta de los datos de la Autoevaluación.

Finalmente, el informe da a conocer al docente el nivel de desempeño resultante de la ponderación de los cuatro instrumentos, así como la decisión tomada por la Comisión Comunal de Evaluación, indicando si ésta ha ratificado o modificado el primero y definiendo así su resultado final en el proceso.

.....

17 Hasta 2009, en lugar de este gráfico se presentaba una tabla de doble entrada que indicaba el nivel de desempeño obtenido para cada una de las ocho dimensiones. A partir de la evaluación 2010 se incorporó el gráfico aquí presentado.

• **Figura 4**
• *Ejemplo de cuadro de retroalimentación de resultados de Portafolio*

Dimensión F: Ambiente de la clase

La información para evaluar esta dimensión se obtuvo del producto 4: Clase grabada.

En esta dimensión un docente competente se caracteriza por:

- lograr que sus alumnos permanezcan focalizados en las actividades, manteniendo normas de convivencia que favorecen el desarrollo de la clase,
- generar instancias equitativas de participación para los alumnos y promover la interacción entre ellos, y
- entregar instrucciones claras junto con monitorear adecuadamente el desarrollo de las actividades.

Descripción de los aspectos logrados

Logra que sus alumnos se mantengan siempre focalizados en las actividades y respeten las normas de convivencia, generando un ambiente de trabajo que favorece el aprendizaje.

Acompaña y monitorea adecuadamente las actividades realizadas por los alumnos, entregándoles instrucciones claras y precisas que orientan su trabajo.

Descripción de los aspectos a desarrollar

Si bien logra que sus alumnos participen de manera equitativa en diferentes instancias de la clase, su desempeño presenta ciertas debilidades, ya que no promueve que éstos interactúen o colaboren entre sí.

En la dimensión F, *Ambiente de la clase*, considerando los aspectos antes descritos, [nombre docente] obtiene el **nivel de desempeño Competente**.

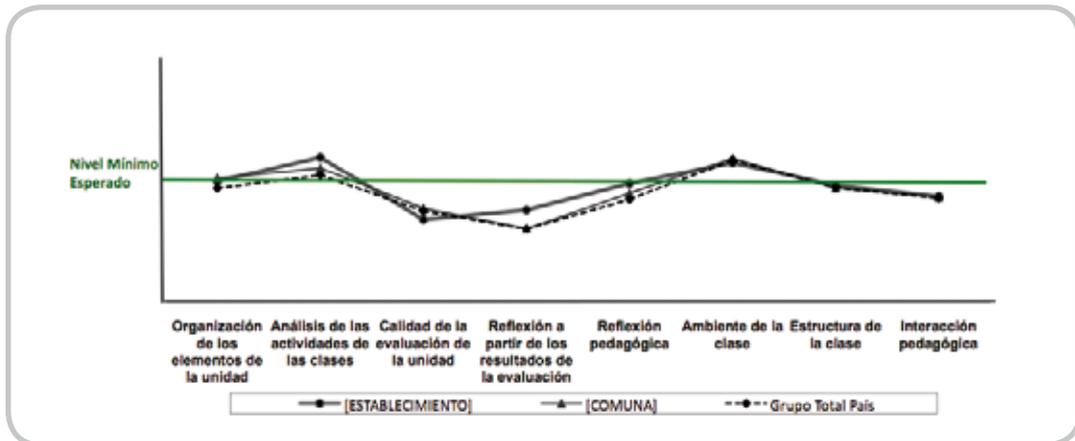
Informe de Resultados para el Equipo de Gestión del Establecimiento Educacional

En cada establecimiento, el equipo de gestión —conformado habitualmente por Director, Jefe de UTP, Orientador e Inspector, entre otros—, recibe un informe que presenta los resultados alcanzados por el conjunto de docentes evaluados de su dotación. El informe señala el desempeño final alcanzado por cada docente, describe en qué consiste cada dimensión evaluada, las características de un desempeño Competente y el promedio obtenido por el grupo de evaluados del establecimiento, para cada dimensión.

Un gráfico permite comparar los promedios obtenidos por los docentes del establecimiento⁽¹⁸⁾ con aquellos de sus pares dentro de la comuna, y los del grupo total evaluado en el país (ver Figura 5).

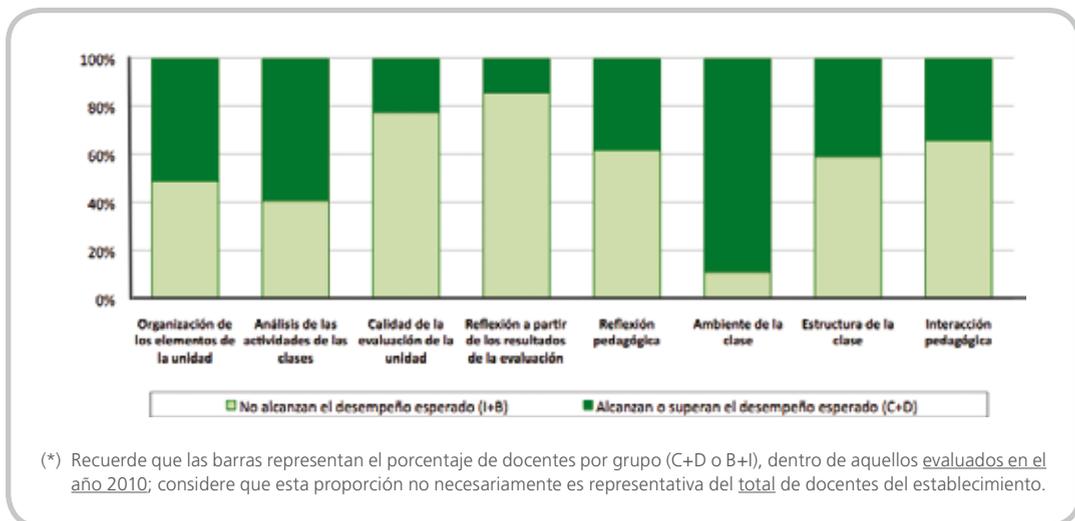
.....
18 Esta información solo se entrega si la cantidad de docentes es igual o superior a tres; de otra forma, se omite esa línea del gráfico, y lo mismo se aplica para el dato de la comuna.

- **Figura 5**
- *Imagen del gráfico comparativo de promedios por dimensión del Portafolio en Establecimiento, Comuna y Grupo Total Nacional*



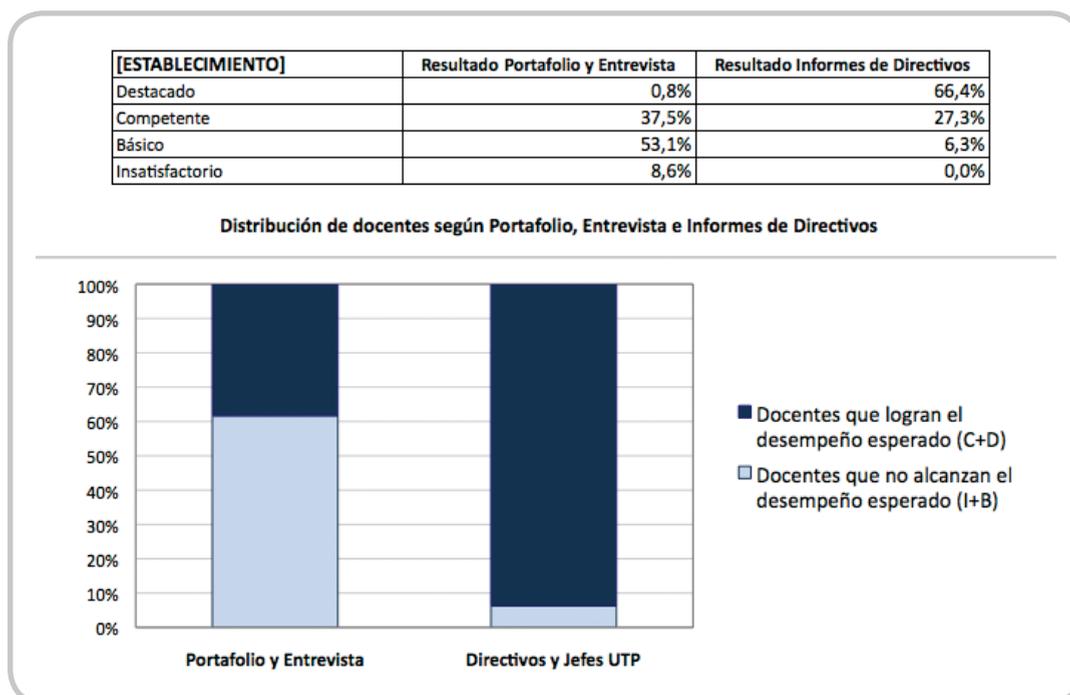
A partir del año 2009, el informe incluye además un gráfico que muestra la distribución de docentes en dos grupos de desempeño: aquellos que cumplen el estándar (obtienen un nivel de desempeño global Competente y Destacado) y aquellos que no lo logran (resultado final Básico o Insatisfactorio), en cada dimensión del Portafolio:

- **Figura 6**
- *Imagen del gráfico de distribución de docentes según si logran o no el estándar esperado—por dimensión del Portafolio*



Un gráfico similar describe la distribución de docentes en dos grupos de desempeño en base a los datos de distintos instrumentos: por un lado el Portafolio y la Entrevista por un Evaluador Par (combinados), y por otro los informes de directivos. La inclusión de este gráfico permite a los Directores y Jefes de UTP visualizar la forma en que su evaluación describe el desempeño de los docentes en relación con el estándar, y complementarla con lo equivalente según los resultados asignados por los Evaluadores Pares en la entrevista y por los correctores del Portafolio.

- **Figura 7**
- *Tabla y gráfico con distribución de docentes según datos de Portafolio y Entrevista (combinados) en Informes de Directivos*



El reporte proporcionado al equipo de gestión de cada establecimiento busca servir de insumo para la toma de decisiones al interior de éste, suministrando información que contribuye a orientar la gestión directiva para implementar iniciativas de apoyo, reconocimiento y desarrollo profesional de sus docentes, entre otras.

Informe de Resultados para el Sostenedor Municipal

Las máximas autoridades de la administración de la educación a nivel comunal (Alcaldes y Coordinadores Comunales) también reciben un informe con los resultados de los docentes de su dotación evaluados en el período. El reporte incluye tablas y gráficos similares a aquellos del reporte para el establecimiento, en la forma de una ficha resumen por cada uno de éstos. Además, presenta algunos análisis adicionales para el conjunto de docentes evaluados en la comuna. Así por ejemplo, se reporta el número y proporción de aquellos que concluyen normalmente su evaluación, la suspenden, son eximidos del proceso y aquellos que, debiendo evaluarse, no lo hicieron. Para los mismos parámetros se proporcionan también los datos del grupo total evaluado en el país.

Entre las funciones de estos informes está la de orientar el diseño de los Planes de Superación Profesional dirigidos a los docentes con resultados Básico e Insatisfactorio¹⁹. Además, en distintas comunas las autoridades locales desarrollan iniciativas propias que, a partir de los resultados de la evaluación, buscan reconocer el mérito de los docentes con un alto desempeño: algunas implementan actos y ceremonias de premiación, otras entregan cartas o diplomas a los profesores bien evaluados o los destacan públicamente a través de entrevistas en medios de prensa local.

.....

19 Para más detalles sobre los Planes de Superación Profesional, ver Cortés & Lagos, en este volumen.

Los Ángeles reconoce a sus docentes con resultado Destacado

A partir del 2007 en Los Ángeles, capital de la provincia de Bío-Bío, como parte de una política comunal se reconoce a todos los docentes que obtienen resultado Destacado. Estos son invitados a un desayuno en el que comparten con el Alcalde y las autoridades educacionales de la comuna, quienes a través de una ceremonia de premiación les entregan un incentivo.

La experiencia ha tenido un impacto altamente positivo para esos profesores, que se manifiestan agradecidos por este reconocimiento público y se motivan para desempeñar un papel especialmente activo en la comunidad educativa local. Además, cuando se realizan concursos públicos para cargos dependientes del DAEM estos docentes tienen prioridad al momento de postular, como otra señal positiva de las autoridades locales ante el buen desempeño demostrado. La Encargada de Evaluación de la comuna menciona de manera entusiasta que “ambas acciones se desarrollan como un reconocimiento que la comunidad local brinda a sus buenos docentes y como un estímulo para aquellos que sienten deseos de superarse”.

Los datos arrojados por la Evaluación Docente también propenden al mejor desarrollo de iniciativas locales para apoyar el mejoramiento del desempeño docente, ya sea a través de sistemas de socialización de buenas prácticas pedagógicas o la coordinación de sistemas de apoyo y tutoría por parte de docentes que han obtenido un mejor resultado hacia sus pares que han evidenciado más debilidades. Por esta vía la Evaluación Docente contribuye a generar círculos virtuosos de colaboración y aprendizaje mutuo al interior de una comuna.

Generando círculos virtuosos en Pucón

En Pucón, localidad de la Región de la Araucanía, el profesor E.C. —que ejerce en una de las escuelas rurales de la comuna— se evaluó por primera vez en el período 2005, obteniendo un nivel de desempeño Básico. A pesar de su reticencia inicial, insatisfecho con su resultado y decidido a mejorarlo, el docente decidió participar en los Planes de Superación Profesional dispuestos por la comuna, los que estuvieron a cargo de dos docentes de la misma localidad que habían alcanzado un resultado Destacado en la evaluación. Estas profesoras, que actuaron como tutoras de sus pares, habían obtenido también la Asignación Variable por Desempeño Individual (AVDI), tras demostrar el dominio de su disciplina en una prueba de conocimientos.

Los Planes de Superación se orientaron especialmente hacia aquellas temáticas que mostraron más bajos resultados en la Evaluación Docente, como la interacción pedagógica y la evaluación del aprendizaje. El docente aprovechó esta instancia para reconocer las falencias en su práctica y se comprometió activamente para superarlas.

Cuatro años después, cuando le correspondió evaluarse nuevamente, el profesor E.C. mejoró sustantivamente su resultado, obteniendo un nivel Competente. Además, de manera voluntaria y simultánea al proceso de Evaluación Docente, postuló a la Asignación de Excelencia Pedagógica y resultó acreditado, accediendo así a un incentivo económico por hasta 10 años.

El proceso de Evaluación Docente, la motivación y compromiso personal del profesor y la eficaz gestión de las autoridades comunales se conjugaron en este caso para producir una combinación virtuosa en beneficio del desarrollo profesional y la calidad del desempeño docente en la comuna.

Desde el punto de vista de la planificación, los informes entregados por la Evaluación Docente a los sostenedores municipales también se han convertido en un referente útil para la definición de metas y planes de acción al interior de cada comuna, como parte del Plan Anual de Desarrollo Municipal o PADEM. Éste es un instrumento de planificación exigido a cada Municipio, que actúa como un elemento ordenador de las prioridades y distribución de recursos en la educación local, y sirve de referente directo para la rendición de cuentas de la gestión municipal. Con frecuencia, los Coordinadores Comunales reportan que los informes de resultados que reciben de la Evaluación Docente los orientan para definir actividades a realizar y formular indicadores de gestión, por ejemplo, en términos de la proporción de docentes que resulta evaluado Competente y Destacado o que mejora su resultado tras haber recibido apoyo, entre otros.

Mejorando prácticas pedagógicas a partir de los resultados de la Evaluación: La experiencia de Mejillones.

En la localidad costera de Mejillones (Región de Antofagasta) desde 2006, las autoridades locales apoyan a los profesores que presentan debilidades, para mejorar la calidad de su desempeño y su resultado en la Evaluación Docente.

Año a año los equipos directivos de los establecimientos, en conjunto con el Departamento de Educación reflexionan acerca de los aspectos que deben mejorar sus docentes con desempeño Básico. Así, deciden áreas de capacitación y actualización, las que son incorporadas dentro de su Plan Anual de Desarrollo Municipal (PADEM). También, por iniciativa local, implementan un sistema de "acompañamiento en aula", empleando una pauta consensuada entre los directivos y docentes participantes, con el objeto de mejorar las prácticas pedagógicas a través del asesoramiento directo al trabajo en la sala de clases.

Dado el éxito que han tenido estas instancias, la comuna decidió llevar a cabo una jornada de capacitación con todos sus docentes. En palabras del Encargado de Evaluación, esta iniciativa comunal tuvo el propósito "de proveer a nuestros profesionales de la educación de instrumentos y herramientas que faciliten el proceso de planificación en las escuelas, así como los contenidos específicos y disciplinarios que permitan abordar todos los aprendizajes esperados del currículum".

Sistemas de difusión y apoyo

Para desempeñar correctamente sus funciones, la Evaluación Docente cuenta con distintos medios que facilitan la difusión de información oficial y el apoyo a los profesores que se evalúan. Entre ellos, destacan el sitio web www.docentemas.cl y el *Centro de Llamados Docentemás*.

a. Sitio web

El sitio web www.docentemas.cl proporciona información oficial sobre el proceso de Evaluación Docente para los profesores que participan en éste o lo harán en el futuro, para los directivos, investigadores y otros interesados en conocer el sistema de evaluación. Para ello, entrega información sobre los criterios e instrumentos con los que se evalúa a los docentes, la calendarización de las actividades del período y datos para contactar a cada Coordinador Comunal, entre otros.

Una sección de *Preguntas Frecuentes* entrega respuesta a consultas recurrentes y otra de *Documentos Descargables* permite obtener en formato digital los instrumentos de evaluación y documentos relevantes, tales como el Marco para la Buena Enseñanza. Además, se proporciona información de interés para el trabajo pedagógico de los docentes: se abordan, entre otros, errores conceptuales frecuentes detectados en los portafolios y se desarrollan ejemplos concretos que vinculan los criterios y descriptores del Marco para la Buena Enseñanza con las prácticas en aula.

Junto con ello, existe un servicio permanente de consultas a través de correo electrónico. Éste se orienta principalmente a los docentes evaluados, pero recibe también consultas de otros actores del sistema escolar vinculados a la Evaluación Docente, tales como Directores, Jefes de UTP y Evaluadores Pares, entre otros.

Anualmente, el sitio web recibe más de 380.000 visitas, concentrándose éstas en momentos claves del proceso como la inscripción de docentes (abril–mayo), la entrega de la Autoevaluación (agosto) y la del Portafolio (octubre). El reporte recogido cada año de los propios evaluados a través del Cuestionario Complementario ha permitido constatar repetidamente que se trata de un medio de información de alto uso por parte de ellos; a modo de ejemplo, el 94% de las educadoras y profesores evaluados en el período 2010 señaló haber visitado el sitio durante el transcurso de la evaluación.

b. Centro de Llamados Docentemás

El Centro de Llamados es un servicio destinado especialmente a responder las consultas de los docentes durante el período de elaboración de la Autoevaluación y el Portafolio. Este servicio funciona durante tres meses y medio, desde la semana inmediatamente anterior al inicio de la elaboración de estas evidencias, hasta la semana siguiente a la fecha de entrega del Portafolio.

Un grupo de profesionales especialmente seleccionados y capacitados por el equipo de MIDE UC está a cargo de dar respuesta a las consultas planteadas principalmente por los profesores en evaluación, quienes constituyen cada año la fuente del 85% a 90% de las consultas recibidas.

Además de prestar un servicio de información para la población evaluada, el Centro de Llamados es una importante fuente de retroalimentación para el propio sistema de evaluación. Todas las consultas recibidas en él son registradas y posteriormente analizadas, para nutrir el diseño del proceso en el período siguiente. Así por ejemplo, las dudas frecuentes se traducen en cambios o ajustes en el Manual Portafolio, las capacitaciones a Coordinadores Comunes y el sitio web, entre otros.

Apoyo tecnológico

El desarrollo de la Evaluación Docente supone una compleja coordinación de múltiples actores, ubicados en distintos lugares del país y responsables de funciones específicas y relevantes para el resultado del proceso. Para monitorear la marcha de éste y resolver prontamente las dificultades surgidas es vital disponer de información precisa y actualizada en cada paso. Desde la inscripción de los docentes que deben evaluarse hasta la elaboración y emisión de los informes de resultados, el flujo de información es clave para garantizar una evaluación equitativa, válida y confiable. Para ello, el equipo de MIDE UC ha desarrollado un sistema computacional de apoyo, accesible a través de internet, que da respuesta a las características y necesidades del proceso.

El sistema computacional proporciona una plataforma que gestiona los datos generados por el proceso, garantizando su precisión e integridad. Además, por el carácter crítico y confidencial que reviste gran parte de la información, cumple con exigentes condiciones de seguridad, incluyendo el acceso restringido mediante contraseñas personales y atribuciones diferenciadas según perfiles de usuario. La cualidad relacional de las bases de datos y la automatización de reglas de validación de éstos permite prevenir y detectar errores involuntarios (por ejemplo, al digitarse los instrumentos), que afectarían el resultado de los docentes. Estos dispositivos son, además, reforzados a través de otros procedimientos de redundancia y verificación, tales como la doble digitación de toda la información crítica.

La tecnología de apoyo desarrollada permite acceder fácilmente a información actualizada en tiempo real sobre distintas variables del proceso, tales como la cantidad de docentes inscritos en cada comuna o la ubicación de su material en un momento específico. También se registran en el sistema: la recepción de materiales por parte de los Coordinadores Comunales, la fecha y hora programadas para la filmación de cada docente y los docentes asignados para entrevista a cada Evaluador Par, entre otros. Asimismo, el uso de este sistema computacional optimiza el control de procesos y resultados asociados a la corrección de instrumentos, proporcionando, por ejemplo, datos sobre la productividad diaria de los Centros de Corrección del Portafolio, los puntajes asignados para cada docente en los distintos indicadores evaluados y el grado de consistencia detectado en las dobles correcciones.

Otra función clave del sistema computacional es la generación de los distintos informes de resultados que proporciona la evaluación. Por ejemplo, durante el período 2008 (aquel que ha concentrado la mayor cantidad de evaluados hasta la fecha) se generaron 16.019 reportes para el trabajo de las Comisiones Comunales, un número igual de Informes de Evaluación Individual para los docentes, 3.679 reportes para los equipos de gestión de los establecimientos y 341 para los sostenedores municipales.

Mejoramiento continuo

Parte integral de la implementación del sistema de evaluación son los mecanismos de recolección de información para promover su mejoramiento continuo, tanto en aspectos psicométricos como en lo que respecta a su operación, difusión e impacto formativo. Para ello, cada etapa es evaluada a partir de la información recogida de distintas fuentes como el Cuestionario Complementario y las consultas recibidas en el servicio de información del sitio web y el Centro de Llamados. Además, periódicamente el equipo de MIDE UC lleva a cabo estudios pilotos, entrevistas y grupos focales y recoge la opinión de los participantes en las actividades de capacitación realizadas. También, vela por la actualización del proceso a través de la realización periódica de revisiones bibliográficas y estados del arte en temas claves como la construcción de rúbricas o la evaluación de desempeño por parte de pares y supervisores; también se asiste a coloquios, seminarios y congresos y se generan instancias de discusión e intercambio profesional con especialistas de otras instituciones y programas dedicados a la evaluación de profesores⁽²⁰⁾.

Como resultado de estas iniciativas, es posible retroalimentar de manera fundada y con base empírica, el diseño de procesos y productos, incorporándose mejoras para el período siguiente, lo que se ha traducido en diversos cambios a lo largo del desarrollo de la Evaluación Docente. Entre éstos destacan aquellos orientados a mejorar año a año los instrumentos de evaluación a partir del análisis métrico de los datos recogidos en períodos previos, las opiniones de los docentes evaluados y los evaluadores y correctores, así como los estudios de pilotaje realizados para poner a prueba versiones preliminares, especialmente del

.....
20 A modo de ejemplo, en abril de 2010, parte del equipo realizó una visita de intercambio profesional con especialistas del programa PACT, *Performance Assessment for California Teachers* en la Universidad de Stanford y la Universidad de California–Santa Bárbara.

Portafolio, incluyendo esto tanto las rúbricas de corrección como el manual que recibe el docente para elaborar su evidencia. Es así como el Manual Portafolio ha sufrido importantes cambios a lo largo del tiempo, merced a las particularidades de los nuevos grupos de evaluación incorporados.

Los cambios en el Manual Portafolio buscan responder al desafío de diseñar un instrumento cuya estructura tenga un carácter profesional y formativo, de modo que la experiencia misma de su elaboración promueva la reflexión del docente y el mejoramiento de su práctica. Además, existe un esfuerzo permanente por entregar instrucciones más claras y pertinentes a la realidad de los distintos docentes evaluados, condición necesaria para la equidad y confiabilidad del proceso evaluativo. Es así como se han desarrollado cinco variantes del manual acorde al nivel o modalidad de evaluación: Educación Parvularia, Primer Ciclo Básico, Segundo Ciclo Básico, Educación Media y Educación Especial. También se han agregado nuevas indicaciones de acuerdo a las prácticas características o más frecuentes en los subsectores o modalidades de incorporación más reciente, por ejemplo: uso de formas de evaluación del aprendizaje basadas no en pruebas o trabajos escritos sino en evaluaciones de desempeño, como es recurrente entre docentes de Educación Física, Artes Musicales o Educación Parvularia. Asimismo, se consideran contextos diversos para la filmación de la clase: dentro del aula común y también en el aula de recursos (docentes de Educación Especial), o bien, en el patio o la cancha del establecimiento (algo usual en el caso de docentes de Educación Física), todo ello con el fin de que la evidencia conserve un carácter auténtico, que refleje los contextos de desempeño naturales y las prácticas cotidianas de los profesores y educadoras evaluados.

El Informe de Evaluación Individual ha sido otro foco de mejoramiento. Este documento se ha enriquecido en la cantidad y calidad de la información que proporciona al docente, con el fin de apoyarlo a través de retroalimentación descriptiva y específica sobre las fortalezas y debilidades identificadas en su evidencia de evaluación. Durante los primeros tres años de implementación, el informe describía el resultado del docente en cada dimensión del Portafolio en forma global, omitiendo así las diferencias posibles de encontrar entre dos profesores evaluados con un mismo nivel en la dimensión, pero con fortalezas y debilidades distintas en los aspectos específicos que aquella incluye. A partir de 2006, en cambio, el informe no solo consigna el resultado global por dimensión, sino que también describe cualitativamente las fortalezas y debilidades detectadas en los distintos aspectos evaluados. Desde el año 2008 se incorporó además una descripción del nivel de desempeño esperado (Competente) para cada dimensión, con el fin de ayudar a los docentes con más bajos resultados a identificar la brecha entre lo observado en su práctica pedagógica y aquello que se espera según lo establecido en el Marco para la Buena Enseñanza. También se diferenciaron expresamente, dentro del reporte, las fortalezas (aspectos evaluados con nivel Competente o Destacado) y las debilidades (aquellos con resultado Básico o Insatisfactorio), y se han mejorado los elementos gráficos que apoyan la presentación de los datos así como el lenguaje utilizado, para hacerlo más claro y comprensible para los docentes.



Tras ocho años de implementación, la Evaluación Docente muestra signos importantes de consolidación: amplia cobertura a nivel nacional, incorporación de los principales niveles, asignaturas y modalidades de enseñanza, fortalecimiento de la normativa que regula el sistema y sus consecuencias, desarrollo progresivo de los instrumentos de evaluación y la retroalimentación que a partir de ellos se entrega a los docentes evaluados, entre otros. También los datos generados por el proceso, así como la evidencia recolectada a partir de él, han impulsado y servido de insumo para la generación de nuevos conocimientos acerca del trabajo de los docentes de aula del sector municipal y las prácticas de enseñanza más frecuentes de observar entre los profesores y educadoras del país.

Al mismo tiempo, desde el punto de vista de su propósito formativo e impacto sobre la calidad del desempeño docente, persisten desafíos, especialmente aquellos relacionados con la apropiación de la evaluación por parte de los actores locales, el óptimo aprovechamiento de la información que ella produce y de los dispositivos de formación asociados, y la plena articulación de este sistema de evaluación con otras políticas educacionales, por ejemplo, referidas al desarrollo de una carrera docente y la gestión directiva a nivel de cada centro escolar.

Como se ha señalado repetidamente, el principal impulsor de las variaciones en el aprendizaje escolar es la calidad de los docentes (Barber & Mourshed, 2007). Disponer de un sistema válido y confiable para evaluar su desempeño es una herramienta clave para promover una cultura de mejoramiento continuo entre los profesores y en el sistema educacional en su conjunto, e impulsar acciones de formación en servicio que impacten efectivamente los aprendizajes de los alumnos y alumnas, mejorando así la calidad de la educación chilena.



4

Resultados: qué dice la Evaluación Docente acerca de la enseñanza en Chile

Yulan Sun · Mónica Correa
Álvaro Zapata · Diego Carrasco

Centro de Medición MIDE UC
Pontificia Universidad Católica de Chile

Resultados: qué dice la Evaluación Docente acerca de la enseñanza en Chile

¿Cómo puede describirse el desempeño de los profesores y profesoras del sector municipal a partir de la Evaluación Docente? ¿Existen diferencias en los resultados arrojados por los distintos instrumentos o de acuerdo al subsector o modalidad en que se desempeñan los docentes? ¿Qué fortalezas y debilidades de sus prácticas pedagógicas se evidencian en las muestras de trabajo que recoge el Portafolio? Estas son algunas de las preguntas abordadas en el presente capítulo.

Se reportan los datos arrojados por la evaluación desde dos perspectivas. En la primera sección se entrega una visión general de los docentes evaluados y sus resultados en ocho años de aplicación (2003 a 2010), mientras que la segunda profundiza el análisis de los datos del Portafolio. Se describe el desempeño alcanzado por los docentes de acuerdo al nivel, subsector y modalidad de enseñanza en que ejercen, y se ilustran las prácticas pedagógicas más y menos logradas a través de algunos ejemplos.

Así, el capítulo busca responder al interés de quienes se desempeñan en la formación de profesores, la investigación y el diseño de políticas educacionales, los actores responsables de la gestión pedagógica en los establecimientos (sostenedores, directores, jefes técnicos, supervisores) y los propios docentes de aula.

.....

El Sistema de Evaluación del Desempeño Profesional Docente (en adelante Evaluación Docente) vigente desde el año 2003, abarca una porción sustantiva del total de profesores en Chile: la dotación municipal, a la que está dirigido el sistema, incluye más de 70 mil docentes de aula, del total de aproximadamente 150 mil que ejercen en el país.

Esta evaluación tiene tanto propósitos sumativos como formativos. En el plano sumativo, sus resultados tienen consecuencias administrativas, tales como la posibilidad de acceder a incentivos económicos en el caso de los desempeños más altos o la obligación de abandonar la dotación docente, entre quienes en forma reiterada no logran demostrar el estándar exigido. En el plano formativo, la evaluación busca apoyar a los profesores para que identifiquen sus fortalezas y debilidades, reflexionen en torno a sus propias prácticas y así las mejoren. También, debiera contribuir a alinear las políticas y estrategias de formación inicial y perfeccionamiento disponibles para los docentes, con las necesidades identificadas a través de la medición.

El presente capítulo busca contribuir a la difusión de los datos recogidos por la Evaluación Docente en ambos planos. La información se ha organizado en dos secciones: en la primera se entrega una visión general de los resultados obtenidos entre 2003 y 2010 a nivel nacional; se da cuenta de los promedios y la distribución del conjunto de profesores evaluados, por nivel de desempeño final y por instrumento. La segunda sección profundiza en el análisis de los datos arrojados por el Portafolio, el instrumento de mayor peso para el resultado final y que reviste importancia crítica para el propósito formativo de la evaluación, pues permite retroalimentar al docente a partir de evidencia directa de su trabajo. Este análisis se basa en los datos provenientes de 55.536 docentes, de todos los niveles, subsectores y modalidades de enseñanza hasta ahora evaluados. Además de datos cuantitativos, referidos a la proporción de profesores que alcanza el desempeño esperado en distintos indicadores, esta sección da cuenta de cómo se entiende un buen desempeño en cada uno de esos indicadores e incluye ejemplos concretos que ilustran algunas de las fortalezas y debilidades observadas con frecuencia en la práctica de los docentes.

Resultados generales

a. Los niveles de desempeño

La Evaluación Docente distingue cuatro niveles de desempeño definidos por Ley. El nivel **Competente** marca el estándar o nivel de logro esperado, señalando un desempeño profesional adecuado que, sin ser excepcional, cumple con lo requerido para ejercer profesionalmente el rol docente, de acuerdo al Marco para la Buena Enseñanza (Ministerio de Educación, 2003). El nivel **Destacado** indica un desempeño que clara y consistentemente sobresale con respecto de lo que se espera, ya sea por la amplitud del repertorio de conductas demostrado por el docente, o bien, por la riqueza pedagógica que se agrega al cumplimiento del indicador. Bajo el estándar, el nivel **Básico** representa un desempeño que cumple con lo esperado solo en forma parcial o irregular, mientras que en el **Insatisfactorio** se evidencian claras debilidades, que afectan significativamente el quehacer del profesor y comprometen las oportunidades de aprendizaje de los alumnos.

Empleando estos cuatro niveles se comunican los resultados de la Evaluación Docente en los distintos reportes que ésta contempla: informes individuales de retroalimentación para los docentes, informes agregados para los equipos directivos de los establecimientos y para los sostenedores municipales, y comunicación de resultados nacionales a la opinión pública, entre otros.

A partir de los resultados obtenidos en los cuatro instrumentos del sistema⁽¹⁾, se determina un nivel de desempeño global para cada docente, aplicando la ponderación definida por norma legal (ver Tabla 1).

• **Tabla 1**
• *Ponderación de los instrumentos de evaluación*

Instrumento	Ponderación
Portafolio (evidencia escrita y clase filmada)	60%
Entrevista del Evaluador Par	20%
Pauta de Autoevaluación	10%
Informe(s) de Referencia de Terceros	10%

Esta ponderación se aplica en todos los casos, excepto para aquellos docentes cuyo resultado vigente es Insatisfactorio, para los que la ponderación del Portafolio aumenta al 80% y el peso de los demás instrumentos se reduce a la mitad.

El nivel de desempeño final que obtiene un docente depende de la decisión tomada por la Comisión Comunal de Evaluación⁽²⁾, una instancia colegiada a nivel local que toma su decisión considerando: (1) los resultados del profesor en los cuatro instrumentos y (2) la información de contexto consignada por el director y jefe de la Unidad Técnico Pedagógica del establecimiento en su evaluación y por el propio docente, ya sea en su Autoevaluación o en la Entrevista por el Evaluador Par.

1 Las características y el proceso de construcción de los instrumentos son abordados en detalle en Flotts & Abarzúa, en este volumen.
2 Para más información sobre cómo se procesan los datos de cada docente y se llega al resultado final, ver Sun, Calderón, Valerio & Torres, en este volumen.

b. Resultados finales

Anualmente la Evaluación Docente reporta resultados que refieren a la cohorte específica de profesores que participaron del proceso en ese período. La composición de estas cohortes es variable en términos del número de docentes y de los niveles, modalidades y subsectores evaluados, así como los municipios en que ellos ejercen. Esta variación se origina en el carácter gradual y progresivo que ha tenido la implementación de la Evaluación Docente desde su inicio (ver Tabla 2).

• **Tabla 2**
• *Número de evaluados por nivel, subsector y modalidad de enseñanza*

Grupos evaluados	2003	2004	2005	2006	2007	2008	2009	2010
Educación Parvularia						4.061	900	289
Primer Ciclo Básico (Generalista)	3.673	920	8.027	4.937	3.137	2.061	5.164	3.269
Segundo Ciclo Básico:		799	2.291	8.544	3.650	5.557	4.419	5.194
Artes Musicales						478	165	108
Artes Visuales						308	124	111
Ciencias Naturales		151	417	1.594	419	345	525	812
Ciencias Sociales		158	487	1.698	536	367	564	779
Educación Física					1.318	343	310	270
Educación Tecnológica						321	167	121
Inglés						1.012	359	245
Lenguaje		241	708	2.602	716	538	952	1.207
Matemática		249	679	2.650	661	509	839	1.299
Religión						1.336	414	242
Educación Media:			347	709	3.626	4.336	3.632	1.697
Artes Musicales						127	270	52
Artes Visuales						143	340	58
Biología					216	507	149	62
Educación Física					319	639	221	102
Educación Tecnológica						110	286	87
Física					106	301	101	56
Filosofía y Psicología						81	245	45
Historia					432	1.063	309	167
Inglés						270	695	169
Lenguaje			176	361	1.174	343	318	396
Matemática			171	348	1.206	265	311	362
Química					173	367	107	59
Religión						120	280	82
Educación Especial:							1.584	612
Trast. de Aprendizaje							1.099	386
Trast. del Lenguaje							485	226
Total Evaluados	3.673	1.719	10.665	14.190	10.413	16.015	15.699	11.061

En el primer año de aplicación solo fueron evaluados docentes generalistas de Primer Ciclo Básico provenientes de 63 de las 345 comunas del país que poseen establecimientos municipales. Posteriormente, se incorporaron en forma gradual las demás comunas (hasta alcanzar cobertura nacional en el 2005), así como distintos subsectores y niveles de enseñanza: Segundo Ciclo Básico, Enseñanza Media, Educación Parvularia y Educación Especial, sucesivamente. Como es posible apreciar en la Tabla 2, esto da a cada cohorte una configuración particular que obliga a mirar con máxima prudencia los resultados de años sucesivos con propósitos de comparación. Las variaciones entre períodos no pueden interpretarse directamente como expresiones de una mejora o deterioro del desempeño en el tiempo, pues reflejan el rendimiento de cohortes distintas, tanto en términos de los individuos específicos que la componen, como de los niveles, subsectores y modalidades de enseñanza en que esos docentes ejercen (ver Tabla 3). Asimismo, a partir de 2007, cada cohorte incluye un número importante de docentes que se re-evalúa tras haberse cumplido los cuatro años de vigencia de su resultado Básico, Competente o Destacado, además de aquellos que deben evaluarse nuevamente por haber obtenido un resultado Insatisfactorio en el año anterior.

• **Tabla 3**
• *Docentes evaluados por Nivel y Modalidad de Enseñanza 2003–2010*

Período	Primer Ciclo Básico	Segundo Ciclo Básico	Educación Media	Educación Parvularia	Educación Especial	Total
2003	100,0%					3.673
2004	53,5%	46,5%				1.719
2005	75,3%	21,5%	3,2%			10.665
2006	34,8%	60,2%	5,0%			14.190
2007	30,1%	35,1%	34,8%			10.413
2008	12,9%	34,7%	27,1%	25,3%		16.015
2009	32,9%	28,2%	23,1%	5,7%	10,1%	15.699
2010	29,6%	47,0%	15,3%	2,6%	5,5%	11.061

Pese a las particularidades antes mencionadas respecto de cada cohorte evaluada, los resultados finales en cuanto a la distribución de los docentes en los cuatro niveles de desempeño, no muestran grandes diferencias entre un período y otro. Como es posible observar en la Tabla 4, predomina en todos los casos la proporción de profesores que obtiene un resultado final Competente (entre 52,4% y 64%, dependiendo del año), seguida por el grupo con resultado Básico (22,8% a 37,2%). Un porcentaje minoritario alcanza un nivel Destacado (6% a 12,1%) y en todos los períodos, la categoría que registra la menor frecuencia es el nivel Insatisfactorio (1,1% a 3,7%).

• **Tabla 4**
• *Resultados finales por período 2003–2010*

Período	Insatisfactorio	Básico	Competente	Destacado	Total Evaluados
2003	3,7%	30,3%	56,6%	9,4%	3.673
2004	3,0%	34,3%	52,4%	10,3%	1.719
2005	3,5%	37,2%	52,7%	6,6%	10.665
2006	2,8%	31,4%	59,0%	6,8%	14.190
2007	2,0%	33,0%	56,6%	8,5%	10.413
2008	1,1%	22,8%	64,0%	12,1%	16.015
2009	1,5%	28,9%	63,1%	6,5%	15.699
2010	2,6%	33,3%	58,1%	6,0%	11.061

La proporción de docentes por nivel de desempeño final es informativa desde el punto de vista de las consecuencias contempladas por la Evaluación Docente⁽³⁾. Aquellos que solo logran un resultado Básico o Insatisfactorio (entre 23,9% y 40,7% del total de evaluados, dependiendo del período) tienen la obligación de incorporarse a los Planes de Superación Profesional (PSP) que organizan las autoridades locales. En el primer caso los docentes tienen la obligación de asistir a estos planes de apoyo hasta que se cumple el período de vigencia de su resultado Básico (dos años, según lo establecido por la Ley 20.501, promulgada en febrero de 2011) y, entonces, deben evaluarse nuevamente. Quienes obtienen un resultado Insatisfactorio, como se ha señalado antes, deben evaluarse en el año inmediatamente siguiente, de modo que su participación en los PSP es de un año (y podría extenderse, dependiendo de su resultado en la nueva evaluación).

Por su parte, los docentes que alcanzan un desempeño final Competente o Destacado (entre 59,3% y 76,1%) tienen la posibilidad de acceder a la Asignación Variable de Desempeño Individual (AVDI), un incentivo económico por hasta cuatro años, tras rendir una prueba de conocimientos disciplinarios y pedagógicos.

Cabe mencionar que, más allá de las consecuencias establecidas normativamente por la Evaluación Docente, existen datos que sugieren que los resultados de ésta podrían vincularse con otros efectos, incidiendo por ejemplo en la permanencia de los docentes en el sistema escolar y en su movilidad laboral. Análisis realizados por el área de Investigación de MIDE UC usando datos de la evaluación 2003-2006 constataron que la tasa anual de salida del sistema municipal de los docentes con resultado Insatisfactorio era sustantivamente superior a aquella de sus pares con resultado más alto: 22% versus 6,7%, respectivamente, no ejercía en el sistema en 2007 (Taut, Santelices & Valencia, 2009). La misma tendencia se observó al incluir datos de la evaluación hasta 2007 y analizar la situación laboral en 2008: 17,9% versus 4,8% (Taut, Santelices & Valencia, 2010). Algo similar se comprueba al realizar un cruce entre el resultado vigente de los docentes evaluados entre 2003 y 2009 y su situación laboral según los datos de la BRP⁽⁴⁾ en el año 2010 (ver Tabla 5).

• **Tabla 5**
• *Situación laboral de docentes según resultado vigente en la Evaluación Docente*

Situación laboral en BRP 2010 (N= 62.870)	Resultado vigente en Evaluación Docente (2003–2009)			
	Insatisfactorio (i)	Básico	Competente	Destacado
Con horas contratadas en Municipal (*)	38,9%	75,4%	84,7%	88,4%
Con horas contratadas en Particular Subvencionado	5,9%	3,3%	2,5%	2,1%
Jubilado (**)	23,0%	7,2%	4,5%	3,6%
No figura en BRP (***)	32,2%	14,1%	8,3%	5,9%
Total	100,0%	100,0%	100,0%	100,0%

(i) Excluye docentes a los que se les asigna Insatisfactorio por haberse negado a la evaluación sin causa justificada.

(*) Incluye docentes que ejercen simultáneamente en establecimientos municipales y particulares subvencionados.

(**) Incluye docentes que han abandonado el sistema escolar por jubilación o han presentado su renuncia anticipada por jubilación. También se incluye un número reducido de docentes fallecidos desde el momento de su evaluación.

(***) De estos docentes cabe suponer que (1) no poseen horas contratadas en establecimientos con subvención o (2) las tienen, pero no poseen título profesional.

3 Para más información sobre ellas, ver Cortés & Lagos, en este volumen.

4 Base de datos de la Bonificación por Reconocimiento Profesional (BRP) de octubre 2010. Esta base reporta la situación contractual de todos los profesores que se desempeñan en establecimientos municipales y particulares subvencionados. La BRP es un beneficio remuneracional que, desde enero de 2007, permite a los docentes percibir un monto mensual por concepto de título profesional y un complemento por concepto de mención, proporcional a sus horas de contrato en establecimientos subvencionados.

Como muestra la Tabla 5, solo un 38,9% de aquellos docentes cuyo resultado vigente en la evaluación es Insatisfactorio se encuentra ejerciendo en establecimientos municipales, proporción que aumenta progresivamente a medida que se asciende en el resultado: 75,4%, 84,7% y 88,4% para Básico, Competente y Destacado, respectivamente.

Así también, es más frecuente encontrar profesionales que cumplen funciones directivas y técnico-pedagógicas entre aquellos docentes que tienen un resultado Destacado en la evaluación⁵, proporción que se reduce cuanto menor es su desempeño. Mientras un 5,7% de los docentes con resultado vigente Destacado se desempeñan como directores o jefes de UTP (Unidad Técnico Pedagógica), esta proporción se reduce a 3,7%, 2,2% y 0,4% entre aquellos con nivel Competente, Básico e Insatisfactorio, respectivamente (ver Tabla 6).

• **Tabla 6**
• *Función desempeñada según resultado vigente en la Evaluación Docente*

Función desempeñada según BRP 2010	Resultado vigente en Evaluación Docente (2003–2009)			
	Insatisfactorio (i)	Básico	Competente	Destacado
Docente	99,6%	97,8%	96,3%	94,3%
Directivo / UTP	0,4%	2,2%	3,7%	5,7%
Total	100,0%	100,0%	100,0%	100,0%

(i) Excluye docentes a los que se les asigna Insatisfactorio por haberse negado a la evaluación sin causa justificada.

Los datos antes expuestos son consistentes con la información recogida a través de los sistemas de monitoreo asociados a la implementación del sistema de evaluación en las comunas. Éstos muestran que, con frecuencia, los resultados obtenidos en la Evaluación Docente son considerados por los sostenedores municipales como un antecedente asociado a oportunidades laborales y profesionales, tales como instancias de promoción interna y concursos públicos para acceder a una mejor situación contractual (cupó de planta en la dotación comunal)⁶.

c. Resultados por instrumento

Desde el inicio de su implementación, los cuatro instrumentos de la Evaluación Docente han arrojado resultados de alta estabilidad en términos de los promedios alcanzados y la distribución de los evaluados en los cuatro niveles de desempeño. Para ilustrar lo dicho, el Gráfico 1 muestra el promedio arrojado por cada instrumento en los últimos cuatro períodos de evaluación (2007–2010), expresado en la escala de 1 a 4 puntos utilizada por el sistema.

Como es posible observar en el Gráfico 1, los resultados de la Autoevaluación exhiben los promedios más altos entre los cuatro instrumentos utilizados, situándose consistentemente muy cerca del puntaje máximo. La evaluación de los directores y jefes técnicos de los establecimientos y aquella de los Evaluadores Pares también muestran promedios relativamente altos y estables, aunque en este caso se registran algunas variaciones entre períodos. Por su parte, el Portafolio ha sido consistentemente el instrumento con más bajos resultados.

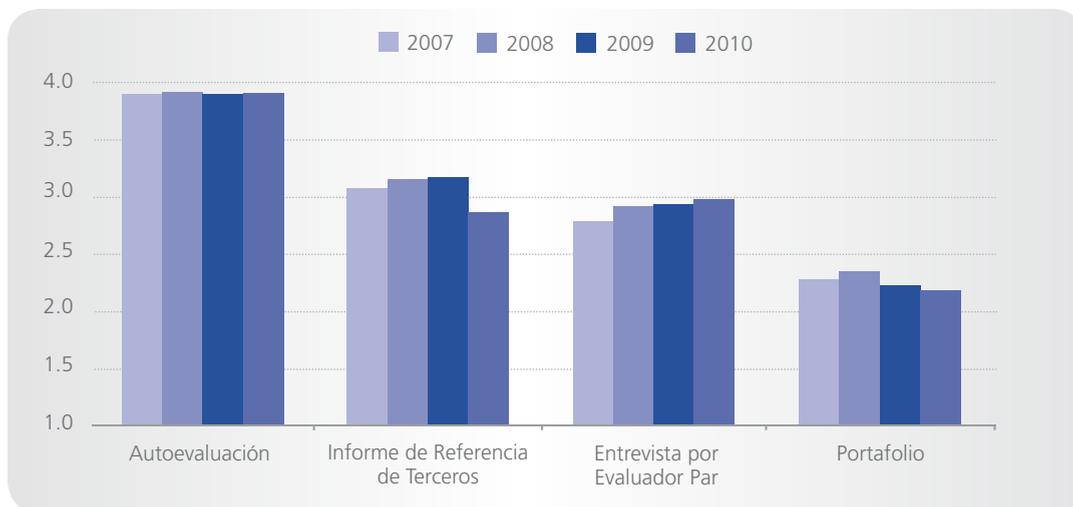
Dentro de las escasas variaciones observadas entre períodos, es posible constatar cierta tendencia hacia el alza en la media de puntajes de la Entrevista por un Evaluador Par. No obstante, la diferencia más llamativa se registra en el promedio de la evaluación realizada por los directivos a través del Informe de Referencia de Terceros. En 2010, la media de este instrumento descendió 0,29 puntos respecto del año anterior. Al

5 Resultados similares se reportan en Taut, Santelices & Valencia, 2010.

6 Algunas de esas iniciativas son abordadas en mayor detalle en Sun, Calderón, Valerio & Torres, en este volumen.

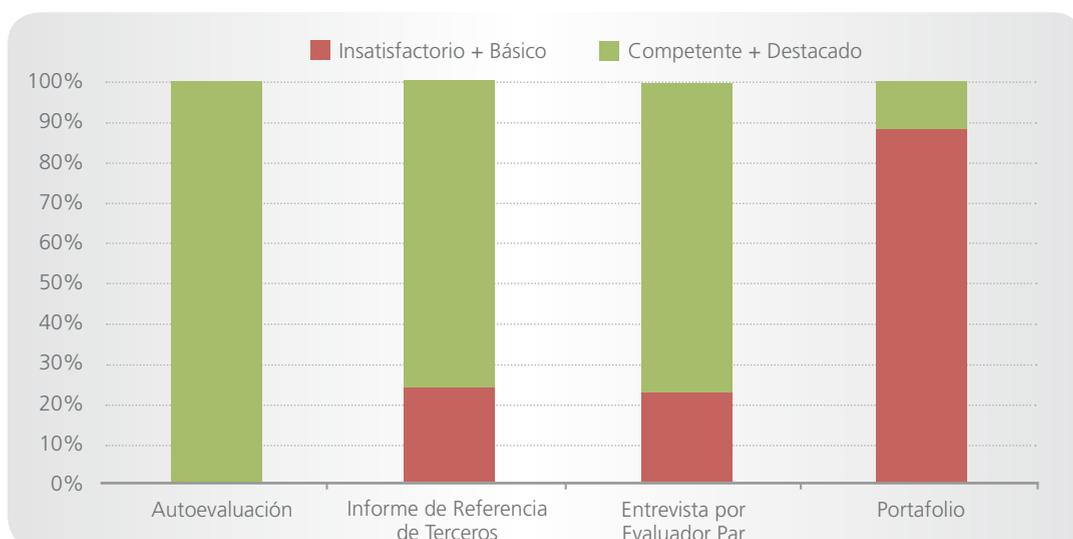
respecto, debe señalarse que en este período se introdujo una modificación en la pauta estructurada que deben utilizar los directivos para llevar a cabo su evaluación. A diferencia de años anteriores, la pauta del informe 2010 incluyó una rúbrica para asignar el nivel de desempeño correspondiente en cada pregunta o indicador evaluado. Además, para calificar con el nivel Destacado en un indicador, se exigió al evaluador justificar el carácter sobresaliente de la práctica del docente en él. Estos cambios se asociaron a un descenso sustantivo en la frecuencia con que este nivel sobresaliente fue asignado: mientras que en 2009 un 53,8% de los indicadores evaluados recibió la calificación máxima, en 2010 esto solo ocurrió con un 21,3% de ellos.

• **Gráfico 1**
 • *Promedios por instrumento 2007–2010*



Las diferencias entre los resultados arrojados por los distintos instrumentos también puede observarse desde el punto de vista de la distribución de los docentes evaluados según su nivel de desempeño (ver Gráfico 2).

• **Gráfico 2**
 • *Proporción de docentes por nivel de desempeño por Instrumento 2010*



Los datos representados en el Gráfico 2 corresponden a la evaluación 2010, pero reflejan la tendencia observada en los ocho períodos de implementación. En todos ellos, los datos de la Autoevaluación ubican virtualmente a todos los evaluados alcanzando el estándar o desempeño esperado. Algo similar, aunque menos acentuado, se observa en la evaluación realizada por directivos y pares. En cambio, el Portafolio ha mostrado consistentemente una distribución inversa, concentrándose la mayor parte de los evaluados en los niveles inferiores de desempeño, por debajo del estándar. La Tabla 7 ilustra lo señalado, a partir de los resultados en los últimos cuatro períodos de evaluación.

• **Tabla 7**
 • *Proporción de docentes según nivel de desempeño por instrumento y período (2007–2010)*

Autoevaluación			
Período	Insatisfactorio + Básico	Competente + Destacado	Total
2010	0,4%	99,6%	100,0%
2009	0,8%	99,2%	100,0%
2008	0,5%	99,5%	100,0%
2007	0,5%	99,5%	100,0%

Inf. Ref. de Terceros			
Período	Insatisfactorio + Básico	Competente + Destacado	Total
2010	24,0%	76,0%	100,0%
2009	12,6%	87,4%	100,0%
2008	16,5%	83,5%	100,0%
2007	19,0%	81,0%	100,0%

Entrevista Ev. Par			
Período	Insatisfactorio + Básico	Competente + Destacado	Total
2010	23,2%	76,8%	100,0%
2009	21,1%	78,9%	100,0%
2008	22,9%	77,1%	100,0%
2007	30,9%	69,1%	100,0%

Portafolio			
Período	Insatisfactorio + Básico	Competente + Destacado	Total
2010	87,7%	12,3%	100,0%
2009	84,4%	15,6%	100,0%
2008	70,6%	29,4%	100,0%
2007	79,2%	20,8%	100,0%

Es pertinente señalar que, dado que poseen características y fuentes distintas, no necesariamente debe esperarse que los cuatro instrumentos muestren patrones de resultados similares, y la recolección de evidencia a través de una variedad de instrumentos y fuentes para evaluar el desempeño docente se considera necesaria para cumplir las condiciones de precisión e imparcialidad (*fairness*) que ésta exige (Isoré, 2009; Joint Committee on Standards for Educational Evaluation, 2009). En particular la participación de múltiples evaluadores es considerada como un elemento clave cuando se trata de juzgar un objeto de evaluación complejo como es el desempeño o la calidad docente, que puede ser visto desde distintas perspectivas y debiera resultar satisfactorio desde todas ellas (Stronge & Tucker, 2003). No por

ello deben ignorarse, sin embargo, las señales entregadas por los datos hasta ahora recogidos en la Evaluación Docente, en cuanto sugieren una posible inflación de puntajes, especialmente en el caso de la Autoevaluación y, en menor medida, en la evaluación de directivos y pares.

Respecto de la Autoevaluación, su inclusión dentro del sistema de evaluación responde al carácter profesional y reflexivo con que se entiende el trabajo docente. Por ello, a través de este instrumento se busca promover que el docente analice y revise sus propias prácticas pedagógicas, evalúe la calidad de su relación con alumnos, apoderados, pares y directivos, identifique sus fortalezas y debilidades y sus necesidades de desarrollo profesional, entre otros aspectos; y es posible que así ocurra en al menos parte de los profesores evaluados. Sin embargo, desde el punto de vista de la calificación que éstos asignan a su propio desempeño, parece difícil revertir la tendencia a registrar puntuaciones altas tratándose de una evaluación con consecuencias y en la que este instrumento tiene una incidencia directa sobre el resultado final (ponderación del 10%). En este contexto, y a la luz de los datos acumulados tras ocho años de implementación, parece razonable re-evaluar la inclusión de este instrumento y/o la forma de ponderarlo para efectos del resultado final de cada profesor. De otra forma, además del riesgo de distorsionar el valor de la autoevaluación como un índice confiable del desempeño docente, podría debilitarse también el ejercicio de reflexión que se busca promover a través de ella.

En cuanto a la evaluación realizada por pares y supervisores directos, algunos estudios señalan cierta tendencia a la “indulgencia” y a sobreestimar la calidad del trabajo docente en el contexto de evaluaciones de altas consecuencias (Prince, Koppich, Morse, Bhatt & Witham, s.d.). También se ha reportado que la fiabilidad del juicio del director se ve seriamente mermada cuando en el ejercicio de su función predomina la dimensión administrativa por encima de la gestión pedagógica (Jacob & Lefgren, 2005, 2006, 2008); esto último no es infrecuente en el caso del sistema escolar chileno⁷). Asimismo, algunos autores informan de ciertos sesgos en las evaluaciones de directivos asociados, por ejemplo, al nivel de conocimiento del profesor o el grado de simpatía personal hacia éste (Harris & Sass, 2007), y se sostiene que los directores son capaces de identificar aquellos docentes que tienen un desempeño extremo (10% a 20% superior e inferior), pero les resulta difícil discriminar entre el resto (Jacob & Lefgren, 2006, 2008).

Pese a lo dicho, existe un alto consenso en considerar que tanto pares como directivos son fuentes relevantes para emitir un juicio acerca de la calidad del desempeño del profesor, aportando una perspectiva diferente que no es posible pesquisar mediante otras fuentes o instrumentos. Basta pensar, por ejemplo, en la oportunidad que tienen los directores y jefes de UTP de conocer en forma directa el trabajo de los profesores de su establecimiento, del que pueden tener además una apreciación longitudinal, sostenida en el tiempo. También, al igual que los Evaluadores Pares, los directivos están familiarizados con el contexto en que el docente ejerce su trabajo: las características del curso y del establecimiento, su proyecto educativo, el cuerpo docente, la infraestructura escolar y los recursos pedagógicos disponibles, entre otros aspectos. Por todo esto, resultan fuentes de evaluación difíciles de sustituir y necesarias para conformar una visión global del desempeño docente. También se ha visto que es posible obtener mediciones más válidas y confiables de pares y directivos cuando ellos reciben capacitación para ejercer este rol (Rockoff & Speroni, 2010). Una capacitación sólida a los evaluadores y asegurar que éstos posean las características y competencias requeridas, aparecen como factores claves para el éxito de cualquier sistema de evaluación de profesores (Isoré, 2009; Heneman III et al, 2006).

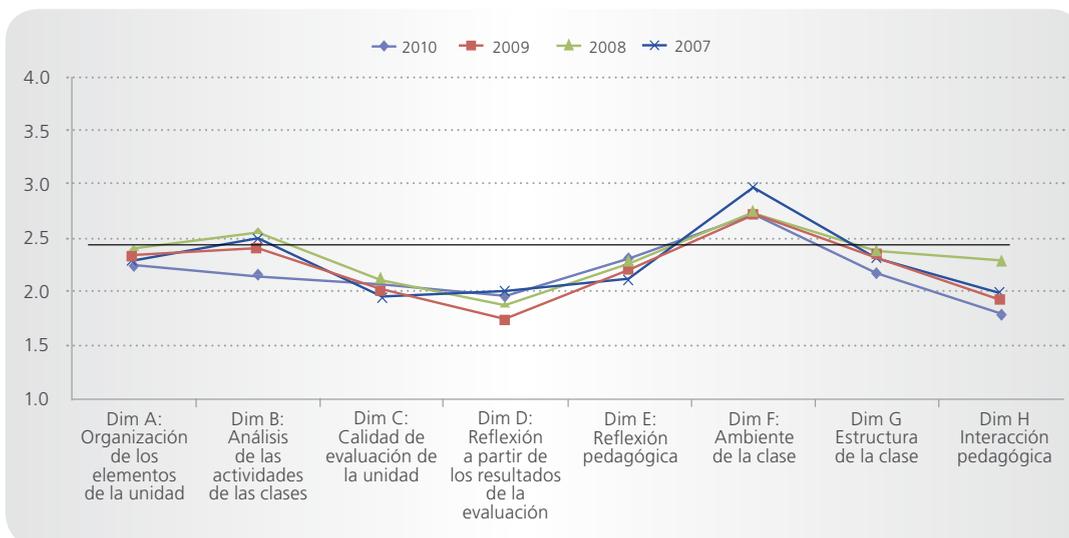
También ha de tenerse en cuenta que pares y directivos forman parte del contexto local en que los docentes llevan a cabo su trabajo, lo que da a su participación en la evaluación un valor virtualmente irremplazable, especialmente en el desarrollo de iniciativas de mejoramiento a partir de los resultados. En particular resulta relevante la participación de los directores y jefes de UTP, por su alta responsabilidad sobre la calidad de la gestión escolar y de sus resultados.

.....

7 Así lo muestran, por ejemplo, los resultados preliminares de un estudio sobre “Liderazgo directivo y calidad de la Educación en Chile”, realizado por un equipo del Centro de Estudios de Políticas y Prácticas en Educación (CEPPE) y la Fundación Chile. Disponible en http://www.mideuc.cl/docs/coloquio_ceppe.pdf

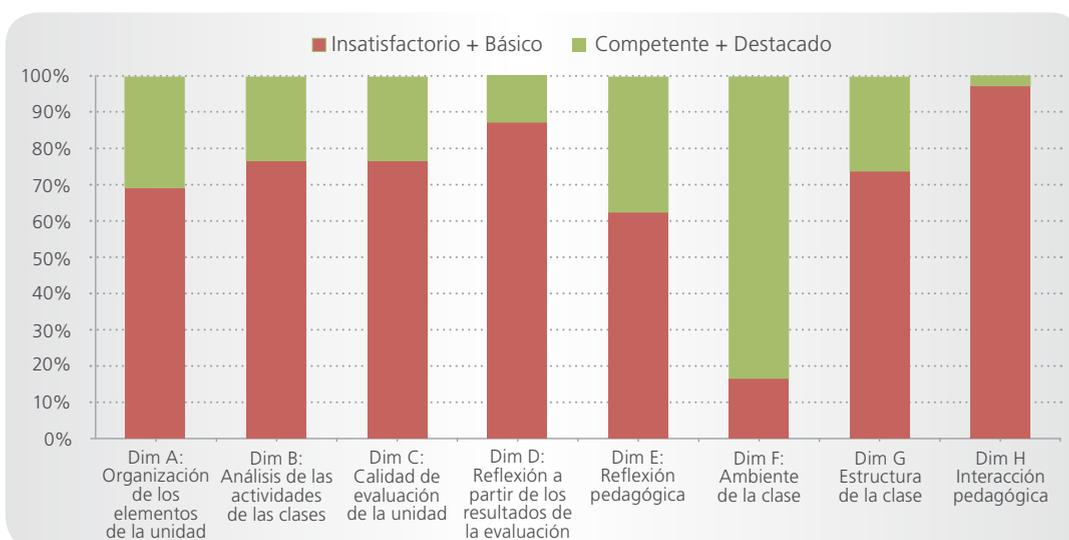
Volviendo a los datos arrojados por los distintos instrumentos de la Evaluación Docente, como se mencionara, el Portafolio resulta ser el instrumento que muestra los más bajos resultados: solo un 10,2% a un 29,4% de los docentes, dependiendo del período, alcanza un nivel Competente o Destacado. Lo mismo se refleja en los promedios de las ocho dimensiones abordadas por este instrumento (ver Gráfico 3).

● **Gráfico 3**
● Promedio por dimensión del Portafolio 2007–2010



Siete de las ocho dimensiones evaluadas muestran un promedio inferior al nivel mínimo esperado (2,5 puntos correspondientes al extremo inferior del rango Competente), lo que se refleja también en la proporción de docentes según nivel de desempeño (ver Gráfico 4).

● **Gráfico 4**
● Proporción de docentes por nivel de desempeño por dimensión del Portafolio 2010

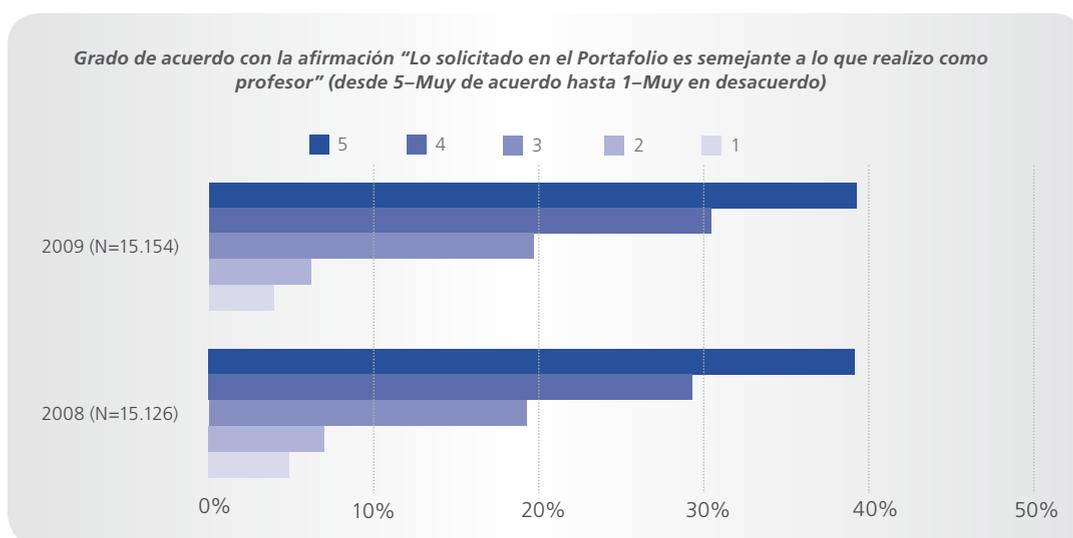


Como es posible apreciar, la dimensión que exhibe un mejor resultado en 2010 es “Ambiente de la clase”, en que 8 de cada 10 docentes logra un resultado Competente o Destacado, mostrando que la mayor parte de ellos es capaz de crear un ambiente favorable para el aprendizaje en la clase filmada, logrando motivar a sus alumnos y comprometerlos con las actividades propuestas.

En cambio, los resultados más bajos se observan en las dimensiones de interacción pedagógica que el docente sostiene con los alumnos en sus clases y de reflexión a partir de los resultados de la evaluación. En 2010 solo un 3,8% de los profesores demuestra una interacción pedagógica de calidad en la clase filmada para el Portafolio. También se constata que solamente un 13,6% de los profesores logra el desempeño esperado a la hora de utilizar la información recogida por la evaluación que aplican a sus alumnos y darles una retroalimentación efectiva a partir de ella.

Cabe señalar que el Portafolio es el instrumento de mayor peso para el resultado final de cada profesor y aquel que reúne las más estrictas condiciones de estandarización entre los cuatro utilizados por el sistema de evaluación⁽⁸⁾. Se trata, además, del único que aporta muestras directas del trabajo del docente y, a juzgar por los datos recogidos en el Cuestionario Complementario que responden los propios evaluados⁽⁹⁾, éstos aprecian su cualidad auténtica al considerar que el tipo de evidencia solicitada es similar a su trabajo cotidiano como docente (a modo de ejemplo, ver los datos 2008 y 2009 en el Gráfico 5).

● **Gráfico 5**
 ● *Opinión de los docentes evaluados acerca de la similitud entre la evidencia solicitada en el Portafolio y su trabajo habitual—2008 y 2009*

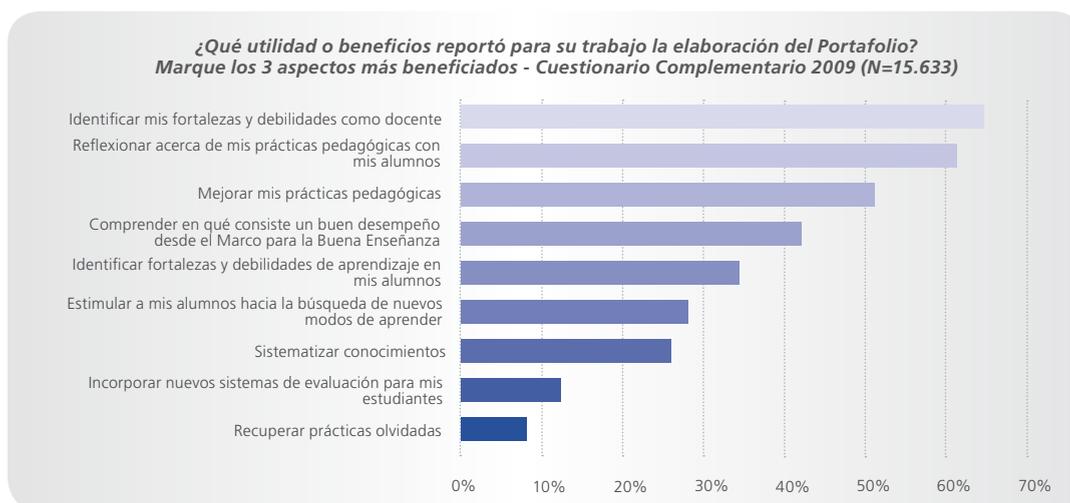


Los datos del Cuestionario Complementario también señalan en forma sostenida a lo largo de los distintos períodos de evaluación, que la elaboración del Portafolio reporta beneficios para el trabajo docente (ver Gráfico 6). Como se aprecia, la mayoría de los evaluados reporta que la realización del Portafolio los ayuda a identificar sus fortalezas y debilidades, reflexionar acerca de las prácticas pedagógicas que utiliza en sus clases y comprender mejor en qué consiste un buen desempeño de acuerdo al Marco para la Buena enseñanza. Estos datos resultan alentadores en cuanto a sugerir que la experiencia misma de realizar la evaluación tendría un valor formativo para los profesores evaluados.

8 Para más detalles sobre el proceso de construcción y corrección del Portafolio, ver Flotts & Abarzúa, en este volumen.

9 Este cuestionario, de respuesta voluntaria, es completado en cada período por más del 98% de los docentes evaluados.

- **Gráfico 6**
- *Opinión de los docentes evaluados acerca de los beneficios que reporta para su trabajo la elaboración del Portafolio—2009*



Resultados del Portafolio: Fortalezas y debilidades de la práctica docente

Como se ha señalado antes, el Portafolio es un instrumento que recolecta evidencia del trabajo cotidiano que realizan los profesores y profesoras, y busca centrarse en aspectos sustantivos y relevantes tanto de su práctica pedagógica al interior del aula como de su quehacer profesional en términos más amplios. Dichos aspectos provienen de los criterios y descriptores señalados en el Marco para la Buena Enseñanza (MBE), que constituyen los estándares de desempeño profesional en los que se fundamenta el sistema de evaluación y sus instrumentos. En base a los descriptores del MBE se crean los indicadores, se define la evidencia que se solicitará a los docentes y se elaboran las rúbricas de corrección de ésta.

Las muestras de trabajo solicitadas en el Portafolio se organizan en dos “módulos” y cuatro “productos”, que incluyen evidencia escrita y audiovisual (ver Figura 1). La evidencia escrita refiere a la implementación de una unidad de 8 horas pedagógicas; se requiere presentar, por ejemplo, una descripción de sus objetivos, las actividades realizadas y los materiales utilizados; presentación del instrumento de evaluación empleado para establecer los aprendizajes logrados por los alumnos al concluir la unidad, y su pauta de corrección; y respuestas a distintas preguntas que exigen al docente reflexionar sobre su trabajo y los resultados obtenidos en éste. La evidencia audiovisual implica la presentación del video de una clase de 40 minutos.

Para asegurar una amplia cobertura de los distintos aspectos que abarca el MBE, así como la comparabilidad de la evaluación, el Portafolio mantiene un alto grado de estabilidad de un período a otro, tanto en el tipo de evidencia que se solicita como en los indicadores evaluados a partir de ella. Al mismo tiempo, es necesario incorporar algunas variaciones en atención a las particularidades de los diferentes subsectores que se han incorporado sucesivamente al proceso. Esto implica asegurar que las muestras de trabajo que se piden en el Portafolio así como los indicadores de evaluación sean pertinentes, sustantivos y factibles de ser medidos en contextos pedagógicos variados¹⁰. También, las variaciones en la evidencia solicitada entre un período y otro contribuyen a garantizar que la evaluación no sea vulnerada mediante la copia de Portafolios bien evaluados en años anteriores.

10 Por ejemplo, la Evaluación Docente debe ser igualmente adecuada para una docente de Inglés que trabaja en aula con alumnos de Enseñanza Media, una Educadora Especial que desarrolla una sesión con tres alumnos en el aula de recursos y un profesor de Educación Física que realiza una actividad deportiva en el patio del establecimiento, entre otros.

• **Figura 1**
 • *Estructura del Portafolio*

MÓDULO 1	MÓDULO 2
Producto 1 Unidad Pedagógica	Producto 4 Clase grabada
Descripción de la unidad pedagógica	Video de la clase grabada
.....
Análisis de las dificultades de los alumnos	Ficha de la clase grabada
.....
Análisis de la unidad realizada	(*) Fotocopias de los recursos de aprendizaje utilizados en la clase grabada
Producto 2 Evaluación de la unidad pedagógica
Evaluación de la unidad y Pauta de corrección	
.....	
Reflexión a partir de los resultados obtenidos en la evaluación	
.....	
Retroalimentación de la evaluación	
Producto 3 Reflexión Pedagógica	

(*) El docente puede adjuntar los recursos de aprendizaje utilizados cuando estime que éstos podrían no verse claramente en el video (por ejemplo, guías de trabajo individual, páginas del texto escolar, imágenes o dibujos empleados por el docente, evaluaciones individuales, etc.).

a. Descripción de los casos analizados

Los análisis que se reportan a continuación abarcan a una vasta proporción de la población de docentes municipales, ya que se cuenta con un total de 55.536⁽¹¹⁾. Los casos reportados corresponden a docentes que cumplen con los siguientes criterios:

- Evaluados entre los períodos 2006 a 2009, provenientes de todos los niveles, modalidades y subsectores⁽¹²⁾ incorporados hasta la fecha. En el caso de los profesores que cuentan con más de una evaluación, se consideró solo su último resultado.
- Docentes cuyo Portafolio no registra anomalías en la presentación de la evidencia (por ejemplo, el docente no entrega la evaluación solicitada en el Módulo 1 o no presenta la filmación de una clase). De esta forma, se asegura que los datos analizados corresponden a resultados de desempeño y no a puntajes totales o parciales afectados por el incumplimiento de instrucciones.

La Tabla 8 detalla la composición del grupo analizado, para cada modalidad, nivel y subsector evaluado.

11 Estas cifras no corresponden al total de evaluados en cada nivel/subsector, debido a que algunos indicadores medidos en el Portafolio varían entre un período y otro. En este caso se ha seleccionado solo los datos de docentes evaluados mediante los mismos indicadores, aunque lo hayan sido en períodos distintos.

12 Por tratarse de datos recogidos hasta el año 2009, en esta sección se empleará la denominación de subsectores vigente hasta esa fecha (previo al ajuste curricular de 2010).

• **Tabla 8**
 • *Composición del conjunto de casos analizados según nivel, subsector y modalidad*

Grupos evaluados	Total
Educación Parvularia	4.922
Primer Ciclo Básico (Generalista)	15.070
Segundo Ciclo Básico:	21.819
Artes Musicales	638
Artes Visuales	428
Ciencias Naturales	2.832
Ciencias Sociales	3.093
Educación Física	1.947
Educación Tecnológica	482
Inglés	1.357
Lenguaje	4.732
Matemática	4.580
Religión	1.730
Educación Media:	12.143
Artes Musicales	395
Artes Visuales	480
Biología	861
Educación Física	1.158
Educación Tecnológica	391
Física	497
Filosofía y Psicología	325
Historia	1.784
Inglés	960
Lenguaje	2.163
Matemática	2.089
Química	640
Religión	400
Educación Especial:	1.582
Trast. de Aprendizaje	1.098
Trast. del Lenguaje	484
Total Evaluados	55.536

b. Resultados por ámbitos evaluados en el Portafolio

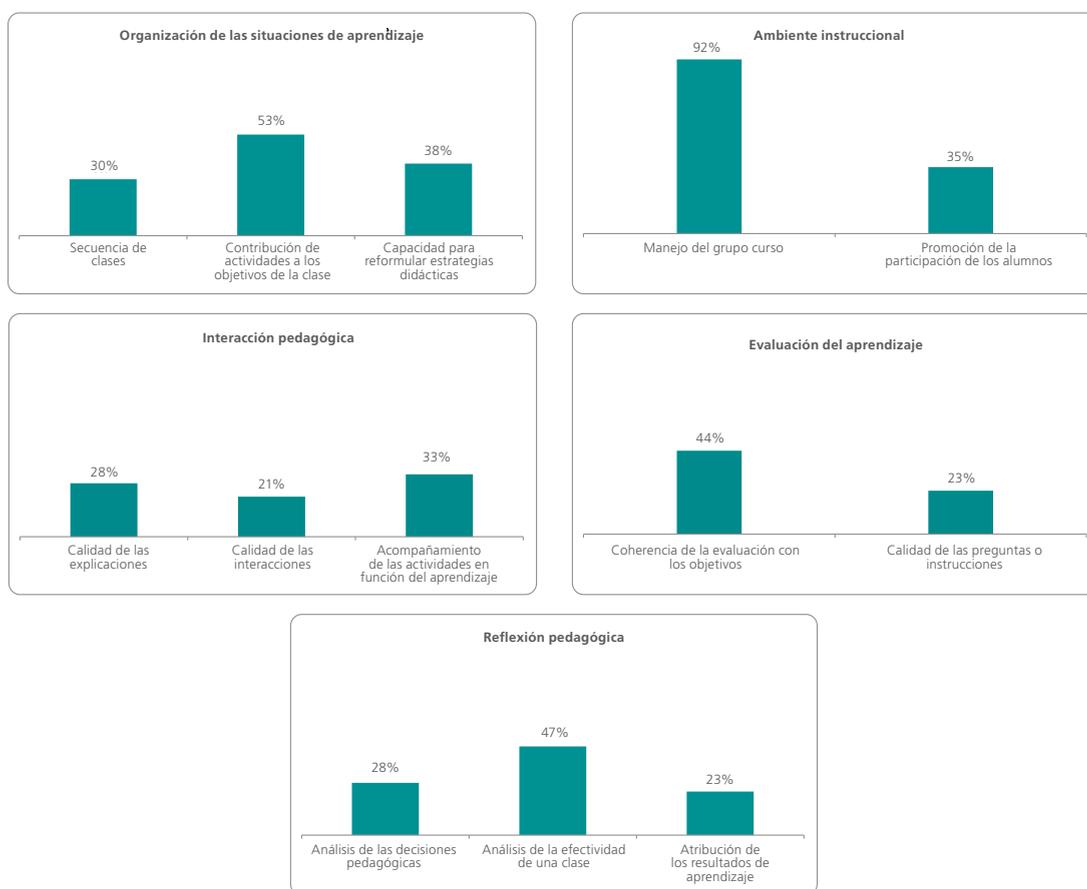
Los análisis reportados consideraron trece indicadores, seleccionados entre aquellos que han sido empleados sin modificaciones en más de un período de evaluación desde 2006 hasta 2009. Para efectos de este capítulo, los indicadores han sido agrupados en cinco ámbitos o temáticas:

- i. Organización de las situaciones de aprendizaje
- ii. Ambiente instruccional

- iii. Interacción pedagógica
- iv. Evaluación del aprendizaje
- v. Reflexión pedagógica

Considerando los resultados de esta muestra, en la Figura 2 se reporta la proporción del total de docentes evaluados, que logra un desempeño Competente o Destacado en cada indicador.

• **Figura 2**
 • *Proporción de docentes que obtiene resultado Competente y Destacado en indicadores del Portafolio*



En lo que sigue se comentará lo más relevante del desempeño de los docentes en cada uno de estos ámbitos. Se hará referencia a aquellos grupos dentro de la muestra que exhiben un desempeño comparativamente superior e inferior en términos de la proporción de docentes que logra el estándar esperado (esto es, al menos un nivel Competente). También se ilustrará, a través de ejemplos, algunas de las debilidades y fortalezas más características o frecuentes en las prácticas pedagógicas observadas en el Portafolio. Cabe hacer notar que, si bien estos ejemplos se basan en evidencia real, en todos los casos se han realizado adaptaciones para asegurar que no correspondan en forma exacta o completa al material entregado por un profesor en particular. Asimismo, los ejemplos cumplen solo un propósito ilustrativo y por lo mismo, no pretenden cubrir la amplia variedad de muestras de trabajo recogidas en los Portafolios y posibles de observar en el desempeño docente.

Organización de las situaciones de aprendizaje

Dentro del Marco para la Buena Enseñanza, la forma en que los docentes estructuran sus clases y organizan las actividades de sus estudiantes se entiende como un elemento fundamental para el proceso de enseñanza-aprendizaje. Es más fácil aprender, recordar y aplicar los conocimientos a situaciones nuevas cuando éstos se encuentran conectados entre sí y estructurados en torno a ideas o temas relevantes, que cuando se entrega información aislada (Anderson, 2004). Un docente efectivo diseña clases y actividades consistentes con los objetivos de aprendizaje que persigue, y las secuencia adecuadamente de modo que faciliten el abordaje de contenidos y el desarrollo de habilidades. También, hace un buen uso del tiempo para la enseñanza y emplea metodologías de trabajo adecuadas a la complejidad de los contenidos y a las características y conocimientos previos de sus alumnos (Stronge, 2007). De esta forma, las actividades que ellos realizan se convierten en experiencias eficaces para el logro de los objetivos.

Otro factor importante para un buen desempeño en este ámbito es la capacidad del profesor para modificar las estrategias pedagógicas planificadas, de acuerdo a los resultados que obtiene en su implementación. Si, por ejemplo, una actividad ha resultado muy compleja para los estudiantes, un docente de buen desempeño es capaz de reorganizarla, incluyendo más ejemplos o ejercicios y entregando más apoyo del previsto; si los alumnos han mostrado más interés del esperado o terminan antes de lo previsto, aprovecha esa motivación para profundizar más en los contenidos, velando así por un buen uso del tiempo en el aula. Todo esto supone un alto nivel de receptividad del profesor hacia las respuestas de los estudiantes y flexibilidad para reaccionar a ellas.

Para evaluar su nivel de desempeño en estos aspectos, en el Portafolio el docente presenta la descripción de una unidad de 8 horas pedagógicas desarrollada con sus alumnos (da cuenta de sus objetivos, actividades realizadas y recursos utilizados, entre otros). También responde a preguntas que ponen en juego su capacidad para proponer estrategias didácticas adecuadas a distintas situaciones enfrentadas con sus alumnos, como por ejemplo, requerimiento de apoyo adicional por aquellos que tuvieron más dificultades en la unidad, necesidad de dividir actividades en unidades de complejidad creciente, incorporación de otros recursos, etc. Además, en la clase filmada, se observa la forma en que ha organizado las actividades para trabajar los objetivos de aprendizaje que se propuso abordar.

A continuación se reportarán los datos de tres indicadores que dan cuenta de la forma en que el docente organiza situaciones de aprendizaje:

- *Secuencia de clases para alcanzar los objetivos de aprendizaje de la unidad;*
- *Contribución de las actividades al logro de los objetivos de la clase filmada;*
- *Capacidad para reformular estrategias didácticas en base a las respuestas y resultados de los alumnos.*

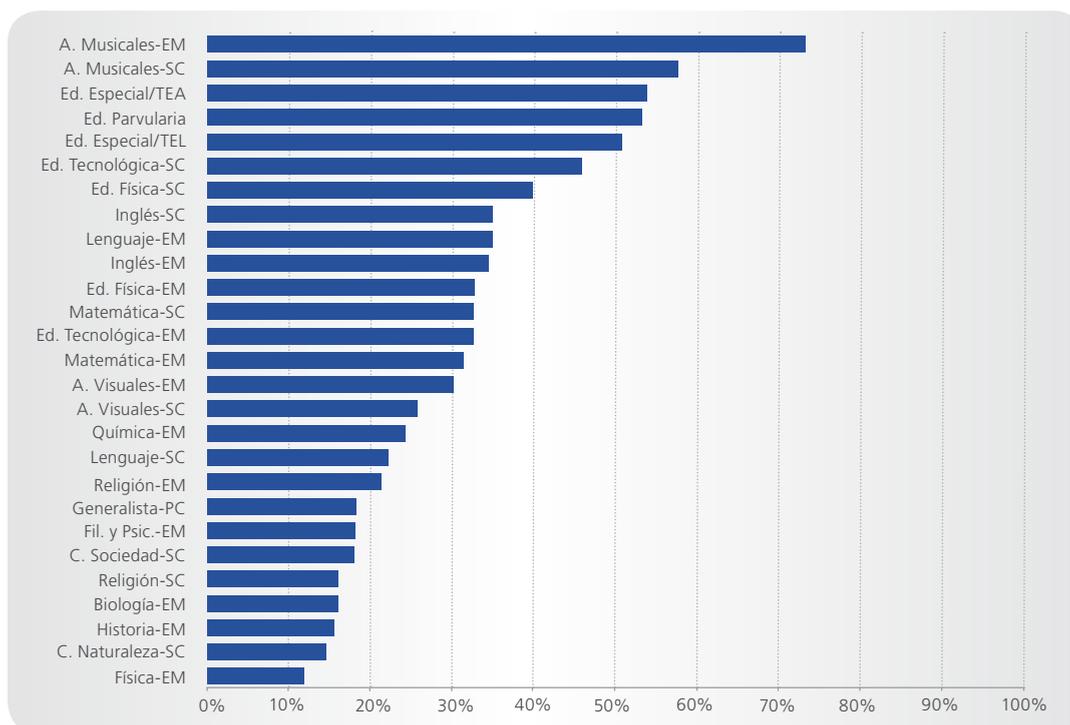
Secuencia de Clases

Este indicador evalúa la capacidad del docente para estructurar una secuencia de actividades y clases que favorezca el aprendizaje. Entre los casos seleccionados, solo un tercio alcanza el nivel esperado, aunque existe un amplio rango de variación desde el 11,8% al 73,1%, dependiendo del grupo evaluado (ver Gráfico 7). Entre aquellos que exhiben un mejor desempeño destacan los docentes de Artes Musicales (73,1% en EM y 57,5% en SC⁽¹³⁾), Educación Especial (TEA y TEL) y Educación Parvularia. En cambio, las proporciones más bajas de docentes que alcanzan el estándar se encuentran entre los evaluados del subsector Física (11,8%), seguidos por aquellos de Comprensión de la Naturaleza (SC), Historia (EM), Biología (EM) y Religión (SC).

.....

13 En adelante, al presentar los resultados para facilitar la lectura, se utilizarán las siguientes abreviaturas: PC= Primer Ciclo; SC= Segundo Ciclo; EM= Educación Media; TEL = Trastorno Específico del Lenguaje; TEA = Trastorno Específico del Aprendizaje.

● **Gráfico 7**
 ● *Proporción de docentes que alcanza un desempeño Competente o Destacado*
 ● *Indicador Secuencia de clases*



Un docente de buen desempeño en este indicador da inicio a su unidad pedagógica favoreciendo en los estudiantes un acercamiento significativo a los aprendizajes; ordena o secuencia las distintas clases de tal manera que a éstos les sea fácil seguirlo y al finalizar la unidad, realiza un cierre que promueve la integración y consolidación de los aspectos centrales trabajados.

En los casos de bajo desempeño, en cambio, es frecuente observar que no existe una adecuada articulación de las clases que componen la unidad, por lo que resulta difícil que los alumnos puedan establecer conexiones entre lo que han estado trabajando o puedan comprender globalmente el sentido de las distintas actividades que realizan en el marco de esa unidad (ver Ejemplo 1). En otros casos, una o más clases abordan habilidades o contenidos para los que se requeriría haber logrado el dominio de otros menos complejos en clases previas.

Otras debilidades detectadas en los docentes con bajo desempeño en este indicador se refieren a la forma en que inician o concluyen la unidad. Respecto de lo primero, se observa, por ejemplo, que al describir la primera clase, el profesor no rescata los conocimientos que los alumnos ya poseen sobre el nuevo contenido ni lo contextualiza en forma adecuada, exigiéndoles que se aboquen de manera inmediata, por ejemplo, a completar una guía de trabajo individual. En cuanto a la forma de concluir la unidad, es frecuente constatar que el docente realiza un cierre adecuado solo para la última clase, refiriéndose únicamente a lo trabajado en esa oportunidad, sin retomar los aspectos centrales de la unidad a modo de consolidación.

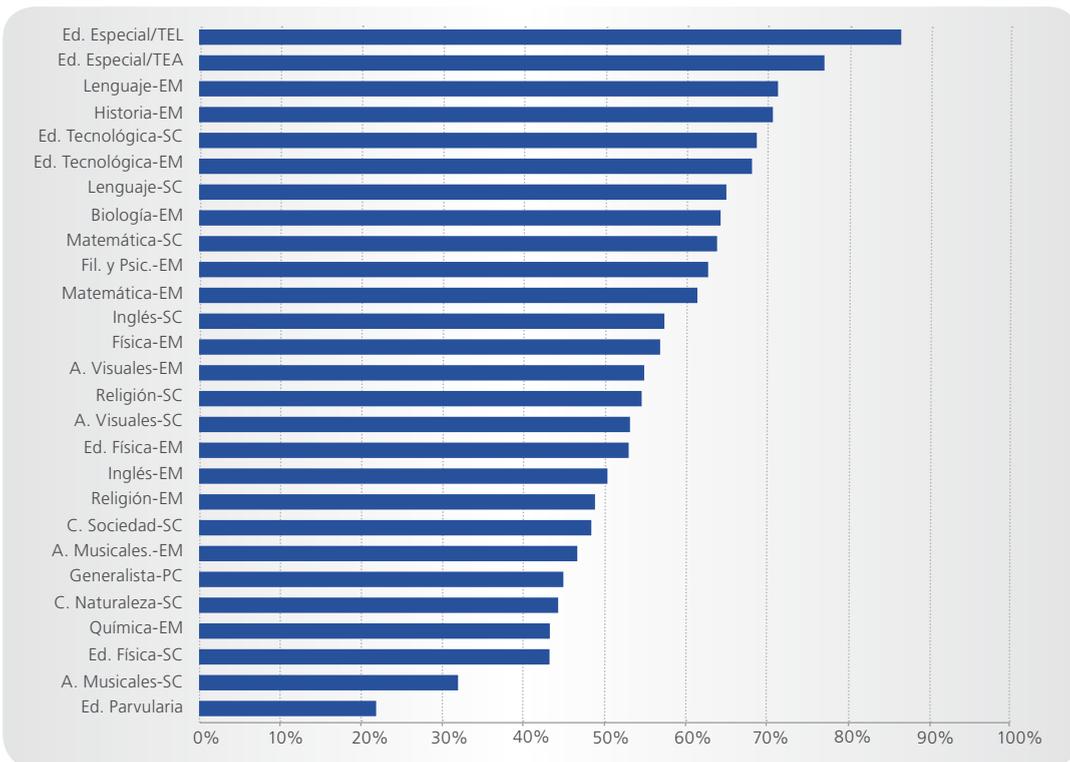
Ejemplo 1: “Época Moderna”	
<p>Un docente de Historia, Geografía y Ciencias Sociales de un octavo básico, describe la unidad en que ha trabajado con sus alumnos “Algunos de los grandes cambios de la época moderna” de la siguiente forma:</p> <p>Clase 1</p> <p><i>Objetivo: Analizar el proceso de Contrarreforma. Entrego a los alumnos una guía de trabajo en la cual, a partir de la información del texto escolar, responden: ¿Qué es la Contrarreforma?, ¿cuáles son las causas que propiciaron la Contrarreforma? Una vez que terminan de realizar la guía se revisan y comentan las respuestas con todo el curso.</i></p> <p>Clase 2</p> <p><i>Objetivo: Conocer los representantes del proceso de Reforma. Organizados en grupo, los alumnos realizan un cuadro comparativo de los principales representantes de la Reforma, según los siguientes criterios: representante principal, país en que se gestó, principales postulados. Al finalizar, se revisa la actividad a través de una puesta en común.</i></p> <p>Clase 3</p> <p><i>Objetivo: Conocer los postulados de algunos exponentes de la Ilustración.</i></p>	<p><i>A cada alumno le entrego distintos textos fotocopiados sobre exponentes de la Ilustración (Montesquieu, Locke, Rousseau). Les pido que subrayen el nombre del autor, el título del texto y la fecha de promulgación. Luego, les pido que escriban en sus cuadernos cuáles son los principales postulados de estos autores.</i></p> <p>Clase 4</p> <p><i>Objetivo: Conocer los postulados de Martín Lutero. Muestro un video con la historia de Martín Lutero, su vida como monje y cómo se desilusiona de la Iglesia a la que pertenecía. Luego del video, los alumnos expresan sus opiniones sobre algunos postulados del luteranismo (eliminación de algunos sacramentos, negación de la adoración a los santos, supresión del latín como idioma oficial de la Iglesia Católica, desconocimiento de la autoridad del Papa como máximo representante de Dios en la Tierra, etc.).</i></p> <p>En este caso no se observa una clara progresión en la secuencia de actividades de la unidad, por ejemplo, no se sitúa a los alumnos en el período, como marco inicial para la clase, tampoco se respeta un orden cronológico en la presentación de los hechos. Por su parte, en la clase final, en lugar de propiciar una consolidación de los aprendizajes centrales vistos en las clases anteriores, el docente retrocede al proceso de Reforma, abordando los postulados de Lutero.</p>

Contribución de las actividades a los objetivos de la clase

Este indicador evalúa, en la clase filmada, la capacidad del docente para diseñar e implementar actividades coherentes con los objetivos de aprendizaje planificados y que permitan cubrirlos.

Del conjunto de casos seleccionados, un 53% logra un desempeño de nivel Competente o Destacado, registrándose diferencias importantes entre los distintos grupos evaluados. La proporción más alta se observa entre las docentes de Educación Especial: 86,4% y 76,9% en TEL y TEA, respectivamente. También muestran proporciones comparativamente altas los profesores de Lenguaje (EM), Historia (EM) y Educación Tecnológica (SC y EM). En cambio, solo un 21,8% de docentes evaluados en Educación Parvularia y un 31,8% de profesores de Artes Musicales (SC), alcanzan el estándar esperado (ver Gráfico 8).

● **Gráfico 8**
 ● *Proporción de docentes que alcanza un desempeño Competente o Destacado*
 ● *Indicador Contribución de las actividades a los objetivos de la clase*



En los docentes de buen desempeño en este indicador se observa que todas las actividades desarrolladas por los estudiantes son pertinentes para los objetivos de aprendizaje y contribuyen a alcanzarlos, y el profesor se fija objetivos acordes al tiempo disponible lo que hace posible cubrirlos adecuadamente (ver Ejemplo 2). En cambio, se evidencian debilidades en el desempeño cuando el docente deja parte de los objetivos sin cubrir y/o incluye actividades que no contribuyen directamente a lograrlos (ver Ejemplo 3).

Ejemplo 2: “La vaca Emilia”

En la sesión filmada para la evaluación, una educadora especial aborda el objetivo: “Reconocer palabras que sirven para expresar cualidades (adjetivos)”.

En el aula de recursos, la docente trabaja con un grupo de 8 alumnos e inicia la sesión recordando: “La semana pasada conocimos palabras que se usan para nombrar. ¿Alguien me puede decir alguna palabra que usamos para nombrar?”. Varios niños responden en voz alta: *mesa, papá, auto*.

“Muy bien” dice la educadora y agrega: “utilizamos palabras para nombrar cosas, personas, animales y también utilizamos palabras que nos dicen cómo son esas cosas, personas y animales, palabras que nos dicen cualidades o características”. Luego pide a los niños que saquen los cojines de un estante y se sienten en círculo porque les va a contar un cuento:

D: Les voy a contar un cuento que se llama “Emilia, la vaca”. Había una vez una vaca llamada Emilia, que vivía en el campo, en un lugar con muchas flores de lindos colores. Emilia siempre estaba contenta porque tenía muchos amigos con los que se juntaba a jugar; era una vaca juguetona y amistosa. Sus mejores amigos eran el caballo, la gallina y la oveja. Emilia tenía un cuerpo blanco con grandes manchas negras. A diferencia de su amigo el caballo, que era forzado y muy veloz, ella era lenta y siempre perdía las carreras con sus amigos...” (continúa el relato).

Al finalizar la docente dice: “Ahora quiero que ustedes me cuenten cómo eran los personajes del cuento, qué características o cualidades tenían. Les voy a ir mostrando unas tarjetas y ustedes me van a ayudar a buscar a la vaca Emilia y a sus amigos”:

D: ¿Será ésta la vaca Emilia? (Muestra una lámina en que aparece una vaca de color negro sin manchas).

As: Noooo!

D: ¿Por qué?

A: No tiene manchas.

A: Es negra.

D: ¿Y cómo era la vaca Emilia?

A: Con manchas!

D: Con manchas, bien, Emilia tenía su cuerpo manchado. ¿Y cuál de éstas creen que es la vaca Emilia? (Muestra dos

tarjetas: una vaca acompañada de varios animales y otra sola).

A: Esa!

D: ¿Por qué?

A: Porque está con sus amigos!

D: Justamente, Emilia tenía muchos amigos, ¿cómo era entonces?...

A: Amistosa.

D: Y como jugaba mucho también sabemos que era...

A: Juguetona.

D: Sí, esa es otra característica de Emilia, ¿y ustedes cómo son?

A: Yo también soy juguetona, me gustaría jugar todo el día (risas), pero también soy estudiosa...

D: Bueno, esa es otra característica tuya. Juguetona, estudiosa... esas palabras nos están diciendo cómo eres. Y además de Carolina, ¿quién más es estudioso?

As: Yooooo (contestan a coro los ocho niños).

D: Qué bien! Tengo muchos alumnos estudiosos. ¿Y qué otras características tienen? Cuéntenme... ¿cómo son ustedes?

A: El profesor de gimnasia me dice que soy travieso... ¿qué es travieso?

El diálogo continúa por varios minutos y surgen nuevos adjetivos que los alumnos usan espontáneamente o con ayuda de la docente para describirse a sí mismos, a sus compañeros o al grupo, tanto física como psicológicamente (por ej. moreno, alta, perseverante, generosa, solidario, risueña, etc.). De esta forma se introduce el concepto de adjetivo.

En el ejemplo, las actividades de aprendizaje que realizan los alumnos permiten cubrir el objetivo propuesto para la sesión: a través de sus preguntas la educadora logra que ellos reconozcan palabras que se utilizan para expresar cualidades (adjetivos). Tanto la narración inicial como las preguntas asociadas al cuento y luego a las características del curso, se orientan directamente a desarrollar el objetivo planteado.

D=docente; A=alumno; As=alumnos.

Ejemplo 3: “Conociendo los números”

En su clase filmada, una educadora de párvulos aborda dos objetivos del Núcleo de Aprendizaje “Relaciones lógico-matemática y cuantificación” de las Bases Curriculares: (1) identificar y asociar números a la cantidad establecida, y (2) desarrollar el lenguaje matemático a través de situaciones cotidianas.

La educadora inicia la clase contando a los alumnos la historia de un payaso que no conoce los números y por eso no sabe cuántos globos tiene. Enseguida los invita a enseñar los números al payaso; para ello, les pide que cuenten del 1 al 10, luego les formula distintas preguntas como “¿Qué número viene después del 4? ¿Qué número es mayor el 6 o el 10?”.

Concluida esta actividad, la docente dice: “*Bueno, hay mucho que ustedes saben de los números. Ahora sí el payaso va a poder contar sus globos. Y ustedes también van a aprender más de los números. Ya vimos los números del 1 al 10, ahora vamos a ver números más grandes, los que siguen del 10...*” y propone a los niños jugar lotería. Reparte cartones con números del 1 al 19, y luego va sacando esferas con números desde una tómbola, las muestra a los alumnos y les pide que digan el número en voz alta. El juego se extiende por cerca de 20 minutos y despierta gran entusiasmo entre los alumnos.

En la última parte de la clase, la educadora indica a los niños que cada uno dibuje un payaso, lo pinte y luego lo recorte y pegue en uno de los muros de la sala.

En esta clase, se observa que la educadora realiza una actividad motivacional inicial (historia del payaso), otra de activación de conocimientos previos (preguntas sobre números del 1 al 10) y una tercera de identificación de los números en un contexto lúdico (lotería). Todas estas son actividades pertinentes y relevantes para los objetivos de aprendizaje que se propuso, pero ninguna aborda un aspecto de ellos: asociar los números a la cantidad establecida. Los alumnos nombran los números, pero no los vinculan a una cantidad o conjunto de elementos (por ej. nunca cuentan los globos del payaso y en la lotería solo relacionan los números con su nombre).

La actividad final, por otro lado, no hace una contribución suficientemente relevante para los objetivos de aprendizaje. Eso habría ocurrido si, además de dibujar y pintar, la docente hubiera incluido elementos directamente vinculados a los objetivos, por ejemplo, indicar a los alumnos que dibujen un payaso con 12 globos o que pinten 6 de un color y 6 de otro.

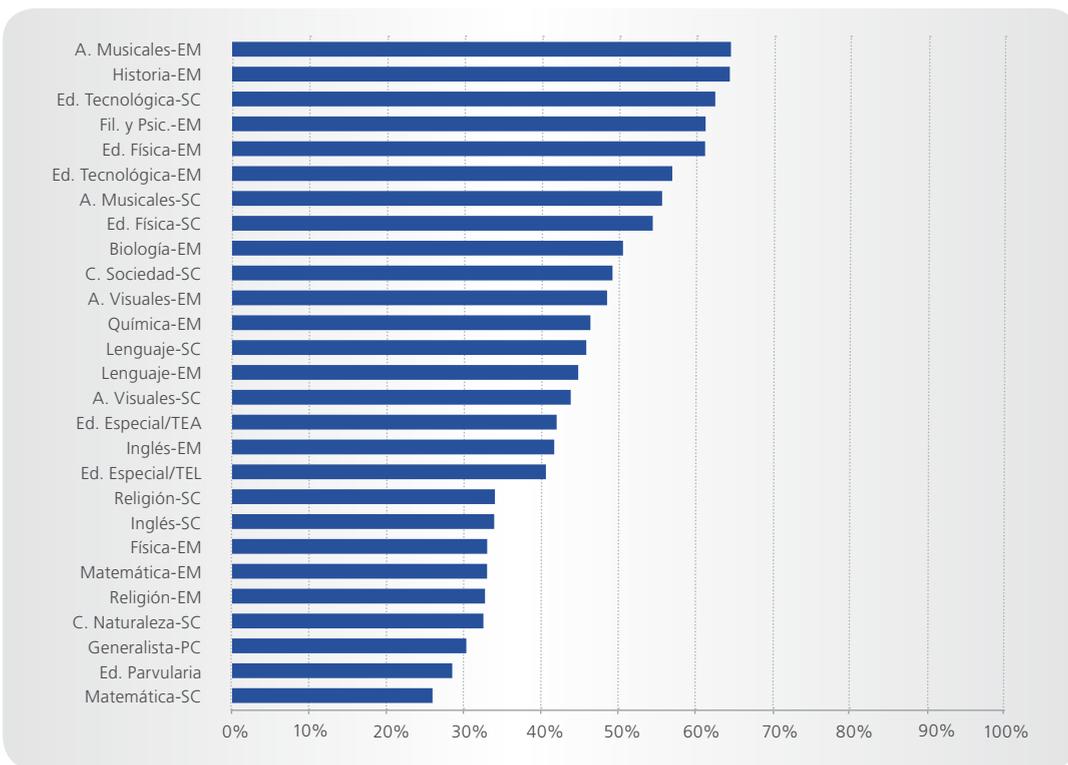
Capacidad para reformular estrategias didácticas

Este indicador evalúa si el docente es capaz de identificar y analizar un contenido (sea éste conceptual, procedimental o actitudinal) en que los estudiantes presentaron dificultades, y de proponer estrategias para abordarlas efectivamente rediseñando o ajustando la metodología de trabajo, las actividades o recursos utilizados, entre otros.

Del total de casos analizados solo 4 de cada 10 docentes evaluados alcanza en este indicador un desempeño Competente o Destacado, observándose diferencias en esa proporción desde el 25,9% al 64,5% entre los distintos grupos evaluados (ver Gráfico 9). Los grupos que muestran un mejor resultado corresponden a los docentes de Artes Musicales (EM), Historia (EM), Educación Tecnológica (SC), Filosofía y Psicología (EM) y Educación Física (EM). En estos casos es frecuente observar que el profesor puede dar cuenta de las dificultades enfrentadas por los alumnos y ajustar la enseñanza, formulando acciones pedagógicas consistentes con su análisis (ver Ejemplo 4).

En cambio, una proporción comparativamente baja alcanza el desempeño esperado entre docentes de Matemática (SC), Educación Parvularia y profesores generalistas de Primer Ciclo Básico. En los portafolios de estos docentes es más frecuente constatar que las nuevas actividades o las modificaciones propuestas, aunque pueden relacionarse con los contenidos, no son suficientes para abordar directamente las dificultades que presentaron los alumnos.

● **Gráfico 9**
 ● Proporción de docentes que alcanza un desempeño Competente o Destacado
 ● Indicador Capacidad para reformular estrategias didácticas



Ejemplo 4: “El proceso productivo”

En el Módulo 1 del Portafolio 2009, una pregunta pide al docente identificar aquel contenido o habilidad frente al cual sus alumnos presentaron mayores dificultades, describirlas e indicar a qué las atribuye y qué haría si tuviera la oportunidad de volver a trabajar ese contenido o habilidad con sus alumnos. Un profesor de Educación Tecnológica responde:

“A mis alumnos les costó mucho comprender los conceptos centrales del proceso productivo, confundían unos con otros y los utilizaban como sinónimos. Pienso que esta dificultad se pudo deber a que, al momento de definir estos conceptos, lo hice de manera muy técnica y abstracta y no pudieron comprender bien qué era lo propio de cada uno.

Para corregir esto, en otra clase les mostraría un power point con imágenes del proceso productivo del chicle o la jalea sin sabor, que fueron los temas que más les interesaron. También iría ejemplificando con imágenes los elementos claves del proceso como: los materiales procesados, el plan de acción y las técnicas de fabricación, para que pudieran ver de manera más concreta a qué se refiere cada uno de esos conceptos.

A medida que fueran entendiendo de qué se trata cada uno, construiríamos en conjunto y utilizando sus propias palabras, un glosario de términos, y lo iríamos

complementando o precisando con los conceptos más técnicos. Así comprenderían mejor los conceptos y los términos, y tendrían un glosario al cual recurrir en caso de dudas”.

En este ejemplo, el profesor identifica claramente una dificultad de los alumnos durante el proceso de aprendizaje y lo vincula con una falencia en la estrategia pedagógica que utilizó (su forma de definir los conceptos fue “muy técnica y abstracta”). Luego, en consistencia con su análisis, el docente propone una nueva estrategia para presentar los conceptos de una manera que resulte más cercana y accesible para los alumnos, recogiendo los temas que les interesaron más y usando ejemplos durante su explicación. También incorpora el uso de imágenes, para ilustrar los conceptos de una manera más concreta. Finalmente, incluye una actividad de definición de conceptos, dando a los alumnos una participación activa (deben elaborar las definiciones en conjunto con el profesor y utilizando sus propias palabras). La nueva actividad propuesta recoge claramente el análisis realizado e incluye una serie de aspectos que favorecen la apropiación de los contenidos por parte de los alumnos (participar en la definición de conceptos, usar sus propias palabras y, al mismo tiempo, precisarlas o completarlas con los términos técnicos).

Ambiente instruccional

El Marco para la Buena Enseñanza incluye el dominio “Creación de un ambiente propicio para el aprendizaje”, que refiere a la capacidad del docente para generar un clima de aceptación, equidad, confianza, solidaridad y respeto. También exige que éste manifieste altas expectativas sobre las posibilidades de aprendizaje de todos sus estudiantes, mantenga una convivencia apropiada en el aula y organice un ambiente de trabajo estructurado.

La investigación y la literatura educacional señalan que el proceso de enseñanza requiere que el profesor sea capaz de crear y mantener en el aula un clima favorable para el aprendizaje (Marzano, Marzano & Pickering, 2003). Esto implica considerar variados aspectos: desde la disposición física adecuada de la sala y los recursos de aprendizaje, hasta la creación de un clima socioafectivo confiable y positivo, en que los alumnos se sientan seguros y respetados. La forma en que el profesor se relaciona con sus estudiantes y el tipo de relación que promueve entre ellos, son claves para generar y mantener un clima adecuado en el aula (Marzano, D’Arcangelo & Sheets, 2004). Un docente efectivo promueve normas de convivencia y comportamiento claras, que favorecen la focalización en el aprendizaje y el buen desarrollo de las actividades, y utiliza estrategias apropiadas cuando esas normas se rompen (Stage & Quiroz, 1997), para permitir un flujo ordenado de la clase y un uso eficiente del tiempo disponible. También, propicia oportunidades de participación para los estudiantes, de modo que éstos se comprometan intensa y activamente en su propio aprendizaje (Wang, Haertel & Walberg, 1997). Permite que todos ellos, no solo los más aventajados, puedan plantear sus preguntas y expresar sus ideas y opiniones, e incentiva a los alumnos a escuchar a sus compañeros con atención y respeto.

La evaluación del desempeño del profesor en este ámbito se realiza a partir de la clase filmada que entrega en su Portafolio (Módulo 2). En el Manual Portafolio, se indica explícitamente al profesor que desarrolle una clase que permita apreciar tanto su desempeño como la participación de sus estudiantes y su interacción con ellos. A continuación se da cuenta de los datos recogidos en dos indicadores que evalúan el desempeño de los docentes en lo referido al *Ambiente instruccional*:

- *Manejo adecuado del grupo curso para favorecer la focalización en las actividades de aprendizaje y el desarrollo fluido de las actividades;*
- *Promoción de la participación de los alumnos durante la clase, a través de la formulación de preguntas, pedir su opinión y/o incorporar sus experiencias e intereses.*

Manejo del grupo curso

En este indicador se evalúa la capacidad del docente para lograr que los estudiantes se involucren en las actividades de aprendizaje y permanezcan focalizados en ellas. Para esto, es importante que durante la clase se respeten las normas de convivencia y comportamiento acordadas y que el profesor logre que los alumnos presten atención a él y a sus compañeros cuando las actividades así lo requieren.

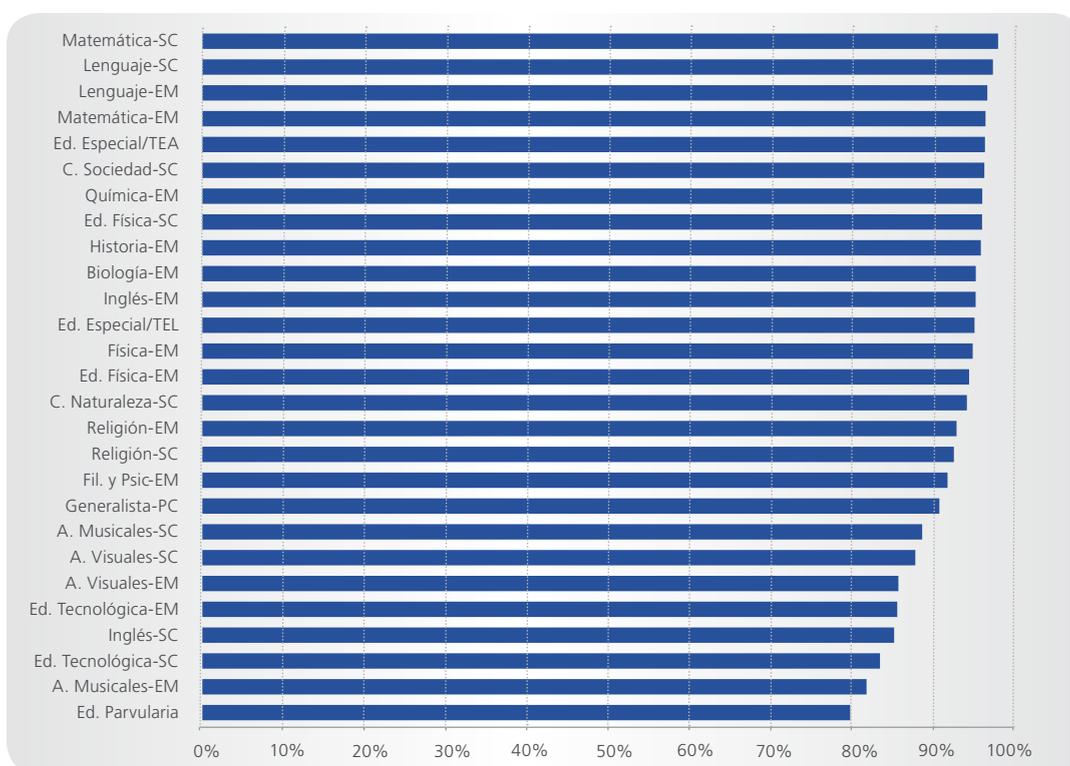
En la gran mayoría de las clases filmadas del portafolio se aprecia un alto desempeño en este indicador: como muestra el Gráfico 10, un 92% de los docentes cuyos datos fueron analizados alcanza el nivel Competente o Destacado, observándose la proporción más baja en Educación Parvularia (79,7%) y la más alta entre los profesores de Matemática (SC) (97,9%).

En general, durante las clases filmadas para el Portafolio los docentes emplean estrategias efectivas para atraer y mantener la atención de sus alumnos, especialmente cuando les entregan instrucciones o desarrollan explicaciones; también permanecen atentos a la dinámica de la clase y reaccionan de manera oportuna y eficaz cuando se producen quiebres en las normas de comportamiento. Esto sugiere que los profesores pueden implementar estrategias efectivas para establecer y mantener un clima que propicie el aprendizaje en sus clases. Esto podría estar influido por el contexto particular en que se lleva a cabo

la clase grabada para el Portafolio. Dada la relevancia que ésta tiene para la evaluación del docente, en la mayor parte de los casos, tanto profesor como grupo curso tienden a mantener un clima de trabajo ordenado y fluido durante la filmación.

En el reducido grupo de docentes que obtiene un resultado Básico o Insatisfactorio es frecuente constatar que el profesor inicia la clase o entrega instrucciones mientras en el aula existe un alto nivel de ruido que dificulta escuchar lo que dice, de modo que un número importante de alumnos no logra atender a las indicaciones necesarias para desarrollar su trabajo. También es frecuente que, debido al uso de prácticas poco eficaces por parte del docente, durante períodos prolongados o en reiteradas ocasiones los alumnos conversan entre sí mientras el profesor u otro alumno se dirigen al curso, lo que perjudica su focalización en las actividades y el aprovechamiento del tiempo de clase para el aprendizaje.

● **Gráfico 10**
 ● *Proporción de docentes que alcanza un desempeño Competente o Destacado*
 ● *Indicador Manejo del grupo curso*



Promoción de la participación de los alumnos

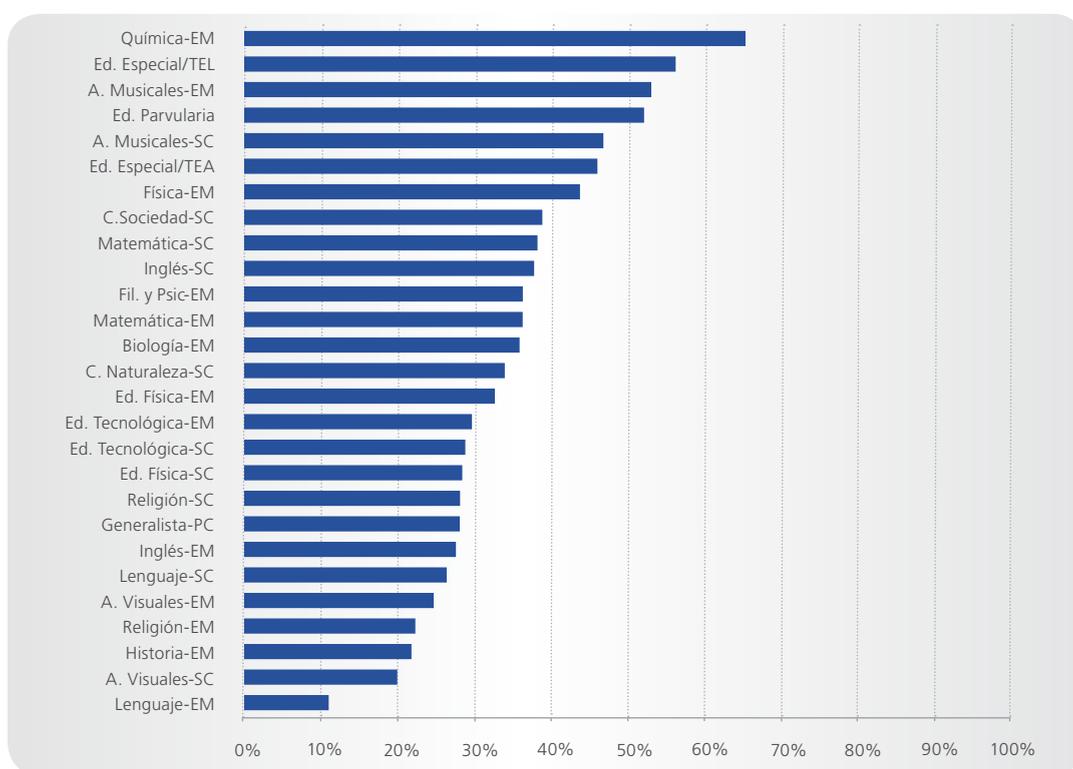
En este indicador se pone en juego la capacidad del profesor para ofrecer oportunidades de participación para todos sus estudiantes, que éstos puedan tomar la palabra en la clase, expresar sus opiniones, plantear y responder preguntas, de modo que asuman un rol activo en su propio aprendizaje y el de sus pares.

Solo un 35% del total de los casos analizados alcanza el nivel de logro esperado en estos aspectos, constatándose diferencias importantes en los distintos grupos evaluados: desde un 11% a un 65,2% (ver Gráfico 11). La proporción que alcanza el estándar esperado es comparativamente más alta entre los docentes de: Química (EM), Educación Especial (TEL), Artes Musicales (EM) y Educación Parvularia. En estos casos, se aprecia en las clases un alto nivel de participación de los alumnos; aun en clases de tipo expositivo, el docente les formula preguntas constantemente para confirmar que están siguiendo sus explicaciones

y para obtener información sobre el nivel de aprendizaje o comprensión que alcanzan. Por su parte, los alumnos demuestran interés por participar y compartir sus opiniones, inquietudes y experiencias, y existen oportunidades equitativas para que todos puedan hacerlo, no solo aquellos de mejor rendimiento.

Los más bajos resultados se observan, en cambio, entre los docentes de Lenguaje (EM): solo 1 de cada 10 alcanza un nivel Competente o Destacado en el indicador, seguidos por los profesores de Artes Visuales (SC), Historia (EM) y Religión (EM). En estos casos, es frecuente observar que los docentes desarrollan clases con escasas instancias de participación para los alumnos: el protagonismo reside total o casi exclusivamente en el profesor, manteniéndose los estudiantes como actores pasivos, con pocas instancias para intervenir, preguntar u opinar. En otros casos, el docente favorece la participación única o preferente de los alumnos más aventajados de la clase.

● **Gráfico 11**
 ● *Proporción de docentes que alcanza un desempeño Competente o Destacado*
 ● *Indicador Promoción de la participación de los alumnos*



Interacción pedagógica

Este ámbito se vincula especialmente con el dominio “Enseñanza para el aprendizaje de todos los estudiantes” del Marco para la Buena Enseñanza. En él se señala que, para ser efectivo, un docente debe ser capaz de organizar situaciones interesantes y productivas que favorezcan la indagación, interacción y socialización de los aprendizajes entre los estudiantes. Estas situaciones deben considerar los saberes e intereses de éstos y proporcionarles recursos adecuados y apoyos pertinentes. Asimismo, en el dominio “Creación de un ambiente propicio para el aprendizaje” se destaca la importancia de las interacciones que ocurren en el aula, tanto entre docente y estudiantes como entre éstos.

La investigación educacional reciente ratifica lo dicho. El aula es un espacio en el que deben generarse oportunidades de aprendizaje y desarrollo para los estudiantes, las cuales dependen críticamente de la capacidad del profesor para ser un mediador efectivo en la construcción de conocimientos (de Corte, 2010). Un docente competente organiza situaciones en las cuales los alumnos tienen la posibilidad de resolver problemas por sí mismos, enfrentar nuevos desafíos, sacar conclusiones, etc., instancias orientadas a favorecer en ellos el desarrollo del pensamiento y de habilidades más complejas. Para ello se requiere, en primer lugar, que el profesor posea un dominio profundo y actualizado de la disciplina que enseña de tal manera que pueda desarrollar explicaciones claras, conceptualmente correctas y proponer situaciones de aprendizaje que consideren los intereses, experiencias y aprendizajes previos de los estudiantes (Bransford, Brown & Cocking, 2000; Schneider & Stern, 2010). También es necesario que sea capaz de fomentar una participación de calidad. Esto implica no solo que los alumnos tengan oportunidad de opinar o responder las preguntas del profesor, sino que sus intervenciones sean aprovechadas para precisar, profundizar y complementar los conceptos o conocimientos abordados (Nystrand & Gamoran, 1991). Una interacción pedagógica efectiva también requiere que el docente esté atento al trabajo y resultados de sus estudiantes, los retroalimente y apoye en forma constante y sistemática para avanzar hacia el logro esperado, por ejemplo, que los observe, les de pistas, modele, etc. (Collins, 2006).

Para calificar el desempeño en el ámbito de la Interacción pedagógica, en el Módulo 2 del Portafolio (clase filmada), el docente entrega evidencia tanto de su propio desempeño como de la participación de los alumnos. Con este fin, al menos durante 10 minutos continuos o espaciados a lo largo de la clase, debe dirigirse al curso para exponer, explicar, sintetizar o profundizar los conocimientos abordados. En las directrices que recibe el docente para realizar esta clase, se le indica que la grabación debe mostrar cómo entrega instrucciones a sus alumnos y acompaña su trabajo, cómo enseña un contenido, habilidad o procedimiento nuevo, entre otros, de modo que sea posible evaluar los indicadores de interacción.

Para evaluar la Interacción pedagógica observada en la clase filmada, se consideran distintos indicadores, tres de los cuales se reportan a continuación:

- Calidad de las explicaciones que desarrolla el profesor para abordar los contenidos y habilidades de aprendizaje de la clase;
- Calidad de las interacciones que promueve el profesor en su diálogo directo con los alumnos, como en la relación que favorece entre ellos;
- Acompañamiento de las actividades en función del aprendizaje, a través del monitoreo y apoyo al trabajo de los alumnos.

Calidad de las explicaciones

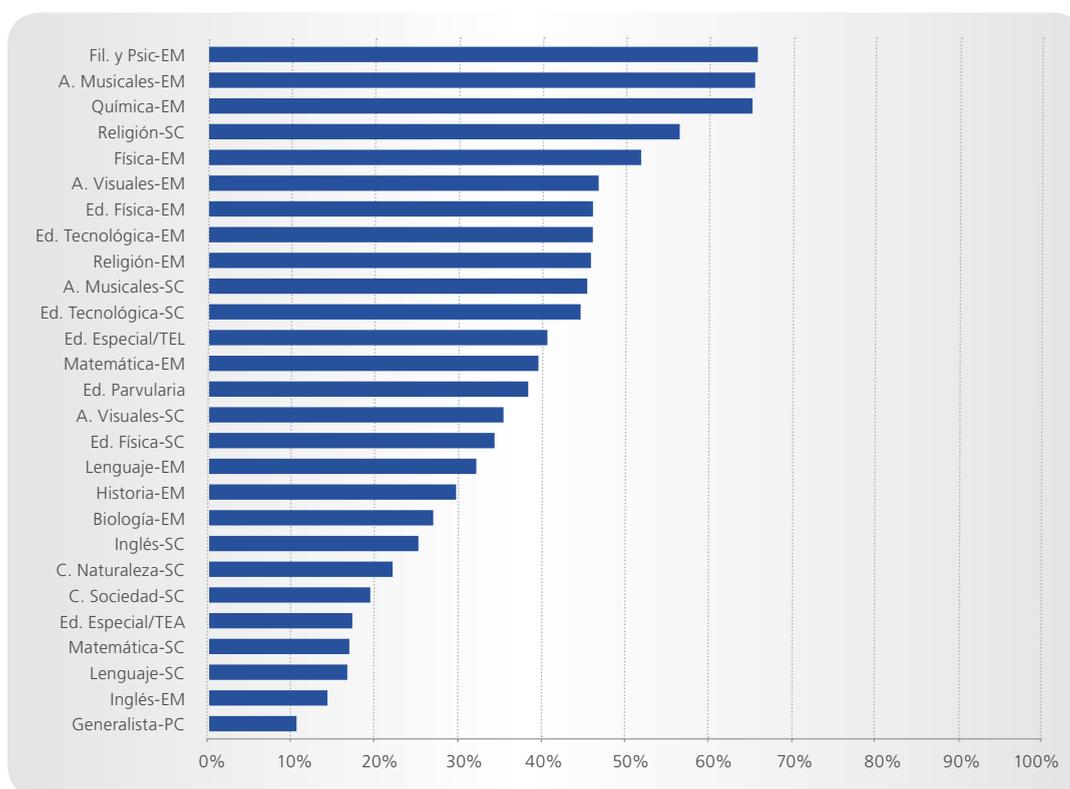
Este indicador se evalúa en instancias en que el docente enseña a sus estudiantes nuevos conocimientos e implica observar cómo vincula esos conocimientos con aprendizajes previos, qué tipo de preguntas formula, qué ejemplos utiliza, la claridad con que expone, modela o demuestra algún contenido o procedimiento, entre otros. A partir de estas situaciones también es posible constatar el dominio que el profesor exhibe de su disciplina.

Los resultados indican que un 28% del total de los docentes evaluados logra el nivel de competencia esperado en este indicador. Los subsectores con una mayor proporción de docentes con buenos resultados corresponden a Filosofía y Psicología (EM), Artes Musicales (EM), Química (EM), Religión (SC) y Física (EM). Frente a esto, llama la atención los bajos resultados entre los docentes generalistas de Primer Ciclo, pues solo 1 de cada 10 logra el nivel esperado. Situación similar ocurre en los subsectores Inglés (EM), Lenguaje (SC), Matemática (SC) y Educación Especial (TEA) (ver Gráfico 12).

Los docentes que alcanzan el desempeño esperado muestran dominio de su disciplina; se observa en ellos el desarrollo de explicaciones claras y completas, utilizan conceptos con rigurosidad y precisión y cuentan con un vasto repertorio de ejemplos para apoyar sus explicaciones. Cuando es necesario, utilizan demostraciones o modelan procedimientos para favorecer en los estudiantes la comprensión y apropiación de los nuevos contenidos o habilidades abordadas. También se observa en ellos capacidad para promover aprendizajes significativos en sus alumnos, recurriendo a sus conocimientos o experiencias así como estimulándolos a establecer conexiones de diferencia, similitud o regularidad, entre otras (ver Ejemplo 5).

En las clases con bajo desempeño se observa en cambio que, si bien los docentes evocan en los alumnos conocimientos o experiencias relacionadas con los nuevos aprendizajes, las relaciones que establecen entre ellas son parciales, superficiales o insuficientes para facilitar la apropiación y comprensión de los contenidos (ver Ejemplo 6).

• **Gráfico 12**
 • *Proporción de docentes que alcanza un desempeño Competente o Destacado*
 • *Indicador Calidad de las explicaciones*



Ejemplo 5: “El arcoíris”

En una clase de Física de Primer año Medio se observa la siguiente interacción entre la profesora y sus estudiantes:

D: ¿Recuerdan que la clase pasada hicimos un experimento en que llegamos a ciertas conclusiones sobre la trayectoria de la luz?

A: ¿Con el experimento de la lata y la vela?

D: Exacto, vimos que la luz se propaga en línea recta y con velocidad constante cuando “viaja” por un medio homogéneo. ¿Se acuerdan qué significa que un medio sea homogéneo?

A: Que es siempre igual, es puro...

D: Justamente. Es el que mantiene sus mismas propiedades en todos sus “puntos”. Ahora, la atmósfera, ¿es homogénea?

A: Sí, es puro aire...

A: No! Hay smog y nubes...

D: A ver, como dice Felipe, no es homogéneo. Cuando llueve, por ejemplo, hay gotas de agua suspendidas. Todos deben haber visto alguna vez que cuando sale el sol después que llueve sale un arcoíris, ¿cierto?

As: Sí! (Contestan a coro varios alumnos).

D: Después que llueve, la luz del sol, que viene viajando en línea recta por el aire, incide en las gotas de agua suspendidas. ¿Qué pasará con la trayectoria de la luz al incidir en las gotas?

A: Sigue recta, pero como pasó por el agua, se ven los distintos colores del arcoíris...

D: ¿Qué dicen los demás? ¿Están de acuerdo con Karina?

A: No, porque ya no es puro aire, ya no sigue recta...

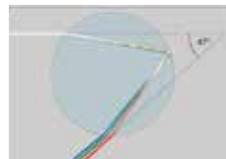
D: Muy bien. La luz, que viene viajando por el aire, penetra otro medio distinto, como serían las gotas de agua, cambiando su dirección y velocidad. Eso se llama refracción. El arcoíris es un ejemplo de refracción.

¿Recuerdan cómo se compone la luz blanca?

A: De distintos colores...

D: Bien! La luz blanca está compuesta por diversas longitudes de onda, cada una asociada a un color. Como ven aquí, en este dibujo, al incidir en la gota de agua, las ondas se refractan y desvían en distintos ángulos, luego se reflejan al interior de la gota y finalmente, al salir, se vuelven a refractar. Esto se llama dispersión cromática y también se puede observar al hacer pasar la luz blanca a través de un prisma triangular.

Luego de esta explicación, los alumnos realizan un experimento en el cual construyen un prisma de agua, utilizando una cubeta con agua y un espejo, buscando la proyección de la luz del sol sobre una pared.



Comparten sus conclusiones respecto de su semejanza con el fenómeno del arcoíris.

En el ejemplo, se observa que la docente recurre a los conocimientos que poseen los alumnos sobre “medio homogéneo” y “luz blanca” como una forma de favorecer la comprensión del nuevo concepto de refracción que abordará. En la clase, a través de preguntas los va guiando para que relacionen este nuevo concepto con lo que aprendieron en una clase anterior a partir de un experimento. A su vez, recoge lo que ellos conocen respecto del fenómeno del arcoíris y utiliza esta experiencia, conocida para todos, para ejemplificar el fenómeno de la refracción.

Ejemplo 6: “El tutti-frutti”

En una clase de Lenguaje y Comunicación de 2º básico, la profesora aborda la estructura de la receta como texto y comienza la clase mostrando una receta a los alumnos:

D: ¿Han visto un texto como éste alguna vez?

As: Sí! (a coro).

D: ¿Dónde?

A: En una revista del diario que hay en mi casa.

D: Ya... ¿Y saben cómo se llama?

A: Es para cocinar!

D: Ya, pero cómo se llama...

A: Es una receta para preparar comida.

D: Muy bien! Es una receta. ¿Qué recetas conocen?

A: El año pasado hicimos un tutti-frutti...

D: Ah, claro! El año pasado ocupamos una receta para hacer un tutti-frutti”. ¿Les gustó hacer eso?

As: Sí! (a coro).

D: ¿Cómo nos quedó?

A: Muy rico!

D: ¿Se acuerdan qué tenía?

Distintos alumnos mencionan: manzana, plátano, duraznos.

D: Eso, tenía muchas frutas ricas ¿cierto? Bueno, ahora yo les voy a contar sobre las recetas. En una receta primero hay un nombre... (muestra en la receta). Luego, viene una lista con los ingredientes que se necesitan

para preparar la receta, acá (lo muestra) y, al final, están las instrucciones que debemos seguir para poder cocinar. ¿Entendieron?

As: Sí (varios)

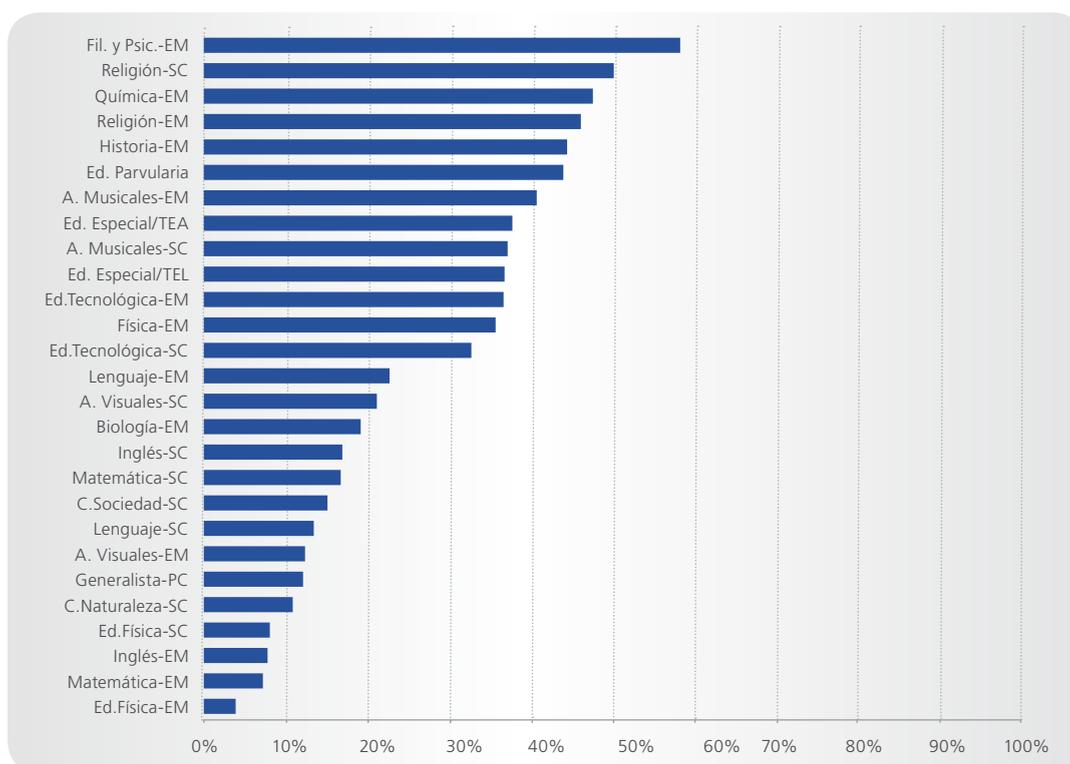
D: Ahora yo les voy a entregar una receta a cada uno, para que pinten las partes que aprendimos con distintos colores.

En este caso, la profesora hace preguntas a los alumnos para activar sus conocimientos previos relacionados con el contenido de la clase (estructura de la receta de cocina), sin embargo, las preguntas no promueven que ellos establezcan relaciones explícitas entre este contenido y lo que ya conocen. Para ello, la docente podría haber preguntado “¿Qué nos dice esto?”, mostrando el nombre de la receta o “¿Para qué nos sirve esta parte?” refiriendo a las instrucciones. Cuando los alumnos mencionan las frutas con que hicieron el tutti-frutti, pudo indicar que esos son los “ingredientes”, mostrarlos en la nueva receta, y entonces preguntarles qué son los ingredientes y ayudarlos a comprender por qué es importante que una receta los incluya, etc. De esa forma los alumnos efectivamente habrían empleado sus conocimientos previos y tendrían más oportunidades para comprender los elementos esenciales del texto de una receta.

Calidad de las interacciones

Este indicador evalúa tanto la calidad de las preguntas que utiliza el docente para fomentar el aprendizaje en aula, como el aprovechamiento que hace de las intervenciones de los estudiantes, ya sea de sus respuestas o comentarios espontáneos. Esto se puede observar en instancias en que el profesor interactúa con un solo alumno, con un subgrupo o cuando se dirige a todo el curso. Los datos dan cuenta de debilidades importantes en la mayoría de los evaluados, encontrándose solo 2 de cada 10 docentes con un desempeño Competente o Destacado. De los casos analizados, únicamente el subsector Filosofía y Psicología (EM) destaca con el más alto porcentaje de docentes que alcanza o supera el estándar (58,1%) mientras que hay varios subsectores que no alcanzan a reunir un 10% de docentes con buen desempeño, situación observada en Educación Física (EM y SC), Matemática (EM) e Inglés (EM) (ver Gráfico 13).

● **Gráfico 13**
● *Proporción de docentes que alcanza un desempeño Competente o Destacado*
● *Indicador Calidad de las interacciones*



En las clases filmadas, se observa que la mayor parte de los docentes establecen interacciones con sus alumnos que no favorecen el desarrollo del pensamiento o no constituyen un aporte significativo para enriquecer su aprendizaje, sea éste de un contenido conceptual o procedimental. Así, por ejemplo, estimulan preferentemente la ejecución mecánica de un procedimiento, o bien, la mera repetición de información a través de preguntas que no exigen elaboración. En otros casos, el docente tiende a indicar la respuesta correcta o a inducirla en lugar de guiar a los alumnos hacia la solución (ver Ejemplo 7). Con esto, se pierde la oportunidad de incentivar el razonamiento de los estudiantes y de aprovechar sus dudas y errores como una fuente de aprendizaje. Los docentes con buen desempeño, en cambio, realizan preguntas que permiten a los estudiantes plantearse hipótesis, indagar y extraer conclusiones, mostrando una actitud receptiva ante las preguntas, ideas o errores de éstos. Así también son hábiles para utilizar las interacciones entre los alumnos como una forma de favorecer y ampliar sus aprendizajes, propiciando que aclaren dudas o desarrollen explicaciones entre ellos (ver Ejemplo 8).

Ejemplo 7: “Metáfora y comparación”

Una profesora desarrolla su clase de Lenguaje en un 6° básico y su objetivo es que los alumnos interpreten distintos fragmentos de poemas y distingan la metáfora de la comparación.

Al iniciar la clase, la docente les comenta que van a trabajar con power point ya que les va a proyectar algunos versos.

D: La clase pasada conocieron lo que era metáfora y comparación, ¿cierto?

As: Sí (contestan a coro los alumnos)

D: ¿Quién me puede decir en qué consiste la metáfora? (Un alumno lee de su cuaderno la definición). ¡Muy bien! ¿Quién puede leer lo que escribimos para comparación?

(Otro alumno lee de su cuaderno). Perfecto! Espero que todos tengan bien anotadas estas definiciones en sus cuadernos.

Ahora vamos a leer algunos versos, ¿qué dice ahí? (Señalando la primera lámina)

A: (lee) Tus ojos son como el crepúsculo dorado.

D: Tus ojos son como el crepúsculo dorado (repite). ¿Dice algo bonito, cierto? ¿Les gusta?

As: Sí!(contestan a coro)

D: ¿Qué es? ¿Una metáfora o una comparación?

A: Una comparación.

D: Ya... bien, una comparación, ¿por qué?

A: Porque usa el “como”.

D: Ya, y ésta que leemos acá, qué es... (“Gracias a la vida que me ha dado tanto. Me dio dos luceros, que cuando los abro, perfecto distingo lo negro del blanco”).

A: Comparación

D: No pues, esta es una metáfora, tienen que fijarse bien para no equivocarse. Fíjense que dice: me dio dos luceros... Me dio dos ojos! ¿Habían escuchado esto antes?

As: Sí (a coro).

A: Es una canción bien conocida.

D: Sí, es bien conocida. A ver, lean este otro: “Me rodearon tus brazos como una enredadera”, ¿qué es?

As: Comparación... (a coro).

D: Muy bien, ven que saben harto. ¿Les gusta la poesía?

A: Sí!

D: ¿Y para qué sirve la poesía?

(Silencio)

D: ¿Para qué sirve la poesía?, ¿quién me puede responder? (Distintos alumnos contestan en forma casi simultánea).

A: Para recitar.

A: Para decir lo que sentimos.

A: Para conquistar a las mujeres (los alumnos ríen).

D: ¿Para qué más?

A: Cuando éramos más chicos nos hacían recitar poesías en los actos!

D: Claro, para el día de la madre, generalmente los niños se aprenden poesías y se las recitan. Bueno, ahora van a trabajar en esta guía.

En este ejemplo se observa que las preguntas que realiza la docente solo le permiten abordar de modo superficial los conceptos de metáfora y comparación con sus alumnos. En lugar de trabajar con ellos el significado de cada elemento les enseña a reconocer mecánicamente la figura literaria “comparación” por la presencia de la palabra “como”. Tampoco aborda claramente el fenómeno de sustitución que subyace a la metáfora, característico de esta figura literaria.

Las preguntas que hace la docente no favorecen una reflexión o razonamiento por parte de los alumnos sino más bien evocan respuestas puntuales y superficiales. Además, desaprovecha las respuestas incorrectas de los alumnos para hacerlos razonar y en su lugar les entrega la respuesta correcta, así como tampoco enriquece las respuestas de sus alumnos. Por ejemplo, cuando ella pregunta ¿para qué sirve la poesía?, no aporta riqueza o contenido a las respuestas que dan los alumnos.

Ejemplo 8: “La libertad”

En un cuarto medio, un docente luego de mostrar un video sobre los derechos humanos, invita a discutir a sus alumnos sobre el concepto de libertad.

D: ¿Qué piensan del concepto de “libertad”? ¿Por qué creen que los seres humanos tenemos este derecho?

A: ¿Y quién dijo que tenemos este derecho? (Alumnos ríen)

D: A ver Claudia, ¿tú crees que los seres humanos no gozamos de libertad?

A: Mmm no tanto!... Tenemos que hacer lo que es correcto hacer... lo que está establecido.

D: ¿Qué piensan los demás?

A: Igual tenemos la capacidad de elegir, de decidir qué queremos hacer.

A: Los jóvenes cada vez tenemos más libertad.

D: ¿Cómo así? ¿Nos podrías dar un ejemplo?

A: Este año salimos de cuarto y damos las PSU y ahí algunos van a decidir entrar a la universidad, a un instituto, otros quizá decidan trabajar.

A: Pero eso no es algo que uno pueda decidir y que se haga tal cual ya que depende de otras cosas como la plata, el puntaje.

D: Claro, en este caso hay varios factores que influyen en esta decisión y quizá, entre ustedes existan distintos niveles de libertad para tomar una decisión como ésta. Veamos, ustedes a esta edad, ¿en qué ámbitos sienten que gozan o cuentan con libertad?

A: Para vestirnos, para pensar.

A: Para elegir a los amigos, a las pololas.

(Alumnos ríen)

A: Sí, tenemos la capacidad de elegir, de pensar; si quiero o no quiero hacer algo.

D: ¿Esa capacidad de elegir o de ser libres de dónde viene?

¿Quién puede agregar algo más de lo que se ha dicho?

A: Yo creo que tenemos la capacidad de darnos cuenta de qué es lo que debemos hacer para lograr nuestras metas y qué queremos hacer.

D: Tienes mucha razón en lo que dices, los seres humanos tenemos conciencia de nuestros actos, eso quiere decir que tenemos algo así como una brújula interna que nos permite ir definiendo qué debemos hacer para lograr nuestras metas. Es por esta razón que los seres humanos somos libres. Ahora, por naturaleza tenemos el derecho a ser libres. Ustedes recién decían que ese derecho a ser libres no se cumple para todos por igual, ¿cierto?

A: No, yo creo que en la práctica ese derecho no se cumple para todos.

D: ¿Qué piensa el resto? ¿Están de acuerdo con su compañero?

A: Yo no estoy de acuerdo, porque depende de uno. Al final si uno se empeña en hacer algo, si se esfuerza en lograr lo que se propone, si busca ayuda, al final lo va a lograr.

D: A ver, tenemos opiniones que expresan fundamentos distintos de la libertad. Lo que vamos a hacer es organizarnos por grupos y, a partir de las temáticas abordadas, haremos un debate sobre la forma en que se expresa la libertad.

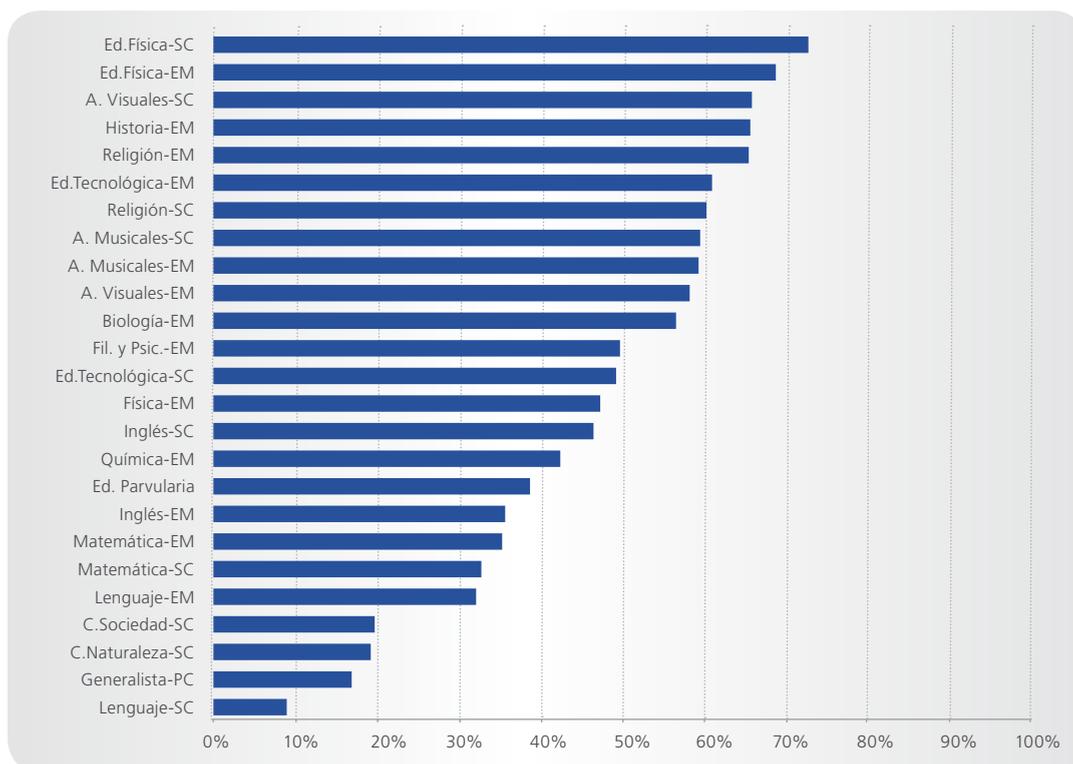
En este caso, el docente logra involucrar a sus alumnos en el tema de la clase, acoge la diversidad de opiniones y, en función de las respuestas de los alumnos, les va haciendo nuevas preguntas, los incentiva a fundamentar, dar sus opiniones, y comentar o rebatir lo que piensan sus compañeros.

Acompañamiento de las actividades en función del aprendizaje

En este indicador se evalúa la forma en que el docente monitorea, supervisa y recoge información durante la clase a partir del trabajo que desarrollan sus estudiantes. Un 33% de los evaluados alcanza o supera el nivel esperado en este indicador. Sobresalen por un alto desempeño los subsectores Educación Física (SC y EM), Artes Visuales (SC), Historia (EM), Religión (EM) y Educación Tecnológica (EM). En todos ellos al menos 6 de cada 10 docentes logra el nivel de desempeño esperado (ver Gráfico 14).

Por el contrario, llama la atención los bajos resultados en el subsector Lenguaje (SC), y en menor medida entre los docentes generalistas (PC), de Comprensión de la Naturaleza (SC) y Comprensión de la Sociedad (SC).

● **Gráfico 14**
 ● *Proporción de docentes que alcanza un desempeño Competente o Destacado*
 ● *Indicador Acompañamiento de las actividades en función del aprendizaje*



Los docentes con buen desempeño en este indicador, durante sus clases se mantienen constantemente alertas al trabajo y a las demandas de los alumnos, exhibiendo estrategias de monitoreo y apoyo acordes con el tipo de actividad o metodología empleada (Ejemplo 9). En clases de tipo expositivo se preocupan de verificar, durante su exposición, que los alumnos vayan comprendiendo los contenidos abordados, observándose momentos o instancias en que solicitan a sus alumnos ejemplificar, relacionar, establecer comparaciones o explicar en sus propias palabras lo expuesto. En clases con una modalidad de trabajo grupal, un docente con buen desempeño se mantiene atento al desarrollo de las actividades, recorre la sala y observa el trabajo de sus alumnos, atiende a las preguntas o necesidades que surgen, precisa o clarifica instrucciones cuando detecta dificultades y apoya el buen funcionamiento del grupo a través de la distribución de roles o sistemas de turnos, entre otros.

En los videos de clases de docentes con desempeño bajo el nivel esperado, se observa que los alumnos desarrollan trabajos individuales o en grupo sin contar con el apoyo ni supervisión del profesor, quien por lo general aprovecha esta instancia para desarrollar otras actividades tales como completar el libro de clases, corregir trabajos, etc. En el caso de clases expositivas, es común observar largas explicaciones o parlamentos sin pausas que permitan levantar información sobre el grado de atención y comprensión de los alumnos (ver Ejemplo 10).

Ejemplo 9: “Zonas de trabajo de básquetbol”

En una clase filmada de Educación Física, se observa a alumnos de 2° Medio en la cancha, reunidos en círculo alrededor de su profesora. Ésta indica que continuarán trabajando en torno al básquetbol: “... usando el sistema de estaciones o zonas de trabajo. Vamos a tener tres zonas y van a ir rotando entre ellas, cada 7 minutos.” Enseguida explica cada zona de trabajo en detalle: menciona el tipo de práctica a realizar, da indicaciones para hacerlas y modela ella misma su correcta ejecución:

- Zona 1: trabajo de pases en parejas, alternando entre pases de pecho, de pique y de sobre cabeza.
- Zona 2: trabajo de bote con *dribling*, alternando mano derecha e izquierda.
- Zona 3: trabajo de marcación en parejas, en que un miembro de la pareja debe tratar de avanzar boteando la pelota y el otro debe intentar obstaculizar su avance y robar la pelota.

Los alumnos se dirigen rápidamente hacia los lugares indicados por la profesora y comienzan a trabajar. Durante la clase, ésta se desplaza permanentemente recorriendo las tres zonas de trabajo. A medida que observa cómo se realizan los ejercicios, va comentando o dando instrucciones que los alumnos incorporan y algunas veces demuestra la ejecución correcta o pide a otros alumnos que modelen el ejercicio.

En la Zona 1 explica al grupo: “Para recibir el balón deben tener los dedos de la mano separados y ligeramente flexionados (demuestra con sus manos). Cecilia, pon las manos como les acabo de decir. Pedro te va a tirar el balón y los demás van a ver cómo ella pone las manos para recibir el balón... ¡Bien! observen sus dedos...”

En la Zona 3 observa a una alumna trabajar y dice: “*El bote debe ser de la cintura para abajo, si es más alto van a perder el control del balón y se los van a quitar más fácil. Probemos... Eso, mantén el bote debajo de la cintura... ¿ven?... (los alumnos asienten)*”

Usando un cronómetro, la profesora lleva el control del tiempo, y avisa a los alumnos cuando se cumplen 7 minutos.

En la Zona 2 los alumnos demoran en ejecutar el ejercicio, pues no saben a quién le corresponde el turno, la profesora se acerca y dice: “*Les quedan solo dos minutos en esta estación, tienen que avanzar más rápido para que todos alcancen a practicar el dribling. Formen una fila y van haciendo el ejercicio en ese orden (los alumnos se forman). Eso... ahora Juan Pablo... Jonathan... Ema. Sigán trabajando así para que alcancen a pasar de nuevo por esta zona.*”

En ocasiones un alumno de otra zona se acerca a la profesora para hacerle una pregunta, entonces ésta se desplaza hasta esa estación y aclara las indicaciones para realizar los ejercicios. También observa a aquellos que realizan mejor el ejercicio y les indica que ayuden a sus compañeros.

En el ejemplo se observa que la docente está atenta a las dudas y dificultades que pueden enfrentar los alumnos y disponible para ayudarlos. Además, muestra una actitud activa de monitoreo, acercándose a las distintas zonas de trabajo y entregando precisiones que permiten corregir errores y mejorar las ejecuciones correctas. Además, se asegura de que todos los alumnos realicen las actividades, entregando indicaciones para facilitar el buen manejo del tiempo.

Ejemplo 10: “El murmullo de los alumnos”

Un profesor de 4º básico inicia su clase filmada formulando preguntas a sus alumnos sobre lo que recuerdan acerca de las figuras geométricas y sus elementos. A continuación, indica que esta clase trabajarán los cuerpos geométricos:

D: Entonces, ¿todos recuerdan los elementos de las figuras geométricas?

As: Sí

D: Bien, entonces hoy vamos a hablar de los cuerpos geométricos (enfatisa al pronunciar). Los cuerpos geométricos son formas que tienen volumen y ocupan un lugar en el espacio, es decir, no son elementos planos como las figuras geométricas. Como tienen volumen, tienen tres dimensiones, alto, largo y ancho, a diferencia de lo que ocurre con las figuras, que solo tienen dos: largo y ancho. Eso ustedes lo saben, ¿verdad? Es más, los cuerpos geométricos están compuestos por figuras geométricas, como este cubo que tengo aquí, que está formado por seis caras cuadradas ¿ven? (muestra, rotando el cubo). Es muy importante que entiendan esta diferencia entre los cuerpos y las figuras geométricas. Entonces, las caras son estas que les mostré en el cubo, las partes planas. Los cuerpos geométricos también tienen aristas, que son estas, donde se cruzan las caras, donde chocan las caras. Y también tienen vértices, puntos donde se cruzan las aristas. ¿Cierto?, ¿queda claro?

As: Sí

D: Entonces ahora ustedes van a desarrollar una guía en la que tienen que identificar los elementos de distintos cuerpos geométricos.

A continuación, el docente entrega a los alumnos una guía de trabajo individual y, mientras éstos la desarrollan, revisa cuadernos en su escritorio, ubicado al frente de la sala.

Los alumnos parecen tener algunas dificultades para el desarrollo de la guía y por largo rato se escucha murmullo entre los grupos, pues se hacen preguntas entre ellos en voz baja. En dos oportunidades, el profesor alza la mirada y hace un gesto pidiendo silencio. En la tercera ocasión dice “*Recuerden que la guía es individual*”. Al finalizar la clase el docente menciona que la próxima clase trabajarán con redes geométricas y que aquellos alumnos que no alcanzaron a terminar la guía deben completarla en su casa.

En el ejemplo, el docente no implementa estrategias de monitoreo que le permitan recoger información sobre el trabajo de sus alumnos y apoyarlo. Presenta nuevos contenidos, pero no verifica que los alumnos estén comprendiéndolos. Mientras desarrolla su explicación realiza pocas preguntas y éstas parecen más bien muletillas (“¿Verdad?”, “¿Cierto?”, “¿Queda claro?”), sin ofrecer un espacio para que los alumnos efectivamente puedan responderlas o plantear sus propias dudas.

Enseguida, durante la actividad de trabajo individual, el docente no verifica si los alumnos han comprendido las instrucciones, tampoco se muestra atento al desarrollo de la guía ni parece disponible a sus requerimientos. Solo se dirige a ellos cuando escucha el murmullo, sin advertir que ello puede ser una señal de dificultad o falta de comprensión por parte de los alumnos. Incluso cuando buscan apoyo en sus compañeros, limita esta iniciativa.

Tampoco en el segmento final de la clase, se genera una instancia para corregir en conjunto la guía; eso habría dado al docente la oportunidad de identificar dificultades y errores de los alumnos, corregirlos, o bien, profundizar y ampliar los contenidos que sí fueron comprendidos por ellos (por ej. ejercitar la identificación de caras, vértices y aristas en objetos en la sala, hacer conexiones entre las figuras que ya conocen y los cuerpos que aparecen en la guía, etc.).

Evaluación del aprendizaje

El Marco para la Buena Enseñanza incluye, como parte de las competencias pedagógicas esperadas en un docente, la capacidad para desarrollar estrategias de evaluación que permitan apreciar el logro de aprendizajes de sus alumnos y retroalimentar sus propias prácticas.

La evaluación constituye una parte esencial del proceso de enseñanza-aprendizaje, no solo por sus propósitos sumativos para certificar los niveles de aprendizaje alcanzados, sino fundamentalmente por su rol formativo y de retroalimentación, tanto para los estudiantes como para el propio docente (Clark, 2008; William, 2010; Shepard, 2001). Levantar información clara y precisa sobre el estado de avance de los alumnos es indispensable para que el profesor pueda orientar su práctica pedagógica y tomar decisiones adecuadas en aspectos tan importantes como los métodos de enseñanza apropiados para un determinado grupo de alumnos o en los contenidos que requieren reforzamiento, ejercitación o profundización (Cross, 1998; McMillan, 2005).

Es fundamental que la evaluación de estudiantes sea congruente con metas importantes de aprendizaje, que se ajuste al contenido evaluado y se encuentre conectado al contexto de aplicación de este contenido particular (Shepard, 2001).

Un docente competente en este ámbito muestra capacidad para diseñar y aplicar instrumentos o actividades de evaluación consistentes con los objetivos curriculares que orientan su trabajo y con las oportunidades de aprendizaje que ofrece a sus alumnos. Además, sus estrategias de evaluación le permiten identificar las habilidades, competencias y conocimientos alcanzados por éstos a partir del uso de instrumentos, pautas y criterios de corrección claros y conceptualmente correctos. De este modo, la información que recoge acerca de los logros de los estudiantes es confiable y precisa.

Para calificar su desempeño en este ámbito, en el Módulo 1 del Portafolio el docente presenta la evaluación que ha aplicado a sus alumnos al concluir la unidad pedagógica. Las evaluaciones presentadas por los docentes varían en su forma en función de las características propias de cada subsector aprendizaje. Así por ejemplo, un profesor que ha aplicado un instrumento de papel y lápiz (como una prueba), adjunta un ejemplar tal como lo administró a sus alumnos. Por su parte, un profesor que evaluó a través de preguntas o instrucciones orales, entrega un registro escrito de las instrucciones o preguntas y una descripción de la tarea solicitada.

Unido a lo anterior, el docente adjunta la pauta de corrección que ha utilizado para evaluar el desempeño de los alumnos, o su equivalente. En el caso de una prueba, entrega un ejemplar de ésta consignando para cada pregunta o ítem la respuesta correcta que demuestra el logro del aprendizaje. En cambio, en el caso de una tarea de desempeño, como la exposición oral de un tema o la ejecución de una destreza física, describe el desempeño correcto esperado; por ejemplo, las características que debe tener la lectura oral de un texto, tales como: lectura fluida, entonación adecuada a los signos de exclamación e interrogación, pausas consistentes con los signos de puntuación, etc.

A continuación se reportan los resultados de dos de los indicadores que se utilizan para determinar las competencias de los docentes para evaluar el aprendizaje de sus alumnos:

- Coherencia de la evaluación con los objetivos de aprendizaje de la unidad;
- Calidad de las preguntas e instrucciones planteadas a los alumnos en la evaluación.

Coherencia de la evaluación con los objetivos

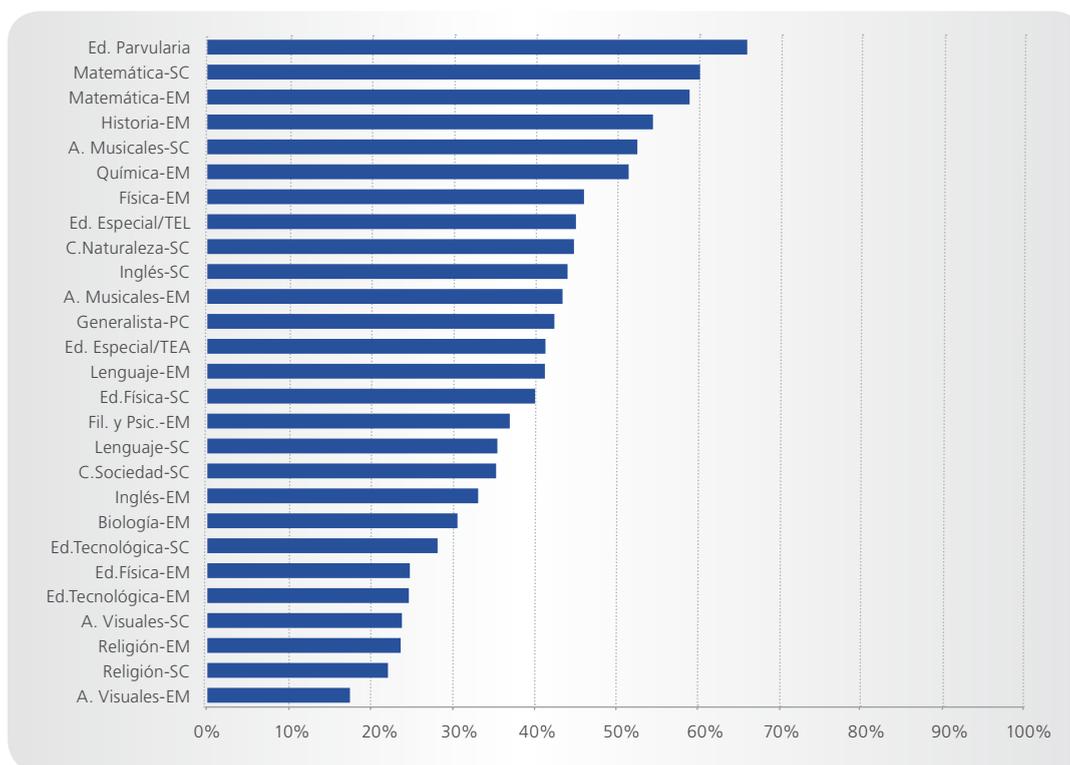
Este indicador evalúa la capacidad del docente para diseñar instrumentos o actividades de evaluación apropiados para recoger información sobre el logro de los objetivos planteados para la unidad pedagógica.

Del total de docentes analizados, el 44% alcanza el estándar esperado. Al observar la proporción de competentes y destacados se distingue preferentemente el nivel de Educación Parvularia, así como Matemática (SC y EM), Historia (EM), Artes Musicales (SC) y Química (EM). En cambio, agrupan una menor proporción de docentes con buen desempeño los subsectores Artes Visuales (EM y SC), Religión (SC y EM), Educación Tecnológica (EM) y Educación Física (EM) (ver Gráfico 15).

Un docente con desempeño Competente diseña y aplica evaluaciones coherentes con los objetivos trabajados con sus estudiantes, las cuales le permiten levantar información sobre todos los aspectos centrales abordados en el período de tiempo evaluado (ver Ejemplo 11).

Las debilidades detectadas en este indicador corresponden a evaluaciones que no son del todo coherentes con los objetivos trabajados, o bien, a evaluaciones que no cubren los contenidos centrales (ver Ejemplo 12).

● **Gráfico 15**
 ● *Proporción de docentes que alcanza un desempeño Competente o Destacado*
 ● *Indicador Coherencia de la evaluación con los objetivos*



Ejemplo 11: “El estirón puberal”

Una docente quiere saber cuánto han aprendido sus alumnos de 7° básico sobre el cálculo de la media aritmética, mediana y moda, y la formulación de conclusiones a partir del análisis de estos estadígrafos. Para ello, les pide que realicen la siguiente guía:

El estirón puberal es un cambio brusco de la velocidad de crecimiento. Este cambio muestra diferencias por género tanto en su cronología como en su intensidad. Observa la tabla con los cambios en las estaturas de los estudiantes de una escuela y realiza las actividades siguientes. (El profesor les entrega una hoja en que aparece la estatura de los niños de un curso en 3°, 5° y 7° básico, en los años 2003, 2005 y 2008 respectivamente).

1. Completa el cuadro con la media, mediana y moda de las estaturas de los estudiantes cada año. Luego, explica los pasos que seguiste para ello.

Estadigráfo/Año	Estaturas estudiantes		
	2003	2005	2008
Media			
Mediana			
Moda			

2. A partir de la tabla, responde: ¿Es siempre el mismo estudiante el más alto del grupo? ¿Cuál es el estudiante más alto cada año?, ¿cuál es el que menos creció entre el 2003 y el 2008?

3. A partir del análisis de los datos responde: ¿Qué puedes concluir respecto del estirón puberal de los estudiantes de esta escuela? Fundamenta tus conclusiones utilizando la media, mediana y moda de sus estaturas cada año.

¿Qué diferencias y similitudes observas al analizar el estirón puberal por género? Fundamenta tu respuesta utilizando la media, mediana y moda de sus estaturas cada año.

En este ejemplo se observa que la prueba aplicada permite conocer en qué medida los alumnos alcanzaron los objetivos que la docente se propuso evaluar. En el ítem 1 los alumnos deben calcular la media, mediana y moda.

Respecto al objetivo de formular conclusiones a partir del análisis de datos, el ítem 2 por sí mismo no permite saber si los alumnos pueden hacerlo, pues solo requiere que comparen a partir de sus cálculos y extraigan información de la tabla. En cambio, la pregunta 3 sí les exige hacer inferencias y extraer conclusiones basadas en el análisis de los estadígrafos.

Ejemplo 12: “Escultura ecológica”

Un profesor trabajó con sus alumnos durante la unidad cinco elementos visuales: líneas, colores, texturas, luces y espacios. En su evaluación final quiere saber si sus alumnos son capaces de utilizar y describir estos elementos en sus propias creaciones y apreciarlos en otras obras. Para ello les propuso la siguiente actividad de evaluación:

“En la evaluación que vamos a realizar ahora vamos a poner en práctica todo lo que hemos aprendido sobre los elementos visuales. Para eso vamos a diseñar de manera individual una escultura ecológica que contenga algunos de estos elementos. Voy a dar las instrucciones paso a paso: Lo primero que van a hacer es elegir un concepto o idea que queramos representar a través de la escultura en torno al tema “Relación hombre-naturaleza”.

Luego van a hacer un boceto de su escultura, considerando que ésta debe mostrar claramente 2 de los elementos visuales que hemos trabajado en clases y que la construiremos en base a materiales del entorno que saldremos a buscar al patio: pueden ser ramas, pasto, hojas, objetos reciclables, entre otros.

No es necesario que la escultura sea exactamente igual al boceto que diseñaron inicialmente, puede que al combinar

los materiales ustedes noten que no se ven armónicos y deban realizar modificaciones. Lo importante es que no se distancien tanto de la idea original y que no pierdan de vista el concepto que ustedes eligieron representar.

Les voy a entregar la pauta con los puntos que consideraré al momento de evaluar, revísenla y si tienen dudas me preguntan.

Comenzaremos a trabajar esta clase y la entrega final será para la sub siguiente. Para ese día deben traer la escultura sobre una base que tenga el nombre de la obra y del autor.

¿Alguna duda? Pueden comenzar a trabajar”.

Si bien la actividad de evaluación que diseña el profesor es clara y orienta muy bien a los alumnos respecto a qué se espera de ellos, no cubre todos los objetivos que se propuso evaluar. Aun cuando le permitirá ver si los alumnos utilizan elementos visuales en sus creaciones, no podrá evaluar si son capaces de describirlos o apreciarlos en otras obras. Para ello, bastaría haberles pedido exponer sus trabajos al curso, y solicitarles escribir una breve crítica respecto al uso de los elementos visuales presente en las distintas esculturas.

Calidad de las preguntas o instrucciones

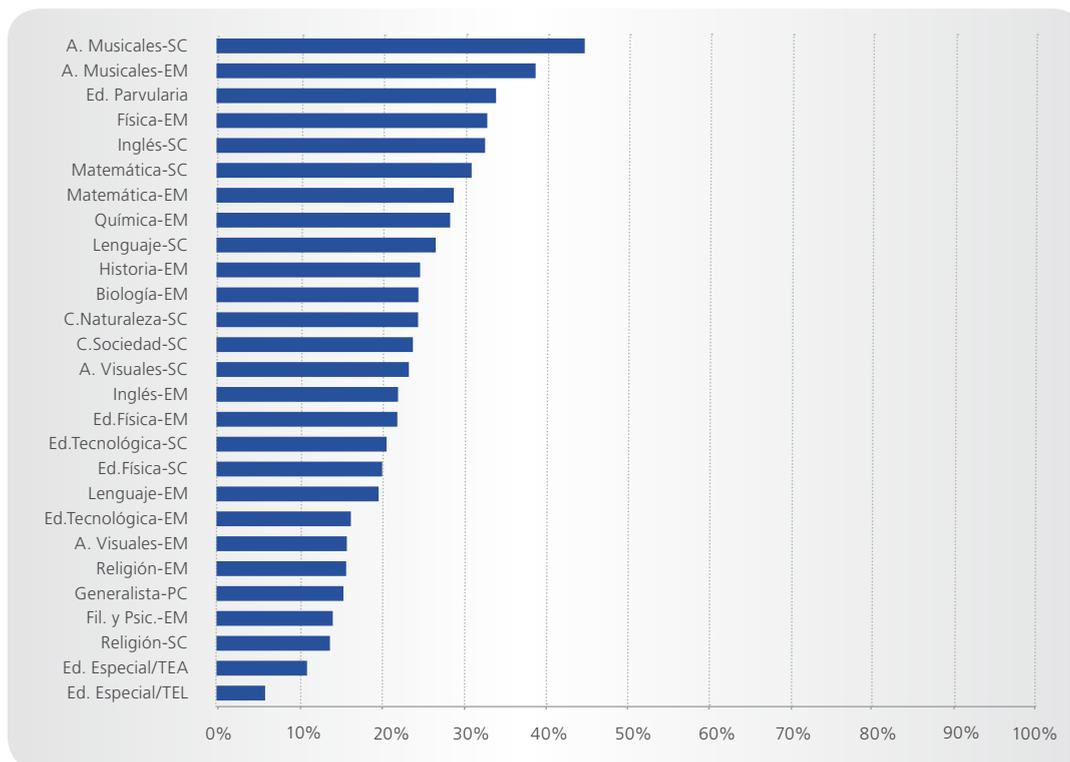
Este indicador evalúa la claridad de las preguntas que los profesores formulan al evaluar a sus alumnos y la precisión de las pautas y/o criterios de corrección que emplean para calificar sus respuestas o desempeños. La importancia de que los alumnos reciban instrucciones claras y completas radica en que ello afecta su ejecución tanto al responder una prueba de papel y lápiz como al desarrollar un ejercicio, elaborar un producto o un trabajo de investigación.

Solo aproximadamente un tercio de los evaluados obtiene el nivel esperado en este indicador y esta proporción es especialmente baja en el caso de Educación Especial (5,9% en docentes de TEL y 11% en TEA). Una situación similar se observa en los docentes de Religión (especialmente SC), Filosofía y Psicología (EM), los profesores generalistas de Primer Ciclo, los de Artes Visuales (EM) y Educación Tecnológica (EM). Cabe destacar que ningún subsector alcanza un 50% de docentes con nivel competente en este indicador (ver Gráfico 16).

Un docente competente diseña pruebas o situaciones de evaluación que permiten a sus estudiantes demostrar lo aprendido. Por ejemplo, en el caso de pruebas con ítems de selección múltiple, tanto las preguntas como las opciones de respuestas son claras y correctas, los distractores son todos plausibles y responden a una misma categoría lógica, de tal manera de no entregar pistas que puedan delatar la opción correcta. En el caso de evaluaciones de desempeño o de productos solicitados, las instrucciones que recibe el alumno indican claramente todos los aspectos que serán tomados en cuenta al momento de evaluar.

Las principales debilidades se reflejan, por ejemplo, en el uso de enunciados ambiguos, incompletos o confusos que pueden afectar las respuestas de los alumnos por falta de claridad sobre la tarea solicitada. Esto perjudica el valor de la evaluación para establecer efectivamente los aprendizajes logrados por los estudiantes, retroalimentarlos y orientar las prácticas pedagógicas del profesor (ver Ejemplo 13).

● **Gráfico 16**
 ● *Proporción de docentes que alcanza un desempeño Competente o Destacado*
 ● *Indicador Calidad de las preguntas o instrucciones*



Ejemplo 13: “La compra”

En su Módulo 1, una docente generalista entrega una prueba de Matemática que incluye, entre otras, la siguiente pregunta:

3. Juan va al Supermercado con \$5.000 ¿Qué cosas de la siguiente lista puede comprar? (Marca con una "x" en el casillero izquierdo)

- 1 kilo de pan = \$ 750
- 1 bebida = \$ 1.100
- 1 kilo de arroz = \$ 650
- 2 kilos de papas = \$ 800
- 1 kilo de detergente = \$ 1.700
- 1 caja de té = \$ 450
- 1 paquete de servilletas = \$ 250

En la pauta de corrección, la docente consigna como única respuesta correcta que el alumno marque los cinco primeros productos del listado, que suman exactamente \$5.000. Sin embargo, para considerar \$5.000 como única respuesta correcta, la pregunta podría haber sido, por ejemplo “¿Qué cosas de la siguiente lista puede comprar para que le alcance justo con los \$5.000?”.

La pregunta planteada, en cambio, también admite como respuestas correctas otras elecciones de productos en tanto no sumen más \$5.000 existiendo varias combinaciones que cumplen esa condición. Este error en la formulación de la pregunta afecta directamente la información que la docente puede levantar sobre los logros de sus alumnos ya que ella puede concluir que sus alumnos no saben sumar correctamente.

Reflexión pedagógica

Dentro de las responsabilidades profesionales de un docente, el Marco para la Buena Enseñanza establece que éste debe reflexionar consciente y sistemáticamente sobre su práctica y reformularla, contribuyendo así a garantizar una educación de calidad para todos sus estudiantes.

La condición profesional del docente así como la efectividad en su trabajo exigen de éste no solo la capacidad de realizar un buen trabajo en el aula, sino también el ser un profesional reflexivo, que hace de su propia práctica un objeto de revisión, análisis y aprendizaje (McKenzie, 2005; Schon, 1990).

Un docente competente es aquel que tiene conciencia de la influencia que ejercen sus acciones y decisiones pedagógicas sobre los resultados de sus estudiantes. Además de saber planificar y desarrollar adecuadamente sus clases, reflexiona críticamente sobre su práctica, pudiendo identificar los aspectos positivos e importantes de mantener, así como aquellos que requieren ser modificados o mejorados.

La condición reflexiva del docente también se expresa en las relaciones que establece tanto con sus pares y superiores al interior de su establecimiento como en el sistema educativo en general. Implica formar parte constructiva del entorno donde trabaja, compartiendo y aprendiendo junto a sus colegas. Un docente que posee una adecuada capacidad reflexiva, analiza con una mirada crítica la efectividad de las actividades implementadas, identificando aspectos de su práctica que han favorecido el aprendizaje de sus alumnos así como decisiones menos acertadas o efectivas. A la vez, identifica los factores que han influido en los niveles de logro alcanzados por sus alumnos en las evaluaciones aplicadas, manteniendo como eje central de su análisis el aprendizaje y sus propias decisiones.

Para calificar el desempeño en el ámbito de la *Reflexión pedagógica*, el docente entrega por escrito un análisis referido a la unidad pedagógica reportada en el Portafolio, a los resultados de aprendizaje de sus alumnos y/o su trabajo docente en general.

A continuación se reportan tres de los indicadores utilizados para evaluar la capacidad de reflexión pedagógica exhibida por los docentes:

- *Análisis de las decisiones pedagógicas* en el contexto de la unidad reportada en el Portafolio;
- *Análisis de la efectividad de una clase* tomando como referencia una que el docente considera especialmente exitosa;
- *Atribución de los resultados de aprendizaje* obtenidos por los alumnos en la evaluación de la unidad informada en el Portafolio.

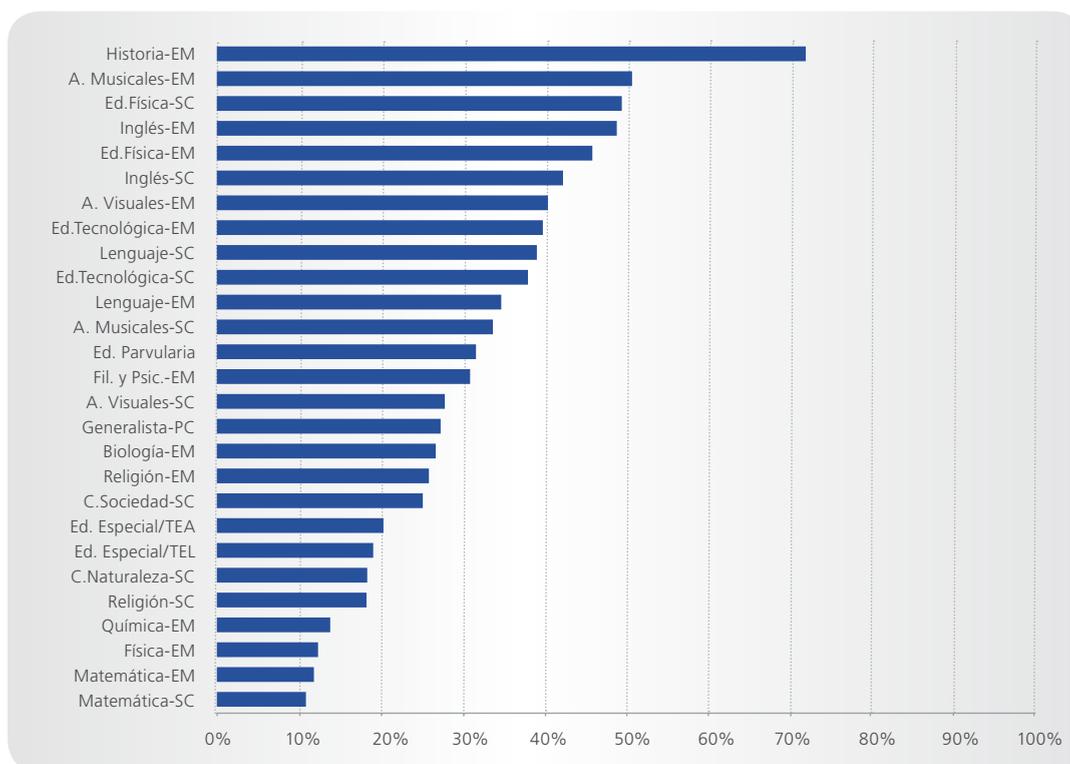
Análisis de las decisiones pedagógicas

En este indicador se evalúa la capacidad del docente para analizar y evaluar las decisiones que ha tomado en el proceso de enseñanza-aprendizaje, en el contexto de la unidad presentada en su Portafolio; debe identificar los aspectos de ésta que han favorecido el aprendizaje de los estudiantes, así como aquellos que aún requiere mejorar, pudiendo aludir en su análisis a la eficacia de las actividades desarrolladas, la secuencia utilizada, los recursos de aprendizaje empleados y las explicaciones entregadas, entre otras decisiones. Del total de casos analizados, el 28% de los evaluados logra el nivel Competente o Destacado, distinguiéndose de forma muy especial el subsector Historia (EM) que reúne a más del 70% de docentes con buen desempeño. Por su parte, los subsectores con mayor proporción de bajos desempeños corresponden a Matemática (SC y EM), Física (EM) y Química (EM) (ver Gráfico 17).

Al analizar la unidad implementada, la mayoría de los docentes no logra demostrar la capacidad reflexiva y crítica esperada para examinar su propio trabajo; si bien rescatan aspectos relevantes del desarrollo de

la unidad y de su propia práctica, lo hacen en forma parcial, superficial o incompleta. Algunos profesores aluden a las actividades y recursos que han resultado favorecedores del aprendizaje, sin embargo muestran debilidades para identificar aquello que interfirió o no aportó al logro de sus alumnos.

● **Gráfico 17**
 ● *Proporción de docentes que alcanza un desempeño Competente o Destacado*
 ● *Indicador Análisis de las decisiones pedagógicas*



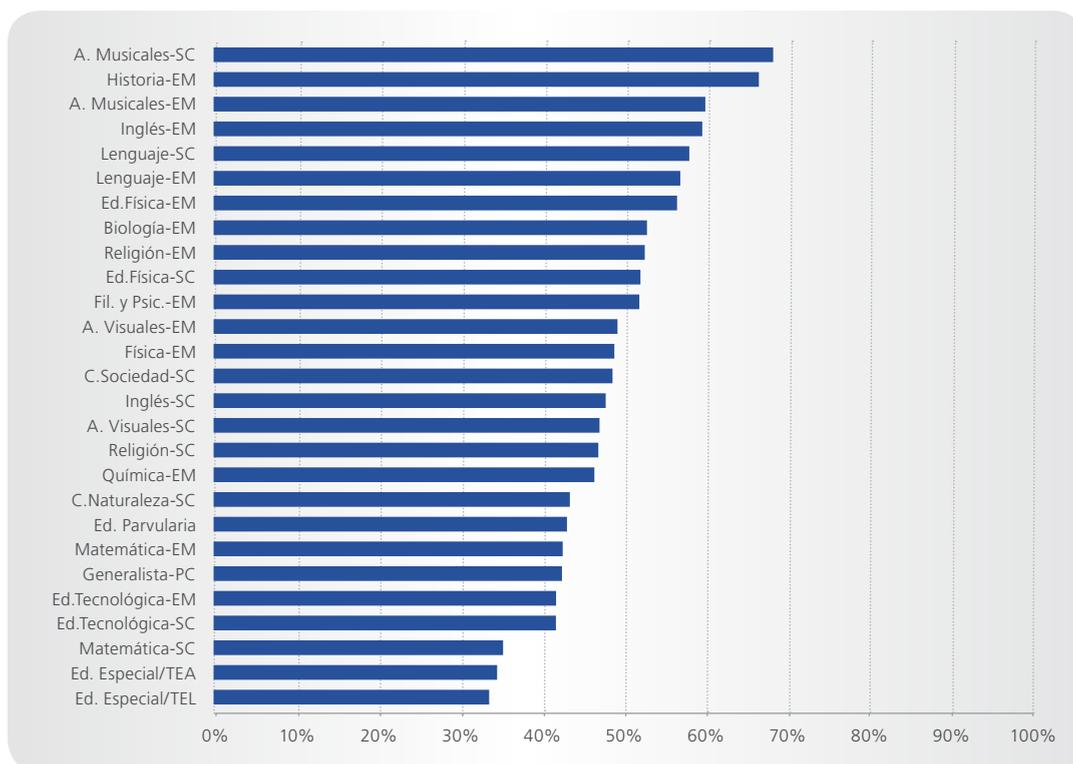
Análisis de la efectividad de una clase

En este indicador se evalúa la capacidad reflexiva del docente en términos de analizar la efectividad o éxito de una clase, considerando como foco central el aprendizaje y la motivación e interés de sus alumnos. Los datos muestran que un 47% de los docentes alcanza un desempeño Competente o Destacado. En el grupo de alto desempeño se ubican los subsectores Artes Musicales (SC y EM), Historia (EM), Inglés (EM) y Lenguaje (SC y EM). En contraste, la menor proporción de docentes con buen desempeño se observa en la modalidad de Educación Especial (TEL y TEA) y en el subsector Matemática (SC) (ver Gráfico 18).

Los docentes con buen desempeño analizan sus clases efectivas o exitosas demostrando clara conciencia de que el principal objetivo del trabajo en clase es que los alumnos aprendan, y valoran la motivación y participación de éstos como variables para lograrlo.

Aquellos docentes que obtienen resultado Básico o Insatisfactorio, por el contrario, no logran situar el aprendizaje como eje fundamental de su análisis, centrándose en aspectos socioafectivos o conductuales, como la solidaridad demostrada por los alumnos al compartir sus útiles o la entretención generada por las actividades, sin vincular esto al desarrollo de nuevos conocimientos o habilidades.

● **Gráfico 18**
 ● *Proporción de docentes que alcanza un desempeño Competente o Destacado*
 ● *Indicador Análisis de la efectividad de una clase*



Atribución de los resultados de aprendizaje

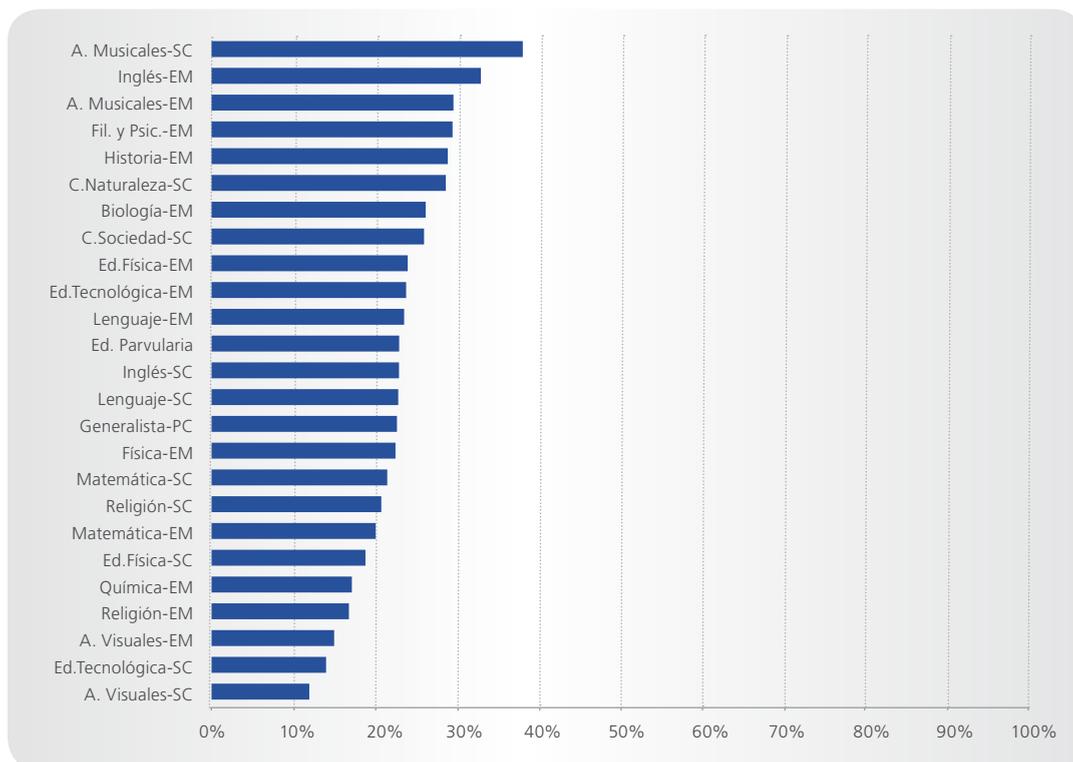
Este indicador refiere a la capacidad del docente para analizar los resultados obtenidos por los alumnos en la evaluación aplicada al término de la unidad, identificando los aspectos logrados y no logrados y los factores que incidieron en ello, incluyendo como principal factor explicativo su propio actuar como docente.

Solo un 23% de los evaluados obtiene un resultado Competente o Destacado. Los desempeños por subsector en este indicador son homogéneamente bajos, destacando comparativamente Artes Musicales (SC) en que el porcentaje de desempeños competentes se acerca al 40% (ver Gráfico 19).

Los docentes que alcanzan el estándar esperado demuestran un mayor conocimiento y una mejor comprensión de los efectos que su práctica tiene sobre los resultados de los alumnos: son capaces de identificar tanto los aspectos logrados como no logrados del aprendizaje y de vincular, en ambos casos, los resultados alcanzados con sus propias acciones y decisiones pedagógicas.

Aquellos docentes que demuestran debilidades en la atribución de los resultados de los aprendizajes por lo general relacionan los éxitos de sus alumnos con su práctica. En cambio, atribuyen las debilidades o falencias de los estudiantes a situaciones externas, tales como las carencias del contexto social de sus familias o la falta de equipamiento o infraestructura del establecimiento, entre otros.

● **Gráfico 19**
 ● *Proporción de docentes que alcanza un desempeño Competente o Destacado*
 ● *Indicador Atribución de los resultados de aprendizaje*



Comentarios finales

Transcurridos ocho años desde su primera aplicación, el proceso de Evaluación Docente incluye ya a la gran mayoría de los docentes de aula del sector municipal, que ejercen a su vez en los niveles, modalidades y sectores de enseñanza de mayor cobertura. En virtud de esto, la evaluación ha permitido disponer de una amplia base de datos, de representación nacional, que informa acerca del desempeño de los profesores y profesoras en nuestro país. Este capítulo ha dado cuenta de parte de esos resultados, con la intención de contribuir a su divulgación entre los actores claves para el desafío de mejorar la educación: los propios docentes, directores y jefes técnicos, sostenedores, formadores de profesores, investigadores, diseñadores e implementadores de políticas educacionales, entre otros.

Como se ha expuesto, los resultados finales de la Evaluación Docente muestran que, año a año, el mayor porcentaje de evaluados (entre 52% y 64%, dependiendo del período) se concentra en el nivel Competente, que marca el estándar de desempeño esperado, en tanto que cerca de un tercio logra solo un resultado Básico. Un porcentaje reducido de profesores (menos del 10% en seis de los ocho períodos) obtiene un resultado Destacado y un porcentaje aún menor (menos del 4%) resulta calificado con el nivel Insatisfactorio. Por otro lado, al desagregarse los datos de este resultado global, es posible observar que existen diferencias ostensibles en los datos arrojados por los distintos instrumentos, exhibiendo la Autoevaluación los promedios más estables y altos (muy cerca del máximo desde el año 2003 a la fecha), seguida por la evaluación de directivos, la de evaluadores pares y por último el Portafolio, que muestra resultados notoriamente inferiores.

Al analizar en mayor profundidad los resultados del Portafolio, se constata que consistentemente los mejores desempeños en este instrumento se refieren al “ambiente instruccional”: en la clase filmada se aprecia que la mayoría de los docentes es capaz de crear un ambiente favorable para el aprendizaje, en que se respetan normas de convivencia que facilitan un trabajo fluido, y logra motivar a sus alumnos y comprometerlos con las actividades propuestas. En cambio, un aspecto que aparece debilitado es la calidad de la interacción pedagógica entre docente y estudiantes. El Portafolio indica que los docentes tienen dificultades importantes para interactuar con sus alumnos de maneras que favorezcan el desarrollo del pensamiento en ellos y los estimulen a plantearse hipótesis, indagar y extraer conclusiones; y lo mismo ocurre en el tipo de interacciones que promueven entre los alumnos durante la clase. También se observan debilidades en competencias asociadas a la evaluación, especialmente en cuanto a utilizar la información recogida para retroalimentar a los alumnos de una manera que efectivamente los ayude a mejorar sus aprendizajes.

En la segunda sección de este capítulo, se han expuesto datos para una selección de indicadores del Portafolio desagregados por nivel, modalidad y sector de enseñanza; y se complementan con algunos ejemplos que ilustran cómo se manifiestan, en situaciones de aula concretas y específicas, las fortalezas y debilidades pedagógicas observadas. En el análisis se constatan, por un lado, aspectos de alta coincidencia, así como peculiaridades y diferencias llamativas en algunos de los indicadores evaluados, según el nivel, modalidad o sector curricular en que ejerce el docente.

También se han incluido algunos datos sobre la relación entre el nivel de desempeño obtenido por los docentes en la evaluación y su situación laboral actual, así como opiniones vertidas por los evaluados en cuanto a beneficios obtenidos en el proceso de evaluación, en particular, al elaborar su Portafolio.

Desde luego el alcance de este capítulo es limitado y la presentación no tiene un ánimo exhaustivo; por el contrario, su intención es servir como invitación a otros actores para recoger alguna(s) de las variadas líneas de análisis e investigación que ofrecen los datos disponibles. Solo a modo de ejemplo, entre ellas están: estudiar qué factores podrían explicar los aspectos que, en forma consistente, resultan mejor y peor evaluados; qué variables subyacen a algunas diferencias observadas según nivel, sector y modalidad de enseñanza; o cuál es la relación entre nivel de desempeño en la evaluación y otras variables como formación inicial del profesor o resultados de aprendizaje de los alumnos; entre muchos otros¹⁴.

Respecto del valor informativo de la Evaluación Docente, cabe recordar que además de datos este programa ha recopilado un caudal —probablemente sin parangón en nuestro sistema escolar— de muestras directas del trabajo de los profesores, incluyendo videos de clases, evaluaciones aplicadas a los alumnos, comentarios y reflexiones acerca de su trabajo, entre otros. Desde la perspectiva de la investigación, este es un insumo de incalculable valor para estudios orientados a caracterizar y analizar en profundidad las prácticas pedagógicas de los docentes. Desde la perspectiva de los profesores, directores y jefes técnicos, los datos de la evaluación en relación con esas evidencias directas permiten comprender mejor cómo es un buen desempeño, observar y analizar las propias prácticas a la luz de esa comprensión e identificar estrategias pedagógicas efectivas hacia las cuales encaminar sus esfuerzos de mejoramiento.

En la medida que los resultados entregados por la Evaluación Docente sean dados a conocer entre los actores involucrados y contribuyan a comprometerlos y orientarlos para mejorar las prácticas de enseñanza-aprendizaje, podrá cumplirse mejor el propósito formativo de la evaluación. Este capítulo ha intentado ser un aporte en esa dirección.

.....

14 Algunos de los temas enunciados ya han sido o están siendo abordados por investigaciones que utilizan los datos y/o evidencias recolectados por la Evaluación Docente. A modo de ejemplo, ver estudios reportados en capítulos 7 y 8 de este volumen; Informes de Resultados Nacionales SIMCE (2007, 2008, 2009); Bravo, Falck, González, Manzi & Peirano (2008); Sección Evaluación e Investigación en www.docentemas.cl.



5

Consecuencias de la Evaluación Docente

Flavio Cortés · María José Lagos

Centro de Medición MIDE UC
Pontificia Universidad Católica de Chile

Consecuencias de la Evaluación Docente

El presente capítulo detalla las consecuencias asociadas a dos diferentes escenarios producto de los resultados de Evaluación Docente. En primer lugar se describe la Asignación Variable por Desempeño Individual, incentivo económico al que pueden acceder, tras aprobar una prueba de conocimientos, los docentes que obtienen un nivel de desempeño Competente o Destacado en la evaluación. Enseguida se abordan los Planes de Superación Profesional, acciones de apoyo implementadas a nivel local para los docentes con resultado Básico e Insatisfactorio. Se profundiza especialmente en estos planes: su cobertura, implementación, los actores participantes y la lógica que les orienta. También se analizan sus principales fortalezas y debilidades en relación a las características que la literatura especializada indica sobre formación docente efectiva; se discute el carácter local (descentralizado) de su diseño y ejecución y su posibilidad de impacto sobre las prácticas docentes.

.....

Este capítulo tiene como objetivo presentar las consecuencias o mecanismos institucionales dispuestos para los docentes de los establecimientos educacionales municipales del país, una vez que han sido evaluados por el Sistema de Evaluación del Desempeño Profesional Docente (en adelante Evaluación Docente)⁽¹⁾

Como se ha señalado en capítulos anteriores, el desempeño de los docentes afectos a la Evaluación Docente pueden ser categorizado en cuatro niveles: Destacado, Competente, Básico e Insatisfactorio. Los dos primeros indican desempeños que corresponden a conductas pedagógicas que cumplen con los indicadores evaluados, mientras los dos últimos revelan falencias evidentes que ameritan acciones de carácter formativo. Consistentemente, las consecuencias que conllevan estos niveles de desempeño son diferentes. Resultados positivos (Competente o Destacado) brindarán a los profesores y profesoras la posibilidad de elevar sus ingresos, mientras que desempeños deficientes (Básico o Insatisfactorio) conducirán a acciones de formación para superar las debilidades constatadas por la evaluación. En el primer caso, los docentes podrán postular a la Asignación Variable por Desempeño Individual (AVDI), un beneficio económico por hasta 4 años; mientras en el segundo, la reglamentación vigente establece que el docente está obligado a participar en acciones de formación remedial denominadas Planes de Superación Profesional (PSP).

En las páginas siguientes, se realizará primero una breve descripción de las características y funcionamiento de la AVDI para después centrar el análisis en la descripción de los objetivos e institucionalidad de la política de formación continua que los PSP buscan llevar a cabo, los actores que en ellos participan, el tipo de acciones de formación que involucran y las dinámicas de expansión y asentamiento de esta política desde sus inicios el año 2005. Esta concluirá con una mirada evaluativa a las fortalezas, debilidades y desafíos que se advierten para que dicha política responda efectivamente al objetivo de mejorar las prácticas docentes.

.....

1 Este sistema, iniciado el año 2003, fue normado jurídicamente al promulgarse la Ley 19.961 sobre Evaluación Docente y aprobarse el Reglamento que lo regula (Decreto Supremo de Educación 192, de 2004).

La Asignación Variable por Desempeño Individual

La Asignación Variable por Desempeño Individual (ADVI) se creó a través de la Ley 19.933 del 12 de febrero de 2004, con el objetivo de reconocer el mérito profesional de los docentes del sector municipal que alcanzan los niveles superiores de desempeño (Destacado y Competente) en la Evaluación Docente, y consiste en un incentivo económico cuyo monto se calcula como una proporción de la Renta Básica Mínima Nacional. Para acceder a él, el profesor que ha obtenido un resultado Competente o Destacado en la Evaluación Docente, debe rendir una prueba de conocimientos pedagógicos y disciplinarios y aprobarla con una calificación Suficiente, Competente o Destacado. Esta prueba solo puede ser rendida en una ocasión durante los 36 meses siguientes a la entrega de los resultados de la Evaluación Docente⁽²⁾ y, por consiguiente, la asignación se extenderá por un período variable, de 2 a 4 años, dependiendo del momento en que se rindió la prueba.

Los diferentes niveles de logro obtenidos en la prueba combinados con los niveles de desempeño de la Evaluación Docente determinan el monto de la AVDI (ver Tabla 1), pudiendo ascender a un 5%, 15% o 25% de la Renta Básica Mínima Nacional (RBMN). Así, por ejemplo, un docente con resultado Destacado tanto en la Evaluación Docente como en la prueba de conocimientos, accederá al beneficio máximo (25%), mientras que —cualquiera haya sido su resultado en la primera— si en la prueba solo logra un nivel Suficiente, solo recibirá el 5%. Cabe señalar que, de no alcanzar el nivel mínimo exigido para la asignación (Suficiente), el docente no percibirá el incentivo, sin que ello afecte el resultado previamente obtenido en la Evaluación Docente.

El valor monetario de la AVDI depende en parte de la situación particular del profesor (por ejemplo, la cantidad de horas contratadas y el valor de la RBMN para el nivel de enseñanza que ejerce). A modo de ejemplo, un docente de Educación Media con jornada completa que recibe la AVDI por el porcentaje máximo, percibirá una asignación anual de aproximadamente \$1.335.000 durante un máximo de cuatro años. Si, en cambio, solo obtiene el beneficio por el porcentaje mínimo, el monto anual será de \$267.089⁽³⁾.

• **Tabla 1**
 • *Porcentaje de Asignación Variable por Desempeño Individual según resultados en Evaluación Docente y en la Prueba de Conocimientos Pedagógicos y Disciplinarios*

		Resultado en prueba de conocimientos pedagógicos y disciplinarios		
		Destacado	Competente	Suficiente
Resultado en Evaluación Docente	Destacado	25% de la RBMN	15% de la RBMN	5% de la RBMN
	Competente	15% de la RBMN	15% de la RBMN	5% de la RBMN

La prueba de conocimientos pedagógicos y disciplinarios es de carácter nacional⁽⁴⁾ y el docente debe rendirla en relación a un subsector (asignatura) específico, que identifica al momento de su postulación. En la prueba, que incluye tanto preguntas cerradas como abiertas, se miden los conocimientos teóricos declarativos que el o la docente debe dominar, para demostrar un profundo conocimiento y comprensión de la disciplina que enseña. Para seleccionar dichos contenidos, se toma como base el marco curricular vigente.

2 A excepción del año 2006, en el cual sólo pudieron postular a AVDI docentes con resultado Competente y Destacado en la Evaluación Docente 2005 y de los años anteriores (2004 y 2005) en los cuales no era posible cumplir esa cantidad de meses debido a que la Asignación entró en vigor en 2004.

3 Además, por el solo hecho de rendir la prueba, el docente tiene derecho a percibir, por una sola vez, un bono de \$50.000.

4 Esta prueba es la misma que se aplica en el programa de acreditación para la Asignación de Excelencia Pedagógica (AEP).

Como se mencionó anteriormente, a partir del año 2007, los docentes con resultado Competente o Destacado en la Evaluación Docente pueden postular a la AVDI hasta 36 meses luego de obtenido dicho resultado, lo cual explica el número acumulativo de postulantes potenciales que se constata entre 2004 y 2009 (ver Tabla 2).

• **Tabla 2**
• *Número de postulantes efectivos sobre potenciales por año 2004–2009*

	2004	2005	2006	2007	2008	2009
Postulantes potenciales por año	2.425 ⁽¹⁾	2.311 ⁽²⁾	6.328 ⁽³⁾	12.881 ⁽⁴⁾	12.836 ⁽⁵⁾	18.268 ⁽⁶⁾
Postulantes efectivos	1.191	1.098	3.237	6.486	4.716	5.671
% Postulantes efectivos sobre potenciales	49,1%	47,5%	51,2%	50,4%	36,7%	31,0%

(1) Docentes con resultado Competente + Destacado (C+D) en la Evaluación Docente (ED) 2003.

(2) Docentes con resultado C + D en la ED 2004 y 2003 (que no hayan rendido AVDI en 2004).

(3) Docentes con resultado C + D en la ED 2005.

(4) Docentes con resultado C + D en la ED 2006, 2005 y 2004 (que no hayan rendido AVDI en 2005 y 2004).

(5) Docentes con resultado C + D en la ED 2007, 2006 y 2005 (que no hayan rendido AVDI en 2006 y 2005).

(6) Docentes con resultado C + D en la ED 2008, 2007 y 2006 (que no hayan rendido AVDI en 2007 y 2006).

Como se observa en la Tabla 2, la proporción de postulantes efectivos a la AVDI ha oscilado entre la mitad y un tercio de los postulantes potenciales. Por otra parte, la cantidad efectiva de asignaciones otorgadas desde el año 2004 al 2009 asciende a 16.577, pudiendo corresponder esta cifra a una cantidad menor de docentes, dado que un profesor puede recibir la AVDI en más de una ocasión (ver Tabla 3). Respecto al monto del beneficio, el 60% de las asignaciones obtenidas corresponde al tramo de logro inferior (5% de la RBMN) y un 38,8% al tramo de 15%, en tanto que un exiguo 1,2% de docentes logra la máxima calificación tanto en la Evaluación Docente como en la prueba AVDI, logrando el tramo más alto (25%).

Los datos también muestran que el período 2004–2005 corresponde a un momento inicial de este programa de incentivos, mientras que 2006 aparece como un momento transicional y desde 2007 se asiste a una etapa de consolidación, en que una cantidad importante de docentes (entre 3.800 y 4.800) percibe este beneficio⁵.

• **Tabla 3**
• *Número acumulado de docentes que perciben AVDI y tramo de asignación por año 2004–2009*

Tramo de asignación	2004	2005	2006	2007	2008	2009	Total
5%	–	–	1.650	2.964	2.273	3.065	9.952
15%	401	261	709	1.827	1.553	1.684	6.435
25%	11	9	14	46	44	66	190
Total	412	270	2.373	4.837	3.870	4.815	16.577

5 No se entregan datos para 2010 pues los resultados de AVDI para ese período no han sido publicados al momento de elaboración de este capítulo.

Los Planes de Superación Profesional

Objetivos y sentido de los Planes de Superación Profesional

El objetivo explícito que asigna el Reglamento sobre Evaluación Docente (Decreto Supremo de Educación 192, de 2004) a los Planes de Superación Profesional (en adelante PSP) consiste en la “superación de las debilidades profesionales que evidencien los docentes con nivel de desempeño Básico e Insatisfactorio”. En todo caso, a renglón seguido, este objetivo es formulado de manera más laxa y se señala que los PSP son un conjunto de acciones de formación docente dirigidas a “favorecer la superación” de esas debilidades⁶. En este sentido la Evaluación Docente tendría un carácter formativo, pues entregaría información que indica en qué aspectos los docentes deben mejorar para superar sus debilidades y cuáles son sus áreas de mayor fortaleza relativa. Basados en la información entregada por la Evaluación Docente, los PSP consistirían entonces en aquellas acciones formativas de aprendizaje y re-aprendizaje de las competencias, conocimientos, habilidades, dominios y criterios establecidos en el Marco para la Buena Enseñanza⁷ (Ministerio de Educación, 2003).

No obstante, no puede desconocerse que la Evaluación Docente también posee un carácter sumativo, pues al categorizar el desempeño docente genera —como se ha señalado antes— consecuencias diferenciales para quienes se sitúan por encima o debajo del desempeño mínimo esperado. Al realizarse los PSP como consecuencia de la Evaluación Docente e inmediatamente después que ésta ha comunicado sus resultados, este carácter sumativo forma parte del contexto en el que han de implementarse las acciones formativas.

La extensión temporal de estas acciones de formación para lograr los objetivos señalados dependerá del plazo que los cuerpos normativos⁸ han establecido para las diferentes categorías de evaluación. En el caso de los docentes con resultado Básico, el proceso evaluativo siguiente ocurre transcurridos cuatro años desde el momento en que se entregó dicha calificación; mientras en el caso de los docentes que se ubican en el nivel Insatisfactorio, su evaluación ocurre al año siguiente. De este modo, las actividades de formación que constituyen los PSP tienen actualmente una duración de un año, para aquellos que se sitúan en el nivel más bajo del sistema de calificación (Insatisfactorio) y de cuatro para quienes se ubican en la segunda posición más baja (Básico). El proceso de re-evaluación y formación ha operado de este modo desde la instalación de la Evaluación Docente y los PSP hasta ahora. Sin embargo, el 8 de febrero de 2011 fue promulgada la Ley 20.051 de Calidad y Equidad de la Educación, que modifica algunos artículos relativos a la Evaluación Docente y los PSP. En particular, esta ley establece que los docentes que han obtenido un nivel de desempeño Básico deberán re-evaluarse al año subsiguiente, esto es, dos años después de su primera evaluación (y no cuatro como ocurría hasta ahora). Consiguientemente, en adelante, los docentes con resultado Básico participarán en actividades de PSP solo por dos años, durante los cuales se espera que mejoren su nivel de desempeño. También, la nueva ley señala que quienes resulten calificados con un desempeño Básico o en forma alternada con desempeño Básico o Insatisfactorio durante tres evaluaciones consecutivas, dejarán de pertenecer a la dotación docente.

6 El mejoramiento de las habilidades y competencias docentes de los profesores tiene como última finalidad contribuir a elevar la calidad de los aprendizajes de sus alumnos y alumnas, pero ello no está indicado en la norma reglamentaria.

7 Para más información sobre el Marco para la Buena Enseñanza ver los capítulos de Bonifaz y de Flotts & Abarzúa en este volumen, que abordan el desarrollo de este marco y su uso en la construcción de los instrumentos de la Evaluación Docente, respectivamente.

8 Ley 19.933 del 12 de febrero de 2004, Ley 19.961 del 9 de agosto de 2004 y Decreto 192 del Ministerio de Educación del 30 de agosto de 2004.

Disposición institucional de los PSP: énfasis en una política de formación radicada en el nivel local

El espacio institucional en que los PSP se asientan y desarrollan es el comunal. La intención que orienta a los PSP es que a nivel comunal se constituyan espacios locales de desarrollo profesional que permitan a los docentes el aprendizaje y/o actualización de las competencias, conocimientos y habilidades establecidas en el Marco para la Buena Enseñanza y, por esta vía, el mejoramiento de sus prácticas docentes. Esta intención de priorizar lo local se advierte en los cuerpos normativos, pues ellos definen que son los departamentos de educación y las corporaciones municipales los facultados para decidir el uso de los recursos destinados para las actividades de formación, de acuerdo a planes que estas entidades han de diseñar, ejecutar, monitorear y, eventualmente, evaluar. Por cierto, se podría argüir que esta radicación de los PSP en el nivel local se desprende de la propia localización del proceso de Evaluación Docente, que se aplica en forma descentralizada correspondiéndole a las Comisiones Comunales de Evaluación, entidades colegiadas conformadas por los evaluadores pares y la autoridad educacional local, una función decisiva en la evaluación de la dotación docente de la respectiva comuna⁹.

Con todo y a pesar de esta localización comunal de la Evaluación Docente, no se desprende de ello una necesidad ineludible de que los PSP se lleven a cabo en dicho nivel territorial. También la escuela o el nivel regional constituyen espacios institucionales que presentan ventajas para llevar a cabo una acción formativa. Se ha sostenido, por ejemplo, que la escuela es un ámbito más cercano y propicio para una formación remedial que enfatice el aprendizaje activo surgido de la reflexión crítica de la práctica docente en el aula y del trabajo formativo entre pares (Wei, Darling–Hammond, Andree, Richardson & Orphanos, 2009). Por otro lado, el nivel regional permite la concentración y utilización, para la formación docente, de recursos escasos que probablemente no se distribuyen de manera homogénea entre todas las comunas de una región. Dadas estas dos alternativas posibles, la opción de ubicar los PSP en el nivel comunal indica que el Ministerio de Educación ha hecho, precisamente, una elección apuntando a generar en las municipalidades espacios en los que se encuentren y dialoguen entre sí diferentes realidades docentes de las que se derivan las necesidades de formación de los profesores (Fullan, citado en Marchesi & Martin, 1998; Núñez, 2000; Pink, citado en Fullan & Stiegelbauer, 1997). Al mismo tiempo, al relevar el ámbito municipal se busca favorecer un asentamiento local de recursos profesionales e institucionales relacionados con la formación continua de los docentes. Estos recursos profesionales debieran tener conocimiento de las realidades educacionales, la especificidad cultural y las características sociales en que los docentes desarrollan su trabajo. Con ello se fomentaría un mejor perfeccionamiento profesional, más relevante y significativo para los docentes involucrados, y con ello aumentarían las posibilidades de lograr un impacto positivo sobre el desempeño docente. En forma adicional, el encuentro de un número importante de docentes pertenecientes a realidades distintas, pero que comparten rasgos comunes, supone un refuerzo de la solidaridad y el sentido de pertenencia a un colectivo profesional, lo que no es irrelevante cuando ello ocurre en un contexto situacional de estrés provocado por la categorización del profesional como deficiente en su ejercicio docente.

Actores y dinámica funcional de los PSP

Las actividades de formación docente que se inscriben en los Planes de Superación Profesional son delineadas en su dinámica funcional por el Reglamento sobre Evaluación Docente. En su Título IX, este cuerpo normativo determina que los municipios y corporaciones municipales que llevan a cabo las acciones de formación reciben los recursos financieros del Ministerio de Educación a través del Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas (CPEIP). Corresponde a este último organismo celebrar convenios con las entidades edilicias para el traspaso de fondos, cumpliendo además un rol de monitoreo y supervisión de las actividades de formación que las municipalidades organicen empleando esos recursos.

9 Decreto 192, artículo 36.

Los montos transferidos por el CPEIP a los municipios por concepto de fondos para actividades de formación en el marco de los PSP, dependerán del número de docentes evaluados con resultado Básico e Insatisfactorio en cada comuna, asignándose un monto a transferir por cada docente evaluado en dichas categorías⁽¹⁰⁾. Como se observa en la Tabla 4, durante el año 2010 este monto ascendió a \$61.000 anuales por cada docente calificado con Básico⁽¹¹⁾ y a \$220.000 anuales por cada evaluado en el nivel Insatisfactorio.

• **Tabla 4**
 • *Monto anual transferido a la comuna por cada docente beneficiario de PSP (2005–2010)*

Año PSP	2005	2006	2007	2008	2009	2010
Docentes con nivel de desempeño Básico	\$50.000	\$50.000	\$50.000	\$56.000	\$60.000	\$61.000
Docentes con nivel de desempeño Insatisfactorio	\$150.000	\$150.000	\$161.650	\$181.000	\$220.000	\$220.000

Aparte de las funciones de transferencia financiera, el CPEIP desempeña un conjunto de responsabilidades: le corresponde una función auditora, expresada en cautelar que los fondos se destinen efectivamente a las finalidades propuestas y, muy relevantemente, asume una función monitorea y evaluadora consistente en verificar que los PSP diseñados e implementados por las municipalidades se compongan de acciones “adecuadas, pertinentes, oportunas y efectivas”. Para ello, el reglamento señala que el CPEIP debe verificar las condiciones en que se desarrollan los PSP y la opinión de los docentes participantes. En caso de que esta verificación concluya que el PSP municipal ha sido deficiente, el CPEIP puede solicitar, previo a la aprobación del siguiente PSP, que éste sea revisado por un profesional experto en formación docente.

Dentro del CPEIP la unidad en la que recae este conjunto de funciones financieras, auditoras, monitoras, supervisoras y evaluadoras de los PSP es el Área de Acreditación y Evaluación Docente. Para llevar a cabo esas funciones, esta unidad ha generado un convenio con la Pontificia Universidad Católica de Chile, en el marco del cual se creó una plataforma informática que brinda soporte administrativo y permite un seguimiento de los PSP comunales⁽¹²⁾.

En el caso de las municipalidades, los responsables funcionales de las actividades de formación de los PSP son los Departamentos de Administración de Educación Municipal (DAEM) o las Corporaciones Municipales de Educación. Cada Departamento o Corporación asumirá la interlocución con el CPEIP de acuerdo a la realidad institucional que la caracterice; en algunos casos el encargado de los PSP será la máxima autoridad de la instancia municipal educacional, mientras que en otros la responsabilidad recaerá en un funcionario subalterno de éste. En términos genéricos, la denominación otorgada para quien tiene la responsabilidad de estas acciones de formación a nivel local es la de “Encargado Comunal de PSP” y el grado de dedicación laboral dependerá del tamaño, complejidad y especialización funcional de la organización municipal, pudiendo este funcionario tener o no dedicación exclusiva a los PSP.

Las funciones de los Encargados Comunales son las de: diseño de los PSP, definición de qué profesional o institución realizará las actividades de formación, monitoreo del desarrollo de estas actividades, administración de los recursos financieros que los PSP demandan y rendición al CPEIP de los gastos incurridos con cargo a los fondos transferidos por éste a la comuna.

10 Se excluyen de ser considerados en la transferencia financiera los docentes que habiendo sido evaluados abandonarán la dotación por causas como retiro o jubilación durante el año que les corresponda realizar PSP, y aquellos que dejarán de desempeñarse en aula, pasando a tener cargos administrativos.

11 Este monto se vuelve a entregar cada año durante un período de cuatro años.

12 Esta plataforma denominada “PSP Online” permite el acceso diferenciado de diferentes actores: el Área encargada del CPEIP, los encargados de PSP en cada comuna del país, los profesionales e instituciones que imparten las actividades de formación y los propios docentes que participan en ellas.

Por su parte, quienes realizan las acciones de formación en los Planes de Superación son denominados "Ejecutores" y pueden ser instituciones o profesionales. Entre las primeras se encuentran instituciones de Educación Superior autónomas así como otras instituciones acreditadas en el CPEIP como formadoras de docentes; mientras que los profesionales que pueden realizar PSP corresponden a: profesores pertenecientes a la Red de Maestros de Maestros, docentes que cuentan con la acreditación de Asignación Excelencia Pedagógica aunque no pertenezcan a dicha red, docentes que han obtenido el máximo nivel de desempeño en la Evaluación Docente ("Destacado") y Jefes Técnicos y otros profesionales de los Departamentos o Corporaciones Municipales de Educación. Son funciones de estos Ejecutores impartir las actividades de formación, dar cuenta de las actividades realizadas a los Encargados Comunales y evaluar a los docentes participantes.

El CPEIP ha buscado orientar que el diseño de las actividades de formación de los PSP sea coherente con las debilidades o falencias identificadas por el proceso de Evaluación Docente y para ello, la plataforma informática que los Encargados Comunales utilizan para realizar este diseño suministra información detallada de los resultados de los docentes de la comuna y sugiere áreas de actividades relacionadas con los déficits detectados en la evaluación. También, la plataforma proporciona funcionalidades para el seguimiento financiero y pedagógico de las actividades realizadas, el registro de las evaluaciones de los alumnos y la recolección del grado de satisfacción usuaria de los asistentes a actividades de los PSP.

Para concluir esta breve descripción de los actores y lógicas de funcionamiento de los PSP, conviene mencionar que éstos consideran distintas modalidades de actividades de formación: a) tutorías o asesorías provistas por profesionales; b) cursos, talleres o seminarios; c) lecturas de bibliografía recomendada; y d) observaciones de clases de docentes destacados.

Los PSP: una dinámica de asentamiento y expansión

Como se ha expuesto anteriormente, la población de docentes beneficiarios de los PSP depende directamente de la Evaluación Docente. En concreto, en un año dado, la cobertura anual potencial de los PSP dependerá del contingente de docentes evaluados bajo el nivel de desempeño esperado (el nivel Competente), a los que se añadirán aquellos con resultado Básico de años anteriores a los que aún no les corresponde re-evaluarse. En la Tabla 5, se presenta el número de docentes evaluados en cada uno de los niveles de desempeño y el subconjunto de ellos que, por su resultado, es beneficiario de los PSP⁽¹³⁾.

● **Tabla 5**
● *Número de docentes por nivel de desempeño en Evaluación Docente (2005–2010)*

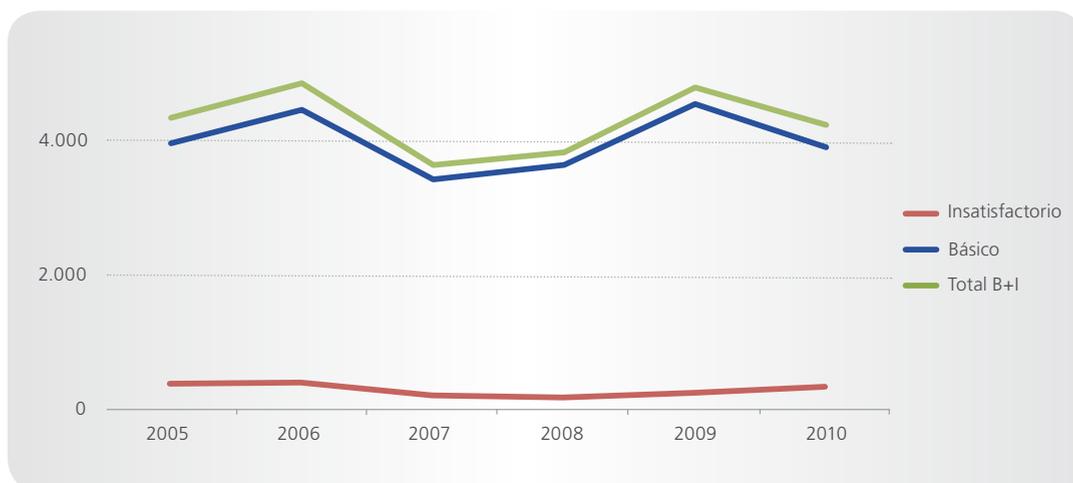
Año Evaluación	2005		2006		2007		2008		2009		2010	
	N	%	N	%	N	%	N	%	N	%	N	%
(A) Insatisfactorio	373	3,5	390	2,8	206	2,0	179	1,1	242	1,5	328	3,0
(B) Básico	3.964	37,2	4.459	31,4	3.432	33,0	3.647	22,8	4.543	28,9	3.905	35,3
Subtotal A+B	4.337		4.849		3.638		3.826		4.785		4.233	
(C) Competente	5.620	52,7	8.371	59,0	5.894	56,6	10.245	64,0	9.899	63,1	6.281	56,8
(D) Destacado	708	6,6	970	6,8	881	8,4	1.944	12,1	1.015	6,5	547	4,9
Total (A+B+C+D)	10.665	100,0	14.190	100,0	10.413	100,0	16.015	100,0	15.699	100,0	11.061	100,0

13 Información válida a marzo de 2011

Como se advierte, la población docente susceptible de realizar actividades en los PSP durante el período 2005–2010 ha tendido a situarse sobre los cuatro mil docentes, registrándose caídas bastante pronunciadas en los años 2007⁽¹⁴⁾ y 2008. En términos comparativos, la proporción de docentes con desempeño Básico e Insatisfactorio respecto del total de evaluados alcanzó su magnitud más alta en 2005 (40,7%), y la menor en el año 2008 (23,9%); en las restantes evaluaciones este porcentaje se sitúa alrededor del 30%. Cabe consignar que el año 2010 el número total de docentes participantes en la Evaluación Docente experimentó una disminución significativa respecto del año anterior (29,5%)⁽¹⁵⁾; sin embargo, la cantidad de evaluados con nivel Básico e Insatisfactorio es similar a la de años anteriores, pues este período registró una proporción comparativamente más alta de ellos (35,9%, la más alta después de 2005 y 2004).

Una representación de estas tendencias se presenta en el Gráfico 1. Éste muestra que, dado el escaso número de docentes evaluados con nivel Insatisfactorio, la población potencial de docentes que pueden seguir actividades de formación en los PSP cada año, está conformada principalmente por docentes con resultado Básico.

● **Gráfico 1**
 ● *Número de docentes evaluados con nivel Básico e Insatisfactorio (2005–2010)*



Conviene recordar que, puesto que hasta ahora los docentes con resultado Básico se evalúan nuevamente solo después de 4 años, la población habilitada para participar en los PSP se engrosa año a año sumándose los docentes calificados bajo el nivel de logro esperado a aquellos con resultado Básico de períodos anteriores. De allí que la población docente beneficiaria de los PSP ha crecido de 1.473 docentes en el año 2005 a 13.586 en 2010 (ver Tabla 6).

Al tenor de los datos presentados en la Tabla 6, es posible sostener que desde el punto de vista de la población beneficiada por los PSP, existe un período de instalación del programa en los años 2005 y 2006, en que ésta bordea los seis mil docentes, mientras que a partir de 2007 se sitúa alrededor de los diez mil. De ellos, como ya se ha señalado, la mayoría (siempre sobre el 93%), son profesores cuyo desempeño ha sido calificado como Básico.

14 En este año la cantidad de docentes evaluados se redujo en un 26% respecto del año precedente; de manera similar, el grupo evaluado con nivel Básico e Insatisfactorio alcanzó a 1.211 docentes, 25% menos que el año anterior.
 15 Esta disminución obedeció fundamentalmente a una menor inscripción —junto con una mayor cantidad de evaluaciones suspendidas— en las regiones afectadas por el terremoto ocurrido el 27 de febrero del mismo año.

● **Tabla 6**
● *Docentes habilitados para participar en los PSP según nivel de desempeño (2005–2010)*

Año PSP ^(*)	2005		2006		2007		2008		2009		2010	
	N	%	N	%	N	%	N	%	N	%	N	%
Docentes con resultado Básico habilitados para PSP	1.394	94,6	5.203	93,2	9.550	95,6	10.474	97,3	12.766	98,1	13.327	98,1
Docentes con resultado Insatisfactorio habilitados para PSP	79	5,4	381	6,8	436	4,4	291	2,7	256	1,9	259	1,9
Total docentes B+I habilitados	1.473	100,0	5.584	100,0	9.986	100,0	10.765	100,0	13.022	100,0	13.586	100,0

Fuente: PSP Online.

(*) El total de docentes habilitados para PSP en un determinado año (por ejemplo, 2007) se compone de aquellos que fueron evaluados con resultado Básico e Insatisfactorio el año inmediatamente anterior (2006, siguiendo el ejemplo), así como de los docentes que obtuvieron nivel Básico vigente en años anteriores (por ejemplo, 2005), a los que aún no les ha correspondido reevaluarse.

Las tendencias ilustradas anteriormente se reflejan también en el número de comunas involucradas en los PSP, el que casi se triplica entre 2005 y 2006, y supera las 330 comunas (de las 346 del país) a partir de 2007 (ver tabla 7). Esto da cuenta de la cobertura nacional que han alcanzado los PSP, ascendiendo desde un 30% del total de comunas en el año 2006, a un 98,8% en 2010.

● **Tabla 7**
● *Comunas que deben desarrollar Planes de Superación Profesional 2005–2010*

Año PSP	2005		2006		2007		2008		2009		2010	
	N	%	N	%	N	%	N	%	N	%	N	%
Comunas que deben hacer PSP	105	30,3	296	85,5	338	97,7	341	98,6	341	98,6	342	98,8
Comunas que no deben hacer PSP	241	69,7	50	14,5	8	2,3	5	1,4	5	1,4	4	1,2
Total de comunas	346	100,0	346	100,0	346	100,0	346	100,0	346	100,0	346	100,0

A pesar de la obligación de las municipalidades de realizar actividades de formación en el marco de los PSP en tanto tengan en su dotación docente profesores evaluados como Básicos o Insatisfactorios, de ello no se desprende que Municipalidades o Corporaciones Municipales efectivamente los lleven a cabo (ver Tabla 8). Esto puede ocurrir porque los planes presentados no son aprobados por CPEIP en razón de su bajo nivel de calidad, o bien, porque aun siendo aprobados, los planes no son ejecutados (ver Tabla 8). Así, el año 2005, de 105 comunas que debían hacer PSP solo 72 los llevaron a cabo (68,6%), mientras que en 2009 eso ocurrió con 299 de las 341 comunas habilitadas para realizar estas actividades de formación (87,7%)⁽¹⁶⁾.

16 Las causas que explican este comportamiento de algunas comunas en relación a aprobar el Plan de Superación y luego no ejecutarlo no han sido exploradas de forma sistemática. La información hasta ahora disponible refiere a la escasa capacidad de gestión de algunas comunas, los escasos recursos asociados cuando el número de beneficiarios en el municipio es muy reducido (asociada a veces a la dificultad de encontrar ejecutores) y la deficiente capacidad de convocatoria a los docentes en algunas comunas, razones todas que parecen incidir en la no ejecución de los planes.

Los datos de la Tabla 8 dan cuenta de que el porcentaje de aprobación de los PSP por parte del CPEIP es alto (sobre el 95% en promedio) y que la ejecución de los planes así aprobados ha crecido significativamente en el período considerado en casi 20 puntos porcentuales (desde 68,6% en 2005 a 87,7% en 2009). Con todo, existe una cantidad variable (de 17 a 64, según el período y 37 como promedio), que teniendo planes aprobados, finalmente no los ejecuta.

• **Tabla 8**
 • *Comunas con PSP aprobado y ejecutado (período 2005–2009)⁽¹⁷⁾*

Año PSP	2005		2006		2007		2008		2009	
	N	%	N	%	N	%	N	%	N	%
Total comunas que deben hacer PSP	105	100,0	296	100,0	338	100,0	341	100,0	341	100,0
Comunas con PSP aprobado	89	84,8*	296	100,0	327	96,7	338	99,1	331	97,1
Comunas con PSP ejecutado	72	68,6*	232	78,4	296	87,6	297	87,1	299	87,7

Fuente: PSP Online.

* Porcentaje en relación al total de comunas que deben realizar PSP.

Además del número de beneficiarios potenciales de los PSP y el número de comunas que ejecuta estos planes, una tercera dimensión que da testimonio de la expansión de cobertura de la política de desarrollo profesional enmarcada en los PSP, es el número y proporción de docentes habilitados, que efectivamente realiza acciones de formación (ver Tabla 9). Para analizar estos datos ha de considerarse que, una vez que el CPEIP ha aprobado el Plan de Superación Profesional presentado por un municipio, es obligación del Encargado Comunal asignar a cada uno de los docentes participantes a al menos una actividad de dicho plan⁽¹⁸⁾. Posteriormente, las instituciones o profesionales encargados de las actividades de formación (“Ejecutores”) inscribirán a los docentes con resultado Básico e Insatisfactorio que les han asignados, en sesiones específicas⁽¹⁹⁾.

A partir de datos registrados en la plataforma informática PSP Online⁽²⁰⁾, la información presentada en la Tabla 9 permite advertir que del total anual de docentes habilitados para realizar actividades de formación de los PSP, el porcentaje que efectivamente concurre a una sesión de formación ha aumentado en los tres últimos años en más de 12 puntos porcentuales. Aun así, en el año 2009 solo la mitad de los profesores que debían asistir a actividades de los PSP realmente lo hacían.

17 Los datos de esta tabla y de la siguiente incluyen solo hasta el año 2009, pues se encuentra en período de consolidación los datos correspondientes al año 2010.

18 Como se mencionara antes, estas actividades pueden ser: tutorías o asesorías provistas por profesionales, cursos, talleres o seminarios; lecturas de bibliografía recomendada; u observaciones de clases de docentes destacados.

19 En los PSP las **actividades** corresponden a un conjunto de acciones formativas referidas a objetivos de aprendizaje específicos y las **sesiones** son las unidades temporales en que se divide el trabajo de dichas actividades. Así, la actividad de una comuna de la Cuarta Región denominada “Estructura de una clase para Educación Media” consta de seis sesiones de una hora cronológica cada una.

20 Estos datos están disponibles solo desde 2007 pues a partir de ese año la plataforma informática PSP Online incorporó funcionalidades para registrar la inscripción de docentes por actividades y sesiones.

• **Tabla 9**
 • *Asignación a actividades y asistencia de docentes a PSP (2007–2009)*

Año PSP	2007		2008		2009	
	N	%	N	%	N	%
Docentes asignados por Encargado Comunal a una o más actividades de PSP	9.747	93,5	10.522	97,7	12.713	99,7
Docentes inscritos por Ejecutor en una o más sesiones de PSP	s/i	s/i	8.156	75,8	10.354	81,2
Docentes que asisten al menos a una sesión de PSP	3.979	38,2	3.903	36,3	6.440	50,5
Total docentes con cobertura para asistir a PSP	10.428	100,0	10.765	100,0	12.753	100,0

La reducción en el número de docentes que asiste a los PSP respecto del número de habilitados proviene en parte de omisiones en los procesos de *asignación a actividades* por parte de los Encargados Comunales y de *inscripción en sesiones* por parte de los Ejecutores; sin embargo, la merma más sustantiva obedece a los docentes inscritos por los Ejecutores en sesiones de PSP que no se presentan a ellas. Si nos concentramos en esta brecha solamente, el porcentaje de no presentación de docentes inscritos en sesiones de formación docente de los PSP ha disminuido de 52,1% en 2008 a 37,8% en 2009⁽²¹⁾.

La explicación de esta significativa tasa de inasistencia tiene al menos dos posibles resortes: por un lado puede haber un sub-registro de los Ejecutores tanto en la inscripción de docentes asignados por el Encargado Comunal, como en el registro de la asistencia misma; por otra parte, es indudable que existe una expresa renuencia de parte de algunos docentes a asistir a sesiones de Planes de Superación, según se ha constatado a través de entrevistas a Encargados Comunales y Ejecutores.

Esta actitud docente podría obedecer a un conjunto de motivos. Una posibilidad es que los docentes no perciban ninguna consecuencia negativa en el corto plazo por la decisión de restarse de las actividades de formación. De hecho, el Reglamento sobre Evaluación Docente señala la obligatoriedad de los PSP, pero no contempla la posibilidad de que el docente de hecho no asista a ellos y, consiguientemente, no establece sanción alguna cuando ello ocurre. Por lo tanto, si existen sanciones o medidas de presión, ellas quedan libradas a la iniciativa específica de cada DAEM o Corporación Municipal. Otra posibilidad es que los docentes no signifiquen los Planes de Superación como actividades de formación relevantes y útiles para su desempeño profesional o que la asistencia a estas actividades sea concebida como una forma de legitimación de una Evaluación Docente cuya validez se cuestiona. Por último, es posible que algunos docentes beneficiarios de los PSP den una connotación negativa a su participación en ellos al no visualizarlos como una oportunidad de desarrollo profesional, sino como un efecto no deseado producto de su bajo desempeño en la evaluación.

Al concluir esta sección queda de manifiesto una política de formación que desde sus inicios en 2005 se ha asentado nacionalmente en su cobertura municipal y que alcanza un número importante y creciente de docentes beneficiarios. Esta expansión de la política se manifiesta también en la magnitud de los recursos transferidos por el CPEIP a las municipalidades para financiar actividades de los PSP; ese monto ha aumentado desde \$25.917.000 en el año 2005 a \$624.826.415 en 2009. Asimismo, del total de recursos del año 2008⁽²²⁾, correspondiente a \$555.600.000, las municipalidades utilizaron el 93,5% de los montos transferidos. No obstante esta expansión de la cobertura, se constata un número minoritario

21 Estos porcentajes se obtienen de la proporción entre docentes con asistencia al menos a una sesión sobre el total de docentes asignados a sesiones, la que al restarse de un 100 por ciento, da la tasa de ausentismo.

22 Último año del que se tiene información del total de recursos utilizados por las municipalidades.

—mas no insignificante— de comunas que no se pliega a esta política y persiste una cantidad importante de docentes que no son alcanzados por las actividades de formación, pese a haber demostrado en la Evaluación Docente un desempeño bajo el nivel esperado.

Efectos de una política de formación continua enmarcada en los PSP

Si la política de formación concebida en el marco de los Planes de Superación Profesional tiene los efectos buscados de superar las debilidades profesionales que evidencian los docentes con desempeño Básico e Insatisfactorio, es algo que, con los datos actuales, no es posible de estimar de manera concluyente. Un efecto inmediato de la superación de dichas debilidades consistiría en que los profesores que asistieron a actividades de PSP obtuviesen un mejor desempeño en la evaluación siguiente. Consiguientemente, en el mediano plazo, su labor docente con los alumnos debería verse favorecida propiciando que éstos exhibieran aprendizajes de mejor calidad. La posibilidad, entonces, de evaluar estas relaciones descansa en un estudio de seguimiento que analice una cohorte de docentes sometidos a la Evaluación Docente y su evolución a lo largo del tiempo. En ausencia de este tipo de datos examinaremos en esta sección algunas características que parecieran favorecer acciones de formación “adecuadas, pertinentes, oportunas y efectivas” y antecedentes preliminares sobre cuál ha sido el desempeño de los profesores a los que les ha correspondido reevaluarse.

El carácter local de las actividades de formación enmarcadas en los PSP hace que un Plan de Superación adecuado y pertinente descansa en buena medida en la capacidad profesional de los DAEM o las Corporaciones Municipales de Educación, los encargados comunales de PSP y especialmente en las instituciones o profesionales que realizan las actividades de formación (Ejecutores).

Por cierto, esta conjunción de factores debe entenderse en el contexto de planes que son llevados a cabo por municipios de muy diferente tamaño, recursos financieros y complejidad organizacional, así como por montos totales transferidos por CPEIP que varían enormemente por el tamaño de la planta docente afecta al programa PSP. Así, por ejemplo, el año 2010 una de las comunas urbanas más importantes del norte del país, con 168 docentes con resultado Básico, recibió una transferencia de \$10.248.000 para realizar actividades de formación, mientras otra comuna de la costa central, con 12 docentes en la misma categoría, dispuso de solo \$732.000. Dadas estas diferencias y la posibilidad de allegar, para las actividades de formación de los PSP, recursos adicionales provenientes de los fondos municipales, es que el respaldo del Alcalde a este tipo de políticas parece ser decisivo. Una autoridad edilicia con una visión positiva de la Evaluación Docente y comprometida con el mejoramiento de la calidad docente en su comuna, presentará mayor probabilidad de respaldar los PSP y que estos se realicen con un alto nivel de exigencia profesional (Santelices, Taut & Valencia, 2008).

Por cierto, lo anterior también depende de contar con una Dirección de Educación Municipal en sintonía con el Alcalde y con un Encargado Comunal que posea la formación y experiencia profesional necesarias para asumir con autoridad técnica la lectura de los resultados de la Evaluación Docente, el diseño de planes, la selección de entidades o profesionales ejecutores de calidad, y una adecuada y eficaz relación de supervisión con dichas entidades.

Un segundo ámbito donde se juegan las posibilidades de contar con buenos Planes de Superación reside en quienes llevan a cabo las acciones de formación: los Ejecutores. La naturaleza institucional de estos profesionales no es un tópico aparentemente relevante. En el estudio ya citado del Centro de Medición MIDE UC sobre Resultados y Análisis de Monitoreo de Planes de Superación Profesional, se detectan excelentes planes que son llevados a cabo tanto por profesores pertenecientes a la Red de Maestros de Maestros, como por docentes de universidades regionales. Lo esencial parece residir en el compromiso de dichos profesionales con las acciones de perfeccionamiento que realizan, su experiencia en el campo

de la educación de adultos, la excelencia profesional expresada en los conocimientos pedagógicos que poseen y la comprensión y dominio del Marco para la Buena Enseñanza (cfr. Santelices, Taut & Valencia, 2008). Conviene señalar que algunos de estos factores, como la experiencia en educación de adultos y la excelencia profesional, coinciden con una revisión de la literatura especializada sobre qué características debiera tener la formación docente para ser efectiva (ver Cortés, Taut, Santelices & Lagos, 2010).

Un tercer tema vinculado a los potenciales efectos positivos de las actividades de formación docente dice relación con la duración y modalidades que asumen los Planes de Superación. La literatura especializada indica que el tiempo de duración del proceso formativo es importante para la consecución de efectos beneficiosos, como también lo es que las actividades de formación utilicen estrategias de capacitación aplicadas a la realidad de los docentes, esto es, actividades de formación en la misma sala de clases y con seguimiento de lo aprendido (Cortés, Taut, Santelices & Lagos, 2010). Considerando los datos registrados en la plataforma informática PSP Online, durante 2008 las actividades de los PSP consistieron mayoritariamente en cursos, talleres o seminarios (76,3% de las actividades planificadas), seguidos muy de lejos por tutorías o asesorías (16,5%), y lecturas recomendadas (3,8%). Una metodología de enseñanza más centrada en el análisis y reflexión de la práctica docente, recomendada por la literatura especializada, como es la observación de clase, solo alcanzó a un 3,3%. En cuanto a la extensión de las actividades de formación se pudo advertir que el 50% de las actividades planificadas en los PSP posee una duración igual o menor a 8 horas. Estos dos últimos datos ponen una nota de duda acerca de cuánto se acercan las actividades de formación enmarcadas en los PSP a un proceso que siga de cerca las recomendaciones para una formación docente exitosa que se pueden encontrar en artículos de investigación científica referidos al tema.

Por otro lado, un aspecto que los PSP logran abordar como efecto no buscado de su realización, es la capacidad de contener a profesionales que han experimentado un fuerte impacto en su autoimagen y autoestima debido al proceso de evaluación. Las entrevistas realizadas a profesionales que llevan a cabo actividades de formación en el marco de los PSP, son coincidentes al respecto: el manejo de expectativas es uno de los primeros aspectos que debe tratarse en los PSP a fin de que los docentes evaluados puedan dar comienzo a un proceso de aprendizaje. Al mismo tiempo los docentes con resultado Básico e Insatisfactorio reportan que "los PSP permiten el encuentro y diálogo con otros docentes que han experimentado situaciones laborales similares y que ello constituye una instancia de apoyo mutuo" (Santelices, Taut & Valencia, 2008).

Aparte de las características que los PSP despliegan, otra forma de sopesar sus efectos es examinar cuales son los resultados en la Evaluación Docente de profesores, tras haber asistido a ellos. A pesar de la heterogeneidad de las actividades de formación contempladas en los Planes de Superación y de las limitaciones anotadas anteriormente, un estudio preliminar (Taut & Valencia, 2011) sobre el rendimiento de docentes con resultado Insatisfactorio en 2008 que se re-evaluaron en 2009, indica que aquellos que asistieron a los PSP alcanzaron un mejor desempeño que quienes no lo hicieron, ya sea que se hubiesen inscrito o no en los planes (ver Tabla 10). De esta forma, un 23,4% de docentes con un resultado Insatisfactorio en la Evaluación Docente 2008, tras haber participado en Planes de Superación Profesional logran un desempeño Competente en su re-evaluación (correspondiente al 2009), en comparación al 0% de docentes evaluados como Insatisfactorios que no han asistido a actividades de formación, y a un 8,3% de los que no han estado inscritos en los PSP.

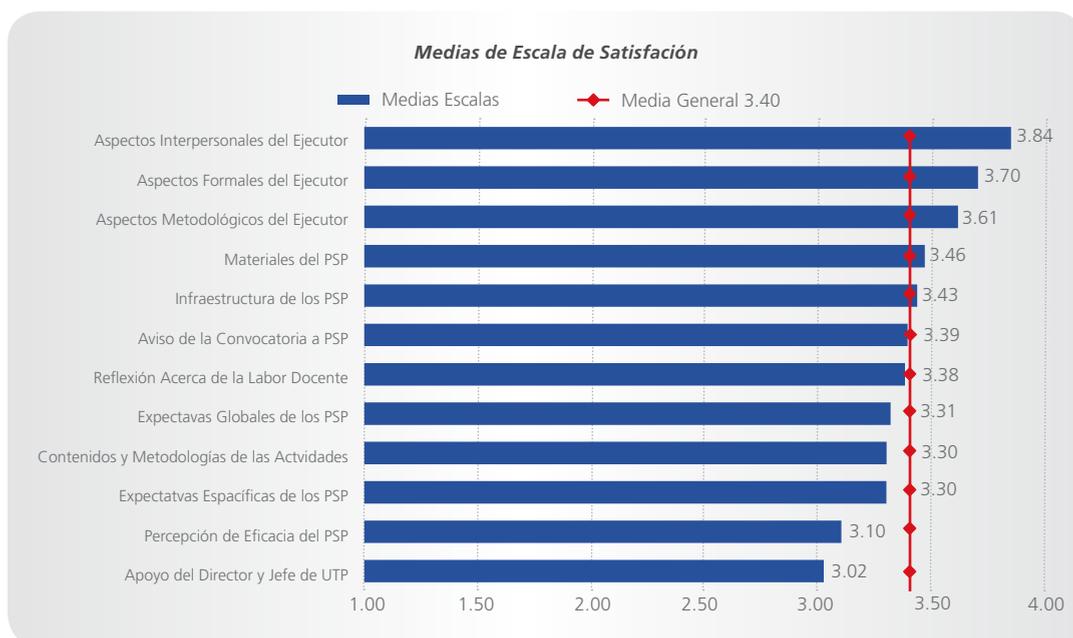
● **Tabla 10**
 ● *Nivel de desempeño en reevaluación (2009) de docentes con resultado Insatisfactorio en la Evaluación Docente de 2008*

	No Inscritos en los PSP (N:12)	Inscritos en los PSP pero no asisten (N:12)	Inscritos en los PSP y asisten (N:64)	Total
Insatisfactorio	16,7%	33,3%	20,3%	21,6%
Básico	75,0%	66,7%	56,3%	60,2%
Competente	8,3%	0,0%	23,4%	18,2%
Total	100,0%	100,0%	100,0%	100,0%

Los datos antes mencionados están lejos de ser concluyentes, particularmente por el reducido grupo de docentes examinado, pero entrega indicios de que las actividades de formación de los PSP podrían estar conduciendo a mejores desempeños en la Evaluación Docente.

Por otra parte, más allá de los factores que favorecerían efectos positivos de los PSP y aun considerando que los docentes con resultado Básico e Insatisfactorio se incorporan, como ya hemos visto, a dichas actividades con reticencia, la evaluación general que los participantes realizan al concluir éstas⁽²³⁾, es positiva (ver Gráfico 2). También la apreciación general positiva es reafirmada en el hecho de que un 83,3% de quienes respondieron la encuesta de satisfacción del año 2009 sostienen que volverían a realizar actividades de los PSP y el 85,9% recomendaría las actividades de los PSP a otros docentes (MIDE UC, 2010).

● **Gráfico 2**
 ● *Promedio en Escalas de Satisfacción entre participantes en los PSP 2009 (N: 1.876)*



23 Encuesta de satisfacción respondida por los docentes participantes en PSP, quienes acceden al sitio PSP Online con una clave personal que garantiza la confidencialidad de sus respuestas. Evidencias de este tipo deben ser miradas con cautela, pues solo un 29,1% de los 6.440 docentes asistentes a PSP en 2009, respondió la encuesta de satisfacción que da origen a este gráfico.

Comentario Final

Al tenor de las cifras y los análisis vistos en este capítulo se puede concluir que el programa de Planes de Superación Profesional desde sus inicios en el año 2005, atravesó una fase de instalación y asentamiento durante los dos primeros años, consolidándose a partir de 2007, hasta alcanzar cobertura nacional, con participación de alrededor del 90% de las comunas, la gran mayoría de ellas con realización de actividades de formación de manera recurrente año tras año. Con esa continuidad se han consolidado las prácticas institucionales y la experiencia de profesionales a cargo de los PSP tanto en el nivel local como en el nivel central, y se han fortalecido los mecanismos informáticos y comunicacionales que dan soporte al programa⁽²⁴⁾. En particular, cabe mencionar el mayor porcentaje de profesores asistentes a actividades de formación de los PSP en relación a lo que ocurría dos o tres años atrás, y la cantidad de comunas que logran diseñar e implementar Planes de Superación Profesional, en franco ascenso a través de los años. Todo lo anterior da cuenta de un programa en clara progresión desde la perspectiva de la cobertura.

Por otra parte, en numerosas municipalidades los Planes de Superación tienen una calidad destacada siendo realizados por profesionales con larga experiencia y acabado conocimiento pedagógico.

Las indagaciones sobre satisfacción de los docentes participantes también muestran rasgos promisorios, reportando una buena opinión general acerca de las actividades de formación y los profesionales que imparten los contenidos. Además, los PSP son apreciados como una instancia que, gracias a los Ejecutores de las actividades de formación, permite generar expectativas de mejoramiento del desempeño profesional. Por último, análisis de los resultados de docentes con desempeño Insatisfactorio que son re-evaluados dan testimonio de que estas actividades tienen un efecto positivo en la enseñanza de habilidades y conocimientos pedagógicos.

En este cuadro aparecen, en todo caso, aspectos que llaman a preocupación. El más llamativo es, sin duda, que pese al mejoramiento indicado en páginas anteriores, aún se manifiesta una alta tasa de inasistencia de docentes con desempeño Insatisfactorio y Básico a los PSP. En este sentido los objetivos indicados por los cuerpos normativos que regulan la Evaluación Docente y sus procesos consecuenciales no se están cumpliendo para una cantidad sustancial de docentes. En la medida que los recursos transferidos desde el nivel central a las comunas dependen del número de docentes a los que les correspondería realizar los PSP (y no de los que efectivamente asisten), podría estarse generando un incentivo para que esta situación se mantenga.

Un segundo elemento preocupante es que los recursos provistos para la realización de actividades en los PSP resultan bastante exiguos. En comunas pequeñas con un número relativamente bajo de docentes evaluados con nivel Básico o Insatisfactorio y sin posibilidades financieras de allegar recursos propios, el proveer de buenos cursos de formación, con una duración adecuada, sistemáticos en el tiempo y con un seguimiento en el aula, parece muy improbable. Una formación docente de calidad requiere necesariamente elevar los fondos que la ley determina para estos efectos. Junto con este mejoramiento, resulta necesario que los aportes desde el nivel central tengan un carácter diferenciado y progresivo, en correspondencia con el nivel de vulnerabilidad de la comuna.

Un elemento adicional a cautelar es la calidad de la formación impartida. Una de las virtudes de los PSP es ser una política con un fuerte acento y autonomía local. Como hemos mencionado, cada municipalidad contrata los Ejecutores que estime convenientes. Sin embargo, no siempre las comunas cuentan con la capacidad técnica para identificar y seleccionar organismos formadores o docentes que ejecuten actividades de calidad, y tampoco es evidente que en comunas pequeñas y distantes de los centros regionales sea factible contar con información completa y detallada acerca de la oferta de formación docente. Al respecto

.....

24 Plataforma informática PSP Online y sitio público www.psp.mineduc.cl

sería conveniente diseñar políticas para la constitución de una oferta de formación más transparente con estándares de calidad reconocidos y públicos que el Estado pueda promover y fiscalizar.

La generación de esos estándares debería contemplar los elementos que la literatura especializada indica como más efectivos en procesos de formación docente y proveer estímulos para que la formación se realice en directa vinculación con la práctica en el aula.

Una política de fijación de estándares para las entidades formadoras habría de complementarse con mecanismos o instrumentos que monitoreen no solo las actividades de los PSP, sino la puesta en práctica de los contenidos aprendidos por parte del docente. La retroalimentación en aula no es parte de las prácticas de enseñanza actuales de los PSP y pareciera esencial para reforzar aprendizajes y potenciar el cambio anhelado.



6

Estudios de validez de la Evaluación Docente

Sandy Taut* · Verónica Santelices**
Jorge Manzi*

* Centro de Medición MIDE UC
Pontificia Universidad Católica de Chile

** Facultad de Educación, Pontificia Universidad Católica de Chile

Estudios de validez de la Evaluación Docente

El presente capítulo describe los estudios de validación de la Evaluación Docente, respondiendo a preguntas tales como: ¿cubre la Evaluación Docente lo señalado por el Marco para la Buena Enseñanza?, ¿existen diferencias significativas en el rendimiento de los alumnos según el resultado obtenido por su profesor en la evaluación?, ¿qué consecuencias positivas o negativas, esperadas y no esperadas, tiene la Evaluación Docente para los profesores, los establecimientos y las municipalidades?

El capítulo comienza describiendo la agenda de investigación que ha implementado desde 2005 el área de investigación del Centro de Medición MIDE UC. Luego se realiza una revisión de la literatura presentando algunas definiciones claves y el pensamiento actual sobre validez. Enseguida se describen los resultados de los estudios de validez, incluyendo algunos desarrollados por otros investigadores que aportan evidencia relevante. El capítulo finaliza planteando una interpretación acerca de la validez de la Evaluación Docente considerando la totalidad de la evidencia presentada.

Agenda de investigación relacionada con validez

Cronbach (1989, p. 151) caracterizó la validación como “un proceso largo, incluso interminable” y Anastasi (1986) postula que casi cualquier información que uno recoge en el proceso de desarrollar o usar una prueba es relevante para su validación. De este modo, Kane (2006) se pregunta, dado que todos los datos son relevantes para la validez, dónde se debería empezar y cuánta evidencia es necesario recopilar para confirmar adecuadamente una afirmación sobre validez. Este autor ofrece algunas recomendaciones, señalando que el proceso de validez debería enfocarse primero en las interpretaciones más comunes de los resultados de una prueba: “Los distintos tipos de evidencia sobre validez son los que apoyan las conclusiones y supuestos centrales de un argumento interpretativo, particularmente aquellos que son más problemáticos” (Kane, 2006, p. 23, traducción de los autores).

Kane reconoce que, en la práctica, la mayor parte de la investigación sobre validez es llevada a cabo por el equipo que ha desarrollado la prueba o medición y, por lo tanto, tiene un sesgo confirmatorio. En la etapa de desarrollo de un programa de evaluación, “es apropiado (y probablemente inevitable) que el equipo que desarrolla la prueba tenga un sesgo confirmatorio, ya que intenta desarrollar el mejor programa posible. De todas formas, en algún punto, especialmente en programas con consecuencias significativas, un distanciamiento apropiado así como un cambio hacia una posición más crítica son necesarios para proporcionar una evaluación convincente de las interpretaciones y usos propuestos” (Kane, 2006, p. 25, traducción de los autores).

En 2005 empezamos la validación del programa de Evaluación Docente, haciendo un esfuerzo por evitar el sesgo confirmatorio mencionado por Kane (2006). Institucionalmente, somos parte del centro universitario que desarrolla la Evaluación Docente, no obstante, desarrollamos nuestro trabajo en forma autónoma y en paralelo al equipo a cargo de su implementación.

Siguiendo las sugerencias encontradas en la literatura, desarrollamos una agenda de investigación comprehensiva sobre la validez de la Evaluación Docente que incluye múltiples tipos de evidencia (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 1999) así como aspectos consecuenciales (Kane, 2006).

En nuestra opinión, las preguntas más relevantes y problemáticas para establecer la validez de la Evaluación Docente son:

- a.** ¿Cubre la Evaluación Docente lo señalado por el Marco para la Buena Enseñanza? Para aportar evidencia a la validez de la evaluación, sus instrumentos, en especial el Portafolio, deben reflejar la mayoría de los aspectos señalados en el documento sobre el cual está fundado;
- b.** ¿Representa el Portafolio diferentes dimensiones (factores) del desempeño docente? Para aportar evidencia de validez del puntaje del Portafolio, los resultados de los docentes deben dar soporte empírico a la estructura de indicadores y dimensiones del Portafolio supuestas en la rúbrica de corrección y los estándares de desempeño docente subyacentes.
- c.** ¿Está el puntaje del Portafolio sesgado por el corrector? Para aportar evidencia de validez del Portafolio, la varianza de los puntajes de éste atribuible a diferencias en la forma en que los correctores asignan puntajes a la evidencia, debe ser pequeña o nula.
- d.** ¿Muestran diferencias en su desempeño en aula los profesores con más alto y más bajo desempeño en la Evaluación Docente? Para aportar evidencia a la validez de la Evaluación Docente, los profesores mejor evaluados debieran mostrar prácticas pedagógicas de mejor calidad que los docentes peor evaluados, al usar instrumentos alternativos para medir aspectos similares y complementarios del quehacer pedagógico.
- e.** ¿Existen diferencias significativas en el rendimiento académico de alumnos de profesores con más alto y más bajo desempeño en la Evaluación Docente? Para aportar evidencia a la validez de la Evaluación Docente, se espera que los alumnos de profesores mejor calificados obtengan mejores resultados en pruebas estandarizadas que los alumnos de profesores con más bajo desempeño.
- f.** ¿Cómo se relacionan el desempeño de profesores evaluados por el programa de Evaluación Docente y el de Asignación de Excelencia Pedagógica (AEP)? Para aportar evidencia a la validez de la Evaluación Docente, los resultados de un profesor evaluado en AEP debiera ser consistente con el obtenido en la Evaluación Docente, especialmente al comparar el resultado de los portafolios de ambos programas.
- g.** ¿Qué consecuencias tiene el programa de Evaluación Docente para los profesores evaluados, establecimientos y municipalidades? Para aportar evidencia de validez, deben existir antecedentes que respalden que el programa de Evaluación Docente tiene consecuencias positivas esperadas, y pocas consecuencias negativas (no esperadas).

La agenda de investigación fue abordada de acuerdo a los recursos disponibles por lo que en un momento se priorizó aquellos estudios que parecían cruciales en relación con la validez del programa de Evaluación Docente. Por ejemplo, decidimos enfocarnos primero en el puntaje final de la evaluación y su validación porque éste tiene consecuencias directas para los profesores evaluados, los establecimientos educacionales y las municipalidades. Respecto a la validación de instrumentos individuales priorizamos el Portafolio, ya que tiene más peso en la determinación del nivel de desempeño final (60%).

Validación y Validez en la Literatura

El concepto de validez ha evolucionado en el tiempo. Kane (2006) entrega un interesante resumen de esa evolución. El autor muestra que luego del desarrollo del modelo de validez de criterio (criterion model) y del modelo de validez de contenido (content model), que surgen en el contexto de la medición para

la selección de personal y la medición educacional respectivamente, en los inicios de los 80 la validación de constructo llegó a ser considerada la base para un modelo unificado de validez (Messick, 1994, 1995). Este modelo unificado incluye todo tipo de evidencia sobre validez: evidencia de contenido y criterio, confiabilidad, generalizabilidad, y todos los métodos relacionados con “testeo” de teoría. Este desarrollo tuvo varias implicancias para la investigación de validación: (1) incrementó la toma de conciencia respecto de la necesidad de especificar la interpretación de los puntajes de una prueba propuesta antes de evaluar su validez; (2) explicitó la necesidad de un programa de investigación, en contraste con un estudio empírico simple; y (3) incorporó la necesidad de considerar interpretaciones alternativas.

Nuestra agenda de validación está basada en una de las fuentes más autorizadas en lo que concierne a la investigación en validación: los Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999). Estos estándares definen validez como “el grado en el que la evidencia y la teoría confirman las interpretaciones de los puntajes, así como los usos propuestos para ellos” (p. 9). Los Estándares concuerdan con Kane (2001), Messick (1994) y otros en que la validación debe enfocarse en la interpretación y usos propuestos de los puntajes de la prueba o medición. Los Estándares diferencian tipos de evidencia de validez, basada en a) el contenido, b) el proceso de respuesta, c) la estructura interna, d) relaciones entre variables, y e) consecuencias.

Debido a su amplio alcance, la actual concepción de validez de pruebas, y en particular el uso de sus resultados para tomar decisiones, también implica una revisión de las consecuencias de las evaluaciones (Linn, Baker & Dunbar, 1991; AERA, APA & NCME, 1999; Kane, 2006). Como Kane (2006, p. 51) dice, “un proceso de medición que informa una decisión [...] se debe evaluar en término de sus efectos o consecuencias”. Otros autores, como Messick (1989, 1995), Linn (1997) y Shepard (1997) concuerdan con Kane y proponen una concepción de validez amplia que incluya una evaluación de ambos, el significado de los puntajes de la prueba y las consecuencias de su uso⁽¹⁾.

Kane (2006) ha sostenido que si un programa de medición juega un rol clave en una reforma para mejorar resultados educacionales, entonces tiene sentido evaluarlo como si fuera un programa educacional, y que las evaluaciones de programa siempre debieran incluir resultados esperados y no esperados de éste. Esto es importante porque, para que los interesados tomen decisiones informadas sobre la efectividad de mediciones con altas consecuencias, necesitan tener información sobre qué tan bien las mediciones logran diferentes propósitos y a qué costo. Asumiendo que hay tanto consecuencias positivas como negativas, las entidades interesadas y los políticos enfrentan la tarea de contrapesar estas consecuencias (Kane, 2006). En concordancia con esta posición, el estudio de la validez consecuencial de la Evaluación Docente ha tenido un lugar destacado en nuestra agenda de investigación.

Estudios de validez implementados hasta la fecha

Esta sección describe brevemente cada uno de los estudios de validación realizados hasta la fecha. Con ellos se intenta responder a las preguntas antes presentadas como guía para nuestra agenda de investigación. Presentaremos los estudios resumidos de acuerdo al tipo de evidencia de validez que abordan: a) contenido, b) estructura interna, c) relaciones entre variables, y d) consecuencias. Las preguntas de investigación relacionadas con el proceso de respuesta están pendientes en la agenda de validación de la Evaluación Docente.

.....

1 Otros investigadores (por ejemplo Popham, 1997) argumentan que mientras las consecuencias son muy importantes de estudiar, no deberían ser consideradas parte del concepto de validez ni de su investigación propiamente tal.

a. Evidencia de validez basada en el contenido

En el año 2005 iniciamos nuestro trabajo estudiando la validez de la Evaluación Docente en términos de su alineamiento con el “Marco para la Buena Enseñanza” (Ministerio de Educación, 2003), documento oficial que sirve de base a este programa de evaluación y sus instrumentos. Las definiciones del Marco para la Buena Enseñanza (en adelante MBE) están categorizadas en cuatro grupos o dominios: Dominio A “Preparación de la enseñanza”, Dominio B “Creación de un ambiente propicio para el aprendizaje”, Dominio C “Enseñanza para el aprendizaje de todos los estudiantes”, y Dominio D “Responsabilidades profesionales”.

El MBE ha sido ampliamente distribuido entre los profesores como un documento adjunto entre los materiales de la evaluación. Este documento tuvo como antecedente directo los estándares para la formación inicial docente elaborados en el marco del Programa de Fortalecimiento de la Formación Inicial Docente (FIDD) del Ministerio de Educación; además, se basó en una amplia revisión de la literatura y la investigación internacional en educación sobre competencias de profesores de aula.

Al analizar el contenido de la rúbrica de puntuación del Portafolio de la Evaluación Docente, constatamos que cubre la gran mayoría de las recomendaciones del MBE: la totalidad de los Dominios B y C, y parcialmente el Dominio A. Respecto de este último, en el Portafolio existe solo una cobertura limitada de indicadores relacionados con prácticas pedagógicas específicas de la disciplina y el conocimiento de contenidos de ésta. Estos aspectos tampoco son medidos directamente por los otros instrumentos, cuyos datos se basan en el juicio u opinión informados por el propio docente, sus supervisores directos y su evaluador par⁽²⁾. Sin embargo, la medición de los conocimientos disciplinarios es abordada por la prueba que deben rendir los docentes interesados en acceder a los beneficios económicos del programa Asignación Variable por Desempeño Individual (AVDI)⁽³⁾. Debido a la naturaleza del Portafolio, éste tampoco evalúa todos los indicadores relacionados con Responsabilidades Profesionales (Dominio D del MBE) cuya evaluación sí se incluye en la Entrevista por el Evaluador Par, el Informe de Referencia de Terceros y la Autoevaluación.

b. Evidencia de validez basada en la estructura interna

Periódicamente se calculan los índices de consistencia interna de los instrumentos empleando el coeficiente alfa de Cronbach. Adicionalmente, se estudia la estructura interna del Portafolio mediante el análisis factorial.

Consistencia interna de los instrumentos de la Evaluación Docente

La Tabla 1 muestra el coeficiente alfa de Cronbach para tres de los cuatro instrumentos entre 2005 y 2008. Los resultados muestran que la consistencia interna de los instrumentos en general ha sido alta y se ha mantenido estable a lo largo del tiempo.

• **Tabla 1**
• *Consistencia interna de los instrumentos de la Evaluación Docente 2005–2008*

Instrumento	2005	2006	2007	2008
Portafolio: Parte escrita	0,81	0,80	0,74	0,78
Portafolio: Filmación de una clase	0,79	0,75	0,70	0,74
Informe de Referencia de Terceros	0,97	0,99	0,98	0,96
Entrevista Evaluador Par	0,80	0,87	0,82	0,81

2 Debe mencionarse que, dentro de la negociación que dio origen al sistema de evaluación, fue consensuado que éste no incluiría la medición directa de conocimientos disciplinarios, enfocándose fundamentalmente en los conocimientos pedagógicos.

3 Para más información sobre AVDI, ver Cortés & Lagos, en este volumen.

Estructura interna del Portafolio

Periódicamente se realizan también análisis factoriales exploratorios al Portafolio de la Evaluación Docente⁽⁴⁾ (Cudeck, 2000; Hoyle & Duvall, 2004). La rúbrica de puntuación para este instrumento contiene 24 indicadores agrupados en ocho dimensiones (ver más abajo). Las primeras cinco dimensiones son evaluadas a partir de la sección escrita del Portafolio (Módulo 1) y las últimas tres con el registro audiovisual de una clase de aula (Módulo 2). Las dimensiones abordan los siguientes aspectos:

- Dimensión A: Organización de los elementos de la unidad
- Dimensión B: Análisis de las actividades de las clases
- Dimensión C: Calidad de la evaluación de la unidad
- Dimensión D: Reflexión a partir de los resultados de la evaluación
- Dimensión E: Reflexión pedagógica
- Dimensión F: Ambiente de la clase
- Dimensión G: Estructura de la clase
- Dimensión H: Interacción pedagógica

La estructura de indicadores y dimensiones del Portafolio es una adaptación de los descriptores y criterios del MBE. Cada dimensión corresponde a la agrupación de un conjunto de tareas pedagógicas afines (indicadores), elaboradas a partir del MBE.

En los análisis factoriales aplicamos el método de extracción de máxima verosimilitud junto a varias reglas de retención de factores (regla de Kaiser, screeplot, interpretabilidad), y comparamos los resultados al utilizar los métodos de rotación Oblimin y Varimax, obteniéndose soluciones factoriales relativamente estables.

Los resultados solo han variado levemente a lo largo de los años; en general se identifican alrededor de cinco o seis factores para el portafolio completo, incluyendo la parte escrita y la clase grabada en video. Juntos, estos factores explican entre un 30% y un 39% de la varianza. Usualmente, tres o cuatro de los factores están asociados con habilidades evaluadas en la parte escrita del Portafolio. Los factores restantes están asociados con la clase filmada y muestran variaciones a través de los años. Respecto de la clase filmada, la mejor solución en términos de correspondencia con las dimensiones preestablecidas fue obtenida en 2006, cuando los tres factores correspondieron a las tres dimensiones que se reportan para ese módulo del Portafolio en los informes de resultados. También, en 2006, para la parte escrita del Portafolio al menos un factor correspondía nítidamente a una de las dimensiones reportadas para esta parte del Portafolio (dimensión C: "Calidad de la evaluación de la unidad").

En síntesis, la estructura factorial identificada no coincide directamente con las ocho dimensiones empleadas para informar el desempeño en el Portafolio. Tampoco coincide con los dominios del MBE. Los análisis del módulo escrito del Portafolio identificaron tres tipos de habilidades: 1) habilidad para diseñar y planificar la enseñanza, a veces dividida en (1a) planificación de la clase y (1b) diseño de actividades pedagógicas; (2) habilidad para construir instrumentos de evaluación; y (3) capacidad reflexiva sobre la propia práctica. Estos factores abarcan típicamente indicadores de más de una dimensión del Portafolio. Al respecto, es importante recordar que en el módulo escrito se informan dos dimensiones asociadas a la planificación, dos a la evaluación y una a la reflexión. En las primeras cuatro dimensiones se combinan indicadores de planificación, implementación y evaluación con otros relacionados con la habilidad de reflexión profesional. La quinta dimensión, en cambio, se centra exclusivamente en la reflexión. En este sentido, los resultados del análisis factorial separan los indicadores relacionados con desempeño de aquellos vinculados con reflexión (lo que sugiere que ambas habilidades son claramente diferenciables), y los reagrupa de una manera diferente a la empleada para reportar los resultados.

.....

4 En el año 2005 también se realizó un análisis factorial a los datos de los demás instrumentos.

Por su parte, el análisis de la clase filmada tiende a quedar representado en uno o dos factores altamente correlacionados (excepto en 2006, cuando fue posible identificar tres factores). Esto sugiere que las habilidades docentes observadas a través del video aparecen interrelacionadas. Pese a ello, tal como se ha podido verificar consistentemente, los promedios de las tres dimensiones reportadas son marcadamente diferentes. Esto indicaría que, aunque los correctores perciben conexiones entre los indicadores, se forman una impresión claramente distinta de cada uno de ellos. Por ejemplo, tienden a hacer una evaluación más favorable de la capacidad de los docentes para manejar los aspectos socio–afectivos (dimensión F) que los aspectos relativos a la interacción pedagógica (dimensión H)⁵.

En su conjunto, estos datos validan parcialmente la forma en que se retroalimenta a los docentes a partir de los resultados del Portafolio, y al mismo tiempo entregan pistas para enriquecer estos reportes a futuro. Por ejemplo, además de identificar las actuales cinco dimensiones del Módulo 1, sería posible retroalimentar el rendimiento de los profesores en aquellos indicadores que implican ejecución de tareas profesionales (planificación y evaluación) versus aquellos que implican reflexión.

Utilizando los datos del año 2005 se realizó un análisis factorial para los instrumentos restantes de la Evaluación Docente: autoevaluación, informe de supervisores (director y jefe de UTP) y entrevista por docente par. Los resultados indican que cada uno de estos instrumentos es uni–factorial, es decir, responde primariamente a una variable latente. De esto se infiere que los profesores, sus supervisores y pares emiten juicios que responden a una apreciación global del desempeño, sin hacer diferenciaciones asociadas a los dominios del Marco para la Buena Enseñanza.

Evidencia de confiabilidad y generalizabilidad

Desde el año 2005 hemos examinado el desempeño de los correctores como una posible fuente de error de medición en la Evaluación Docente en tres ocasiones. En 2005, calculamos la confiabilidad entre jueces, y llevamos a cabo un estudio de generalizabilidad que incluyó a los correctores como una faceta. En 2008, 2009 y 2010 realizamos solo estudios de generalizabilidad.

Nuestros hallazgos en 2005 mostraron que el promedio de confiabilidad entre jueces (IRR) era de 0,61 tanto para el Módulo 1 como para el Módulo 2. Como referencia comparativa, el Estudio del National Research Council (Hakel, Koenig & Elliot, 2008) acerca del National Board for Professional Teaching Standards (NBPTS) reporta una confiabilidad media de las tareas de 0,69. Mientras las dimensiones B a G mostraron índices de IRR entre 0,61 y 0,70, las dimensiones A y H mostraron índices más bajos (0,51 y 0,52 respectivamente).

Los estudios de generalizabilidad permiten determinar el porcentaje de variabilidad en los puntajes obtenidos que es atribuido a (1) diferencias “reales” entre los profesores en su desempeño en el Portafolio; (2) diferencias sistemáticas atribuibles a los correctores, y (3) error “residual”, el cual combina un término de interacción junto a otras fuentes de error no especificadas. En los tres años en que hemos realizado estudios de generalizabilidad, entre un 25% y 50% de la varianza ha sido atribuible a diferencias entre profesores–portafolios. La varianza atribuible a efecto de los correctores ha representado porcentajes pequeños (entre un 3% y 10%) con sus valores específicos variando según dimensión y subsector.

Los coeficientes de generalizabilidad de estos estudios han variado entre 0,31 y 0,76 dependiendo de la dimensión del Portafolio, año y subsector analizado. En 2009 el índice G promedio para el Portafolio fue 0,69 y se observaron diferencias no significativas entre la parte escrita y la filmación de una clase (0,72 y 0,71 respectivamente).

5 Un mayor detalle de los promedios obtenidos por los docentes en las ocho dimensiones del Portafolio aparece en Sun, Correa, Zapata & Carrasco, en este volumen.

A lo largo de los años el diseño e implementación de los estudios de generalizabilidad se ha mejorado y expandido permitiendo un análisis más representativo de subsectores (asignaturas), centros de corrección, y niveles de enseñanza. Mientras que en 2005 se contó solo con observaciones suficientes para evaluar la corrección de cerca de 140 portafolios, todos de profesores generalistas, en 2009 los datos disponibles permitieron incluir una variedad de subsectores, y se pudieron analizar también diferencias en los índices de generalizabilidad entre centros de corrección y niveles de enseñanza (más de 3.000 portafolios fueron doble-corregidos ese año). En 2009 el promedio del índice de generalizabilidad de las dimensiones del Portafolio varió entre 0,65 y 0,75, dependiendo del centro de corrección, y entre 0,64 y 0,70 dependiendo del nivel de enseñanza. Cabe mencionar que no todos los centros corrigieron los mismos subsectores o niveles, por lo que no fue posible incluir estas variables como facetas adicionales del estudio de generalizabilidad.

En un intento por superar la limitación de estudiar solo una faceta (corrector) impuesta por el diseño del proceso de corrección del Portafolio, se diseñó e implementó un estudio de generalizabilidad separado del proceso de corrección real de Evaluación Docente y, por lo tanto, con tiempos y prácticas regidas específicamente por los objetivos del estudio. Este estudio se focalizó en el módulo escrito de 20 portafolios correspondientes al subsector Matemática del Primer Ciclo Básico. En base a los estudios anteriores se definieron las variables a estudiar: faceta corrector (12 correctores) y faceta tiempo (2 ocasiones con siete días de diferencia); además se comparó la generalizabilidad de la faceta corrector en dos modalidades distintas de corrección: (1) corrección por profesor (se corrige el módulo escrito de un profesor en su totalidad) y (2) corrección por dimensión (se corrige la totalidad de la dimensión A para todos los portafolios asignados al corrector, luego la totalidad de la dimensión B, y así sucesivamente).

Los resultados del estudio indican que la mayor cantidad de varianza explicada en el puntaje final del módulo escrito del Portafolio es atribuible a diferencias reales en el nivel de desempeño (48%), seguido por aquellas atribuibles a correctores (31%), y error (17%). El estudio muestra que la faceta tiempo (ocasión) no explica un porcentaje significativo del puntaje final obtenido. Al analizar la varianza de los puntajes por dimensión, los resultados del estudio muestran que en este nivel de desagregación se reduce la proporción de varianza explicada por diferencias reales entre los portafolios (54% el mayor y 20% el menor, por dimensión), aumentando aquella atribuible a error sistemático de los correctores.

El estudio de decisión (D-study)⁶ indica que en las condiciones actuales de corrección del módulo escrito del Portafolio, es decir, un corrector en una ocasión, el G-Index es de 0,73 y el Phi-Index es de 0,43. Estos datos indican que la generalizabilidad es adecuada en lo que respecta al ordenamiento de los puntajes finales (decisiones relativas), pero que el actuar de los correctores no alcanza niveles óptimos para el caso en que las decisiones deben basarse en el puntaje observado por sí mismo (decisiones absolutas). Si bien el estudio tiene limitaciones debido al pequeño número de correctores, portafolios y subsectores evaluados, y a que no reproduce completamente las condiciones reales de corrección, los resultados muestran que las facetas tiempo y método de corrección no afectan los puntajes finales obtenidos por los profesores. Además, observamos que la consistencia entre correctores es mayor cuando éstos evalúan aspectos más globales que específicos.

Conviene mencionar que algunas de las recomendaciones discutidas en Myford y Engelhard (2001) en el contexto de moderada confiabilidad del sistema de medición del National Board for Professional Teaching Standards (NBPTS) ya han sido implementadas en la Evaluación Docente en Chile; por ejemplo, recolectar información detallada acerca del desempeño de los correctores presentado en tablas o gráficos, y hacerlo en tiempo real de modo que los supervisores puedan reforzar el entrenamiento a correctores específicos, de manera individual y temprana en el proceso. No obstante, aún pueden enriquecerse estas acciones. Algunas modificaciones a considerar son: que los correctores evalúen portafolios conocidamente heterogéneos durante el proceso de capacitación con el objetivo de

6 Un D-study permite estimar índices de generalizabilidad en situaciones hipotéticas de combinaciones de facetas (Brennan, 2001; Shavelson & Webb, 1991).

establecer el uso que hacen de la escala de puntuación; avanzar hacia la doble-corrección de la totalidad de los portafolios (no solo el 20% como ocurre hoy); reducir el número de dimensiones evaluadas y explorar el uso de una rúbrica de corrección holística.

Estudio de Sesgo

Dentro de este ámbito, hasta ahora se han llevado a cabo análisis preliminares a partir de los datos del Portafolio 2009 y utilizando el procedimiento de Mantel-Haenszel para ítems politómicos y análisis de invarianza de medición. De esta forma se ha comparado el desempeño de ítems entre docentes de establecimientos rurales y urbanos y entre hombres y mujeres, sin encontrarse evidencia de sesgo en contra de los grupos focales (docentes de escuelas rurales y hombres). En el futuro se profundizará el estudio de sesgo de ítem en la Evaluación Docente.

c. Evidencia de validez basada en relaciones entre variables

Existen varios estudios que han aportado evidencia sobre la validez de la Evaluación Docente basada en las relaciones entre variables. El primero de ellos es un estudio de validez convergente que compara prácticas pedagógicas de docentes con alto y bajo desempeño en la Evaluación Docente. Otro grupo de estudios conecta el desempeño docente con rendimiento de los estudiantes medido a través del Sistema de Medición de Calidad de la Educación (SIMCE). El tercero es un estudio que conecta el resultado de los docentes en la Evaluación Docente con el resultado que obtienen en el programa Asignación de Excelencia Pedagógica (AEP), programa al que también se postula por medio de un portafolio basado en el Marco para la Buena Enseñanza.

Estudio de las prácticas pedagógicas de docentes con alto y bajo desempeño

Durante el 2006 se realizó el primer estudio de validez convergente de la Evaluación Docente. El estudio revisó si este programa identifica (y, consecuentemente, recompensa o castiga) correctamente a los profesores con alto —o bajo— desempeño. Se recogieron datos en profundidad del desempeño de una muestra de 58 profesores, evaluados en 2005 ya fuera como Destacado (grupo 1) o Insatisfactorio (grupo 2). Luego se comparó el desempeño de los profesores del grupo 1 (N=32) con el grupo 2 (N=26). La evidencia recogida incluyó:

- a.** Progreso en el logro académico de los estudiantes, medido a través de la aplicación al comienzo y final del año escolar de una prueba estandarizada alineada con el currículum.
- b.** El registro de la observación de 3 clases (de dos horas pedagógicas cada una) realizadas por el profesor durante el año escolar. En esta pauta de observación se evaluaron elementos como uso del tiempo, estructura de la clase, estimulación del pensamiento crítico, presencia de errores conceptuales, comportamiento de los alumnos, adaptabilidad y flexibilidad pedagógica.
- c.** Evaluaciones de expertos de una carpeta con material de enseñanza del profesor para una unidad curricular de dos semanas de duración, incluyendo aspectos de planificación, enseñanza, evaluación y muestras del trabajo de los alumnos. También se pidió a los docentes responder un cuestionario de antecedentes de contexto escolar y de reflexión docente.
- d.** Puntajes de los profesores en una prueba de conocimientos disciplinarios y pedagógicos (la prueba de AEP).

Los docentes evaluados como Destacados e Insatisfactorios presentaron diferencias significativas en la mitad de los indicadores estudiados y observamos diferencias en la dirección esperada, aunque no significativas, en la otra mitad. Encontramos diferencias especialmente relevantes (en término de tamaños de efecto) relacionadas con el tiempo efectivo dedicado a la enseñanza de contenidos durante la clase

("time on task"), la estructura de la clase, el comportamiento de los estudiantes, y el material que utiliza el profesor para la evaluación del aprendizaje de éstos (ver Tabla 2). También encontramos correlaciones significativas entre nuestros resultados y los puntajes de la muestra en tres de los cuatro instrumentos que conforman la Evaluación Docente.

Datos longitudinales sobre el aprendizaje de los alumnos de docentes con resultado Destacado versus Insatisfactorio, analizados usando modelos lineales jerárquicos (HLM), mostraron que el desempeño de los profesores en la Evaluación Docente es un predictor significativo del logro de sus alumnos al final del año escolar, controlando por el logro de éstos al comienzo del año.

En resumen, los resultados del estudio dan apoyo a la validez de la Evaluación Docente en los grupos extremos de las cuatro categorías de desempeño (detalles en Santelices y Taut, 2011; Taut y Santelices, 2007).

• **Tabla 2**
• *Tamaños de efecto t-test entre docentes Insatisfactorios y Destacados*

Instrumento	Indicador / subescala	Cohen d
Prueba de conocimiento pedagógico y disciplinario	Preguntas de alternativas múltiples	0,58
	Preguntas abiertas	0,48
Observaciones de clase	Proporción de tiempo con menos de 75% de alumnos en la tarea	0,98
	Proporción de tiempo con más de 95% de alumnos en la tarea	0,86
	Estructura de la clase	1,11
	Enseñanza especialmente estimulante	0,62
	Conducta de los alumnos adecuada	1,05
	Materiales instruccionales (evaluación holística)	0,46
Carpeta con materiales de enseñanza	Diseño de evaluación de los alumnos	1,21
	Desempeño de los alumnos (indicador continuo)	0,58
	Diseño de evaluación y desempeño de los alumnos (evaluación holística)	0,95
	Reflexión de la propia práctica (evaluación holística)	0,75

Conexión entre Evaluación Docente y rendimiento de los estudiantes

Otro tipo de evidencia en relación a la validez de la Evaluación Docente surge a partir de distintos estudios que han analizado la conexión entre el desempeño de los profesores en la evaluación y el rendimiento de los alumnos en el SIMCE. Estos estudios han vinculado el resultado de los profesores con el de sus alumnos de manera directa (profesor-curso) y de manera colectiva (en función del desempeño de un grupo de docentes de la escuela).

Un estudio conducido por investigadores del Centro de Microdatos y de MIDE UC (Bravo, Falck, González, Manzi & Peirano, 2008) indagó la relación entre el desempeño en la Evaluación Docente y el puntaje SIMCE promedio a nivel de profesor y curso, a partir del análisis de funciones de producción educacional, demostrando que mejores resultados de evaluación del primero se asocian a mejor resultado de los alumnos en SIMCE. Cabe destacar que esta relación persiste aún controlando estadísticamente por antecedentes individuales, socioeconómicos y de rendimiento escolar, tales como: educación de los padres, repitencia del alumno, género, tamaño del curso y puntaje SIMCE previo de la escuela. Este estudio indica que la diferencia estimada en el puntaje SIMCE entre alumnos con docentes de nivel Destacado versus Insatisfactorio es equivalente a 10,25 puntos (0,2 desviaciones estándar), comparable a la diferencia de rendimiento entre alumnos cuyos padres cuentan con Educación Superior versus aquellos que solo cuentan con Enseñanza Media.

Al estudiar la relación entre el desempeño del profesor y el desempeño de sus alumnos en el SIMCE, se observa que aquellos estudiantes expuestos a una mayor cantidad de docentes bien evaluados obtienen un mejor desempeño en la prueba, y esto ocurre en todos los niveles en los que se aplica SIMCE (4° Básico, 8° Básico, 2° medio y 4° medio). También vinculando datos de profesores y sus cursos de manera directa⁷⁾, la diferencia entre tener un 50% o más de profesores bien evaluados versus ninguno es de 15 puntos en SIMCE, tanto en 4° como en 8° Básico del año 2007, y de 12 y 15 puntos respectivamente en 4° Básico y 2° medio en 2008 (Ministerio de Educación, 2008; 2009).

A nivel colectivo, análisis utilizando modelos lineales jerárquicos (HLM) preparados para el Banco Interamericano de Desarrollo muestran que los resultados de Evaluación Docente de los profesores contribuyen en torno a un 1,76% y 2,76% a explicar las diferencias de desempeño de los alumnos de 4° Básico en el SIMCE de Lenguaje y Matemática, respectivamente, una vez incorporadas variables socioeconómicas individuales, del hogar, y de administración de la escuela (Manzi, Strasser, San Martín, & Contreras, 2007). Se observa adicionalmente que la Entrevista del Evaluador Par, la grabación de la clase y la parte escrita del Portafolio están asociadas positiva y significativamente con rendimiento en el SIMCE. Este estudio muestra que la calidad global de los docentes de la escuela, no solo la del docente frente a la clase, es también importante para el rendimiento de los alumnos.

A través de estos estudios se configura un cuerpo emergente de evidencia que apoya la conexión positiva entre logro de los estudiantes y desempeño profesional según la Evaluación Docente, sea a nivel individual o estudiando el efecto colectivo del desempeño docente sobre el rendimiento de los alumnos. En su mayor parte, estos estudios destacan por su intento por controlar diferencias en el rendimiento atribuibles a factores sociodemográficos de los estudiantes (a excepción de los análisis presentados en los informes de resultados nacionales en SIMCE). En la medida que exista disponibilidad de datos de tipo longitudinal en el sistema educacional, debiera ser posible conectar el desempeño de los profesores con el progreso o aprendizaje de los alumnos; esto mejoraría sustantivamente la calidad de la evidencia disponible para estudiar validez de la Evaluación Docente usando el desempeño de sus alumnos como criterio.

Convergencia entre Evaluación Docente y Asignación de Excelencia Pedagógica

El programa de Asignación de Excelencia Pedagógica (AEP) certifica la excelencia pedagógica de profesores a través del desempeño en un portafolio y una prueba de conocimientos disciplinarios y pedagógicos. El programa ha sido implementado desde el 2002 y a él pueden acceder todos los docentes de escuelas que reciben subvención del estado. En el caso de los profesores del sector municipal, ellos deben obtener un desempeño Competente o Destacado en su siguiente Evaluación Docente para mantener la acreditación AEP. Tanto AEP como la Evaluación Docente se basan en el Marco para la Buena Enseñanza e incorporan un portafolio como parte de sus instrumentos de evaluación. La información disponible para profesores que han sido evaluados por ambos programas constituye una herramienta valiosa en el proceso de validación del programa de Evaluación Docente. Esto es así al menos en los inicios de ambos programas, entre 2002 y 2006, antes de que pudiera manifestarse un proceso de auto-selección entre quienes postulan a la AEP, en base al conocimiento de su resultado en la Evaluación Docente.

Entre 2002 y 2006 se cuenta un total de 739 profesores evaluados por ambos programas. Los resultados arrojaron que la gran mayoría (93%) de los profesores certificados por el programa AEP provenían de altos niveles de desempeño de acuerdo a la calificación de la Evaluación Docente: 67% había obtenido un resultado Competente y 26% fue Destacado. Del mismo modo, al comparar los profesores con buenos resultados en la Evaluación Docente con sus pares de desempeño más bajo, observamos que los primeros

7 El reporte nacional de resultados SIMCE incluye desde 2007 una sección especial de análisis que relaciona el desempeño de los estudiantes con el número de docentes bien evaluados que han realizado clases a dichos alumnos (profesores con desempeño Competente, Destacado y Acreditados AEP).

tienden a postular, y tienen más probabilidades de ganar la AEP, que aquellos que obtuvieron las dos categorías más bajas de desempeño de la Evaluación Docente.

Adicionalmente, se observa una correlación positiva moderada entre el rendimiento del Portafolio de la Evaluación Docente y del Portafolio de la AEP ($r=0,33$). La relación entre ambas mediciones podría estar limitada por diferencias temporales en las postulaciones a uno y otro programa (separadas por uno o más años) y por diferencias en la naturaleza de la participación en cada uno de ellos: mientras la Evaluación Docente es obligatoria con estímulos económicos difíciles de alcanzar directamente (por ejemplo, la prueba AVDI no es parte de la Evaluación Docente), la participación en AEP es voluntaria y además del estímulo económico directo existe un reconocimiento social que podría motivar a los docentes a participar.

Se observa una asociación de similar magnitud entre el puntaje final en la Evaluación Docente y el Portafolio de la AEP ($r=0,36$) y relaciones débiles (menores a 0,3) entre los puntajes de los otros tres instrumentos de la Evaluación Docente (Entrevista Evaluador Par, Informe de Referencia de Terceros y Autoevaluación) y el Portafolio de la AEP. Asimismo, se detectó una relación débil entre los puntajes de los instrumentos de la Evaluación Docente y la prueba de conocimientos disciplinarios y pedagógicos de la AEP.

En definitiva, concluimos que hay moderada evidencia convergente respecto de la validez de la Evaluación Docente cuando se considera el desempeño en la AEP como criterio de validación.

d. Evidencia de validez basada en las consecuencias

Investigamos las consecuencias del proceso de Evaluación Docente en el marco de un proyecto FONDECYT⁸ de tres años (2008–2010)⁹. En el estudio se utilizaron metodologías mixtas: en primer lugar, analizamos bases de datos existentes que dan cuenta de los efectos de la Evaluación Docente, por ejemplo, en términos de la participación en el programa de desarrollo profesional “Planes de Superación Profesional” (PSP) y en el programa de incentivos “Asignación Variable de Desempeño Individual” (AVDI). En segundo lugar, investigamos consecuencias esperadas y no esperadas a nivel municipal (local), de escuela y de profesor evaluado, mediante entrevistas personales y discusiones grupales. A continuación presentamos los hallazgos y resultados de los estudios realizados.

La teoría subyacente a la Evaluación Docente

El estudio sobre las consecuencias de la Evaluación Docente se inició elaborando una teoría subyacente sobre cómo este programa debía funcionar y cuáles eran sus efectos esperados. Para esto, entrevistamos a profesionales del Ministerio de Educación relacionadas con el diseño del programa, y a representantes del Colegio de Profesores y de la Asociación de Municipalidades que participaron del proceso de negociación inicial del programa. En base a estas entrevistas desarrollamos diagramas y listas de consecuencias esperadas para cada grupo de interesados (el Ministerio, el Colegio de Profesores y la Asociación de Municipalidades), analizando sus similitudes y diferencias. En este primer análisis identificamos consecuencias mencionadas al menos por dos de los grupos de interesados y usamos éstas para la siguiente fase del estudio: la investigación empírica de las consecuencias esperadas y no esperadas de la Evaluación Docente (para más detalles ver Taut, Santelices, Araya & Manzi, 2010).

Se distinguen consecuencias de la política a nivel municipal y a nivel individual. Las consecuencias con mayor nivel de acuerdo entre los grupos de interés para el nivel municipal son:

.....

8 FONDECYT es un fondo de investigación concursable ofrecido por la Comisión Nacional de Investigación Científica y Tecnológica (CONICYT), Gobierno de Chile.

9 Taut, S., Araya, C., Santelices, V. & Manzi, J. (2008–2010) Estudio de Validez Consecuencial de los Sistemas de Evaluación de Desempeño y Excelencia Docente. Proyecto FONDECYT 1080135.

- a. promover el reconocimiento social de buenas prácticas docentes;
- b. entregar herramientas para subsanar las prácticas de los docentes con desempeño débil;
- c. proveer de información relevante para la contratación y despido de profesores;
- d. promover la colaboración entre pares para el fomento de buenas prácticas docentes;
- e. mejorar las perspectivas laborales al entregar incentivos económicos asociados al buen desempeño individual;
- f. entregar un diagnóstico de la calidad docente para la toma de decisiones.

A nivel individual los grupos de interés coinciden en identificar dos consecuencias esperadas: (1) generar un cambio en las prácticas débiles y (2) promover la mantención de buenas prácticas al generar un reforzamiento interno de las fortalezas individuales. A más largo plazo los entrevistados coinciden en que la política de evaluación debería resultar en una mejora en el aprendizaje de los estudiantes de los profesores evaluados.

Consecuencias de la Evaluación Docente a nivel municipal

Este estudio se centra en las consecuencias esperadas y no esperadas de la Evaluación Docente a nivel municipal, de acuerdo a la teoría subyacente elaborada por los involucrados en el diseño del programa. En esa lógica, la evaluación empírica de resultados informa sobre la existencia o no existencia de los efectos y usos de la Evaluación Docente respecto de los procesos de contratación y despido de profesores, los planes de desarrollo profesional y las estrategias de reconocimiento de los docentes.

Para el estudio se entrevistó a 19 autoridades de educación local (municipal), provenientes de diez municipalidades. Los resultados indican que el programa ha alcanzado algunas de sus metas, mayoritariamente respecto de los mecanismos de entrega de recompensa a docentes con alto rendimiento (localmente y a través del programa AVDI), el apoyo a profesores con dificultades mediante el programa de desarrollo profesional PSP y la inclusión de resultados de la evaluación en el proceso de toma de decisiones en el ámbito educacional municipal.

También se encontraron consecuencias no esperadas de relevancia. La consecuencia negativa no esperada más extendida se relaciona con las emociones negativas que el proceso de evaluación produce en los profesores (p. ej., miedo, ansiedad, sentimiento de inseguridad laboral) y la rebelión y resistencia hacia la evaluación, las que fueron enfrentadas por los actores municipales implicados en el programa especialmente durante los primeros años, y han ido disminuyendo paulatinamente. La consecuencia positiva no esperada más importante se refiere a los mecanismos de apoyo que las municipalidades han desarrollado para ayudar a los docentes a preparar y sobrellevar el proceso de evaluación (ver Santelices, Taut, Araya & Manzi, 2009).

Consecuencias de la Evaluación Docente a nivel de escuela

En entrevistas con directivos (N=57) de 30 escuelas en 10 municipalidades se investigó las consecuencias que la Evaluación Docente ha tenido en las escuelas en términos de efectos y usos. A nivel institucional, por ejemplo, se incrementó el trabajo en equipo, los procesos de reflexión interna sobre la base de los resultados de la evaluación, y la validación externa de la calidad de la escuela, producto de los resultados obtenidos en el programa.

Desde la perspectiva de los directivos, a nivel de profesor hay efectos mixtos. En todas las escuelas del estudio los entrevistados reportaron emociones negativas y sobrecarga de trabajo producto del proceso de evaluación. Sin embargo, los profesores que obtuvieron resultados satisfactorios presentaron reacciones positivas. También, la resistencia inicial de los profesores ha ido disminuyendo en el tiempo. Aún más

importante, el proceso de evaluación ha propiciado la revisión de las prácticas docentes y la internalización de los estándares de evaluación. Explicaciones tentativas para las diferencias observadas entre escuelas, en cuanto a los efectos y usos de la evaluación, se relacionan con el rol de líder pedagógico que asume el director y con su opinión sobre la legitimidad de la evaluación (más detalles en Taut, Santelices, Araya & Manzi, 2011).

Percepciones de profesores que se niegan a participar en la Evaluación Docente

El propósito de este estudio fue comprender por qué algunos profesores de escuelas municipales se rehúsan abiertamente a participar en la Evaluación Docente y describir sus percepciones acerca de ésta. Entendimos la “rebeldía” a partir de los significados construidos por los profesores y elaboramos un modelo explicativo de la conducta rebelde, entendida ésta como negarse a la evaluación cuando al docente le corresponde realizarla.

La metodología incluyó nueve entrevistas en profundidad que se analizaron de acuerdo a los procedimientos propuestos por la Grounded Theory (Strauss & Corbin, 1990). Nuestros hallazgos indican que distintos factores inciden en la conducta de rechazo activo a la evaluación por este grupo de profesores, entre ellos las emociones negativas respecto del proceso (por ejemplo, el temor a obtener un bajo resultado), aspectos de la cultura docente, además de percepciones negativas como la falta de legitimidad de la Evaluación en términos de sus criterios, instrumentos y obligatoriedad, y la falta de información sobre el sistema de evaluación (ver más detalles en Tornero, 2009).

El programa de incentivos AVDI

El análisis de los datos de profesores que participaron en este programa de incentivos asociado a la Evaluación Docente entre 2004 y 2006 indica que, del total de profesores calificados como Destacados y Competentes habilitados para participar de la Asignación Variable por Desempeño Individual (AVDI), solo un 46% rindió la prueba. Para obtener este incentivo, los docentes deben aprobar una prueba de conocimientos específicos del subsector (asignatura). El resultado de la prueba puede ser Destacado, Competente o Suficiente, y combinado con el resultado de la Evaluación Docente (Competente o Destacado) determina el monto a percibir. El beneficio más alto recae en el profesor que obtiene un resultado Destacado en ambas evaluaciones, quien recibe un bono equivalente al 25% de la Renta Básica Mínima Nacional (RBMN). En el caso del profesor con resultado Competente en la Evaluación Docente, obtendrá un bono equivalente al 15% de la RBMN si su resultado en la prueba fue Competente o Destacado, y desde 2006 tiene derecho a recibir un bono del 5% si fue solo Suficiente⁽¹⁰⁾.

En todos los casos el bono se paga durante un período de hasta cuatro años, hasta que el profesor deba enfrentar un nuevo proceso de evaluación. Considerando el período estudiado, alrededor de la mitad (54%) de quienes rindieron la prueba ganaron algún tipo de bonificación. Es importante mencionar que de estos profesores, en promedio, solo un 23% accedió al bono del 15% de la remuneración y solo un número muy reducido (N=32) accedió al bono de 25% (ver Tabla 3).

Concluimos que —a pesar de los cambios introducidos en 2006— el incentivo monetario ofrecido a través de AVDI está siendo subutilizado por los docentes. Datos cualitativos arrojados por grupos focales indican que los profesores perciben que (1) el monto del bono es muy bajo⁽¹¹⁾, (2) la prueba constituye una barrera difícil, la cual además es deslegitimada por su mera existencia, y (3) hay falta de información respecto del programa.

10 Además, se entregan \$50.000 solo por el hecho de rendir la prueba.

11 Aunque la fórmula de cálculo es compleja y depende de la configuración de pago de cada docente, equivale aproximadamente a un sueldo más al año. Para más información sobre AVDI, ver <http://www.avdi.mineduc.cl/>

• **Tabla 3**
 • *Participación y resultados AVDI 2004–2006*

Participación AVDI	2004	2005	2006	Total
Número de Habilitados	2425	3182	6329	11936
No participa	1234	2084	3089	6407
Participa	1191	1098	3240	5529
Tasa de participación	49%	35%	51%	46%
Resultados AVDI				
No gana AVDI	871	828	865	2564
Gana AVDI	320	270	2375	2965
Tasa Ganadores AVDI (sobre total prueba)	27%	25%	73%	54%
Tasa de Competentes y Destacados en prueba AVDI (sobre total prueba)	27%	25%	23%	23%
Gana 25% de la RBMN	9	9	14	32

El programa de desarrollo profesional PSP

Los Planes de Superación Profesional (PSP) son obligatorios para los profesores que obtienen un nivel de desempeño Insatisfactorio o Básico. Cada municipalidad recibe un monto fijo de dinero por cada profesor de bajo desempeño para organizar planes de desarrollo profesional para dichos docentes. Se analizó la información proveniente de la base de datos en línea del programa PSP desde el año 2007 al 2009 para conocer las principales características de estas instancias de capacitación y la participación de los profesores en ellos. La información analizada incluye: tasas de participación de profesores, contenido y tipos de perfeccionamiento, y la satisfacción del docente respecto del programa al que asistió.

Los resultados mostraron que solo el 41% del total de docentes que debía participar efectivamente lo hizo. En 2008, las acciones de desarrollo profesional de los PSP tuvieron con la más alta frecuencia, la forma de talleres o seminarios (76%), seguidos de trabajo individual guiado por un supervisor o mentor (17%), en tanto que solo un 3% consistió en observaciones de aula y retroalimentación. En cuanto a los temas abordados, la mayoría de los PSP desarrolló acciones referidas a “evaluación de aprendizaje de los estudiantes” y la “planificación de clases en aula”. Por otro lado, datos de encuestas de opinión aplicadas a los docentes que participan en los PSP, muestran altos niveles de satisfacción con el programa, aunque solo un porcentaje menor responde estos cuestionarios (menos de un tercio) y un 15% señala que no participaría nuevamente.

En términos de impacto, evidencia preliminar proveniente del cruce de datos de Evaluación Docente 2009 y PSP 2009 indica que los docentes Insatisfactorios que participaron de las actividades de PSP en 2009 alcanzaron en promedio mejores resultados en su re-evaluación (obligatoria y durante el mismo año), que sus pares que no lo hicieron (ver más detalles en Cortés, Taut, Santelices & Lagos, 2011).

Consecuencias en la trayectoria laboral de los docentes Insatisfactorios

Realizamos dos análisis de la trayectoria laboral de los profesores con desempeño Insatisfactorio usando para ello la base de datos de la Bonificación de Reconocimiento Profesional (BRP) del Ministerio de Educación y los resultados de la Evaluación Docente. En el primer estudio se analizó la trayectoria de docentes evaluados entre 2003 y 2007 con datos de BRP 2008, incorporándose en el segundo los datos de la evaluación 2008 y de BRP 2009.

Ambos estudios muestran una tasa de salida del sistema municipal mucho más alta entre los docentes con desempeño Insatisfactorio en la evaluación, lo que persiste analizando por grupo etario. En 2009, dicha tasa correspondió a un 31,9% para los docentes con nivel Insatisfactorio, comparada con un 11,2% de sus pares con resultado Competente o Destacado (ver Tabla 4). La misma tendencia se observa utilizando datos de 2008 y 2007. Esto implica que los docentes con el más bajo desempeño en la Evaluación Docente salen del sistema municipal con anterioridad a resultar evaluados con Insatisfactorio tres veces consecutivas¹².

Los resultados de este estudio también muestran que los docentes con mejor desempeño logran acceder a cargos directivos en mayor medida (ver Taut, Santelices & Valencia, 2010). Los factores asociados a las decisiones laborales de los profesores bien evaluados versus mal evaluados podrían ser estudiados en profundidad en el futuro.

• **Tabla 4**
 • *Estado de docentes evaluados entre 2003 y 2008 en el sistema educacional en 2009*

	Alguna vez Insatisfactorio	Alguna vez Competente	Alguna vez Destacado	Alguna vez Competente o Destacado
Activo Municipal	68,1%	88,3%	92,1%	88,8%
Fuera del sistema municipal	31,9%	11,7%	7,9%	11,2%
Total	100,0%	100,0%	100,0%	100,0%

Conclusiones: Un juicio preliminar acerca de la validez

Este capítulo ha presentado un resumen de la evidencia disponible para formular un juicio acerca de la validez de la Evaluación Docente. La mayoría de dicha evidencia proviene de nuestra agenda de investigación, a la que se le han sumado algunos estudios relacionados. También hemos presentado el marco general desde el cual conceptualizamos la validez, en consistencia con la conceptualización actual de este constructo aceptada en el ámbito de la investigación educacional.

Cuando se analiza la evidencia de validez recolectada a la fecha, encontramos en primer lugar que el proceso de Evaluación Docente da cuenta de diferencias reales entre profesores con más alto y más bajo desempeño. El estudio de prácticas pedagógicas encontró diferencias sustantivas (medido por tamaños de efecto) entre el desempeño de docentes con resultado Insatisfactorio versus Destacado, respecto de importantes aspectos de su labor pedagógica en aula. Adicionalmente, el estudio mostró que tres de los cuatro instrumentos utilizados en la Evaluación Docente (especialmente el Portafolio) se relacionan moderadamente con otros instrumentos utilizados en la investigación para recoger información sobre la calidad del desempeño docente. Queda pendiente estudiar la validez de categorías adyacentes de la Evaluación Docente, por ejemplo, de Insatisfactorio versus Básico, o de Básico versus Competente.

Por otra parte, nuestros resultados comparando docentes evaluados por dos programas similares (Evaluación Docente y AEP), ambos basados en el Marco para la Buena Enseñanza y utilizando portafolios, entregan evidencia adicional positiva respecto de la validez convergente.

.....
 12 Hasta 2010 la ley establecía que tras cumplir tres evaluaciones consecutivas con resultado Insatisfactorio, el profesor debe abandonar la dotación docente, número que se reducirá a dos a partir de la evaluación 2011, tras promulgarse un cambio en la normativa legal (Ley 20.501 de febrero de 2011).

Otro criterio importante de la validez de la Evaluación Docente es el rendimiento que alcanzan los alumnos de los profesores evaluados. En el estudio de prácticas pedagógicas, que utilizó datos longitudinales de los estudiantes, se encontró que los alumnos con docentes de alto desempeño tienden a lograr mejores resultados que sus pares con profesores de bajo desempeño. También otros estudios que utilizan datos del SIMCE han mostrado correlaciones positivas entre desempeño docente y logro de los alumnos. Es deseable complementar estos datos con más estudios basados en mediciones longitudinales que permitan relacionar de manera más directa ambas variables.

Otro aspecto de la validez es la cobertura del Marco para la Buena Enseñanza en los instrumentos de la Evaluación Docente. Nuestros análisis muestran una apropiada cobertura de los contenidos. Queda pendiente una validación de ese aspecto por expertos educacionales externos.

La evidencia de la consistencia interna de los instrumentos de la Evaluación Docente es satisfactoria, aunque los análisis factoriales sugieren que sería conveniente revisar la estructura interna del Portafolio y, de manera correspondiente, aquella de las rúbricas de corrección y los reportes de resultados. Por ejemplo, se podría simplificar la estructura, reduciendo el número de dimensiones evaluadas (y, por lo tanto, reportadas).

Estudiamos también el efecto no deseado del corrector sobre los puntajes del Portafolio en la Evaluación Docente. Encontramos que ese efecto es generalmente bajo, aunque heterogéneo entre distintos subsectores y dimensiones del Portafolio. Al mismo tiempo, los índices de confiabilidad son comparables a los que se reportan para sistemas de evaluación docente similares (ver Hakel, Koenig & Elliot, 2008). Una recomendación que emana de los estudios realizados en este ámbito es aumentar la doble-corrección para la totalidad de los portafolios, y no solo para un 20% de ellos como ocurre en la actualidad; los datos indican que de esta forma se podrían mejorar sustantivamente los índices de generalizabilidad.

Finalmente, investigamos las consecuencias esperadas y no esperadas de la Evaluación Docente para distintos actores involucrados. Constatamos que, a nivel municipal, se utilizan los resultados de la Evaluación Docente para la toma de decisiones, por ejemplo, la elaboración de perfeccionamiento docente (PSP) y el reconocimiento local de buenos docentes. Este efecto fue intencionado por quienes participaron en el diseño de la evaluación. Un aspecto positivo no esperado, en tanto, es el apoyo pedagógico y psicológico de preparación y contención para profesores con bajo desempeño. A nivel de los establecimientos escolares, también observamos que el proceso de evaluación produce efectos mayoritariamente positivos, por ejemplo, promoviendo la colaboración entre pares y la reflexión sobre los resultados. A nivel individual, en cambio, los efectos son mixtos. Existe una percepción general de sobrecarga laboral producto de la evaluación e intensas reacciones emocionales negativas por parte de los docentes evaluados. No obstante, los profesores generalmente reconocen que existen también importantes beneficios del proceso de evaluación, sobre todo de la elaboración del Portafolio, en términos de la revisión y actualización de sus prácticas.

El trabajo realizado hasta hoy estudiando la validez de la Evaluación Docente evoca las palabras de Koretz (2008): "...la validación es, a menudo, presentada como una dicotomía: la conclusión en base al desempeño en un instrumento es válida o inválida. Desafortunadamente, la situación (en la vida real) es menos clara. La validación es un continuo (...) Algunas conclusiones están mejor respaldadas que otras, pero la evidencia con la que generalmente se cuenta es limitada, por lo que es necesario 'asegurar' nuestras apuestas" (p. 219, traducción de los autores). De allí que, el modo en que se juzguen los resultados aquí presentados acerca de la validez de la Evaluación Docente, es una cuestión de perspectivas y valores (Herman & Baker, 2009). Mientras algunos podrían opinar negativamente del programa por la sola mención de uno o dos resultados negativos, otros podrían considerar esto como algo "esperable" y al mismo tiempo subyugarlos a los resultados positivos. Y una tercera posición insistirá en la necesidad de completar la información pendiente con nuevos estudios antes de emitir un juicio conclusivo acerca de la validez del sistema de evaluación.

A pesar de todas estas precauciones, concluimos que hay evidencia inicial que lleva a un juicio general positivo respecto de la validez de la Evaluación Docente. La evidencia más determinante para emitir este juicio es la relación entre los resultados de la evaluación y otras variables de criterio, además de las consecuencias positivas (esperadas y no esperadas) observadas en distintos actores involucrados en el proceso. Sin embargo, hay aspectos de la evaluación que merecen revisión, lo que no resulta sorprendente ni inesperado, dada la complejidad que encierra medir el desempeño docente (Berliner, 2005; Ingvarson & Rowe, 2008). Entre los aspectos a mejorar creemos que está, como primera prioridad, la revisión de la estructura interna del Portafolio y la doble-corrección del 100% de éstos.

Para concluir, recogiendo la idea de Cronbach de que la validación es un proceso “largo, incluso interminable” (1989, p. 151), estamos conscientes de que siempre habrá más y mejor evidencia que recoger y analizar, y nuevos estudios que realizar.



7

Desempeño docente: relaciones con antecedentes de los profesores y su contexto

Edgar Valencia • Jorge Manzi

Centro de Medición MIDE UC
Pontificia Universidad Católica de Chile

Desempeño docente: relaciones con antecedentes de los profesores y su contexto

Este capítulo tiene por propósito mostrar la evidencia disponible que relaciona el desempeño de los profesores en la Evaluación Docente con antecedentes individuales y del contexto en el que ejercen la enseñanza. Entre los primeros se consideran elementos de su formación (inicial y en servicio), antecedentes laborales y creencias sobre la práctica docente. Entre los antecedentes de contexto se incluyen características del establecimiento educacional en que ejerce y de su entorno local como tamaño, recursos y condiciones socioeconómicas del municipio.

Se reportan resultados utilizando dos enfoques complementarios: datos provenientes de análisis descriptivos y resultados de investigaciones que han utilizado modelos lineales jerárquicos. Con ambos enfoques, se busca entregar una visión comprensible y precisa de las relaciones existentes entre los antecedentes previamente mencionados y el desempeño docente. El capítulo busca propiciar discusiones entre los propios docentes, autoridades y centros de formación acerca de cómo fortalecer la calidad de la enseñanza, sea a través de dispositivos de apoyo para los actuales o futuros docentes, las condiciones de ejercicio de la profesión u otros.

Introducción

Que cada estudiante pueda acceder a profesores y profesoras de calidad es un aspecto clave para el mejoramiento educativo. Sin embargo, esta aspiración está subordinada al conocimiento de los elementos que explican cómo un docente va a desempeñarse frente a sus alumnos. En este capítulo se va a revisar qué antecedentes se relacionan con los resultados de los profesores cuando participan en la Evaluación Docente. La pregunta detrás de los análisis que se presentan se relaciona con una cuestión importante: la identificación de factores que permitirían predecir un mejor desempeño. Esto contribuye con elementos para un mejor diagnóstico de la calidad docente, y además, es clave para ofrecer condiciones que estimulen el desarrollo de buenos docentes, tanto en su formación como a lo largo de la carrera profesional.

Antes de comenzar, es necesario establecer consideraciones sobre dos conceptos claves en este capítulo: Primero, qué se va a entender por calidad docente y por qué la Evaluación Docente daría cuenta de ella, y segundo, qué aspectos específicos podrían ser considerados para explicar la calidad docente, y consecuentemente el desempeño de los profesores en la evaluación.

En cuanto a lo primero, la calidad docente puede ser entendida al menos de dos maneras. Siguiendo la distinción realizada por Fenstermacher & Richardson (2005) y retomada por Ingvarson & Rowe (2008), es útil distinguir entre enseñanza efectiva (*successful teaching*) y buena enseñanza (*good teaching*). Metodologías como la de Valor Agregado, que intentan estimar el efecto del profesor sobre el progreso de los estudiantes, representan uno de los desarrollos más importantes para evaluar la enseñanza efectiva (ver por ejemplo, Raudenbush, 2004; Sanders & Rivers, 1996; Wright, Horn & Sanders, 1997). Sin embargo, la pertinencia de esta metodología para evaluar docentes está siendo objeto de un amplio debate en el

contexto norteamericano (Baker et al., 2010; Papay, 2010). En contraste, la Evaluación Docente chilena se basa en estándares de desempeño y por tanto, apunta a evaluar la buena enseñanza (ver Flotts & Abarzúa, en este volumen).

Sobre el segundo punto, es importante mencionar que los antecedentes de la calidad tratados en este capítulo están relacionados con una gran cantidad de investigación educativa, principalmente referida a efectividad docente. En estos estudios, el supuesto básico es que existen ciertas cualidades individuales del profesor que se asocian a incrementos en el desempeño de sus alumnos. Estas investigaciones conducidas comúnmente en EE.UU. han considerado antecedentes como la formación, experiencia y ciertos atributos individuales de los docentes (Darling–Hammond, 2000; Darling–Hammond, Berry & Thoreson, 2001; Darling–Hammond, Wei & Johnson, 2009; Wayne & P. Youngs, 2003; Rice, 2003; Wilson, Floden & Ferrini–Mundy, 2001). Otros estudios han utilizado medidas directas de prácticas docentes como la evaluación del National Board for Professional Teaching Standards (NBPTS) (Goldhaber, Perry & Anthony, 2004). Investigaciones como la referida al NBPTS y otras consideran también antecedentes del establecimiento educativo y del distrito como importantes para comprender la calidad docente. En consecuencia, se ha considerado también otro grupo de atributos, aquellos relacionados con el contexto de la enseñanza (Ding & Sherman, 2006; Heneman & Milanowski, 2004; Odden, 2004; Robalino & Körner, 2005).

Bajo el marco anterior, este capítulo presenta análisis y resultados de investigaciones que dan luz acerca de la relación entre antecedentes individuales de los profesores y del contexto de la enseñanza, con el desempeño obtenido en la Evaluación Docente. En el primer grupo de antecedentes se encuentran algunas características personales del profesor, como el sexo y la edad, antecedentes de formación inicial y en servicio, antecedentes laborales, condiciones de trabajo y expectativas. Entre los antecedentes del contexto se encuentran características de la escuela y el municipio como el tipo de entorno (rural–urbano), el tamaño y características socioeconómicas.

Explorar las relaciones entre estos antecedentes y el desempeño puede resultar en algo extremadamente complejo. Por ejemplo, una vez analizada la relación entre el tipo de entorno del establecimiento (rural versus urbano) y el desempeño docente, cabría preguntarse si dicha relación se modifica o mantiene al considerar el nivel socioeconómico, el tamaño o la disponibilidad de recursos de la escuela.

Para abordar este desafío, en este capítulo se utilizan dos perspectivas metodológicas complementarias. En primer lugar, se presentan los resultados de un conjunto de análisis descriptivos. Esta primera perspectiva es útil a modo introductorio debido a la simplicidad y la facilidad con la que se observa cómo el desempeño docente varía según los distintos antecedentes individuales y de contexto considerados. En la segunda parte del capítulo se presentan los resultados de una serie de investigaciones que han abordado las relaciones entre antecedentes y desempeño usando una familia de modelamiento estadístico llamado Modelos Lineales Jerárquicos. Esta perspectiva introduce varias correcciones a los análisis descriptivos, las que serán detalladas más adelante en este capítulo.

Análisis descriptivo de los correlatos del desempeño docente

Los análisis que se presentan se basan en los resultados de desempeño de los docentes participantes en la Evaluación Docente durante los años 2005, 2006 y 2007. En dicho período se recogió información para más de 33.000 profesores de Primer Ciclo, Segundo Ciclo y Enseñanza Media, incluyendo aspectos claves de su formación y trabajo, por medio de un cuestionario complementario aplicado en el proceso de evaluación. También se ha utilizado información acerca de las condiciones del establecimiento y la comuna (municipio) proveniente de las bases de datos del Sistema de Medición de Calidad de la Educación (SIMCE), el Sistema Nacional de Información Municipal (SINIM) y el programa Asignación de Excelencia Pedagógica (AEP).

Como se ha señalado anteriormente, dos son los niveles de análisis implicados en la comprensión del desempeño docente. En primer lugar, se presentan correlatos con antecedentes de orden individual y posteriormente se presentan las relaciones entre desempeño y antecedentes de contexto.

Características de los docentes evaluados

Dentro de las características generales del grupo de docentes analizado puede señalarse que dos tercios de ellos son mujeres y la edad típica de los profesores se encuentra en el rango de 46 a 60 años. La gran mayoría de los evaluados se formó en carreras diurnas, aunque con una duración heterogénea (entre 5 y 10 semestres). Además, un 42% señala haber cursado postítulos o diplomados; un 4,1%, programas de magíster; y un 0,1%, estudios de doctorado. Asimismo, un 69% ha participado en los Programas de Perfeccionamiento Fundamental (ver Tabla 1).

Un 61% de los docentes indica contar con más de 10 años de experiencia en el subsector en que fue evaluado (ver Tabla 2). También los datos muestran que, en promedio, los profesores tienen 34 horas lectivas semanales (DE=7,69) y 28 alumnos por sala (DE=11,19).

Otro aspecto estudiado fue la obtención de la Asignación de Excelencia Pedagógica (AEP), encontrándose que solo un 2,3% del grupo total analizado había logrado esta acreditación en forma previa a su participación en la Evaluación Docente.

Relación entre desempeño docente y antecedentes individuales

Características personales y antecedentes de formación

En este grupo de antecedentes se han considerado atributos como el sexo y la edad junto con características básicas de la formación inicial.

De la Tabla 1 se observa en primer lugar que las profesoras no solo son mayoría en la dotación docente, sino que una mayor proporción de ellas logran un desempeño Competente o Destacado: 64% vs. 57% de los profesores.

En términos de rango etario, entre los 41 y 45 años se encuentra la mayor proporción de docentes Competentes o Destacados, observándose una leve tendencia creciente en el desempeño hasta dicho rango, y luego una caída sostenida a medida que la edad progresa. Esto es consistente con la literatura que reporta que el desempeño en función de la edad presenta en general una forma de U invertida (Sturman, 2003).

Con respecto al tipo de formación inicial, se observa una mayor proporción de docentes con desempeño Competente o Destacado entre quienes asistieron a carreras diurnas (65%), por sobre otro tipo de planes, por ejemplo, vespertino (55%). También se constata que los docentes formados en carreras más extensas obtienen mejores resultados en la Evaluación Docente.

En cuanto a la formación en servicio, quienes reportan haber asistido a este tipo de actividades obtienen mejores resultados: un 67% de los que cursaron post-títulos o post-gradados alcanza un nivel Competente o Destacado, versus un 57% de quienes no lo hicieron. Lo mismo ocurre con la participación en los Programas de Perfeccionamiento Fundamental: 64% alcanza el desempeño Competente o Destacado versus 56% de quienes no participaron.

● **Tabla 1**
 ● *Desempeño docente según características personales y de formación⁽¹⁾*

	Proporción de Docentes	Resultado en Evaluación Docente	
		Insatisfactorio o Básico	Competente o Destacado
Sexo			
Mujer	71%	36%	64%
Hombre	29%	43%	57%
Edad			
25 o menos años	1,0%	34,1%	65,9%
26 a 30 años	4,4%	36,8%	63,2%
31 a 35 años	5,8%	33,0%	67,0%
36 a 40 años	7,1%	33,4%	66,6%
41 a 45 años	13,0%	31,4%	68,6%
46 a 50 años	19,0%	34,7%	65,3%
51 a 55 años	27,0%	40,0%	60,0%
56 a 60 años	18,1%	42,9%	57,1%
61 a 65 años	4,0%	53,0%	47,0%
66 o más años	0,5%	62,7%	37,3%
Modalidad de la carrera			
Diurna	69,5%	34,9%	65,1%
Vespertina	23,3%	45,1%	54,9%
Otra	7,2%	50,8%	49,2%
Duración de la carrera			
4 semestres o menos	7,4%	44,0%	56,0%
5 a 6 semestres	31,7%	39,7%	60,3%
7 a 8 semestres	32,9%	38,4%	61,6%
9 a 10 semestres	28,0%	32,9%	67,1%
Participación en actividades de formación continua			
Sí	43,8%	33,1%	66,9%
No	56,2%	42,6%	57,4%
Participación en Planes de Perfeccionamiento Fundamental (PPF)			
Sí	69,2%	36,1%	63,9%
No	30,8%	43,7%	56,3%

En resumen, es más probable que quienes son mujeres y están en la medianía de su carrera alcancen un mejor desempeño en la Evaluación Docente. A su vez y de acuerdo a lo previsible, entre quienes poseen una formación inicial de mayor duración y participan de actividades de formación continua se observan más altos resultados en la evaluación.

1 En la primera columna se indica el porcentaje de docentes con cada atributo o característica analizada. La segunda columna muestra la proporción de esos docentes que no cumple con el estándar exigido (es decir, con resultado Insatisfactorio o Básico), y la tercera, aquel que sí lo cumple (Competente o Destacado).

Antecedentes laborales y condiciones de trabajo

En este grupo de antecedentes se ha considerado los años de experiencia en el subsector, la condición de docente de excelencia (acreditación en el programa AEP), la cantidad de horas lectivas y el número de alumnos en sala. En atención a estos importantes antecedentes es posible contrastar algunas suposiciones. Una de ellas es que los docentes con mayor cantidad de años de experiencia y aquellos acreditados en el programa AEP debieran mostrar un mejor desempeño en la Evaluación Docente. A su vez, quienes poseen una mayor carga de horas lectivas y trabajan en cursos más grandes debieran obtener resultados más bajos, en la medida que esas condiciones podrían ser menos favorables para el buen ejercicio de la docencia.

Respecto a la experiencia en el subsector, se observa que, efectivamente, en la medida que ésta aumenta, mayor es la proporción de docentes con desempeño Competente y Destacado. El grupo con mayor experiencia tiene una proporción de docentes Competente y Destacado de 63%, mientras que en el grupo con menos experiencia dicha proporción baja a 55% (ver Tabla 2).

En cuanto a la condición de excelencia (acreditación en AEP previa a la Evaluación Docente), se observa una mayor proporción de docentes con resultado Competente y Destacado entre quienes poseen la acreditación (90%) versus quienes no la poseen (61%), siendo este el atributo que muestra más diferencia entre profesores de distinto desempeño. Esta información indica, además, que los docentes de excelencia muestran mayoritariamente un desempeño consistente con su condición de tales en la Evaluación Docente, lo que sugiere un aporte a favor de la validez de ambos sistemas de evaluación.

● **Tabla 2**
● *Desempeño docente por antecedentes laborales y condiciones de trabajo*

	Proporción de Docentes	Resultado en Evaluación Docente	
		Insatisfactorio o Básico	Competente o Destacado
Años de experiencia en el subsector			
1 a 2 años	11,7%	44,6%	55,4%
3 a 5 años	12,0%	41,5%	58,5%
6 a 10 años	15,9%	36,7%	63,3%
Más de 10 años	60,5%	37,1%	62,9%
Cantidad de horas lectivas			
Hasta 24 horas	5,7%	37,8%	62,2%
Entre 25 y 30 horas	27,9%	37,8%	62,2%
Entre 31 y 36 horas	21,0%	34,6%	65,4%
Más de 36 horas	45,4%	37,8%	62,2%
Número de alumnos por curso			
Hasta 10 alumnos	8,6%	46,8%	53,2%
Entre 11 y 20 alumnos	15,9%	39,6%	60,4%
Entre 21 y 30 alumnos	26,5%	37,5%	62,5%
Entre 31 y 40 alumnos	34,4%	35,5%	64,5%
Más de 41 alumnos	14,6%	36,3%	63,7%
Acreditados AEP			
Sí	2,3%	10,1%	89,9%
No	97,7%	39,1%	60,9%

Respecto a las horas lectivas, los datos sugieren la ausencia de una relación entre este atributo y el desempeño en la Evaluación Docente. Y, en el caso del tamaño del curso, la relación exhibe una dirección opuesta a lo esperado: profesores con cursos más numerosos muestran mejor desempeño que aquellos con cursos más pequeños (ver Tabla 2).

Expectativas del profesor

A pesar de no ser un antecedente de la naturaleza de los anteriores, se ha incluido adicionalmente el análisis de cómo varía el desempeño docente según las expectativas de los profesores sobre la escolaridad de sus alumnos². El efecto de las expectativas de los profesores ha sido estudiado largamente en educación y ocupa un lugar prominente a la hora de explicar las prácticas pedagógicas y el desempeño de los alumnos. Expectativas más altas se asocian generalmente a mejores prácticas y mayor efectividad del profesor o profesora (OECD, 2009; Bandura, 1997; Coll, 2001).

En el grupo analizado, la gran mayoría de los docentes cree que sus alumnos terminarán la Educación Media o lograrán estudios superiores técnico o universitario, y los pocos que declaran expectativas más bajas (inferior a Enseñanza Media Completa) muestran un desempeño inferior al resto de los evaluados. Así, por ejemplo, la proporción de docentes Competentes o Destacados entre quienes sostienen expectativas de logro de Enseñanza Media es 63%, versus un 53% cuando la expectativa es de Enseñanza Media incompleta y 46%, para el caso de Enseñanza Básica completa (ver Tabla 3).

• **Tabla 3**
• *Desempeño docente según expectativas de escolaridad para sus alumnos*

	Proporción de Docentes	Resultado en Evaluación Docente	
		Insatisfactorio o Básico	Competente o Destacado
Expectativas escolaridad alumnos			
Enseñanza Básica Incompleta	0,3%	51,4%	48,6%
Enseñanza Básica Completa	1,9%	54,4%	45,6%
Enseñanza Media Incompleta	7,7%	46,5%	53,5%
Enseñanza Media Completa	45,4%	36,6%	63,4%
Nivel Superior o Técnico	36,8%	32,2%	67,8%
Educación Universitaria	7,8%	31,3%	68,7%

Desempeño docente y antecedentes de contexto

En este apartado se presenta cómo los antecedentes de contexto se vinculan con el desempeño de los profesores en la Evaluación Docente. En primer lugar se muestra el cruce entre éste y condiciones objetivas como tamaño y recursos. Hacia el final de la sección se analiza cómo varía el desempeño docente según las expectativas de escolaridad de los alumnos que sostienen los padres del establecimiento³.

.....
2 Este aspecto fue consultado en el Cuestionario Complementario respondido por los docentes que participaron en la Evaluación Docente 2006 y 2007.
3 Datos provenientes de cuestionarios aplicados por el SIMCE.

A modo de contextualización, el 29% de los docentes participantes de la Evaluación Docente realiza clases en escuelas ubicadas en zonas rurales, mientras que el 71% restante se desempeña en zonas urbanas (ver Tabla 4). La distribución de profesores evaluados según grupo socioeconómico (GSE del SIMCE⁽⁴⁾) del establecimiento muestra que un 22% de los docentes ejerce en establecimientos con GSE bajo, un 49% en establecimientos con GSE medio bajo, un 25% en establecimientos con GSE medio, y solo un 3% se desempeña en establecimientos con GSE medio alto. No hay docentes evaluados en establecimientos con clasificación GSE alto, pues no existen establecimientos municipales con dicha clasificación.

• **Tabla 4**
• *Desempeño docente según antecedentes de contexto*

	Distribución	Insatisfactorio o Básico	Competente o Destacado
Ubicación geográfica			
Urbano	71,0%	37,0%	63,0%
Rural	29,3%	41,9%	58,1%
Nivel Socioeconómico			
GSE A Bajo	22,3%	38,2%	61,8%
GSE B Medio Bajo	49,6%	40,1%	59,9%
GSE C Medio	25,0%	35,3%	64,7%
GSE D Medio Alto	3,1%	32,6%	67,4%
Nivel Recursos Tecnológicos			
Bajo	14,4%	45,7%	54,3%
Medio	48,2%	39,8%	60,2%
Alto	37,3%	33,9%	66,1%
Tamaño del Establecimiento			
Hasta 250 alumnos	26,6%	42,2%	57,8%
Hasta 500 alumnos	17,4%	37,5%	62,5%
Hasta 750 alumnos	20,3%	38,0%	62,0%
Más de 750 alumnos	35,8%	36,3%	63,7%
Expectativa media de los padres sobre escolaridad de los hijos			
Menos de 10 años	0,6%	55,7%	44,3%
Menos de 12 años	7,2%	41,6%	58,4%
Menos de 13 años	19,8%	42,5%	57,5%
Menos de 14 años	35,6%	40,3%	59,7%
Menos de 15 años	22,3%	34,4%	65,6%
15 años o más	14,5%	30,9%	69,1%

Cuando se compara el resultado de los docentes según la ubicación de la escuela, se encuentra un porcentaje ligeramente mayor de docentes Competentes y Destacados en contextos urbanos (63%) comparado con contextos rurales (58%). Asimismo, los profesores que ejercen en establecimientos de los dos grupos socioeconómicos más altos tienen un desempeño relativo mejor que sus pares de establecimientos de nivel más bajo: la proporción de docentes con desempeño Competente y Destacado en establecimientos

4 El grupo socioeconómico (GSE) es un indicador calculado por el Ministerio de Educación en el marco del SIMCE que clasifica a las escuelas tomando en consideración el nivel de ingresos y nivel educacional de las familias, junto con el índice de vulnerabilidad de establecimiento (que es calculado por la Junta Nacional de Auxilio Escolar y Becas, JUNAE).

con GSE bajo y medio bajo es 62% y 60%, respectivamente, mientras que el porcentaje sube a un 65% y 67% en establecimientos GSE medio y medio alto. Esto representa diferencias pequeñas por GSE (en una tendencia a mejor desempeño en establecimientos de más alto GSE).

Otro antecedente del contexto indagado fueron los recursos disponibles al interior de la escuela. Para ello se utilizó un índice que da cuenta del nivel de equipamiento tecnológico dentro de las escuelas, basado en la disponibilidad de computadoras, proyectores y videograbadores en el establecimiento. El promedio de este índice, que varía de 0 (total ausencia) a tres (total disponibilidad), es 2,2 (DE=0,78). Al ordenar a las escuelas según esta medida y distribuir las en tres grupos (bajo, medio y alto nivel de recursos tecnológicos) se observa que a mayor disponibilidad de recursos tecnológicos, también hay una mayor proporción de docentes con buen desempeño en la Evaluación Docente (Competente o Destacado). En establecimientos con bajo nivel de recursos tecnológicos esta proporción es de 54%, mientras que en establecimientos con nivel medio es de 60%, y en establecimientos con nivel alto es 66%.

En consideración al tamaño del establecimiento no es posible apreciar una relación con desempeño. Si bien existe una menor proporción de docentes Competentes y Destacados en las escuelas más pequeñas, no se observa un patrón a medida que aumenta el tamaño.

También se estudió la relación entre desempeño docente y expectativas de los padres sobre la escolaridad de sus hijos. Dichas expectativas podrían dar cuenta indirectamente de diversas formas de apoyo de la familia a la labor del establecimiento y del profesor. El promedio de estas expectativas es 13 años de escolaridad (DE=1,62), que representa un año más que la Enseñanza Media Completa. Según se observa en la Tabla 4, solo una minoría de los docentes trabaja en establecimientos cuyos padres creen que sus hijos no terminarán la Enseñanza Media y, en aquellas escuelas en que los padres tienen expectativas más bajas, existe claramente una menor proporción de docentes Competentes y Destacados. Cuando las expectativas de escolaridad de los padres son menores a 10 años la proporción de docentes Competentes y Destacados es 44%, porcentaje que sube a un 58% cuando esas expectativas son de 12 años. Y cuando las expectativas de los padres son mayores a 15 años de escolaridad, el porcentaje de docentes Competentes y Destacados alcanza el 69%. Esta es una interesante tendencia que será retomada posteriormente en este capítulo.

En esta sección se presentaron análisis que ilustran diferencias en el desempeño de los docentes según diversos antecedentes individuales y contextuales. Al evaluar éstos independientemente, sin tomar en consideración otros atributos, aparecen diferencias a favor de las docentes mujeres, los profesores y profesoras de edad media, aquellos formados en modalidades diurna y de mayor duración, con participación en actividades de formación continua y mayor experiencia específica en el subsector en que se evalúa. Adicionalmente, quienes mostraron mayores expectativas académicas sobre sus alumnos obtuvieron también un mejor desempeño en la Evaluación Docente. Sobre los elementos de contexto, obtienen mejor resultado los docentes en establecimientos urbanos y con mayor nivel de recursos tecnológicos. Adicionalmente, no existe una tendencia que conecte el grupo socioeconómico del establecimiento con el desempeño docente. Y, por último, se constata que hay más docentes mejor evaluados en escuelas donde los padres tienen expectativas más altas sobre la escolaridad que alcanzarán sus hijos.

Análisis multinivel de los correlatos del desempeño docente

La sección anterior entregó una aproximación descriptiva a la relación entre diversos antecedentes individuales y de contexto con el desempeño docente. Sin embargo, las limitaciones propias de un análisis solamente descriptivo restringen el alcance de sus resultados. Existen al menos tres limitaciones que obligan a mirar los resultados ya expuestos con precaución. En primer lugar, esos análisis no permiten establecer relaciones de causalidad; por ejemplo, no es posible afirmar con certeza que un mayor tamaño de la clase cause mejor desempeño en los docentes ni que el hecho de participar en actividades de formación explique los mejores resultados. Hipótesis alternativas también son válidas: los mejores docentes podrían ser asignados a clases más numerosas y estar más interesados por actividades de formación. Una segunda limitación es que algunos de estos antecedentes están vinculados entre sí en alguna medida; por ejemplo, la elección de modalidad de la carrera con su duración, o el nivel socioeconómico del establecimiento con su entorno (rural o urbano). En ese sentido, sería conveniente estimar la contribución única de cada atributo al desempeño del docente. En tercer lugar, tal como se sugiere desde la investigación educacional, es necesario controlar estadísticamente que los profesores compartan influencias en común por el hecho de trabajar en un mismo establecimiento y municipio.

A través de los estudios que se presentan más adelante, se intenta abordar en parte estas limitaciones. La primera de ellas, referida a la causalidad, es compleja de resolver considerando la extrema dificultad y el alto costo de diseñar y realizar estudios experimentales en educación. Teniendo en mente esta limitación, los *modelos lineales jerárquicos* (Bryk & Raudenbush, 2001), llamados también *modelos multinivel* (Snijders & Bosker, 1999) o *modelos con efectos mixtos* (Singer, 1998) permiten abordar la segunda y tercera limitación: estudiar la contribución única de cada atributo simultáneamente, y considerar la dependencia entre las mediciones de docentes en un mismo contexto. Este último punto es lo que distingue a los modelos lineales jerárquicos de otros métodos de estimación, como el de mínimos cuadrados ordinarios (regresión múltiple).

La consideración de usar modelos lineales jerárquicos pasa por el interés de superar los problemas asociados con el análisis de regresión múltiple bajo una estructura multinivel o jerárquica, tal como la que se tiene al considerar datos de contexto como en el caso de las mediciones educacionales, dado que los profesores están siempre agrupados en establecimientos educacionales⁵, y éstos en unidades mayores como los municipios, en el caso chileno. Siguiendo a Luke (2004), en un orden teórico, problemas en los que conceptualmente existen datos e hipótesis jerárquicas en naturaleza, debieran analizarse utilizando aproximaciones que se hagan cargo de dicha particularidad. Ignorarla, explica este autor, lleva a problemas de inferencia como la falacia ecológica o la falacia atomista.

En un orden estadístico, existen dos razones para emplear modelos lineales jerárquicos. Primero, bajo una estructura de datos jerárquica utilizar una regresión múltiple viola el supuesto de independencia de las observaciones. El segundo problema es que, las estimaciones de los coeficientes de regresión múltiple se aplican de la misma manera a todos los contextos. En ambos casos se pueden obtener resultados inadecuados.

Sin pretender entrar en mayores detalles técnicos y estadísticos, de paso ya descritos en cada investigación en particular, se presentan los resultados de tres estudios que han hecho uso de modelos lineales jerárquicos para investigar las relaciones entre desempeño docente con antecedentes individuales y de contexto. Si bien estos estudios comparten una misma aproximación metodológica, difieren en las especificaciones de los modelos, el conjunto de antecedentes considerados, y las poblaciones de profesores evaluados. Cada una de estas investigaciones aporta evidencia sobre la importancia de lo individual por encima de lo contextual, y orienta sobre cuáles de los antecedentes en los análisis ayudan de mejor manera a comprender los altos y bajos desempeños de los profesores en la Evaluación Docente.

.....

5 Para efectos de la evaluación cada docente está asociado a solo un establecimiento.

Investigación con datos de docentes generalistas de la Evaluación Docente 2005

La primera de las investigaciones en utilizar la aproximación multinivel con resultados de desempeño de los profesores en Chile es un estudio realizado con datos de 5.164 docentes generalistas de la Evaluación Docente de 2005 (Eisenberg, 2008). El modelo que se reporta a continuación⁶ consideró conjuntamente antecedentes del nivel individual del profesor y la escuela (juntos elementos individuales y de contexto), y en un segundo nivel, antecedentes de la comuna (solo contexto). Entre los hallazgos del estudio se encuentra que la mayor parte del desempeño de los docentes es atribuible a características individuales o de la escuela. Sin embargo, un porcentaje no irrelevante (12%) del desempeño docente puede ser atribuido a aspectos de la comuna. Este resultado apoya el argumento de que es razonable considerar los aspectos del contexto al momento de analizar los resultados de los profesores en la Evaluación Docente, idea que se verá apoyada en los siguientes estudios comentados en este capítulo.

Sobre los antecedentes individuales específicos que estarían conectados con los resultados de la Evaluación Docente, en el estudio de Eisenberg (2008) los antecedentes de formación inicial (duración de la carrera y modalidad) y continua (participación en los Programas de Perfeccionamiento Fundamental, la formación de post-grado), se asociaron de manera positiva con el desempeño (formación inicial en modalidad diurna y más extensa, así como participación en actividades de formación continua se asocian a mejor desempeño). Otro de los hallazgos fue que el número de años enseñando en el subsector también se conectó de manera positiva con desempeño. En cambio, los docentes con mayor experiencia general enseñando (correlacionado altamente con edad) obtuvieron resultados más pobres, lo que se acentuó en contextos rurales comparado con urbanos. Se encontró además que las mujeres mostraron un mejor desempeño que los hombres.

Como se mencionó, a nivel individual (del profesor) se incluyeron variables que dan cuenta de correlatos a nivel de la escuela, y se encontró que el nivel socioeconómico se asoció positivamente con el desempeño (GSE más alto, mayor desempeño), aunque muy débilmente. En tanto, el promedio SIMCE de la escuela mostró una asociación positiva con el desempeño docente. Eisenberg también estudió el tamaño de la clase y la ubicación del establecimiento (urbano-rural), sin encontrar asociaciones significativas con desempeño.

Sobre las variables incluidas a nivel de la comuna, se determinó que aquellas cuya administración educacional depende de un Departamento de Administración de Educación Municipal (DAEM) (en comparación con una Corporación), las que tienen un nivel socioeconómico más bajo, y las que exhiben mayores puntajes SIMCE comunal mostraron docentes con mejor desempeño. Aunque, nuevamente debe señalarse que la fuerza de las asociaciones es relativamente débil.

Investigación con datos de docentes evaluados entre 2005–2007

Uno de los aspectos que no pudo ser implementado en el estudio de Eisenberg (2008) fue analizar un modelo estadístico con una estructura jerárquica que incluyera los tres niveles: individual, escuela y municipio. Debido a limitaciones de los datos disponibles, dicho estudio realizó un análisis modelando estadísticamente aspectos individuales y escolares en conjunto, como si pertenecieran a un mismo nivel, y en otro nivel los aspectos vinculados al municipio. Un análisis considerando por separado estos tres niveles jerárquicos (individual, escuela y municipio) se encuentra en la investigación de Valencia (2009), que utilizó datos de docentes de Primer Ciclo, Segundo Ciclo y Enseñanza Media en diversos subsectores. Este estudio confirma que un porcentaje mayoritario de las diferencias de desempeño son atribuibles a diferencias individuales, y muestra un incremento en el peso del contexto: se encontró que un 11% de las diferencias en el desempeño son atribuibles a aspectos del establecimiento y un 8% a aspectos de la comuna. Esto implica que el peso del contexto alcanza un 19% (ver Tabla 5).

.....

6 Se hará alusión al principal de los modelos analizados.

Cabe mencionar que Valencia (2009) confirma la mayoría de las relaciones encontradas por Eisenberg (2008), a pesar de usar datos de distintos años y diferentes especificaciones de los modelos estadísticos. Los hallazgos de la investigación más reciente tienen que ver con un efecto de la edad del docente, encontrándose una relación cuadrática (forma de U invertida), esto es, un incremento inicial y luego un decrecimiento del desempeño con la edad, tal como indicaban también los análisis descriptivos presentados en la primera parte de este capítulo. Esto es distinto de lo encontrado por el estudio de Eisenberg usando experiencia general (altamente correlacionada con edad), que mostró una asociación negativa (sin embargo, no se modeló cuadráticamente).

Valencia (2009) también incluyó otros antecedentes laborales. El número de horas de contrato del docente mostró un comportamiento similar al de edad (en forma de U invertida) observándose un mejor desempeño en los docentes con 35 a 39 horas contratadas, lo que sugiere la existencia de un punto óptimo en ese rango. Por otro lado, el número de alumnos en la sala se asoció directamente con desempeño docente (más alumnos, mejor desempeño), mientras que el número de escuelas en las que el docente trabaja mostró una relación negativa. Finalmente, otro de los aspectos individuales estudiados, el obtener certificaciones profesionales (en este caso la Asignación de Excelencia Pedagógica) aparece conectado directamente con desempeño docente.

Es conveniente enfatizar que estos resultados están controlados estadísticamente por el nivel y subsector en que se evalúa el docente, antecedente que marcó diferencias importantes en el desempeño. Los docentes generalistas estuvieron asociados a los resultados más bajos, mientras que aquellos de Lenguaje y Matemáticas de Segundo Ciclo, y los de Historia en Enseñanza Media se asociaron a un desempeño más alto.

Un hallazgo interesante de la investigación es que el nivel socioeconómico de la escuela no mostró relación con el desempeño de los docentes, lo mismo que la ubicación del establecimiento (urbano o rural) y su tamaño (número de alumnos). Las variables de la escuela estudiadas y que se asociaron de manera positiva con desempeño de los profesores son: el nivel de recursos tecnológicos del establecimiento (computadores, proyectores, impresoras, etc.), las expectativas de los padres sobre los años de escolaridad de los alumnos y el número de docentes con AEP del establecimiento; así, mayores recursos, más altas expectativas y mayor número de docentes con AEP se asociaron a un mejor promedio de Evaluación Docente en la escuela. Cabe señalar que este estudio no incluyó el promedio SIMCE de los alumnos, porque pareció pertinente estudiar la conexión en el sentido inverso, esto es, el desempeño docente como antecedente del rendimiento de los alumnos (tal como se presenta en Taut, Santelices & Manzi, en este volumen).

Finalmente, entre las variables que dan cuenta de condiciones a nivel de la comuna, se examinó el tipo de administración (DAEM o Corporación), el gasto por alumno en educación, el tamaño de la comuna y el porcentaje de escuelas privadas en la misma, sin que se encontrara conexión con desempeño docente en ninguna de ellas. Sí se encontró una asociación directa con el índice de pobreza de la comuna (de la encuesta de caracterización socioeconómica, CASEN), y el porcentaje de escuelas particulares subvencionadas: mayor índice de pobreza comunal y mayor proporción de escuelas subvencionadas se asociaron a un mejor desempeño docente en la comuna. Cabe consignar que todas las relaciones encontradas en este y los demás estudios se interpretan manteniendo constante el efecto de los demás antecedentes.

● **Tabla 5**
 ● *Efectos fijos de los antecedentes individuales y de contexto sobre el desempeño docente*

Efecto	coef	s.e.	sig.
Intercepto	2,22	0,05	***
Se evalúa en 2005	0,01	0,01	
Se evalúa en 2006	0,02	0,00	**
Se evalúa en 2007	0,00	.	
Nivel individual			
Sexo (mujer)	0,07	0,00	***
Edad (cuadr.)	0,00	0,00	***
Lenguaje segundo ciclo	0,13	0,01	***
Matemáticas segundo ciclo	0,12	0,01	***
Ciencias Naturales segundo ciclo	0,04	0,01	***
Ciencias Sociales segundo ciclo	0,02	0,01	**
Lenguaje enseñanza media	0,04	0,01	**
Matemática enseñanza media	0,02	0,01	*
Biología enseñanza media	0,06	0,02	**
Física enseñanza media	0,07	0,03	**
Historia enseñanza media	0,12	0,02	***
Química enseñanza media	0,04	0,02	
Generalista primer ciclo	0,00	.	
Duración de la carrera	0,00	0,00	*
Modalidad de la carrera (diurna)	0,04	0,00	***
Tiene postítulo o posgrado	0,04	0,00	***
Participa en PPF	0,02	0,00	***
Experiencia subsector: 1–2 años	-0,07	0,01	***
Experiencia sub.: 3–5 años	-0,05	0,01	***
Experiencia sub.: 6–10 años	-0,03	0,00	***
Experiencia sub.: 10 o más años	0,00	.	
Horas de contrato (cuadr.)	0,00	0,00	***
Horas lectivas sobre contrato (%)	0,00	0,00	
Número de alumnos	0,00	0,00	***
Trabaja en 2 o más escuelas	-0,06	0,01	***
Tiene acreditación AEP	0,15	0,01	***
Nivel escuela			
Escuela es urbana	0,00	0,01	
Matrícula total escuela	0,00	0,00	
GSE de la escuela	0,01	0,01	
Expectativa académica padres	0,02	0,00	***
Número docentes AEP	0,02	0,00	***
Índice de Recursos tecnológicos	0,02	0,01	**
Nivel comuna			
Municipio tiene DAEM	-0,01	0,02	
Gasto por alumno en educación	0,00	0,00	
Índice de pobreza CASEN	0,00	0,00	***
Porc. de escuelas part. subv.	0,00	0,00	***
Porc. de escuelas privadas	0,00	0,00	
Total de población de la comuna	0,00	0,00	***

Extraído de Valencia (2009); * p < 0,05; ** p < 0,01; *** p < 0,001

Investigación con datos de docentes evaluados en 2006–2008

Un tercer estudio exploró en mayor detalle los tópicos investigados por los dos estudios ya presentados, incorporando como antecedentes del profesor aspectos motivacionales y creencias; además, este estudio agrega nuevos elementos del contexto en la sala de clases y realiza un análisis diferenciando por nivel de enseñanza (Santelices, Valencia, Taut & Manzi, 2010; 2011).

Los autores modelaron los datos de desempeño de 1.422 docentes que fueron evaluados en 4° u 8° año básico. Al igual que en el estudio anterior, el desempeño de los docentes se estudió bajo una estructura de tres niveles considerando diferencias individuales, entre escuelas y entre comunas. Los hallazgos vuelven a confirmar que las diferencias individuales son las que en mayor medida ayudan a explicar los resultados de desempeño de los docentes. Sin embargo, una proporción considerable puede ser atribuida a diferencias en la escuela (estimado en un 9%) y en la comuna (15%) a nivel general, que suman en total un 24%, proporción mayor a la encontrada en los dos estudios precedentes. Un hallazgo destacable es que la parte atribuible al contexto varía según el nivel de enseñanza considerado, y existe heterogeneidad en la importancia relativa de la escuela y municipio; ésta es siempre mayor al 20% y alcanza incluso un 45%.

Entre las nuevas variables exploradas a nivel individual incluidas en esta investigación y conectadas con el desempeño de los docentes, los resultados indican que las expectativas del profesor sobre los años de escolaridad de sus alumnos se relacionan positivamente con su desempeño en la Evaluación Docente (más altas expectativas, mejor desempeño). Del mismo modo, la mayor percepción de auto-eficacia en la asignatura que el docente imparte y en la que es evaluado también se asocia a mejor desempeño. Por otra parte, las expectativas laborales futuras (mantenerse como profesor, abandonar el profesorado, abandonar la educación) no mostraron una conexión con el desempeño del docente.

A su vez, algunas de las variables de la sala de clases incorporadas al estudio, como disponer de biblioteca de aula y el ingreso familiar promedio del curso, mostraron una relación positiva, aunque pequeña, con desempeño. También fue incluido el grado de cobertura de la materia por parte del profesor, el nivel educacional promedio de los padres de la clase, y las expectativas de años de escolaridad promedio de los padres de la clase, los que mostraron relación con el desempeño docente.

Respecto a las variables incluidas a nivel de la escuela, se consideró la ubicación del establecimiento, el nivel socioeconómico de la escuela (GSE del SIMCE), la cantidad de recursos tecnológicos y el número de docentes certificados (AEP) del establecimiento. Estos antecedentes de contexto mostraron el mismo patrón de resultados que en el estudio de Valencia (2009), en el que el GSE no se relacionó con desempeño mientras que el acceso a recursos tecnológicos y la cantidad de docentes de excelencia se conectan positivamente. Sobre las variables incluidas para comprender el efecto de la comuna se incluyó la matrícula inicial de los establecimientos de la comuna (número de alumnos, otra medida de tamaño) y un índice de recursos económicos que señala la contribución municipal al total del ingreso en educación. Éstas resultaron relacionadas de forma inversa con desempeño, esto es, que en las comunas con menor número de alumnos y con menor contribución al ingreso municipal en educación se estimó un desempeño docente más alto. También se incluyó en el nivel comuna el tipo de administración de educación y el porcentaje de pobreza de la comuna, ambas no mostraron conexión con desempeño docente.

Conclusiones sobre la relación entre desempeño y antecedentes

Durante este capítulo se ha mostrado evidencia que ayuda a entender cómo ciertos aspectos individuales y contextuales aparecen conectados al desempeño de los profesores participantes en la Evaluación Docente. Esto es importante para clarificar qué elementos están efectivamente conectados con la práctica docente y cuáles no, y con esto ayudar a informar decisiones que contribuyan a potenciar el quehacer docente.

Es conveniente consignar que tanto los análisis descriptivos como las investigaciones citadas tienen la limitación de no poder concluir sobre relaciones de causalidad. Esto quiere decir fundamentalmente que ninguno de los análisis e investigaciones anteriores permite establecer de manera categórica que dichos antecedentes causen cambios en el desempeño de los docentes. Otro aspecto a considerar es que en las investigaciones revisadas las magnitudes de las conexiones entre los antecedentes y el desempeño docente son más bien pequeñas. Sin embargo, a pesar de las limitaciones mencionadas, estas relaciones muestran ser altamente consistentes entre las distintas investigaciones, y además son en alto grado coherentes con la evidencia acumulada en contextos diferentes al nuestro (Darling–Hammond, Berry & Thoreson, 2001; Darling–Hammond & Youngs, 2002; Goldhaber & Brewer, 2000; Goldhaber, Perry & Anthony, 2004; Wayne & Youngs, 2003)⁷.

El aspecto más importante revisado en este capítulo y que se desprende de las investigaciones que usan modelos lineales jerárquicos, es el hecho de que un considerable porcentaje de las diferencias en el desempeño de los docentes puede ser atribuido a condiciones en sus establecimientos y comunas, es decir, aspectos del contexto. Esto implica básicamente que distintos elementos de la escuela y comuna, como los recursos materiales, la organización y la cultura, pueden ayudar a explicar el desempeño de los docentes al interior de la sala de clases. Este hallazgo tiene numerosas consecuencias. Una de ellas, consiste en la necesidad de fortalecer el rol de apoyo al trabajo de aula del docente desde la dirección escolar y municipal. Sin embargo, se requiere de mayor evidencia empírica relativa a la Evaluación Docente y de las características y elementos específicos de la estructura y organización escolar que pueden ser promovidos para incrementar la calidad de las prácticas pedagógicas de los profesores en servicio.

Al intentar explicar las diferencias de desempeño entre los docentes a partir de antecedentes personales, los distintos análisis y estudios revisados muestran que los hombres y aquellos docentes en las últimas etapas de su carrera son quienes requieren de mayor apoyo en su trabajo de aula. Por otra parte, aspectos claves de la formación inicial como la modalidad y duración de la carrera aparecen conectados de manera importante con desempeño docente, indicando que la formación inicial juega un rol en el desempeño de los docentes. Del mismo modo —e independiente de la dirección de la relación causal—, la conexión existente entre el desempeño docente y la participación en actividades de formación en servicio, refuerza la vital importancia de la formación continua en el marco de una carrera profesional docente orientada a mejorar la calidad de la enseñanza.

De los antecedentes laborales estudiados por otra parte, la importancia de una mayor experiencia en el subsector permite sugerir, entre otras cosas, fortalecer la inducción inicial a la espera de que el docente adquiera las competencias y habilidades necesarias para el ejercicio docente en su subsector específico.

Los estudios con modelos lineales jerárquicos dan evidencia de consistencia a los resultados de AEP y Evaluación Docente en la línea de lo esperable. En primer lugar, los docentes de excelencia muestran una alta consistencia en su resultado de evaluación, aun cuando fueron evaluados en distintas condiciones y mediante programas diferentes. Aunque es justo señalar, que en esto tiene un rol importante el hecho de que, obtener un resultado Básico o Insatisfactorio, implica perder automáticamente la condición de

7 También pueden ser consultadas las investigaciones presentadas en este capítulo para obtener un listado más completo de artículos y antecedentes relacionados con desempeño en particular y calidad docente en general.

excelencia para quienes ya la tienen. Por otra parte, los estudios muestran que una mayor cantidad de docentes de excelencia en el establecimiento está asociada a mejores resultados de Evaluación Docente a nivel escuela. Si bien es un resultado sujeto a verificación futura, sugiere un efecto de pares e invita a estudiar y estimular mecanismos por el cual docentes de excelencia o de alto rendimiento apoyen y transfieran habilidades al resto de los docentes en su comuna.

Otro tema muy importante revisado en este capítulo es cómo las creencias y expectativas se conectan con desempeño docente. Aquí, las creencias sobre la escolaridad de los alumnos tanto propias (del mismo docente), como de los padres a nivel del establecimiento se conectan positivamente con el desempeño de los profesores. Lo mismo ocurre con el sentido de auto-eficacia en el subsector (cuán preparado se siente para enseñar), en el que mayor confianza se asocia a mejor desempeño. Si bien el estudio de las creencias y expectativas aparece con mucho desarrollo en otros contextos de investigación, la Evaluación Docente ofrece la oportunidad de estudiar más profundamente las conexiones de ellas con las prácticas pedagógicas directamente.

También es importante mencionar el rol de las condiciones socioeconómicas del establecimiento. Al contrario de lo podría suponerse, la evidencia actual apunta a que no existe una relación entre nivel socioeconómico y desempeño docente. Es decir, una vez controlado por los antecedentes personales del profesor, los profesores en escuelas con menor nivel socioeconómico no obtienen peor desempeño que los docentes en escuelas con alto nivel socioeconómico. A nivel de comuna se observa incluso una pequeña tendencia en la que los docentes de comunas más dependientes financieramente o con mayor índice de pobreza tienen profesores de mejor desempeño. Sin embargo, cuando se consideran los recursos con que cuentan apoderados y el establecimiento educacional, tanto en la sala de clases como a nivel general, sí hay evidencia de una relación positiva con desempeño docente. En otras palabras, el acceso a recursos, se presume para el aprendizaje pues es una medida indirecta, se asocia a mejor desempeño docente. A pesar de estar considerando solo la disponibilidad de recursos y no sus formas de uso, ello ya es indicativo de que existen condiciones básicas en el establecimiento que favorecen o entorpecen la labor docente, incluso cuando las características de los alumnos y el entorno son similares.

Todo lo reportado anteriormente abre un espacio para profundizar en la investigación sobre los condicionantes de la calidad docente en Chile utilizando datos de Evaluación Docente. Los hallazgos presentados aquí tienen un carácter inicial, a la espera de poder ser ampliados, confirmados u objetados por otros estudios que puedan utilizar y vincular datos de la Evaluación Docente con otras fuentes de información o incluso datos levantados intencionadamente. A modo de ejemplo, el uso de pruebas para evaluar conocimiento pedagógico y disciplinario al comienzo de la carrera con el Programa Inicia puede indicar si la calidad de la formación inicial recibida se conecta con el desempeño docente en el futuro. También la re-evaluación obligatoria de los docentes luego de cuatro años ofrece la oportunidad de estudiar longitudinalmente los efectos de algunos de los antecedentes estudiados, de modo de someter a prueba la dirección de las relaciones enunciadas.

Las investigaciones y análisis presentados en este capítulo representan un primer paso y una de las tantas alternativas para intentar comprender la relación entre los antecedentes de los docentes, y del contexto en que ejercen, sobre su desempeño. Los estudios descritos han puesto énfasis en el efecto de lo contextual sobre el desempeño, descartando el efecto de la condición social del alumnado. El mejor y mayor conocimiento de los antecedentes vinculados a los resultados de la Evaluación Docente posibilita una comprensión más cabal de las prácticas pedagógicas en el sistema municipal chileno que beneficia a todos sus actores, principalmente profesores y autoridades educacionales que interpretan y hacen uso de los resultados de la Evaluación Docente.



8

Investigación educacional
a partir de la Evaluación
Docente

Microgénesis de la enseñanza: Zoom en el modo en que los profesores presentan contenidos disciplinarios⁽¹⁾

Carlos Cornejo · David Silva · Himmbler Olivares
Escuela de Psicología, Pontificia Universidad Católica de Chile.

Introducción

Todo profesor debe entregar una serie de contenidos disciplinarios a sus alumnos. Los modos de presentar estos contenidos, sin embargo, varían sensiblemente según el profesor y la disciplina. En este estudio indagamos si existen diferencias pesquisables en los modos de presentación de contenidos según el nivel de desempeño obtenido por el profesor en el Sistema de Evaluación del Desempeño Profesional Docente (en adelante Evaluación Docente). Para contestar esta pregunta, realizamos un análisis detallado de la situación interaccional entre el profesor y sus estudiantes dentro de la sala de clases. Técnicamente, este tipo de análisis es interaccional y microgenético. Es *interaccional* por cuanto el asunto de interés no está localizado en los procesos mentales particulares del profesor ni del estudiante en la enseñanza/aprendizaje, sino en los procesos emergentes acaecidos en la situación interaccional. En este sentido, el presente estudio es sintónico con la investigación de las últimas décadas acerca de prácticas pedagógicas en situaciones interaccionales reales en sala de clases (Coll & Edwards, 1997; Edwards & Mercer, 1987; Mercer & Littleton, 2007; Mercer, 2010). Por otra parte, es un análisis *microgenético* debido a su foco en la dinámica en tiempo real de la construcción colectiva de un significado común (Rosenthal, 2004).

En síntesis, el presente trabajo asume el estudio de las prácticas educativas que ocurren en el aula, enfatizando por igual el carácter interaccional del aprendizaje como la dimensión de significado implicada en el encuentro profesor–estudiante.

Antecedentes teóricos relevantes

En este estudio registramos aquellas expresiones lingüísticas que, producidas ya sea por el profesor o sus estudiantes, generan una construcción sustantiva en el sentido o significado de un determinado contenido disciplinar. A esa unidad de análisis la denominamos *expansión de sentido* (en adelante ES). Para cada una de esas unidades es posible identificar un asunto disciplinar y un *modo de presentación* de éste dentro de la situación pedagógica. El asunto es el *qué* de la ES, en tanto que el modo de presentación se refiere a cómo el asunto es introducido. Por ejemplo, supongamos que un estudiante expresa verbalmente “si cortamos una torta en 10 pedazos, cada pedazo corresponde a 1/10 de la torta”. En tal caso, el asunto de la ES es “Fracción 1/10”; su modo de presentación, es la ejemplificación.

En la presente investigación distinguimos tres modos de presentación: ostensión, metaforización y adscripción. En la *ostensión*, un asunto es presentado a través de un caso o situación particular que permite comprender lo que se quiere dar a entender. Un caso típico de ostensión es el ejemplo. En la *metaforización*, el asunto es insinuado a través de una presentación que, si bien en términos objetivos corresponde a otro asunto, figuradamente permite su aparición. Casos típicos de metaforización son las comparaciones y las metáforas mismas. Por último, en la *adscripción*, el asunto es presentado a través

1 Esta investigación fue financiada por el Programa Bicentenario de Ciencia y Tecnología CIE 01 CONICYT a través del Centro de Estudios de Políticas y Prácticas Educativas (CEPPE) de la Pontificia Universidad Católica de Chile. Agradecemos al Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas (CPEIP) del Ministerio de Educación y al Centro de Medición MIDE UC, por las facilidades prestadas para el uso de material audiovisual y resultados recogidos por el Sistema de Evaluación del Desempeño Profesional Docente. Email para correspondencia: cca@uc.cl

de categorías o características generales relacionadas con él de un modo claro y preciso. Ejemplos de adscripción son definir, describir o enumerar las características de un determinado asunto.

Método

Revisamos un total de 138 clases de matemáticas de 5° y 6° año de Enseñanza Básica en Chile, grabadas en el año 2004 para la Evaluación Docente (Manzi, 2007b; Manzi, Preiss, González & Sun, 2008). Todos los profesores cuyas clases fueron analizadas habían firmado previamente un consentimiento informado para el uso de su evidencia de evaluación con fines de investigación.

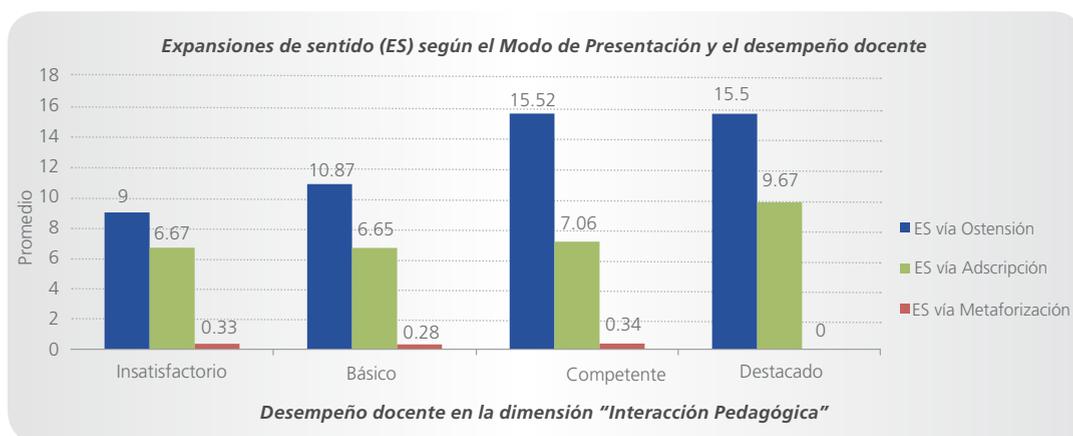
Para cada clase videograbada se analizó el número y tipo de ES desplegados. Se codificó además el asunto y su modo de presentación. Posteriormente esos datos fueron correlacionados con el desempeño alcanzado por el profesor en la Evaluación Docente. Los videos fueron revisados por evaluadores capacitados en codificar modos de presentación y articulación de los contenidos. La revisión de los videos contempló los 10 primeros minutos de cada clase.

Resultados

Se observó un promedio de 20.49 ES (DE=11,91) y 10.67 asuntos (DE=6,93). La clase con mayor cantidad de *expansiones de sentido* (ES) sumó 56, mientras que la clase con menor cantidad de ES no registró ninguna dentro de los 10 minutos iniciales. La clase con mayor cantidad de asuntos tuvo 38, mientras que la clase con menor cantidad no presentó ninguno.

Respecto a los Modos de Presentación, los profesores realizan en promedio 13,16 ES vía *ostensión* (DE=9,19), lo que equivale, en promedio, a un 63,85% (DE=22,35%) de las ES de los 10 primeros minutos de cada clase (rango: 40–0). Las ES vía adscripción aparecieron en promedio 7,05 veces (DE=6,00) en los primeros 10 minutos, lo que equivale a un porcentaje promedio de 34,41% (DE=22,49%) de las ES de cada clase (rango: 28–0). Por último, las ES vía metaforización aparecen solo en 21 de las 138 clases analizadas, siendo por lo tanto un fenómeno bastante escaso. Esto equivale al 1,73% (DE=5,97%) de las ES de cada clase. En el Gráfico 1 se puede observar que el promedio de la cantidad total de verbalizaciones vía ostensión y vía adscripción aumenta al agrupar a los profesores según su resultado en el Módulo 2 (clase filmada) del Portafolio de la Evaluación Docente, y en cada una de las tres dimensiones evaluadas en este módulo.

● **Gráfico 1**
 ● Promedio de *Expansiones de Sentido* según el nivel de desempeño en la clase filmada de la Evaluación Docente



El número de ES presenta una correlación positiva baja, pero significativa con el desempeño global en la clase filmada ($r_s=.282$, $p=.000$). El número de ES correlaciona también con las dimensiones "Interacción Pedagógica" ($r_s=.296$, $p=.000$), "Estructura de la clase" ($r_s=.243$, $p=.002$) y "Ambiente de la Clase" ($r_s=.220$, $p=.005$), en las que se agrupan los indicadores evaluados en el Módulo 2 (clase filmada) del Portafolio. Por su parte, el número de asuntos también presenta una correlación positiva con el desempeño global en la clase filmada ($r_s=.274$, $p=.001$), así como con las dimensiones "Interacción Pedagógica" ($r_s=.273$, $p=.001$), "Ambiente de la Clase" ($r_s=.249$, $p=.002$) y "Estructura de la clase" ($r_s=.207$, $p=.007$). En la Tabla 1 se resumen estas y otras correlaciones.

● **Tabla 1**
 ● *Resumen de correlaciones observadas entre Expansiones de sentido (ES) y nivel de desempeño en la clase filmada de la Evaluación Docente*

	Resumen de correlaciones			
	Desempeño global en clase filmada	Ambiente de la clase	Estructura de la clase	Interacción pedagógica
N de ES	,28**	,22**	,24**	,30**
N de Asuntos	,27**	,25**	,20**	,27
N de ES vía ostensión	,315**	,214**	,282**	,327**
N de ES vía adscripción	,205**	,171*	,175*	,199*
N de ES vía metaforización	,000	,062	,022	,003

(*) $p < 0,05$

(**) $p < 0,01$

Conclusión

La comparación entre profesores efectivos y no efectivos muestra que los docentes de buen desempeño tienen mayor variedad de formas de presentar los contenidos a sus estudiantes, así como mayor frecuencia de intervenciones que articulan dichos contenidos. En otras palabras, los buenos profesores hablan más de contenido técnico y, además, abordan una mayor variedad de aspectos de los contenidos que presentan. Adicionalmente, estos profesores integran explícitamente los diferentes aspectos abordados en sus clases y muestran un equilibrio óptimo entre los distintos modos de presentación de los contenidos: el profesor de alto rendimiento en la Evaluación Docente presenta una relación aproximada de tres ejemplos por cada definición que entrega.

¿Qué actividades realizan los docentes de NB1 para enseñar a leer en situación de evaluación docente? Enfoques tras las prácticas pedagógicas⁽¹⁾

Viviana Galdames

Facultad de Educación, Universidad Alberto Hurtado.

Lorena Medina

Facultad de Educación, Pontificia Universidad Católica de Chile.

Ernesto San Martín

Departamento de Estadística, Pontificia Universidad Católica de Chile.

Rosa Gaete

Facultad de Educación, Universidad Alberto Hurtado.

Andrea Valdivia

Becaria Postdoctorado Centro de Estudios de Políticas y Prácticas en Educación (CEPPE), Facultad de Educación, Pontificia Universidad Católica de Chile.

Introducción

El trabajo presenta los resultados de la primera fase de un proyecto centrado en las teorías implícitas y paradigmas, disciplinares y pedagógicos, de profesores responsables del aprendizaje inicial del lenguaje escrito. La pregunta central de esta fase es qué acciones y actividades realizan 90 profesores de 1° y 2° año de Enseñanza Básica que participaron en el Sistema de Evaluación del Desempeño Profesional Docente (en adelante Evaluación Docente) el año 2008, para enseñar a leer y escribir en sus salas de clase, y si dichas acciones se relacionan con concepciones, modelos o estrategias existentes en el área.

El problema

La deficiente calidad de la comprensión lectora de los estudiantes de todos los niveles del sistema escolar se constituye como uno de los principales desafíos tanto para los profesores como para la política pública. Los resultados de Lenguaje en el Sistema de Medición de la Calidad de la Educación (SIMCE) en 2007 muestran que el 40% de los niños y niñas se ubica en el nivel inicial de lectura. Ahora bien, desde una perspectiva de equidad, el problema es aún mayor, pues el 53% de los estudiantes de 4° básico de los estratos socioeconómicos bajo y medio-bajo se ubicaron en dicho nivel. Estos y otros datos muestran la necesidad de explorar las prácticas de enseñanza de lectura en el aula y preguntarse por las posibles causas del problema.

Las concepciones y teorías de los profesores incluyen su pensamiento acerca de cómo los niños desarrollan habilidades para la literacidad; el deber ser o rol del profesor en este proceso, y cómo sienten que deberían implementarse esas prácticas en la sala de clase (Charlesworth, Hart, Burts & Hernández, 1991; De Ford, 1985). A partir de estos hallazgos, resulta plausible el supuesto de que las creencias de los profesores impactan sus prácticas de aula (Fenstermacher, 1994; Kagan, 1992; Richardson, Anders, Tidwell & Lloyd, 1991). Hindman & Wasik (2008) señalan que se sabe poco de las creencias de los profesores acerca de cómo los niños aprenden a leer, cómo debe enseñárseles, y cómo esas creencias pueden relacionarse con factores claves de su experiencia de formación y práctica profesionales. Por tanto, esta investigación aborda un tema relevante, aunque poco explorado, tanto a nivel nacional como internacional.

1 Este trabajo corresponde a los resultados obtenidos en la primera fase de investigación del Proyecto CIEO-01-CONICYT "Teorías implícitas y paradigmas disciplinares de los docentes responsables del aprendizaje del lenguaje escrito en NB1", que se inscribe en la línea de Prácticas e interacciones de aula del Centro de Estudios de Políticas y Prácticas en Educación (CEPPE), de la Pontificia Universidad Católica de Chile.

Objetivo de esta etapa

Caracterizar las prácticas pedagógicas de los docentes responsables del aprendizaje inicial del lenguaje escrito.

Diseño metodológico

El estudio tiene una metodología mixta organizada en tres etapas. La primera, cuyos resultados se dan a conocer en este artículo, corresponde a la caracterización de las acciones docentes en las salas de clase a partir de la observación de sus prácticas de enseñanza. La segunda etapa se enfoca en las teorías implícitas subyacentes a sus prácticas de aula y en su reflexión sobre las mismas. La tercera etapa se centra en el análisis interpretativo y la triangulación de la información.

Muestra

De una población de 1.182 videos de docentes de 1º y 2º básico que participaron en la Evaluación Docente en el año 2008, siendo evaluados en el subsector Lenguaje y Comunicación, se construyó una muestra por cuotas con selección aleatoria de 90 videos. Los criterios de selección fueron: zona geográfica del país (norte, centro y sur), nivel de enseñanza (1º y 2º básico) y resultado obtenido en la Evaluación Docente.

Instrumento: Pauta de observación de la clase de lectura y escritura

Se visualizaron 90 videos de 40 minutos cada uno utilizando una pauta cerrada de observación directa de la enseñanza de la lectura y la escritura iniciales, construida en base a preguntas dicotómicas. El instrumento fue elaborado y piloteado por un equipo experto y recoge un amplio número de indicadores que describen la presencia o ausencia de actividades. Las observaciones fueron sometidas a análisis de acuerdo interjueces, obteniéndose para un 93% de los indicadores un acuerdo de 0,7 o superior.

Principales hallazgos

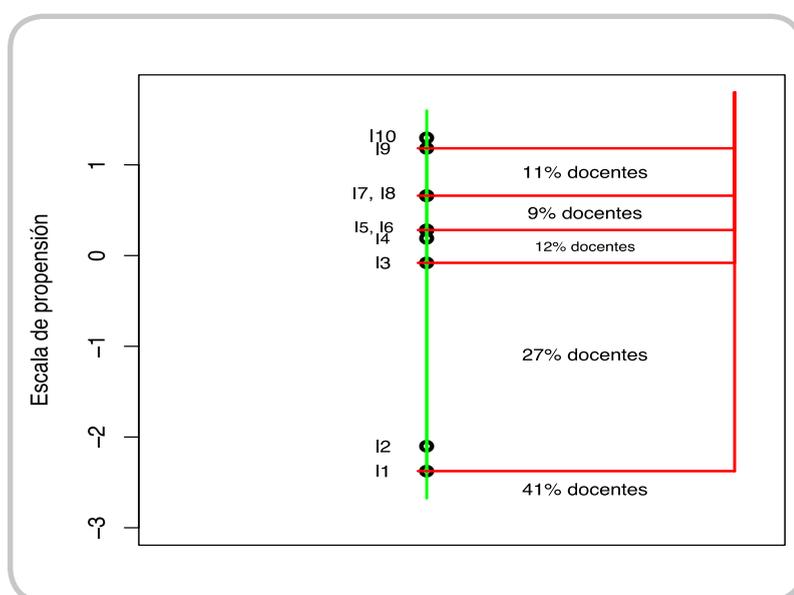
A partir de análisis descriptivos básicos, se pudo apreciar que las actividades más frecuentes correspondieron a profesores que utilizan textos (72%), lo cual es un avance interesante respecto a las prácticas más tradicionales; sin embargo, leen (55%) y formulan preguntas de carácter explícito (54%) a los alumnos, lo cual sigue dejando al docente como protagonista de la clase.

Mediante análisis factoriales exploratorios se obtuvieron cinco factores, de los cuales el factor 1 (que representa actividades centradas en la lectura) captura la mejor variabilidad individual de los profesores. Las actividades centradas en la lectura que conforman dicho factor son: preguntas sobre información explícita (I-1); profesor lee en voz alta (I-2); dar opiniones o puntos de vista (I-3); dar juicios o valoración de lo leído (I-4); preguntas sobre información implícita (I-5); anticipar la lectura con las ilustraciones (I-6); preguntas para que estudiante fundamente (I-7); anticipar texto por título (I-8); inducir texto a partir de formato y estructura (I-9); y realizar secuencia de hechos a partir de texto (I-10).

Cada docente está, por lo tanto, caracterizado por la eventual realización de cada una de estas 10 actividades a lo largo de la clase filmada. Dichas actividades son, en consecuencia, indicadores de un rasgo individual de cada docente, rasgo que caracteriza la forma en que éste realiza una clase centrada en la lectura. Estos rasgos o características individuales permiten *distinguir* a los docentes entre ellos.

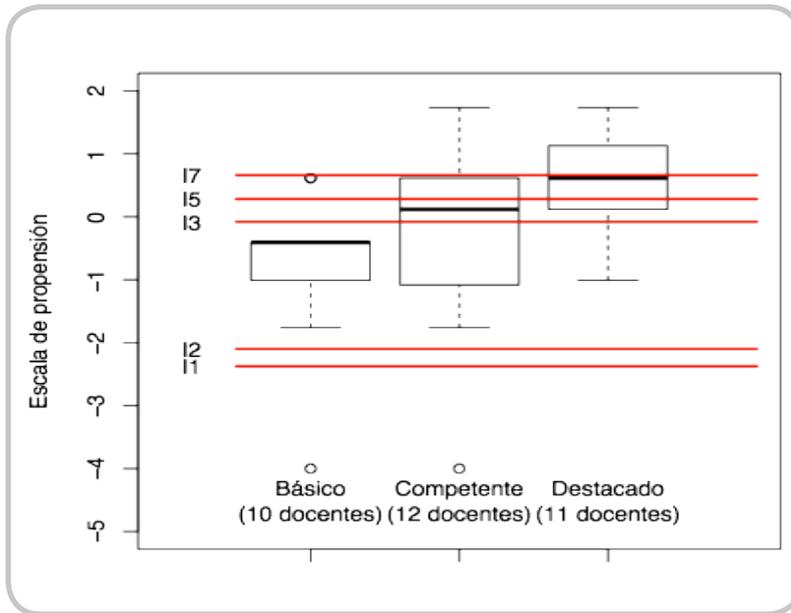
Esta distinción se lleva a cabo por medio de un modelo Rasch (San Martín, González & Tuerlinckx, 2009) que permite representar simultáneamente (esto es, en una misma escala intervalar) los rasgos individuales y los indicadores asociados a las actividades. Esta representación, ilustrada en la Figura 1, permite extraer las siguientes conclusiones: un 27% de los profesores realiza tareas de lectura con preguntas textuales, apelando a la capacidad de los niños para identificar informaciones explícitas, y además, leen textos que sus alumnos escuchan. Las actividades de lectura con preguntas que requieren de habilidades más complejas como manipular información implícita, realizar procesos inferenciales o realizar juicios de valor sobre lo leído, presentan baja frecuencia. Solo un 12% de los profesores realiza preguntas abiertas que impliquen opiniones o puntos de vista sobre textos leídos, y solo un 9% realiza preguntas abiertas que impliquen un juicio o valor.

• **Figura 1**
 • Representación Simultánea “Docente–Actividades de Lectura”



La Figura 2 muestra el cruce de las actividades antes señaladas con la Evaluación Docente de aquellos profesores que centraron su clase en actividades de lectura (N=33). Se aprecia una diferencia estadísticamente significativa (p -valor=0,014) entre los *básicos* y los *destacados*, mientras que la diferencia entre los *competentes* y los *destacados* es significativa al 8%. Se puede observar además que la mitad de los profesores evaluados como destacados, realiza actividades de lectura que requieren respuestas fundamentadas por parte de los estudiantes (indicador I7). Además, es claro que solo los profesores *destacados* realizan actividades que implican desafíos cognitivos más complejos para los niños (I3, I5, I7), y la pendiente va decayendo en la medida en que la evaluación también lo hace.

••• **Figura 2**
 ••• *Comparación de Docentes agrupados por Niveles de Evaluación Docente*



Conclusiones

Los resultados descriptivos muestran que un porcentaje importante de los profesores observados opta por centrar la enseñanza del lenguaje escrito en el eje de lectura y aborda esa enseñanza utilizando textos completos, aunque asumiendo el protagonismo de la clase; además de formular preferentemente preguntas que reclaman de los estudiantes la identificación de información explícita en el texto.

Los análisis factoriales y de diferencias individuales (obtenidas con el modelo Rasch) proporcionan pistas para diseñar el estudio cualitativo que se constituye como la siguiente fase de este estudio, mediante el uso de los datos para la selección de submuestras en función de preguntas específicas de investigación que sirvan al propósito de hacer emerger las teorías implícitas de los docentes, así como la reflexión que hacen sobre sus prácticas.

Uso del tiempo e interacciones profesores–alumnos en la sala de clases

Sergio Martinic

Facultad de Educación, Pontificia Universidad Católica de Chile.

Introducción

La asignación y uso del tiempo en la escuela es un factor importante en la calidad de los resultados educativos. Se trata de un recurso escaso cuya disponibilidad y gestión afecta la práctica y vida cotidiana de profesores, alumnos y familias. Por ello, cualquier política en relación al tiempo escolar es objeto de una amplia discusión. Así lo demostró la compleja tramitación e implementación de la ley de Jornada Escolar Completa, que se aplica en Chile desde 1997, y la discusión sobre aumento de número de horas de lenguaje y matemáticas para mejorar resultados en el Sistema de Medición de la Calidad de la Educación, SIMCE.

Las investigaciones sobre el tema no son concluyentes en cuanto al impacto que tiene el aumento del tiempo asignado en los aprendizajes. En efecto, diversos estudios demuestran solo una pequeña relación positiva entre el tiempo asignado (en cualquier forma de medición) y el logro de los estudiantes (Cotton, 1989; Slavin, 1996). Las evaluaciones realizadas de la Jornada Escolar Completa en el caso de Chile confirman esta tendencia (Bellei, 2009).

Sin embargo, esta relación cambia sustantivamente cuando se analiza el tiempo asignado desde el punto de vista del uso o del tiempo comprometido en la tarea por parte del estudiante. La investigación en este caso es clara y demuestra que el tiempo que invierte el alumno en el aprendizaje que el profesor suscita, es predictivo del rendimiento en la misma materia (Walberg, 1986; Delhaxhe, 1997; Bellei, 2009). Así entonces, el tiempo comprometido con la tarea tiene impacto en los resultados y guarda estrecha relación con la calidad de la interacción pedagógica de profesores y alumnos en la sala de clases (Aronson, Zimmerman & Carlos, 1999).

Objeto, material y método

El año 2006 iniciamos un estudio que abordó las interacciones en la sala de clases y el uso o gestión del tiempo que realizan profesores y alumnos de Segundo Ciclo de Enseñanza Básica y de Primer Ciclo de Enseñanza Media⁽¹⁾.

El estudio analizó 40 videos en los que se registran las clases presentadas por profesores de Enseñanza Básica de los sectores Lenguaje y Matemáticas. Se definieron y validaron categorías las que fueron aplicadas a su vez en un segundo estudio que consideró otras 29 clases grabadas por profesores evaluados en el año 2008⁽²⁾. Contamos así con una muestra de 69 casos y que da cuenta del mismo número de horas lectivas observadas.

Las clases son analizadas como una interacción comunicativa y se siguió un procedimiento de observación continuo, realizando marcas cada 4 segundos durante los 40 minutos promedio de duración de cada una de ellas. Se ocupó el software Videograph para la clasificación de los videos en función de las categorías definidas y previamente validadas a través de una doble codificación.

.....

- 1 S. Martinic (2006–2008) Proyecto FONDECYT N°1060715 "Representaciones sociales y gestión del tiempo en la sala de clases de Establecimientos con Jornada Escolar Completa".
- 2 S. Martinic (2008–2011) Proyecto "Interacciones Pedagógicas y Uso del Tiempo en la Sala de Clases", Centro de Estudios de Políticas y Prácticas en Educación (CEPPE), Pontificia Universidad Católica de Chile.

Principales resultados

Organización social de la clase

El tiempo de la clase fue segmentado en grandes periodos o secuencias definidas como *inicio*, *desarrollo* y *cierre*. Cada segmento se caracteriza por funciones y tareas específicas y la distinción entre una y otra se realiza a través de “marcadores” lingüísticos y/o gestuales del profesor. En los periodos de inicio y cierre distinguimos, a su vez, un momento de *pre-inicio* y otro de *pre-cierre*, los que dan cuenta de actividades particulares de preparación o de transición. Del mismo modo, al interior de la fase de *desarrollo* se distinguen secuencias de acuerdo a la unidad temática o el tipo de actividad que se desarrolla en el tratamiento de la materia.

Los profesores observados dedican, en promedio, un total de 3 minutos para el *pre-inicio* e inicio de la clase propiamente tal, lo que corresponde al 4,2% del tiempo total de clase. La fase de desarrollo ocupa 36,8 minutos representando el 92,2% del tiempo total de la clase. Llama la atención que la preparación y el cierre propiamente tal ocupen en promedio menos de 2 minutos, lo que equivale al 3,5% del tiempo de la clase. Muchos profesores no anticipan un cierre pedagógico adecuado y se observa una finalización sin transición o preparación.

Tiempo y habla en profesores y alumnos

El primer acto de organización de la clase es el turno de la palabra que, en el contexto de la clase, es regulado por el profesor. Como se observa en la Tabla 1, en el conjunto de clases observadas las intervenciones del profesor ocupan la mayor parte del tiempo (52%) y éstas se dedican, centralmente, a tareas instruccionales. Por lo general, el docente dirige su habla a la clase como un todo (más del 45%, tanto en Lenguaje como en Matemáticas), dedicándose a exponer materias o explicar procedimientos (36%).

En segundo lugar, destaca el tiempo clasificado como silencio (10,6%) y que corresponde, principalmente, a la realización de actividades por parte de los alumnos (*seat work*), quienes responden a instrucciones o enmarcamiento realizado por el profesor. Por último, en un 11,6% del tiempo de la clase, se observa a los alumnos realizando intervenciones que responden, principalmente, a preguntas e instrucciones del docente. Asimismo, se observa que en las clases de Lenguaje los alumnos tienen más tiempo de intervención verbal; en Matemáticas, en cambio, hay una mayor proporción de tiempo en silencio, con los alumnos realizando tareas y ejercicios enmarcados por el profesor. Sin embargo, estas diferencias no son estadísticamente significativas.

• **Tabla 1**
 • *Tiempo de habla según actor observado.*

Actor	Materia	Media	D.S.	Total	
				Media	D.S.
profesor	Lenguaje	49,754	13,1150	52,196	13,6299
	Matemática	54,569	13,8868		
un alumno	Lenguaje	12,430	8,9675	11,590	7,8976
	Matemática	10,774	6,7306		
varios alumnos	Lenguaje	21,087	19,6395	19,316	17,9370
	Matemática	17,596	16,2149		
ruido	Lenguaje	5,171	11,6101	5,621	13,4347
	Matemática	6,058	15,1576		
silencio	Lenguaje	10,668	15,7016	10,554	14,4327
	Matemática	10,442	13,3150		
medios audiovisuales	Lenguaje	,890	3,0243	,723	2,8805
	Matemática	,561	2,7680		

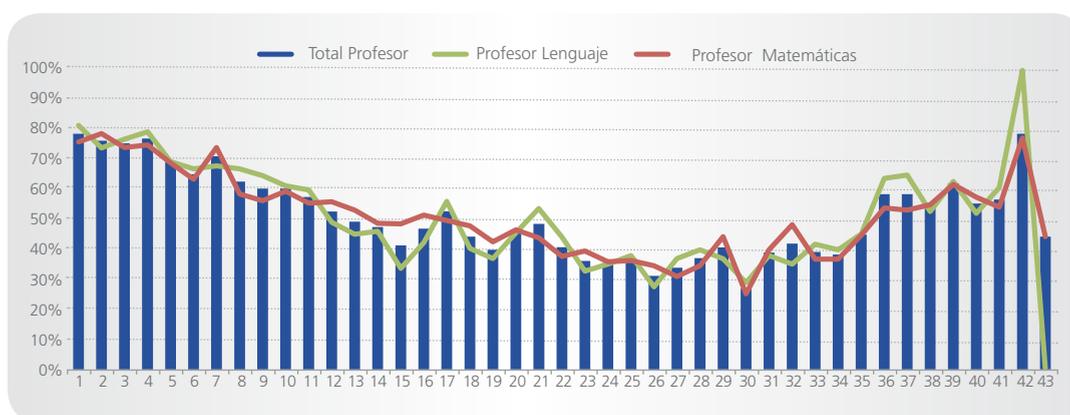
n=69

En términos generales se constata un estilo de enseñanza en el que predomina una pedagogía directiva con poco diálogo y conexión con los estudiantes tal como ha sido demostrado en otros estudios (Martinić & Vergara, 2007).

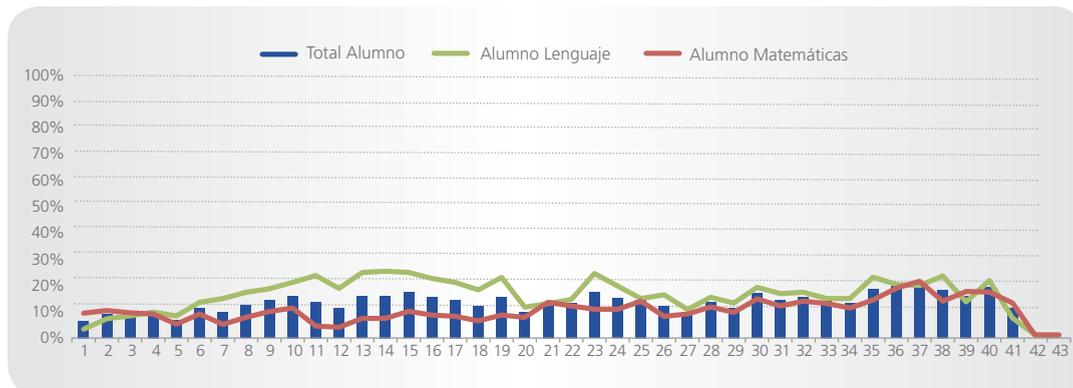
Cambios en el transcurso de la clase

La participación y el protagonismo que tienen alumnos y profesores cambia a medida que transcurre la clase. En los gráficos 1 y 2 se puede observar el tiempo promedio que ocupan las intervenciones de cada uno, como hablante principal, a lo largo de la clase.

• **Gráfico 1**
 • *Tiempo de habla de Profesor en Lenguaje y Matemáticas (en %)*



• **Gráfico 2**
 • *Tiempo de habla de un Alumno en Lenguaje y Matemáticas (en %)*



Como se aprecia en los gráficos 1 y 2, los profesores son el hablante principal durante la mayor parte de los primeros siete minutos de la clase, bajando esta proporción en forma importante entre los minutos 12 y 18, para aumentar nuevamente en los últimos cinco, en la fase de preparación y de cierre propiamente tal. Se observan interesantes diferencias entre los profesores de Lenguaje y de Matemáticas. En promedio las intervenciones de estos últimos ocupan más tiempo que las de sus pares de Lenguaje. Solo en el minuto 27 de la clase los docentes de Matemáticas disminuyen sus intervenciones para que aumenten de los alumnos, recuperando nuevamente la palabra en el minuto 31 con bastante protagonismo.

Este perfil del tiempo que ocupa el profesor se distingue claramente de la intervención de los alumnos, quienes adquieren mayor protagonismo en el desarrollo de la clase (minutos 21 a 23) y en la fase de preparación del cierre (minutos 34 a 37). También se observa que la participación de alumnos tiene más importancia, en términos de uso del tiempo, en las clases de Lenguaje que en las de Matemáticas.

Tipo de intervención de profesores y alumnos

La mayor parte del tiempo se concentra en instrucciones o información relacionada con procedimientos a seguir. Ello es particularmente importante en el caso de los docentes de Matemáticas, en que un 30,3% del tiempo de la clase es ocupada por el habla del profesor, versus el 26,7% en el caso de sus pares de Lenguaje. La exposición de contenidos de la disciplina propiamente tal ocupa cerca del 14% del tiempo de habla de los profesores de Lenguaje y de Matemáticas.

En cuanto a las intervenciones de los alumnos, se observa que la mayor parte del tiempo de éstos se ocupa en realizar tareas o trabajar en función de actividades instruidas por el profesor. El 15,2% del tiempo de las intervenciones de los alumnos en Lenguaje corresponde a respuestas a preguntas del profesor. En Matemáticas este fin corresponde al 13,4% del tiempo de sus intervenciones.

Conclusiones y discusión

Las clases observadas son altamente estructuradas y los profesores tienen un alto protagonismo y concentración del tiempo. Los alumnos, en general, tienen un rol pasivo con intervenciones que ocupan poco tiempo en el conjunto de la clase. Se observa una gestión efectiva del tiempo por parte del profesor, predominando el tiempo instructivo: los alumnos atienden y ocupan parte importante del tiempo realizando las tareas que instruye el profesor.

Se constata que los alumnos tienen poca participación e intervenciones que den cuenta de un involucramiento activo. En los intercambios predomina el habla del profesor y la estructura de diálogo descansa en las preguntas que éste realiza a los alumnos. Sin embargo, tanto en Lenguaje como en Matemáticas las preguntas producen respuestas con baja elaboración: los alumnos se limitan a repetir o aprobar afirmaciones y contenidos enunciados por el profesor. Una estructura similar se observa en Mares, Guevara, Rueda, Rivas & Rocha (2004) y Preiss (2009).

Esta estructura de diálogo es problemática si se asume que el aprendizaje es un hecho interactivo y que el lenguaje es la herramienta de apropiación del conocimiento y desarrollo cognitivo. Por otra parte, el tiempo de la clase y su estructuración puede limitar la participación de los alumnos y con ello la motivación y el compromiso del estudiante.

Patrones instruccionales en Chile: La evidencia de la Evaluación Docente

David Preiss

Escuela de Psicología, Pontificia Universidad Católica de Chile.

Usando la evidencia audiovisual recogida por el Sistema de Evaluación del Desempeño Profesional Docente (en adelante Evaluación Docente), he intentado elucidar si las escuelas públicas en Chile cuentan con un patrón instruccional identificable en la enseñanza de Matemáticas y de Lenguaje en el Primer y Segundo Ciclo de Enseñanza Básica, y si este patrón es compatible con los criterios de buena enseñanza avanzados por las ciencias del aprendizaje.

Un patrón instruccional se identifica por la existencia de un alto grado de semejanza en clases realizadas por distintos docentes en diferentes escuelas de un país. Más específicamente, un patrón instruccional implica que estas clases tienen ciertos rasgos recurrentes comunes, y una organización similar de esos rasgos a lo largo de la clase. Las semejanzas se hacen evidentes a partir de la observación de los procesos instruccionales en el aula, y, por efecto de contraste con otros países, indican cual es la firma de una cultura en el aula (Givvin, Hiebert, Jacobs, Hollingsworth & Gallimore, 2005; Preiss, 2009, 2010; Radovic & Preiss, 2010; Stigler & Hiebert, 1999). Presumiblemente, estos patrones estarían basados en ideales relativos al proceso instruccional y serían específicos para la cultura de un país.

La hipótesis sobre la existencia de patrones de enseñanza nacionales tomó fuerza a partir de las videoencuestas asociadas a los estudios TIMSS 1995 y TIMSS-R 1999 (Hiebert et al., 2003; Stigler, Gonzales, Kawanaka, Knoll & Serrano, 1999). Tal como sugiere Manzi (2007a), la Evaluación Docente ofrece una oportunidad única para responder esta pregunta en Chile, en virtud de sus características:

- a. Este programa genera una base de información documental y audiovisual acerca de las prácticas de aula sin precedentes en nuestro país,
- b. la información proporcionada por los profesores cumple con los requerimientos éticos propios de la investigación educacional dado que todos los participantes firman voluntariamente una forma de consentimiento informado,
- c. la información de la evaluación, aunque focalizada en el sector municipal, tiene representatividad nacional.

Para responder esta pregunta, he concentrado nuestra atención en dos dimensiones de la sala de clase: estructura y discurso.

En el ámbito de **estructura**, he publicado resultados sobre la enseñanza de Lenguaje en Primer Ciclo Básico y la enseñanza de Matemáticas en Segundo Ciclo Básico. El primer estudio que realizamos (Preiss, 2005, 2009) se focalizó en la enseñanza del Lenguaje en Primer Ciclo y mostró que, independientemente del nivel de competencia de los docentes, las clases observadas tenían una estructura similar: la mayor parte del tiempo, la atención pública estaba focalizada en el profesor, cuando los estudiantes ejecutaban trabajo independiente éste era individual y no grupal (aun cuando el arreglo físico de los pupitres favorecía el desarrollo de actividades grupales), y el tipo de contenido favorecido en clases era la práctica de una habilidad lingüística por sobre el ejercicio de actividades de lectura o escritura completa.

En el mismo estudio identifiqué un tipo de actividad que parece ser propio de las clases chilenas y que denominé *juego instruccional*. En otro estudio sobre estructura realizado en una muestra de clases de Matemáticas de Segundo Ciclo (Preiss, 2010), mostré que el tipo de resolución de problemas favorecido en estas clases se basa en la práctica repetida, guiada por el profesor, de problemas concurrentes. Tal como en Estados Unidos, los profesores chilenos presentaban a los estudiantes definiciones y procedimientos

que ellos debían practicar posteriormente. Por otro lado, tal como observamos en Lenguaje, la mayor parte del trabajo independiente observado involucraba trabajo individual en vez de trabajo grupal. Tomando como referencia la nomenclatura utilizada para describir las clases norteamericanas, describí estas clases como la *apropiación privada de términos y procedimientos*. Un estudio reciente ha reproducido estos hallazgos para la enseñanza de Matemáticas de Primer Ciclo y ha mostrado que el patrón observado es independiente del género, edad y competencia de los Docentes (Preiss et al, 2011).

En el ámbito de **discurso**, el estudio ya citado sobre la enseñanza de Lenguaje mostró que, independientemente del nivel de competencia de los docentes, las preguntas y seguimientos de los profesores a las intervenciones de los estudiantes se orientaban a la regulación de la clase o al chequeo de información. Más específicamente, encontré que, durante los diez primeros minutos de la clase, los docentes producían en promedio, menos de 3 preguntas que indujesen a sus estudiantes a elaborar contenido, y menos de 3 *seguimientos* que reformularan metacognitivamente las contribuciones de sus alumnos. Pese a este formato común, reporté una correlación pequeña, pero directa entre la evaluación de desempeño del profesor y la cantidad de preguntas y seguimientos que tenían mayor probabilidad de comprometer a los estudiantes en un pensamiento metacognitivo.

En otro estudio similar, Radovic y Preiss (2010) estudiaron clases de Matemáticas de Segundo Ciclo Básico de escuelas públicas filmadas en Chile⁽¹⁾, y en base al análisis descriptivo de las frecuencias con la que los profesores utilizaban distintos tipos de preguntas y seguimientos y a la asociación de estas con actividades de la clase y desempeño del profesor, identificaron un patrón de interacción relativamente compartido, caracterizado por preguntas cerradas, escasa participación de los estudiantes y seguimientos de bajo potencial metacognitivo. Se evidenció que, al ejecutar actividades de práctica de contenidos matemáticos, los profesores produjeron más preguntas de aplicación y que los profesores mejor evaluados produjeron mayor cantidad de preguntas abiertas y sobre experiencias personales.

Aun cuando la evidencia audiovisual de estos estudios ha sido obtenida en un contexto de evaluación, es altamente probable que en sus elementos substantivos estas clases reflejen patrones culturales. El proceso instruccional es altamente estructurado y los profesores suelen reposar en sus modos habituales de enseñar para ejecutar su actividad, aun cuando esta actividad esté interferida por la observación de externos o por el uso de una cámara de video (Stigler, Gallimore & Hiebert, 2000).

Los resultados que hasta aquí hemos encontrado sugieren que tanto como una buena base disciplinaria, nuestros profesores necesitan una buena base de conocimiento pedagógico del contenido (Shulman, 1986, 1987). Del mismo modo, el desarrollo de estándares para la formación inicial de profesores y los Planes de Superación Profesional (programa de apoyo al desarrollo profesional de los docentes que obtienen bajos resultados en la evaluación)⁽²⁾, deben atender no solo a los contenidos, sino que integrar a éstos tanto el aspecto metacognitivo-individual como el aspecto dialógico grupal del proceso de enseñanza aprendizaje. Los estándares y planes de perfeccionamiento no deben estar limitados a qué contenidos deben enseñar los profesores, sino integrar esos contenidos con los modos en que serán presentados en el aula, integrados en el desarrollo cognitivo individual de los estudiantes y transferidos al diálogo de pares entre ellos.

Futuras iniciativas de investigación debieran considerar aquellos profesores, en ejercicio, que muestran un patrón instruccional distinto y pueden constituir un modelo de cambio para sus pares, especialmente los más jóvenes. Sería conveniente, por lo mismo, reconocer e identificar profesores que se apartan de la media, estudiarlos en profundidad y analizar cuan diseminables son sus prácticas. El modelo de “lesson study” puede ser adecuado para desarrollar iniciativas de esta naturaleza (Fernández & Yoshida, 2004).

.....

1 Estas grabaciones de clases también provenían de la evidencia recogida por la Evaluación Docente.

2 Para más información sobre este programa, ver Cortés & Lagos en este volumen.



9

Comentarios

Felipe Martínez⁽¹⁾ · Pedro Ravela⁽²⁾
Gilbert Valverde⁽³⁾

(1) University of California, Los Angeles.

(2) Universidad Católica del Uruguay.

(3) State University of New York, Albany.

La evaluación docente en Chile: Perspectivas sobre validez

José Felipe Martínez

Graduate School of Education and Information Studies
University of California - Los Angeles

Introducción

A principios de la década del 2000 comenzó a desarrollarse en Chile el Sistema de Evaluación del Desempeño Profesional Docente (en adelante Evaluación Docente), aprovechando experiencias anteriores en países como México y Cuba, y otras en regiones y distritos escolares en los Estados Unidos y otros países. Después de una década de operación, el sistema chileno se ha convertido en una referencia en el medio internacional, en un contexto de interés creciente en la implementación de programas de este tipo. Los capítulos anteriores de este libro dan muestra de la complejidad de la Evaluación Docente, la fuerza conceptual que guió su desarrollo, y la sofisticación técnica que implica su operación y posterior validación.

Mi intención en este espacio no es abordar en detalle los aspectos sustantivos y metodológicos de un sistema de medición de esta complejidad, sino realzar algunos que considero pueden ser particularmente relevantes, dado lo que el sistema chileno ya muestra acerca de las fortalezas y retos de la evaluación docente en general, y de las características técnicas de algunos de sus componentes en particular. En el primer apartado se presentan como referencia la experiencia mexicana y algunos ejemplos recientes en los Estados Unidos. La siguiente sección examina los principales problemas conceptuales y metodológicos que enfrenta la evaluación docente, utilizando el sistema chileno y los resultados presentados en este volumen como referencia. Esta sección presta atención particular a dos temas que pueden adquirir mayor importancia en el futuro, a saber, la integración de indicadores que capturan aspectos diversos del quehacer docente en la evaluación y, en particular, el creciente interés que despierta el uso de pruebas estandarizadas de aprendizaje de los alumnos para evaluar a los maestros. Finalmente se abordan algunas consideraciones sobre los retos de la evaluación docente en el plano internacional.

Contexto internacional de la evaluación docente

México

A principios de la década de 1990 se implementó en México el programa *Carrera Magisterial*, uno de los primeros sistemas nacionales de evaluación y estímulos docentes implementados a nivel mundial. El programa se diseñó para evaluar la labor docente con base en seis factores ponderados diferencialmente: el grado de estudios (15%), antigüedad (10%) y actualización docente (17%), desempeño profesional (equivalente al informe de terceros en Chile con un peso de 10%), preparación profesional (examen de conocimientos, 28%), y finalmente el aprovechamiento escolar de los alumnos (20%). Los factores se norman en cada estado por lo que el programa no evalúa a los maestros en base a puntajes mínimos satisfactorios, sino en relación al grupo participante en cada uno.

En 2004, a quince años de la creación de Carrera Magisterial, la RAND Corporation fue comisionada para realizar una evaluación comprensiva del programa, misma que reportó una serie de problemas de medición y diseño que limitan considerablemente su efectividad (Santibáñez et. al. 2007). Entre los hallazgos destaca el que las pruebas de maestros y alumnos, aunque de amplia cobertura en términos curriculares, se enfocan primordialmente a niveles muy bajos de demanda cognitiva (i.e., memorización

y procedimientos). El sistema de construcción es robusto en el caso de las pruebas del docente, pero no cumple estándares mínimos en el caso de las pruebas de alumnos, lo que no sorprende si se considera el volumen de la muestra de sujetos y reactivos en cada caso⁽¹⁾.

La medición de las prácticas docentes en aula también evidencia limitaciones serias; éstas se miden mediante un procedimiento de evaluación colegiada que resulta en un pronunciado efecto de *techo* que limita grandemente su utilidad como medida sumativa (la mediana es 9,1 puntos de un máximo de 10). Igualmente notable es la ausencia de un vínculo directo de la medida a un modelo explícito de práctica docente que diera valor formativo a la información obtenida (lo que contrasta con el caso chileno en que el sistema de evaluación está anclado en el Marco para la Buena Enseñanza).

En suma, la evaluación de RAND cuestiona la capacidad de Carrera Magisterial de identificar a los *mejores maestros*, y justifica en cierta medida el escepticismo entre los docentes que dudan que un mayor esfuerzo de su parte se traducirá necesariamente en mejores puntajes y acceso a incentivos. Por lo anterior, no es sorprendente que los resultados indiquen un impacto mínimo del programa sobre el rendimiento de los alumnos, mismo que además desaparece una vez que los docentes son incorporados al sistema⁽²⁾. Una constante que resulta particularmente preocupante en un sistema de esta envergadura es la ausencia de estudios de validez que justifiquen la interpretación de los puntajes de los indicadores usados en Carrera Magisterial, y su combinación para formar puntajes promedio. Tampoco se conocen estudios que investiguen sesgos e inequidades aunque la experiencia mexicana sugiere que estos existen en la práctica (Santibáñez & Martínez, 2010).

Ejemplos recientes en Estados Unidos

La evaluación docente es sin duda el principal tema de política educativa en la Unión Americana en los últimos cinco años, como lo demuestran las provisiones para la evaluación de la calidad docente en la nueva iniciativa federal *Race to the Top* (US Department of Education, 2010), y la implementación reciente o en curso de sistemas con ese propósito en ciudades como Denver (Professional Compensation Program, ProComp), Washington D.C. (DCPS Effectiveness Assessment System for School-Based Personnel), Chicago (Teacher Advancement Program), y Los Angeles (Supporting all employees).

El proceso que resultó en el desarrollo de un sistema de este tipo en el distrito escolar de Los Angeles (LAUSD), es interesante por lo que revela acerca de la problemática y los beneficios y peligros de la evaluación docente de alto impacto. Hacia finales de la década pasada este distrito ya estudiaba activamente la implementación de un sistema para evaluar a sus maestros. Sin embargo, en agosto de 2010, el diario Los Angeles Times difundió un estudio sobre la efectividad de los docentes basado en los puntajes de los alumnos de primarias públicas, lo que causó controversia al identificarse allí con nombre, apellido (y, en algunos casos particularmente dramáticos, fotografía y nota periodística) a los maestros, junto con indicarse la efectividad estimada de cada uno. Lo anterior precipitó la actividad al interior del distrito, que en 2011–12 estrenará un sistema piloto que en teoría se refinaría para su implementación definitiva en 2013. Para desarrollar este sistema se está siguiendo un proceso amplio que se alimenta de otras experiencias recientes en Estados Unidos y ejemplos a nivel internacional como el chileno. El modelo de evaluación que se busca se basa en indicadores múltiples, con fuerte anclaje sustantivo en un *modelo de enseñanza–aprendizaje* explícito y detallado que le confiere un carácter principalmente formativo⁽³⁾. Los principales indicadores de evaluación se obtendrían mediante observación en aula, con la posibilidad

-
- 1 A partir de 2005 se utiliza la nueva prueba nacional de aprovechamiento de los alumnos (ENLACE) si hay una disponible, pero se desarrollan aún pruebas especiales en grados y materias que no cubre la prueba nacional.
 - 2 A pesar de las limitaciones técnicas detalladas en la evaluación de RAND, es importante notar que el hallazgo de efectos mínimos o nulos coincide con el de otros estudios, y que los expertos en general son cautos en cuanto al efecto de los incentivos en el desempeño docente (Springer et al, 2010; Dee & Keys, 2004).
 - 3 El modelo utilizó como base al conocido *Marco de Enseñanza* de Danielson (Danielson, 1996) adaptado para su uso local mediante una consulta amplia que incluyó maestros, administradores, expertos y otros actores.

de incorporar en el futuro encuestas a alumnos y directores; también se contempla incluir el aprendizaje de los alumnos como parte de la evaluación mediante el uso de modelos de valor agregado, pero con un peso menor y solo en forma complementaria a la información proporcionada por los otros indicadores.

Otra experiencia importante es el estudio MET (*measures of effective teaching*)⁽⁴⁾, el mayor a la fecha en investigar los problemas relativos a identificar y medir la efectividad. Desde 2009 y hasta 2012 el estudio recolectará una amplia variedad de medidas de efectividad docente en miles de aulas en la Unión Americana; éstas incluyen observaciones en aula, pruebas de conocimientos pedagógicos, y encuestas de maestros y alumnos, así como el aprendizaje de los alumnos, estimado mediante modelos de valor agregado. Por su escala y el cuidado técnico con que se desarrolla, este estudio es uno de los más importantes realizados hasta la fecha y seguramente ejercerá gran influencia en el área de efectividad docente en los próximos años. Sin embargo, el mismo ejemplifica también los riesgos de una concepción de la evaluación excesivamente centrada en el uso de pruebas estandarizadas del alumno. Los investigadores proponen un *modelo de validación* que, aunque a primera vista podría parecer amplio, al aceptar un gran rango de indicadores de enseñanza basados en protocolos de observación diversos, en realidad resulta limitado, puesto que reduce el problema a determinar si “el protocolo de observación rutinariamente asigna la notas más altas a los maestros cuyos estudiantes logran el mayor crecimiento en términos de aprovechamiento” (MET, 2010a, 2010b). Es decir, la validez de un indicador de desempeño docente se define explícitamente en función última de su correlación con el aprendizaje del alumno. Este modelo de validación parece innecesariamente restrictivo y en todo caso no coincide con el concepto de validez ampliamente aceptado basado en fuentes diversas de evidencia (Kane, 2006; AERA, APA & NCME, 1999).

Consideraciones metodológicas

Confiabilidad y Validez

La evaluación docente involucra un ejercicio de medición particularmente complejo y multidimensional, sin embargo, existen para él estándares profesionales y criterios técnicos claramente delineados y generalmente aceptados por la comunidad internacional (AERA, APA & NCME, 1999). La evaluación de profesores es un ejemplo prototípico de medición de alto impacto que, por ello, se rige por criterios psicométricos referidos a la confiabilidad de las medidas y la validez de las inferencias realizadas a partir de ellas. En la práctica, sin embargo, el trabajo técnico suele entenderse como limitado al estudio de las propiedades estadísticas de las pruebas (estabilidad, confiabilidad, error estándar, dimensionalidad) y la información que genera, en el mejor de los casos, se considera como elemento secundario en la revisión de los resultados sustantivos del programa. Esto contrasta con la acepción más amplia y moderna de validez (Kane, 2006) desde la que resulta claro que la tarea técnica no solo implica ofrecer evidencia de que se ha identificado a los mejores maestros, sino que también exige proveer una base empírica que ayude a explicar lo que significa ser un *mejor maestro*.

En la Evaluación Docente chilena parece entenderse la validez en este sentido amplio, como indican los múltiples estudios que han investigado aspectos técnicos y conceptuales referentes a la confiabilidad de los indicadores y la validez de las inferencias últimas⁽⁵⁾. Lo mismo sugiere la conceptualización sofisticada que se evidencia en el estudio de las consecuencias de la evaluación. Esta merece mención especial al haber sido objeto de un extenso debate en los últimos años: mientras algunos autores proponen integrar las consecuencias del uso de la evaluación directamente al modelo de validez como un elemento más de evidencia (Shepard, 2005), para otros ello representa una complicación conceptual innecesaria y de

.....
4 Este proyecto es una iniciativa de la fundación Gates con la colaboración de universidades e instituciones como las universidades de Harvard y Stanford y la corporación RAND.

5 Ver Taut, Santelices & Manzi, en este volumen.

dudosa utilidad práctica (Mehrens, 2005). Sin importar el modelo conceptual que se prefiera, es claro que las consecuencias de la evaluación deben investigarse y considerarse cuidadosamente, particularmente cuando implican decisiones de alto impacto (AERA, APA & NCME; 1999) para prevenir, y en su caso detectar, posibles efectos perniciosos del programa de evaluación.

Sin olvidar lo anterior, la validez de un sistema de evaluación docente se enfrenta fundamentalmente al reto que representa la multidimensionalidad intrínseca de un constructo que requiere combinar múltiples indicadores en un marco único y coherente que facilite la toma de decisiones. Independientemente de los desafíos técnicos y conceptuales que representa la medición de los productos de una tarea tan compleja como la educativa por separado, la verdadera frontera metodológica a que nos enfrentamos es la que representa combinar estos índices que capturan aspectos disímiles de un constructo global que denominamos *efectividad o calidad docente*. La medición adquiere así su verdadera dimensión práctica y política, ya que, sin importar la sofisticación conceptual y técnica de los modelos estadísticos utilizados, éstos se concretizan en juicios de valor de naturaleza más concreta, unidimensional y, si se quiere, coloquial. A saber, al clasificar el desempeño de un docente como *destacado (o insatisfactorio)* transmitimos el juicio de valor central que pretende el sistema, pero las propiedades técnicas de esta clasificación no son tan bien conocidas como las de sus componentes, ni se siguen de forma directa o predecible de éstos—es difícil derivar un error estándar que capture directamente la incertidumbre estadística que existe al clasificar a un docente en un nivel de desempeño bajo este sistema de indicadores múltiples.

Es revelador notar que una buena parte del trabajo metodológico en estos temas se origina en décadas anteriores y con frecuencia presenta una complejidad técnica menor comparada con los modelos estadísticos y psicométricos que utilizan muchos sistemas de evaluación modernos. Ello por supuesto no refleja carencias metodológicas en el trabajo, sino más bien la naturaleza misma de un problema que es solo parcialmente estadístico e involucra también consideraciones no estadísticas acerca de la importancia y apropiada combinación de los diversos indicadores. Mehrens (1989) ofrece una perspectiva metodológica general para la combinación de indicadores por métodos *clínicos* o *estadísticos*. Este autor sugiere que siempre es preferible la combinación estadística de indicadores, pero agrega un matiz importante: cuando existe un criterio cuantitativo único, varios métodos de regresión permiten obtener combinaciones de indicadores de máximo poder predictivo. Sin embargo, cuando no existe una medida criterio necesariamente se deberán usar otros criterios de tipo *lógico* para combinar y valorar los indicadores (y en especial para decidir de qué manera “los excesos en un indicador pueden servir para compensar deficiencias en otro”, Mehrens, 1989 p.332). Mehrens incluso enfatiza que se debe considerar la posibilidad de simplemente no combinar los indicadores cuando el propósito fundamental de la evaluación es generar un efecto formativo.

Ponderación de indicadores

Relacionado con lo anterior, la asignación de pesos a los diversos indicadores involucrados inevitablemente incluirá consideraciones políticas y prácticas, pero debe también considerar la evidencia técnica disponible y sobre todo reflejar el modelo conceptual en que se basa el sistema.

El Portafolio incluido en la Evaluación Docente chilena ofrece un ejemplo interesante porque el peso asignado a éste se corresponde con el cuidadoso proceso que se sigue en el diseño e implementación, tanto en la elaboración del instrumento como del sistema de calificación que lo acompaña, y el estudio riguroso de sus propiedades psicométricas. El Informe de Referencia de Terceros ofrece otro ejemplo interesante porque su menor peso parecería justificado dadas las limitaciones técnicas que ha evidenciado este instrumento en años anteriores. En capítulos previos se ha mencionado que la evidencia recogida por la propia evaluación se utilizó para realizar cambios y mejoras a este instrumento⁽⁶⁾; entre ellos, los cambios realizados en 2010 resultan particularmente interesantes porque, al requerir una justificación

.....

6 Ver Flotts y Abarzúa, en este volumen.

escrita al evaluador cuando asigna el nivel Destacado, reafirma un modelo en que el director tiene un papel importante como fuente de información para la evaluación y para la retroalimentación al docente⁷⁾. Lo anterior no solo se justifica empíricamente (el funcionamiento inadecuado de un instrumento no implica una desvalorización de la información que éste busca aportar), sino que evidencia la fortaleza conceptual al interior del sistema. Si los cambios realizados logran mejorar la calidad de la información —como sugieren los resultados preliminares⁸⁾— en principio podría considerarse asignar un peso mayor a este instrumento para reflejar su importancia dentro del marco conceptual⁹⁾.

Caso similar es la pauta de autoevaluación del maestro. Como en el caso de México y otros sistemas norteamericanos, este instrumento enfrenta en Chile un efecto de techo en el que la inmensa mayoría de docentes obtiene los puntajes más altos¹⁰⁾, resultando en información de dudosa calidad y uso muy limitado. Frente a ello, cabría considerar una variedad de alternativas que van desde incorporar la autoevaluación en la *reflexión sobre el quehacer docente* que ya incluye el Portafolio (cuyo peso es de un 60% del resultado final), hasta hacer cambios y adaptaciones para mejorar el instrumento actual—por ejemplo para alejarlo de la autoevaluación directa asociada a los niveles de desempeño usados por el sistema de evaluación, enfocándolo en su lugar al reporte de prácticas, actitudes y percepciones relacionadas con el desempeño y efectividad del docente. Aunque estos dos extremos representan opciones de cambio operativamente muy distintas entre sí, como se notó anteriormente tienen en común el que ninguna abandona la idea de utilizar el proceso de autoevaluación como una herramienta valiosa para la mejora educativa.

Evaluación docente basada en pruebas estandarizadas

Las graves distorsiones e inequidades que se pueden generar al comparar docentes (o escuelas) en base a puntajes agregados brutos, han limitado su uso para propósitos evaluativos en el pasado. La extensión del uso de pruebas estandarizadas y el auge reciente de modelos como los de *valor agregado* (VAM)¹¹⁾ representan opciones interesantes, pero también encierran riesgos si no son entendidos adecuadamente. La lógica coloquial de sistemas de evaluación docente empleando modelos de valor agregado es sencilla y poderosa: al enfocarse este método en el progreso del estudiante de un año a otro, se eliminan los problemas de inequidad asociados al peso de la composición socioeconómica del aula, puesto que cada alumno ejerce como su propio control¹²⁾.

A pesar de esta simpleza aparente, los riesgos que supone la evaluación docente basada en los puntajes del alumno (aún en la versión sofisticada que ofrecen los modelos de valor agregado) son bien conocidos y se detallan en múltiples estudios y en reportes firmados por paneles que agrupan a los investigadores más distinguidos en la materia (e.g. Baker et. al, 2010). Entre los problemas más serios están: la imposibilidad de asignar interpretaciones causales a los efectos estimados (Rubin, Stuart & Zanutto, 2004); la posibilidad de inequidades, ya sea que se utilicen o no ajustes para tomar en cuenta las características de los alumnos (Briggs & Domingue, 2011); la imprecisión estadística que impide diferenciar a la mayoría de docentes evaluados entre sí (Harris & McCaffrey, 2010); y la inestabilidad de los estimadores que pueden variar significativamente con el tiempo (Goldhaber & Hansen, 2008) o dependiendo de la medida utilizada (incluso tratándose de sub-puntajes en una misma prueba; ver Lockwood et. al. 2007). Irónicamente, la

.....

7 Esto contrasta con el caso de Estados Unidos donde con frecuencia se considera al director como fuente inferior de información con nula base conceptual y evidencia empírica limitada a sesgos en la distribución y correlaciones débiles con la prueba estandarizada de turno.

8 Ver Sun, Correa, Zapata & Carrasco, en este volumen.

9 Modificar el peso de este instrumento no implica un cambio paralelo a las atribuciones de decisión operativa y de tipo laboral con que se faculta al director; ambos, indicador y decisiones existen en universos de validez independientes. Por otro lado, un cambio sí podría influenciar la proporción de docentes que puede acceder al programa de incentivos AVDI.

10 Ver Flotts y Abarzúa, en este volumen.

11 Modelos de Valor Agregado en su sigla en inglés (Value-added models).

12 El poder discursivo de estos modelos se evidencian en la rápida incorporación de que han sido objeto en el lenguaje político: junto al maleable concepto del bienestar del alumno los modelos VAM son utilizados por actores que enarbolan propuestas diametralmente opuestas.

mención repetida de argumentos sólidos, pero de alta complejidad técnica tiende a generar una cierta insensibilización e incluso escepticismo, particularmente en círculos periodísticos y políticos, en los que el efecto “*si no lo entiendo no puede ser tan importante*” adquiere particular fuerza¹³.

Más aún debe notarse que, aun dejando de lado el debate teórico–estadístico, la operación de un sistema de evaluación basado en los resultados de los estudiantes en pruebas estandarizadas enfrenta limitaciones prácticas importantes. El ejemplo reciente en Los Angeles parece relevante: en el debate nacional que siguió a la publicación de estimadores de valor agregado por el diario Los Angeles Times se abordaron muchos de los puntos técnicos y políticos antes mencionados, pero se perdió de vista el hecho fundamental de que una tercera parte de los docentes en el distrito (más de seis mil) no fueron evaluados al no contarse con datos suficientes en la base de datos. A nivel nacional se estima que un modelo de evaluación docente de este tipo alcanzaría como máximo al 50% de los docentes en el mediano plazo, y una proporción menor en secundaria (Baker, et. al. 2010), debido a la dificultad y costo de implementar pruebas censales anuales en todos los grados y materias, además de muestras insuficientes o inestables, transferencias laborales, y otros factores. El excluir una proporción importante de docentes de la evaluación podría poner en entredicho la lógica misma del modelo de mejora.

Por último, la evidencia empírica a nivel internacional no sugiere de ninguna manera que la evaluación de maestros en base a pruebas estandarizadas constituya una solución *milagrosa* o siquiera particularmente prometedora a los problemas de desigualdad y calidad educativas. Como ocurre con muchas políticas basadas en la “lógica de la bala de plata”, la evidencia revela resultados inferiores a los deseados en cuanto a la medida objetivo y efectos contraproducentes en otros sentidos importantes. Lo anterior puede resultar desalentador, pero no debería sorprender si se considera que este tipo de políticas suelen combinar expectativas exageradas y una visión simplista de problemáticas mucho más complejas. Como establece la “Ley de Campbell”, la posibilidad de consecuencias negativas no deseadas en política social se incrementa en la medida en que se asigna peso primordial a indicadores susceptibles de ser distorsionados o incluso corrompidos (e.g. Nichols & Berliner, 2005). Es claro que la rendición de cuentas centrada en pruebas del alumno puede generar distorsiones de este tipo, que van desde el reduccionismo excesivo en cuanto a objetivos y prácticas docentes, hasta la *inflación* de puntajes como resultado de enseñanza orientada a las pruebas, o incluso fraude. Por lo anterior la literatura especializada es enfática y consistente en cuanto a no recomendar el uso de pruebas aplicadas a los alumnos como factor único o incluso dominante en la evaluación de docentes (Braun, 2005; Baker et. al. 2010); en la práctica los resultados de los alumnos tienen poco peso en decisiones de alto impacto que involucran al maestro, aun en países que cuentan con sistemas de pruebas de gran cobertura.

El uso de pruebas estandarizadas puede constituir una herramienta valiosa para la mejora educativa en la medida en que se valore y combine adecuadamente la información recolectada y se le den usos sumativos o formativos apropiados—es decir, en la medida en que las inferencias derivadas sean *válidas*. Aunque es posible desarrollar pruebas que permitan usos tanto sumativos como formativos, ello requiere de recursos adicionales, concesiones en el diseño, o ambos; en general, una prueba no puede ofrecer el mismo poder diagnóstico a nivel del alumno, el docente, la escuela y el municipio o estado. Por ello, al plantear la incorporación de resultados de los alumnos a un modelo de evaluación docente, se debe tener claro que el foco y nivel de medición e inferencia definen los parámetros de validez relevantes. Una prueba diseñada para ofrecer un diagnóstico del aprendizaje de los alumnos a nivel regional, difícilmente podrá ofrecer también poder diagnóstico para identificar las prácticas y factores docentes de mayor efectividad; pretender lo contrario resultará predeciblemente en que ninguno de los propósitos se cumplan a cabalidad, además de ser cuestionable en un contexto en el que existen medidas más ricas para determinar el desempeño en aula—como el portafolio docente o el informe de terceros (en su versión 2010), en el caso chileno.

.....

13 Esta tendencia refleja también la naturaleza abstracta de las consecuencias de una evaluación poco confiable, que aunque serías para el docente en cuanto a ingresos o seguridad laboral, son relativamente invisibles para el ojo público, por lo menos inicialmente. Finalmente, en el caso norteamericano existe también un creciente rechazo al papel de los sindicatos magisteriales en el marco de debates ideológicos provocados por la grave crisis económica.

Consideraciones finales

Chile está marcando camino en la evaluación docente a gran escala, al combinar la sofisticación técnica de los indicadores con un trabajo cuidadoso para examinar la validez, utilidad y consecuencias de las inferencias que de ellos derivan. Vista desde fuera, esta experiencia provee además una excelente oportunidad para el estudio de indicadores de desempeño docente en los que el criterio (los puntajes en pruebas estandarizadas) no se asocia aún a decisiones individuales de alto impacto que distorsionen gravemente los resultados. Esto contrasta con otros modelos como el norteamericano, donde las pruebas no solo son el criterio dominante, sino con frecuencia el único. En esta lógica circular, la validez de cualquier medida o instrumento que capture aspectos del desempeño docente se define explícitamente en términos de su correlación con los puntajes del alumno en pruebas estandarizadas (como demuestra el ejemplo del estudio MET). Los estudios del sistema chileno adquieren, por tanto, mayor importancia en el contexto internacional al abrir una ventana no siempre disponible en otros lugares (ver e.g. Taut, Santelices & Manzi, en este volumen; Santelices & Taut, 2011; Taut, Santelices & Valencia, 2010; Manzi, Strasser, San Martín & Contreras, 2007).

El peso preponderante que se da a la evidencia que proviene del portafolio del maestro en la Evaluación Docente chilena refleja la fuerte base conceptual y carácter formativo de esta evaluación, pero también una decisión explícita de anteponer la riqueza y calidad de la evidencia al uso de la *fuerza bruta* estadística con base en pruebas estandarizadas. El cuidado con que se desarrolla el portafolio y su sistema de calificación constituyen un ejemplo que, siendo por supuesto perfectible, se equipara con otros a nivel internacional, como los portafolios del programa del National Board for Professional Teaching Standards (NBPTS) o del Performance Assessment for California Teachers (PACT). Considero que dedicar esfuerzo y recursos a mejorar la calidad de la información sobre desempeño docente obtenida mediante el portafolio y la evaluación del director ofrece a Chile (y a cualquier país) una promesa igual o mayor de mejora en principio, que el aumento en la cobertura de las pruebas estandarizadas que se administran a los alumnos. En la medida en que logre mantener y mejorar la calidad y el balance de sus indicadores y evitar la tentación reduccionista común en otras latitudes, el modelo chileno de evaluación docente seguirá constituyendo un ejemplo y un caso de estudio importante para tomadores de decisiones e investigadores en otros países.

Aportes para pensar las políticas de evaluación docente

Pedro Ravela

Instituto de Evaluación Educativa
Universidad Católica del Uruguay

Las reflexiones que siguen están referidas tanto al libro que las contiene como a las alternativas de política de evaluación docente, actualmente en debate en Chile. Parten de un conocimiento bastante detallado del caso chileno, que he estado siguiendo durante los últimos cuatro años, así como de mi reflexión y estudio sobre el tema. Están organizadas en tres grandes secciones. La primera está destinada a proponer algunas ideas eje en torno al mejoramiento del desempeño docente y de las prácticas de enseñanza. En la segunda sección me enfoco en la discusión de aspectos centrales de las políticas de evaluación docente. En la tercera formulo en forma rápida algunos desafíos de carácter técnico, que entiendo debería encarar el sistema de evaluación docente de Chile en los próximos años.

¿Evaluación de los docentes o evaluación de la enseñanza?

La evaluación de la docencia comenzó a estar en el centro de la agenda educativa en la última década. Los esfuerzos de reforma educativa en los 90 estuvieron orientados hacia otros aspectos del sistema educativo: libros de texto y bibliotecas de aula, descentralización y autonomía para la gestión educativa, evaluación externa de los logros educativos de los alumnos, inversiones en infraestructura orientadas a la ampliación de la cobertura en nuevos tramos de edad (por ejemplo, la educación inicial) o a la ampliación del tiempo educativo (por ejemplo, escuelas de jornada completa), entre otras. Este tipo de acciones, si bien importantes, tuvieron un impacto inferior al esperado, por una razón principal: los modos de enseñar al interior de las aulas no fueron mayormente modificados. Progresivamente se fue tomando conciencia de algo bastante elemental y muchas veces declarado: mientras no se produzcan cambios en los modos de enseñar, difícilmente mejoren los aprendizajes. Esta toma de conciencia se tradujo en una creciente percepción de tres desafíos fundamentales para las políticas educativas:

- evaluar el desempeño de los docentes;
- construir una base de conocimiento empírico acerca de qué es enseñanza efectiva;
- desarrollar dispositivos más efectivos de formación docente (tanto inicial como en servicio).

En nuestra región es poco lo que se ha avanzado en estos temas. La mayor parte de los sistemas educativos de la región continúa funcionando con procedimientos de supervisión propios del siglo XIX, basados en la visita al aula de un supervisor o director y la posterior elaboración de un informe, sin criterios ni instrumentos explícitos. Este es el caso, por ejemplo, de Argentina y Uruguay. En muchos países ni siquiera se realizan este tipo de visitas, con lo cual la mayoría de los docentes trabajan en forma aislada, sin ningún tipo de escrutinio externo sobre sus prácticas. Algunos países, como Perú y Ecuador, han comenzado por algunas cuestiones más elementales, como evaluar los conocimientos de los docentes que ingresan a la profesión. En Chile, el programa INICIA hace lo propio con los egresados de las carreras de pedagogía. Otra experiencia que está siendo explorada en países como Perú y El Salvador, es la construcción de la figura del acompañante didáctico o asesor pedagógico. En varios países se han impulsado sistemas de reconocimiento a los buenos docentes a través de premios (por ejemplo, Maestro 100 Puntos en Guatemala, Educador Nota 10 en Brasil, Compartir al Maestro en Colombia)⁽¹⁾ (Grupo de Trabajo sobre Estándares y Evaluación PREAL, 2009). Si bien es muy importante reconocer e incentivar a los buenos docentes, la debilidad de

.....
1 También podría entrar en esta categoría la Asignación por Excelencia Pedagógica (AEP) de Chile, aunque tiene características diferentes, porque tanto la evaluación que se realiza, como la cantidad de docentes premiados es de mucho mayor envergadura en este país.

este último enfoque es que, por lo general, no incluye una estrategia destinada a mejorar la práctica del conjunto de la profesión.

Solamente Cuba y Chile han emprendido esfuerzos sistemáticos, continuos y de gran escala para evaluar el desempeño de sus docentes. Si bien diferentes, ambas políticas comparten un enfoque que combina una finalidad formativa (mejorar los modos de enseñar) con un sistema de consecuencias en términos de incentivos y reconocimiento para los docentes destacados y exigencia de esfuerzo de recuperación para los docentes cuyo desempeño es insatisfactorio, incluyendo la posibilidad de tener que abandonar el aula en caso de no mejorar el desempeño.

Dos ideas principales querría proponer como supuestos de partida para las reflexiones que siguen. En primer lugar, que no deberían esperarse cambios sustantivos en los aprendizajes si antes no ocurren cambios en los modos de enseñar. Es bastante probable que algunas mejoras se produzcan como resultado de cambios positivos en las condiciones culturales y económicas de la población socialmente desfavorecida. Sin embargo, si los estudiantes de mejor situación cultural y económica de la región tienen un desempeño muy por debajo de sus pares de los países con mejores resultados en PISA, esto obedece a diferencias en la enseñanza, no únicamente a las condiciones socioculturales.

La segunda idea de base es que las prácticas de enseñanza constituyen un sistema cultural, de allí la enorme dificultad que existe para modificarlas. Stigler y Hiebert (1999) han analizado este punto de manera exhaustiva, a partir de la filmación y análisis de clases de matemática en una amplia gama de países, en el marco del estudio internacional conocido como TIMSS. La enseñanza es una práctica cultural, un sistema de creencias y guiones construido a lo largo de la vida de cada docente y reforzado en forma cotidiana por múltiples actores:

“Los métodos que usan para interactuar con sus alumnos no están determinados por su formación, sino por la cultura en la que enseñan (...) son prácticas heredadas de las generaciones anteriores, no inventadas (...) la mayor parte de lo que ocurre en el aula está determinado por un código cultural que funciona, en algunos sentidos, como el ADN de la enseñanza. Por esta razón cambiar a los maestros no produce automáticamente cambios en la enseñanza (...) es preciso darse cuenta de cuáles son las rutinas culturales que gobiernan la vida en el aula, cuestionar los supuestos de estas rutinas, y trabajar para modificarlas durante un largo tiempo. Implica reconocer que los detalles de lo que los maestros hacen —las preguntas específicas que formulan, el tipo de tareas que proponen a los estudiantes, las explicaciones que dan— son cosas fundamentales para el aprendizaje de los estudiantes (...) y que todos estos detalles son opciones que los maestros hacen...” (Stigler & Hiebert, 1999, p. xii–xiii, traducción del autor).

Sin embargo, afirman los autores, la mayor parte de los esfuerzos de política educativa se enfoca en cambiar a los maestros, no en cambiar la enseñanza (*“teachers rather than teaching, is the critical factor”*). Desde esta perspectiva, el desafío central para las políticas educativas (y, por tanto, para las políticas de evaluación docente), es modificar los modos de enseñar, lo cual implica un cambio cultural en una profesión, más que cambios en características individuales de los docentes. Para ello, es necesario moverse de la visión de la enseñanza como una actividad solitaria, propia de cada docente, hacia una visión de la enseñanza como una actividad profesional abierta a la observación, el estudio y el mejoramiento colectivos.

Introducir cambios en los modos de enseñar es un proceso de largo plazo, y trabajo continuo y cotidiano. No es algo que pueda lograrse en períodos cortos ni que produzca rápidamente impactos observables. Por esta razón suele ser un tema complicado para los políticos, que normalmente necesitan mostrar resultados en períodos cortos de tiempo, demasiado cortos para el modo en que la educación funciona. Partiendo de estos supuestos clave, las reflexiones que siguen se apoyan en la idea de que los tres

desafíos planteados al inicio deben estar claramente alineados entre sí, es decir, la política de evaluación del desempeño docente debe estar articulada con una agenda de investigación sobre las prácticas de enseñanza y con un esfuerzo sistemático de formación en servicio de los docentes, orientado a modificar dichas prácticas.

Decisiones clave en una política de evaluación de docentes

En el contexto de las ideas expuestas, en este apartado abordaré tres aspectos centrales que toda política de evaluación docente debe definir con claridad:

- ¿Qué es lo que debe ser objeto de evaluación?
- ¿Qué tipo de consecuencias debería tener la evaluación y qué usos dar a sus resultados?
- ¿Quiénes y cómo deben llevar adelante los procesos de evaluación?

¿Evaluar el desempeño de los docentes o evaluar los resultados de sus alumnos?

Tal vez sea esta la primera disyuntiva que una política de evaluación docente debe enfrentar y resolver: ¿se pondrá el foco de la evaluación en la observación y valoración del desempeño del docente, es decir, de su trabajo, o, por el contrario, se asumirá que el principal indicador de la calidad de su desempeño son los logros educativos de sus estudiantes? La cuestión ha sido objeto de múltiples debates y experiencias, especialmente en los Estados Unidos. En el caso de Chile fue resuelta en favor de la primera opción, a través de un interesante proceso de construcción de acuerdos entre el Ministerio de Educación y el Colegio de Profesores, según se da cuenta en el Capítulo 1 de este libro. De cualquier modo, entiendo importante agregar algunos argumentos en favor de dicha opción.

La opción de utilizar los resultados de los estudiantes fue defendida inicialmente a partir de un argumento muy simple: no sabemos qué es buena enseñanza y, dado que los educadores no se ponen de acuerdo en ello, una buena definición sería “buena enseñanza es aquella que logra que los estudiantes aprendan”. Con esto se resuelve, aparentemente, el problema. No es necesario evaluar ni medir el desempeño de los docentes, alcanza con medir los aprendizajes de sus estudiantes. Si estos son satisfactorios, el docente es bueno.

El primer problema que debió enfrentar esta postura fue de carácter técnico. Sabemos que el aprendizaje de los estudiantes depende de múltiples factores, además de lo que haga el docente: de su origen social, de sus condiciones de vida, del apoyo que reciba en el hogar, de la calidad de los docentes que tuvo en años anteriores, del nivel y motivación del grupo del que forma parte, de los recursos con que cuenta su centro educativo, por mencionar algunos. La discusión se centró entonces en el problema técnico para la atribución de efectos al docente, que puede resumirse de la siguiente manera: dado que el aprendizaje de los estudiantes es resultado de múltiples factores, ¿cómo construir un dispositivo de medición que permita distinguir en forma válida y confiable el aporte al aprendizaje que puede ser atribuido a la labor del profesor? De allí surgen los modelos de valor agregado (Glazerman, Loeb, Goldhaber, Staiger & Raudenbush, 2010).

Más allá de los problemas de carácter técnico, cuya discusión excede el alcance de este texto, mi convicción es que la evaluación del docente a partir de los resultados de sus estudiantes tiene dos problemas principales.

El primero es de carácter filosófico: los seres humanos, incluidos los estudiantes, tenemos libre albedrío. Podemos elegir qué hacer y en qué invertir nuestro tiempo y nuestra energía. Por tanto, el que un estudiante aprenda o no lo que su profesor le enseña, depende tanto de la calidad de la enseñanza impartida por el profesor como de la motivación del estudiante para aprender (además de otros múltiples

factores que enumeramos someramente más arriba). Si un estudiante no está interesado, si no cree que la educación es importante para su presente y para su futuro, si no está dispuesto a destinar más que algunos minutos a la semana a estudiar (algo que ocurre con buena parte de los adolescentes), difícilmente pueda producirse el aprendizaje. Por cierto, una de las habilidades de un buen docente es motivar a sus alumnos. Pero el resultado de sus mejores esfuerzos por hacerlo nunca está garantizado. Es más fácil en la primaria y mucho más difícil durante la cada vez más prolongada adolescencia de nuestros tiempos. El aprendizaje depende fuertemente del entorno social, de la motivación familiar y de las actitudes colectivas hacia el estudio, que se generan en la sociedad y en el grupo de pares.

Imaginemos a un médico especializado en salud cardiovascular que hace un excelente trabajo con sus pacientes: los diagnostica en forma apropiada, les explica las características de la enfermedad y los riesgos que corren, les indica un conjunto de acciones a seguir (medicación, dieta, ejercicios, entre otros). Por diferentes motivos, parte importante de sus pacientes hace caso omiso a sus recomendaciones: unos no tienen tiempo para hacer ejercicio, otros no modifican su dieta, otros no logran abandonar el hábito de fumar. ¿Evaluáramos negativamente el desempeño del médico si una parte importante de sus pacientes tiene accidentes cardiovasculares? No necesariamente. La práctica médica puede ser excelente, pero los pacientes deben poner su parte. Por tanto, el desempeño del médico es una cosa, el resultado de su labor es otra. Algo similar ocurre en educación, con el agregado de que aprender es una actividad bastante más compleja, tanto para el docente como para el estudiante, que seguir un conjunto de indicaciones médicas.

Lo dicho no implica, en absoluto, que no exista relación entre enseñanza y aprendizaje. Esperamos que los estudiantes de un buen docente aprendan más y mejor. Por eso es importante, en primer lugar, observar la práctica docente y, por otro lado, estudiar si lo que estamos considerando como buena enseñanza tiene efectos positivos sobre los aprendizajes (aspecto sobre el cual hay una primera incursión en el Capítulo 6 de este libro).

Digamos pues, a modo de síntesis, que la buena enseñanza es condición necesaria, pero no suficiente para asegurar el aprendizaje de los estudiantes. Esta es una razón por la cual los resultados no deberían tomarse como el indicador principal de un buen desempeño.

La segunda razón para optar por la primera alternativa es de tipo práctico. Si el objetivo perseguido es mejorar la educación, lo cual inevitablemente pasa por mejorar los modos de enseñar, evaluar exclusivamente en función de los resultados de los alumnos deja a los profesores inermes y desmotivados. Deben adivinar qué es lo que está fallando. Saben que sus alumnos no han aprendido, pero no saben por qué. El supuesto de base es que la docencia es una tarea más o menos sencilla y que lo importante es ponerle "ganas" y esfuerzo. Por otra parte y más importante aún, más allá de los efectos individuales sobre los docentes, este enfoque priva a la profesión docente como tal de la oportunidad de aprender.

En cambio, evaluar el desempeño docente a partir de una definición conceptual sobre qué es buena enseñanza, aporta información sobre qué es lo que hay que mejorar y una base común para el desarrollo de la profesión. Y, sobre todo, permite articular la evaluación docente con los otros dos grandes desafíos planteados inicialmente: acumular evidencia empírica sobre buenas prácticas de enseñanza y desarrollar programas de formación en servicio.

Esta es la dirección en la que se ha encaminando la implementación del sistema de evaluación docente en Chile, de lo cual este libro es testimonio. Es interesante notar que en esta misma dirección se está encaminando una parte importante de la discusión en los Estados Unidos. A partir de 2010 la Fundación de Bill Gates está invirtiendo 335 millones de dólares en una iniciativa dirigida a introducir la filmación de clases a gran escala en ese país, comenzando por 3.000 docentes voluntarios de siete grandes distritos escolares (Dallas, Denver, Charlotte–N.C., Hillsborough County - Florida, Memphis, Nueva York y Pittsburgh) (Dillon, 2010b). La iniciativa está dirigida a subsanar uno de los problemas de los modelos de valor agregado que son utilizados en cientos de distritos en los Estados Unidos: la

evaluación del docente en un único dato estadístico que sirve para posicionarlos en una escala, pero que no da cuenta de su desempeño (Dillon, 2010a). Para junio de 2011 está prevista la codificación de unas 24.000 horas de clase, lo que requerirá entrenar cientos de codificadores.

Es interesante notar que el sistema de evaluación docente de Chile ya se ha anticipado en este camino, con un esfuerzo organizativo y logístico encomiable (si bien hay aún mucho por hacer en cuanto al aprovechamiento del material recogido y al refinamiento de los instrumentos utilizados). Otro aspecto interesante del emprendimiento de la Fundación Gates es que no se habla de un sistema de evaluación docente, sino de un sistema para evaluar la enseñanza en el aula (*"a better system for evaluating classroom instruction"*) (Dillon, 2010a).

¿Para qué evaluar? El problema de las consecuencias de la evaluación docente

Un segundo problema crucial que todo proceso de evaluación debe abordar y resolver de alguna manera más o menos inteligente, es el relativo al de las consecuencias y usos de sus resultados. Existe una tensión inevitable entre la alternativa de una evaluación puramente formativa y orientada al aprendizaje por parte de los evaluados, por un lado, y una evaluación con consecuencias directas sobre los evaluados, por otro. En el aula, esta es la diferencia entre la evaluación orientada a calificar a los estudiantes y la evaluación que forma parte del proceso de aprendizaje, sin consecuencias de calificación. La primera tiene la virtud de que obliga a los estudiantes a realizar un esfuerzo, pero la debilidad de que este esfuerzo puede fácilmente transformarse en una simulación dirigida a satisfacer al profesor, pero no a aprender. Viceversa, la segunda tiene la virtud de que puede promover un aprendizaje más auténtico y automotivado, pero la debilidad de que no existe obligación para el estudiante y no todos tienen la motivación intrínseca requerida.

Esta tensión está vinculada a una contradicción inherente a los seres humanos: todos necesitamos exigencias externas y, al mismo tiempo, espacios para crecer en forma autónoma.

El problema que se plantea en la evaluación docente es que sus consecuencias determinan fuertemente la actitud de los profesores hacia la misma. Si la evaluación tendrá consecuencias, el docente evitará mostrar debilidades e, incluso, podrá llegar a simular comportamientos apropiados a lo esperado como buen desempeño (algo que puede lograrse en algunos aspectos como, por ejemplo, copiando partes del portafolio, pero no en otros como, por ejemplo, respondiendo en forma apropiada a las preguntas de los estudiantes en el aula). Si la evaluación no tiene consecuencias directas y es exclusivamente un sistema de orientación para el desarrollo profesional, el docente se sentirá más libre para mostrar sus dificultades y aprender de las mismas. Pero la ausencia de consecuencias puede dar lugar a que buena parte de los docentes simplemente ignoren el sistema. Es una contradicción intrínseca a la condición humana.

El sistema de evaluación docente chileno está encaminado a lograr una articulación razonable entre un sistema formativo (que es lo que predomina en el discurso que sustenta el sistema y los acuerdos fundacionales del mismo) y un sistema con consecuencias (el hecho de ser calificado públicamente como un docente insatisfactorio, básico, competente o destacado es por sí mismo una consecuencia fuerte, más aún si hay consecuencias económicas o de estabilidad laboral).

En este marco y a partir de la lectura de este libro, querría formular algunas sugerencias:

- Antes de ampliar el uso de los resultados de la evaluación docente con fines de tipo contractual y laboral, es importante profundizar en los estudios de validez y confiabilidad del sistema, en el sentido de asegurar que la diferencia entre categorías adyacentes (insatisfactorio/básico; básico/competente y competente/destacado) reflejan diferencias reales y lo hacen con un nivel de precisión adecuado.
- Un desafío central es la implementación de un sistema de formación en servicio y orientación a los docentes más amplio, más claramente alineado con el Marco para la Buena Enseñanza (MBE)

y con un foco en la modificación de las prácticas de enseñanza. La situación actual en este sentido parece preocupante: el 76% de los Planes de Superación Profesional (PSP) consisten en cursos, talleres y seminarios, en tanto que el 50% de los mismos tienen una duración inferior a ocho horas. Partiendo de las reflexiones iniciales, considero bastante improbable lograr un cambio sustantivo en los modos de enseñar a través de este tipo de actividades. Se requiere de una labor mucho más sistemática y continua en el tiempo, con un foco principal en el análisis de las propias clases. Los videos de clase deberían ser una herramienta fundamental en los procesos de formación, al igual que el contacto con docentes destacados.

- Un riesgo fundamental a evitar es que los PSP se constituyan en una herramienta para mejorar específicamente el resultado de la evaluación (léase, aprender a hacer un buen portafolio) y pierdan de vista el objetivo más amplio de mejorar el desempeño docente en la línea del MBE. Uno de los problemas a analizar, en este sentido, es si no sería más apropiado organizar la construcción de los resultados de la evaluación en torno a las dimensiones e indicadores del MBE, en vez de hacerlo en torno a los instrumentos. Retomaré este tema en la tercera sección.
- La construcción de un buen sistema de apoyo a los docentes excede las responsabilidades del sistema de evaluación y corresponde al Ministerio de Educación y a las municipalidades. Esta es una tarea que requiere combinar capacidades locales y apoyo central, aspecto que es analizado en el Capítulo 5 de este libro. En dicho capítulo, luego de analizar las ventajas y problemas de los enfoques centralizados (mayor acumulación y especialización, pero distancia de la práctica real) y de los enfoques basados en la escuela (ceranos a las prácticas de enseñanza pero con escasos recursos y especialización), se concluye sugiriendo una estrategia de radicación de los PSP en el nivel comunal. En mi opinión, es necesaria una estrategia que combine adecuadamente aportes y monitoreo desde el nivel central con ejecución en el nivel comunal, pero sin perder una base de trabajo en cada centro educativo, por dos motivos principales. En primer lugar, la apertura y análisis de lo que ocurre al interior de cada aula debería ser un elemento central de la estrategia de formación y mejora profesional. En segundo lugar, es importante propiciar el intercambio entre docentes de todas las categorías y evitar la conformación de “clubes de desplazados” en el nivel comunal (que sería el resultado de reunir únicamente a los docentes evaluados como insatisfactorios y básicos). Es importante que aquellos que resultaron mal evaluados se sientan parte de un cuerpo profesional que puede ayudarlos a mejorar su desempeño.
- Finalmente, en materia de consecuencias, creo que el sistema educativo debería aprovechar de un modo específico y cuidadosamente planeado, la capacidad de los docentes destacados, encomendándoles nuevos roles y funciones sin retirarlos completamente del aula. Más que incentivos económicos asociados al mero hecho de haber logrado la categoría de destacado, debería remunerárseles de una manera estimulante pero como contrapartida del desempeño de estos nuevos roles, como la evaluación de pares, la tutoría y acompañamiento didáctico y la formación en servicio de otros colegas.

¿Quién debe evaluar? La tensión entre lo central y lo local

Si se quiere construir un espacio profesional docente, en el que se trabaje sobre la base de un cuerpo de conocimiento compartido, y si se concibe la mejora de la enseñanza como el resultado del mejoramiento del conjunto de la profesión docente —en lugar de concebirla como resultado de esfuerzos individuales—, tanto la evaluación docente como los esfuerzos de superación deben tener un importante componente de carácter central y común al país. Descentralizar fuertemente la evaluación docente implica un serio riesgo de crear mundos profesionales de diferente calidad, asociados a la capacidad y recursos de las comunas.

Por tanto, si bien una parte importante de la evaluación debe ocurrir en el nivel local, dado que allí están los agentes con un conocimiento más directo y cotidiano de la labor de cada docente, simultáneamente, debe buscarse un equilibrio adecuado entre este componente local y un componente externo o central.

Una segunda cuestión importante a tener en cuenta es que cada actor debe evaluar aquello que está en condiciones de valorar de manera válida y confiable. Un director de escuela puede valorar mejor que nadie el grado de cumplimiento de sus obligaciones por parte del docente, pero difícilmente pueda valorar los aspectos didácticos de su trabajo, sobre todo si implican conocimientos propios de una disciplina. Aquellos aspectos del desempeño docente cuya evaluación requiere conocimiento especializado, es mejor que tengan un componente central, dado que requieren una cierta escala para tener acumulación de conocimiento y experiencia como para hacerlo bien.

En tercer lugar, las consecuencias de la evaluación también son importantes a la hora de pensar quién debe evaluar qué. Como surge claramente de lo expuesto en este libro, cuando la evaluación tiene consecuencias, la autoevaluación deja de ser pertinente (nadie se autoevaluará en forma negativa si de ello se derivará una categoría de desempeño de carácter público), al tiempo que la evaluación de pares o personas cercanas debe ser cuidadosamente pensada. Los superiores jerárquicos están sujetos a dos riesgos opuestos: el exceso de generosidad, por un lado, porque no quieren perjudicar a colegas y docentes conocidos, y la arbitrariedad, por otro, porque la evaluación puede servirles para premiar y castigar a las personas según su grado de afinidad.

Por estas razones, es importante distinguir entre la descentralización de los usos de los resultados y la descentralización del proceso de evaluación mismo. Claramente, los directivos locales (municipales y escolares) deben tener un importante poder de decisión respecto a qué hacer a partir de los resultados de la evaluación. Y deben tener una participación importante, pero acotada en el proceso mismo de la evaluación, estructurada en torno a aquellos aspectos del desempeño docente que pueden valorar con propiedad. Dicho con otras palabras, la principal razón para descentralizar aspectos del proceso mismo de evaluación debería ser la valoración de los aspectos del desempeño docente que solamente pueden ser observados en forma adecuada a nivel local y a partir del contacto personal y directo con el docente. Pero debe evitarse el riesgo de establecer estándares diferenciados para el desempeño docente.

Los acuerdos con la profesión docente y la continuidad de los mismos como cimiento de la legitimidad de un sistema de evaluación docente

Para cerrar la sección política, quiero destacar la importancia del proceso de acuerdos entre el Ministerio de Educación y los sindicatos, probablemente sin precedentes en la región. Un sistema de evaluación docente duradero y efectivo solamente puede construirse sobre la base de un acuerdo con la profesión docente (y en cualquier profesión).

También es destacable que se trata de un acuerdo complejo efectivamente implementado, lo cual también es poco común en la región. Primero, porque suele ser muy difícil lograr este tipo de acuerdos. Segundo, porque normalmente se trata de acuerdos, o bien sobre aspectos mucho más concretos y puntuales, o bien absolutamente generales y vagos. Tercero, porque la mayoría de los acuerdos suelen quedar en el papel. En este caso, hay un acuerdo programático, relevante, complejo y que fue llevado a la práctica con éxito. Un gran logro y un gran ejemplo para la región.

Los desafíos técnicos de un sistema de evaluación docente

La lectura del libro es entusiasmante por el esfuerzo realizado en el desarrollo del sistema de evaluación docente y por la calidad de éste. Sin embargo, en tanto investigador y evaluador especializado en el tema, uno no puede sino pensar en nuevos desarrollos y aspectos que podrían ser mejorados. En el poco espacio que me resta intentaré resumir algunas de estas propuestas, organizadas en torno a dos grandes ejes.

Un diseño más sistemático y externo para la evaluación del sistema de evaluación

La evaluación, monitoreo y mejoramiento permanentes de un sistema de la envergadura del que se ha logrado implementar, no debería apoyarse únicamente en estudios particulares e investigaciones realizadas por los propios miembros de la Unidad Técnica a cargo de la implementación. Algunas ideas que fueron surgiendo durante la lectura del libro son las siguientes:

- Estudiar las diferencias entre categorías adyacentes de docentes (insatisfactorios/básicos, básicos/competentes, competentes/destacados) y no entre destacados e insatisfactorios como se ha hecho, dado que es obvio que habrán de encontrarse diferencias entre estas dos últimas categorías (salvo que todo el sistema de evaluación fuese muy malo). Es crucial contar con evidencia empírica acerca de en qué medida el instrumental aplicado clasifica bien entre categorías adyacentes, en que las diferencias pueden ser mucho más sutiles.
- En estrecha relación con lo anterior, sería interesante pensar en estudios dirigidos a una descripción más profunda de los modos de enseñar de los docentes pertenecientes a las distintas categorías, especialmente de los destacados, y su efecto en el aprendizaje de sus estudiantes, combinando evidencia de tipo cuantitativo y cualitativo. El Capítulo 4 aporta alguna información interesante sobre estos aspectos, tomada de los portafolios. Parece interesante pensar en estudios complementarios que impliquen la visita y estudio en mayor detalle de la práctica cotidiana de las diferentes categorías de docentes. Más allá de la cuestión de la validez y confiabilidad de la evaluación, este tipo de estudios aportaría información relevante para los PSP y para los esfuerzos de mejoramiento de las prácticas de enseñanza.
- El Capítulo 6, destinado a dar cuenta de los estudios de validez realizados, debería conducir a una agenda concreta de estudios para la validación y mejora permanente del sistema. Tal como se afirma en dicho capítulo, la cuestión de la validez nunca es definitiva ni taxativa, sino una cuestión de grados y de perfeccionamiento progresivo. A modo de ejemplo, sería interesante contar con estudios más detallados acerca de cómo operan en la realidad las Comisiones Comunales de Evaluación. La pregunta central a responder sería la siguiente: ¿Habría diferencias entre dos grupos aleatorios de comisiones comunales, uno sometido a observación externa y otro no? Otro conjunto de estudios debería estar orientado a los usos y efectos, deseables y no deseables, de los resultados derivados del sistema de evaluación.
- Es muy importante el papel explícito de monitoreo de los PSP asignados al Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas (CPEIP). Llevarlo a cabo requiere algo más que sistematizar datos sobre los gastos, la asistencia y satisfacción de los docentes, y las modalidades que asumen los PSP. Requiere investigación sistemática orientada a responder preguntas como las siguientes: ¿qué hacen los docentes durante su participación en los PSP?; ¿en qué tipo de experiencias formativas participan?; ¿qué impacto tiene la participación en los PSP en los modos de enseñar en las aulas, en el aprendizaje de los estudiantes y en la siguiente evaluación de cada docente participante? Responder a estas preguntas requiere de un estudio de seguimiento, que analice una cohorte de docentes categorizados como insatisfactorios o básicos y compare su evolución en función de si participaron o no en un PSP, por un lado, y en función del tipo de PSP en el que participaron, por otro. Debería, además, incluir un análisis detallado de su evolución en distintas dimensiones e indicadores del MBE, no solo su categorización global. La variedad y descentralización de los PSP constituye una oportunidad para construir un gran laboratorio de experimentación sobre formación en servicio y orientación para el desarrollo profesional. Pero para ello se requiere de un conjunto planificado de estudios adecuadamente diseñados.

Sobre la construcción del índice

Un segundo conjunto de ideas gira en torno al modo en que se construyen los resultados de la evaluación. De acuerdo a lo descrito en este libro, la asignación final de cada docente a una categoría es el resultado de cuatro instrumentos: la autoevaluación, la entrevista de pares, el informe de referencia de terceros y el portafolio. En los tres primeros casos, se otorga a cada docente una valoración en cada una de las cuatro dimensiones del MBE, como resultado de promediar la valoración de distintos indicadores. En el caso del portafolio, la valoración sigue un procedimiento distinto, organizada en torno a dimensiones que el instrumento permite evaluar a través de sus diferentes componentes. Finalmente, se promedian todas las valoraciones, con un peso diferente para cada instrumento. La Comisión Comunal de Evaluación puede ratificar o modificar en forma justificada el resultado obtenido por este camino (que no es otra cosa que la construcción de un índice bastante sofisticado).

Mi sugerencia es que sería importante revisar este modo de proceder (si bien reconozco que una cosa es pensarlo en la teoría y otra es modificar un sistema de trabajo que está instalado), a partir de dos criterios: i) el reporte final debería estar organizado en torno a las dimensiones e indicadores del MBE, no de los instrumentos; ii) a cada instrumento y agente de evaluación solamente debe pedírsele aquello que puede evaluar de manera válida y confiable.

Intentaré explicar las implicancias de estos dos criterios.

En primer lugar, creo que no tiene sentido que los tres primeros instrumentos se propongan dar cuenta de las cuatro dimensiones del MBE. A lo sumo pueden hacerlo de algunos de sus criterios. Un director no puede valorar adecuadamente si el modo en que un profesor evalúa los aprendizajes en historia es apropiado, aunque sí puede evaluar mejor que nadie el cumplimiento de las obligaciones funcionales. El sujeto evaluado no puede autoevaluarse en un contexto de evaluación con consecuencias directas (salvo que su autoevaluación sea, a su vez, objeto de análisis y calificación por otros evaluadores, como forma de valorar su capacidad de autocrítica y autopercepción de fortalezas y debilidades). A partir de una clase filmada en video no puede valorarse el ambiente de clase, porque es obvio que la disciplina será buena, salvo en casos extremos. En cambio, el modo de conducir una clase, de preguntar, de atender a las preguntas de los alumnos, de explicar un tema, de usar la pizarra, entre otros muchos aspectos, no son tan afectados por la presencia de la cámara, dado que no pueden ser imprevistos ni simulados.

Todo esto está muy bien presentado en el Capítulo 2 de este libro, en el que se explica cómo, a partir del MBE con sus criterios y descriptores, se llega a los instrumentos.

Mi sugerencia es realizar el proceso inverso a la hora de calificar a los profesores y asignar categorías de desempeño, es decir, en lugar de promediar los resultados por instrumento, asignar valores a cada docente en los distintos descriptores del MBE que cada instrumento puede evaluar en forma válida y confiable (el peso de cada instrumento en la valoración de cada dimensión y criterio, puede realizarse en este paso). Porque lo importante es que el docente mejore su desempeño en las dimensiones del MBE, no en los instrumentos de evaluación.

Escapa a mis posibilidades estimar las dificultades de esta propuesta. La ventaja que tendría sería alinear el reporte de resultados con el MBE. Esto permitiría entregar a cada docente y a las instancias encargadas de las acciones de superación posteriores, información sobre los aspectos específicos del MBE en los que cada docente necesita mejorar. Pondría el foco de atención en mejorar los desempeños en función del MBE y no tanto en mejorar el desempeño en los instrumentos específicos. El objetivo del docente que obtiene un resultado Básico o Insatisfactorio no debería ser aprender a producir buenos portafolios, sino a mejorar su desempeño en concordancia con el MBE.

La buena enseñanza y la agenda abierta de la evaluación docente ante un recurso escaso

Gilbert A. Valverde, Ph.D.

Institute for Global Education Policy Studies

Department of Educational Administration and Policy Studies, School of Education

University at Albany, State University of New York

La buena enseñanza: un recurso escaso

Pocas cosas son tan importantes para la calidad de la educación como la buena enseñanza. La preponderancia de la evidencia no solo indica el impacto de una enseñanza de calidad sobre el desempeño estudiantil, sino que muestra también que es un beneficio acumulativo. Se ha demostrado que por cada año en que un estudiante cuenta con un buen docente, aumenta el beneficio de su impacto. Inclusive, en algunos casos a nivel mundial, se ha encontrado que estudiantes que provienen de hogares de bajos ingresos y que cuentan con buenos docentes por 3 años seguidos o más, alcanzan niveles de desempeño promedio similares a los de sus pares que provienen de hogares de clase media (Glazerman, Loeb, Goldhaber, Staiger & Raudenbush, 2010; Hanushek, Kain, O'Brien & Rivkin, 2005; Nye, Konstantopoulos & Hedges, 2004). Asimismo, se ha demostrado, una y otra vez, a nivel global (Schmidt et al., 2001; Schmidt, McKnight, Valverde, Houang & Wiley, 1997; Schmidt, Raizen, Britton, Bianchi & Wolfe, 1997), en América Latina (Valverde & Näslund-Hadley, 2010), y en el caso específico de Chile (Ramírez, 2006) que el comportamiento de los docentes, las decisiones que toman, los recursos didácticos que ponen en juego en sus aulas, y el dominio personal de los ámbitos del conocimiento que enseñan, tienen una relación intrínseca con los niveles de desempeño de sus estudiantes.

La buena enseñanza, por lo tanto, podría ser uno de los más potentes correctivos de inequidad en la distribución de oportunidades para la vida. No obstante, como muchos otros recursos importantes, la buena enseñanza es escasa.

Su escasez en América Latina es alarmante. Aunque la mayoría de los docentes de la región tiene el nivel de capacitación requerido por los sistemas nacionales de educación —por lo general un título de un instituto de formación docente o de una universidad (LLECE, 2010)—, las evidencias sugieren que muchos docentes no están adecuadamente preparados y ofrecen muy pocas oportunidades para que sus alumnos aprendan las destrezas en las materias escolares necesarias para ejercer con plenitud sus derechos ciudadanos y competir con éxito en la economía. En América Latina, como hemos reportado en otras oportunidades (Valverde & Näslund-Hadley, 2010; Valverde, 2009), las evidencias muestran serios déficits en la calidad de la docencia, que se acumulan y presentan un cuadro crítico.

En Perú, por ejemplo, un estudio de estudiantes de sexto grado en 22 escuelas públicas de Lima mostró que los estudiantes solo lograron resolver menos de la mitad de los ejercicios de matemática que copiaron. La evidencia encontrada en sus cuadernos también indicó que los docentes ponen excesivo énfasis en los temas del currículo nacional que son menos exigentes en el aspecto cognitivo. El mismo estudio descubrió que es común encontrar errores en los libros de ejercicios de los estudiantes que no habían recibido ninguna retroalimentación de parte de los docentes, o lo que es peor, una retroalimentación errónea (Cueto, Ramirez & Leon, 2006).

Una investigación internacional efectuada en aulas de matemática de sexto grado en Argentina, Colombia, Costa Rica, El Salvador, Guatemala, Perú, México y Uruguay, encontró que las prácticas de evaluación que los docentes utilizan durante la clase son extremadamente débiles. A los estudiantes y a sus familias se les da muy poca retroalimentación formativa en esas evaluaciones, y las calificaciones asignadas a los estudiantes son arbitrarias y, en su mayor parte, carentes de sentido (Ravela, 2009).

En Panamá y Costa Rica, un estudio comparativo de docentes de tercero y séptimo grado, que incluyó la grabación en videocinta de lecciones de matemáticas y la administración de un instrumento de evaluación que medía el dominio de la matemática que tenían los docentes y su conocimiento de la pedagogía de esta disciplina, puso al descubierto que los maestros tienen deficiencias preocupantes en ambas áreas (Sorto, Marshall, Luschei & Carnoy, 2008).

En la República Dominicana, una evaluación reciente del dominio matemático de los docentes y los rendimientos en matemáticas de sus alumnos, encontró que los primeros exhiben debilidades extraordinarias en su conocimiento del contenido correspondiente al cuarto grado de la educación básica (por ejemplo: solo cerca de la mitad de los docentes de cuarto grado en las provincias más importantes, Santiago y Santo Domingo, reconocieron que la fracción común $1/2$ es mayor que $1/3$). Asimismo, el estudio mostró debilidades comparativamente mayores en aquellas áreas de las matemáticas que también resultaban difíciles para los alumnos: por ejemplo la proporcionalidad, las fracciones comunes y decimales, elementos de estadística y probabilidades (Luna, Valverde & Roncagliolo, 2009).

Estudios de caso en Colombia indican que los docentes a veces no tienen una conciencia crítica de sus carencias en matemática o del efecto que ellas tienen sobre los estudiantes en sus aulas. En el caso de los profesores de álgebra de secundaria en Colombia, la evidencia muestra que parte de ellos tiene concepciones erróneas y carencia de conocimientos que son un obstáculo prominente en su labor docente. Al mismo tiempo, y de modo abrumador, los docentes responsabilizan a factores institucionales y/o contextuales por los bajos niveles de desempeño de sus alumnos (Agudelo-Valderrama, Clarke & Bishop, 2007).

La evidencia en Chile concuerda con el cuadro latinoamericano general: un estudio internacional marcó puntos de referencia de oportunidades para los estudiantes en el aprendizaje de la matemática y del impacto de estas oportunidades sobre el desempeño de los estudiantes en el TIMSS 1998/1999 frente a jurisdicciones con características económicas similares, pero con un desempeño académico superior (Corea del Sur, Malasia, Eslovaquia y las Escuelas Públicas del Condado de Miami/Dade en Estados Unidos). Un análisis estadístico multinivel (HLM)¹ mostró que la calidad de la enseñanza tuvo un impacto positivo sobre el rendimiento en matemáticas en todos esos países, pero la experiencia de los estudiantes en Chile se centró más en el aprendizaje de memoria y tuvieron menos oportunidades para aprender una matemática más exigente en lo cognitivo que sus pares en otros sistemas educativos (Ramírez, 2006).

La búsqueda de políticas para remediar la escasez

Evidentemente la escasez de una buena enseñanza es crítica y en el mundo entero se ha puesto en juego una variedad de iniciativas para favorecer la enseñanza de calidad (Hanushek, Kain, O'Brien & Rivkin, 2005; Ingersoll, 2007; Ochs, 2007).

En primer lugar, en muchos países, se han explorado políticas con respecto a la preparación inicial de docentes y a su formación continua en servicio. También se han usado políticas con respecto a la certificación o entrega de una licencia para enseñar. Aunque tales políticas pueden ser relevantes para incrementar la calidad de la enseñanza, la evidencia sugiere que son necesarias, pero ciertamente no suficientes (Boyd, Goldhaber, Lankford & Wyckoff, 2007; Boyd, Grossman, Lankford, Loeb & Wyckoff, 2008; Hanushek, Kain, O'Brien & Rivkin, 2005).

El problema en el ámbito latinoamericano con respecto a la preparación inicial de docentes es quizás más complejo que en otras regiones del mundo. Por ejemplo, en aquellos países en la región donde existen pruebas de postulación a la educación terciaria, se ha reportado con mucha frecuencia que el promedio obtenido en esas pruebas por parte los postulantes a carreras en ciencias de la educación es más bajo, a

.....

1 *Modelos lineales jerárquicos* en su sigla en inglés (Hierarchical Linear Models).

menudo significativamente, que los de quienes postulan a otras carreras². Pero esto es solo un indicador, otras miradas a la profesión docente han encontrado que en Perú los maestros que enseñan en el sector público poseen menor nivel educativo que aquellos que enseñan en el sector privado (Díaz & Saavedra, 2000); que los estudiantes de docencia en República Dominicana y Venezuela provienen de familias con menor nivel educativo que sus pares en otras carreras, y que en Chile las notas en el colegio de los estudiantes de educación son menores a las declaradas por estudiantes de otras profesiones (Navarro, 2002).

La calidad de la preparación de docentes, una vez admitidos a los programas de formación, también es baja a menudo. Como se señala en una importante revisión de la profesión docente en siete países de América Latina auspiciada por PREAL:

“...los casos examinados muestran un importante déficit de calidad en los conocimientos disciplinares enseñados en las instituciones de formación docente, así como escasa articulación con el conocimiento pedagógico y con la práctica docente. Todavía persisten importantes problemas a resolver en lo que respecta al material de apoyo, al desarrollo y la planificación curricular, a la supervisión de las prácticas docentes y a la preparación de los formadores de formadores”. (Vaillant & Rossel, 2006, p. 28).

Claramente el repertorio de políticas usualmente implementadas es insuficiente y muchos países han adoptado, o están considerando, políticas de evaluación de docentes para complementar o suplir sus políticas preexistentes (Isoré, 2009; Santiago & Benavides, 2009).

La evaluación docente como incentivo para la buena enseñanza

Quizás no sea controversial aseverar la relación entre la buena enseñanza y el logro académico. No sucede lo mismo, sin embargo, con respecto a la medición y el monitoreo de la calidad de la docencia. Esto es cierto en el mundo entero, y especialmente en América Latina, una región en la cual las formas tradicionales de evaluación del trabajo docente han sido débiles. Generalmente las evaluaciones de docentes llevadas a cabo por directores de escuelas, inspectores o personal técnico de los Ministerios, y similares, son referidas a criterios vagos, y a menudo se efectúan mecánicamente para cumplir requisitos formales, o se hacen deficientemente (Vaillant & Rossel, 2006). Todos estos tropiezos indican un interés limitado en la evidencia como base para el monitoreo de la calidad de la enseñanza y es frente a esta tradición regional que se introducen los nuevos esfuerzos de evaluación nacional de docentes.

La reticencia, en especial por parte de los maestros, frente a evaluaciones más rigurosas del trabajo docente, no es exclusiva ni mucho menos de América Latina.

Una encuesta reciente en EE. UU. (Coggshall, Ott & Lasagna, 2010) puede ser aleccionadora para América Latina. Esta encuesta indicó que los propios docentes exhiben importantes diferencias de opinión con los diseñadores de sistemas de evaluación docente y entre sí mismos, dependiendo especialmente de sus años de servicio. El debate, con obvios paralelos a la discusión en Chile, comienza por cómo definir en qué consiste la buena docencia y cómo medirla. En la encuesta a la que nos referimos, el indicador de calidad de enseñanza con mayor valoración por parte de docentes es el nivel de compromiso que exhiben los estudiantes con las actividades de la clase. Los años de experiencia importan mucho a la hora de considerar las opiniones de docentes. Entre todos los posibles indicadores de calidad presentados en la encuesta, el menos aceptado por los docentes es el uso del desempeño de sus estudiantes en pruebas estandarizadas como componente de la evaluación de la calidad docente. Este indicador es rechazado por más de la mitad de los docentes con 5 años o menos de experiencia en

.....

2 Curiosamente, aunque esto se me ha reportado a menudo en entrevistas en facultades de educación, ministerios y en oficinas administrativas universitarias, no he encontrado estudios al respecto.

el aula; sin embargo, solo un 32% de aquellos con más experiencia (20 años o más) están de acuerdo con sus colegas más jóvenes sobre este punto.

No solamente hay objeciones de principio por parte de los docentes a la evaluación de las competencias para enseñar. Investigaciones recientes han encontrado que algunos esquemas de evaluación docente han producido un incremento en los salarios de los maestros, sin efectos positivos sobre la calidad de la docencia. Estos estudios inclusive han encontrado que algunos esquemas de evaluación conducen a problemas en la retención de buenos docentes y a la exclusión de minorías étnicas y lingüísticas del magisterio (Angrist & Guryan, 2008; Boyd, Lankford, Loeb & Wyckoff, 2005). Esto es especialmente preocupante dado que es precisamente a los docentes más capaces a quienes se quiere estimular. Asimismo, ocurre que las minorías étnicas y lingüísticas tradicionalmente son más perjudicadas por las inequidades en la distribución de oportunidades educativas y su exclusión de la profesión docente podría precisamente exacerbar estas inequidades.

Frente a las situaciones que hemos descrito aquí, existen algunas lecciones que podemos derivar a la hora de pensar en políticas de evaluación docente deseables para América Latina.

Preguntas para América Latina

Detallar todos los retos que enfrenta la región al considerar o implementar sistemas de evaluación docente sobrepasa una contribución de este tipo—este libro reporta con detalle las decisiones que se han tomado en Chile y da cuenta del curso de su implementación. Como se observa, son muchos los detalles técnicos, los acuerdos políticos, y los recursos necesarios para desarrollar, refinar y extender sistemas de evaluación de este tipo.

Sin embargo, cualquier lista de preguntas fundamentales que debe abordar un sistema de evaluación docente debe incluir al menos las siguientes:

- ¿Con respecto a qué criterios se evaluará a los docentes?
- ¿Qué modalidades de levantamiento y análisis de evidencias son apropiados?
- ¿Qué papel jugará —o no— la evaluación de los conocimientos de los docentes?
- ¿Cómo se reconocerá que el fin último de la buena enseñanza es el aprendizaje?

Con respecto a los criterios de evaluación, uno de los retos que enfrenta todo tipo de medición desde el comienzo, es la definición de los criterios a los que se referirán las mediciones. En América Latina es relativamente nueva la idea de formular criterios de calidad empíricamente verificables, y esta perspectiva inclusive encuentra importantes grupos de detractores que la descartan como una orientación mecánica o “gerencialista”, cosa que ha ocurrido también en otros países, por ejemplo el Reino Unido (Mahony & Hextall, 2001). Sin embargo, la alternativa es prolongar una tradición de vaguedad e imprecisión en los criterios de calidad que de alguna manera se asocia a una noción de la enseñanza como **vocación** (semejante e históricamente asociada a la vocación religiosa) y no como **profesión**. En definitiva, las nuevas herramientas de evaluación docente, al dar primacía a la formulación de criterios de calidad empíricamente verificables, se orientan hacia la idea de *profesión docente* y no de *vocación docente*. Pese a importantes esfuerzos de resistencia a esta idea, tal parece que la política educativa en América Latina favorece en la actualidad la profesionalización del docente y los sistemas de evaluación también debieran verse como herramientas en ese proceso. Se discute mucho en la región acerca de cuán deseable es crear algo llamado una “cultura de evaluación”, y los avances en este sentido han sido lentos. Una razón que podría explicar los magros logros en este sentido es que el problema no reside necesariamente en la falta de una *cultura de evaluación*, sino más bien en la ausencia de una tradición de valorar y usar la evidencia empírica en la toma de decisiones en la política social o en la toma de decisiones en la administración

de establecimientos educativos. Desarrollar un sistema de evaluación docente, lo mismo que un sistema de evaluación estandarizada de estudiantes, implica aceptar la noción de que es importante verificar el cumplimiento de objetivos o criterios empíricamente, y que estas evidencias juegan un papel en la toma de decisiones. Esto es sumamente novedoso en América Latina en lo que respecta a la valoración del trabajo docente, y su aceptación a nivel regional es aún problemático.

Esclarecer los criterios de calidad y evaluarlos con transparencia favorece el buen curso de la evaluación y potencia sus posibilidades de estimular, favorecer, y recompensar la buena enseñanza. Sin embargo, también hace más posible la resistencia hacia su uso.

La enseñanza también es compleja y desafía la simplificación en unos cuantos objetivos e indicadores sencillos. Debido a ello, las discusiones acerca de los criterios de calidad son muy importantes. Aunque es posible que no haya grandes desacuerdos acerca de la idea central de este artículo —que la buena enseñanza importa mucho— si los habrá acerca de las formas de especificar cuáles son las características verificables de la buena enseñanza.

Volviendo a la encuesta a los docentes en EE.UU. mencionada anteriormente, queda claro que hay indicadores de la calidad valorados por los docentes (y muchos otros actores) que son difíciles de medir mediante pruebas o encuestas. Éstas involucrarían metodologías cualitativas como observación de aulas, videos, o similares. Lograr niveles aceptables de validez y confiabilidad en ese tipo de levantamientos de información es difícil y hacen falta considerables recursos. El objeto mismo de la medición en este caso también supone decisiones importantes. ¿Se pondrá atención únicamente en el maestro o la maestra? ¿Se evaluarán las decisiones que toman los docentes en términos del uso de recursos en la clase? ¿Se medirá también la relación docente–alumno o el “nivel de compromiso” de los estudiantes con las actividades de clase? ¿Cuántas clases requiere una muestra para dar suficiente sustento a inferencias acerca de cuán bien enseña o no el docente?

El sistema de evaluación docente chileno que hemos estudiado en este libro, propone una serie de soluciones a estos dilemas. En cuanto a los criterios de evaluación, sus instrumentos están alineados con una serie de criterios claros reunidos en el Marco para la Buena Enseñanza y, por tanto, se fundamenta en un modelo de la enseñanza explícito. Adopta también un modelo de “cuerpo de evidencias” en el cual múltiples tipos de evidencias, con respecto a distintos indicadores se consideran en conjunto. Estas evidencias, tomadas a partir de portafolios, pautas de autoevaluación, entrevistas de evaluadores pares e informes de referencia de directivos son considerados en conjunto, de acuerdo a una ponderación por tipo de evidencia que forma parte del Reglamento sobre Evaluación Docente. El liderazgo en materia de evaluación docente que Chile asume para América Latina es palpable en la descripción de los esfuerzos técnicos —y políticos— en estas áreas. Evidentemente, los desafíos técnicos y políticos de la evaluación no se resuelven “de una vez”: un sistema sólido de evaluación se compromete a refinar, extender y corregir sus prácticas técnicas y políticas en forma continua. De tal modo, la Evaluación Docente chilena será un referente obligatorio para América Latina y el mundo por mucho tiempo.

Muy importantes también son las preguntas que Chile está aún en proceso de contestarse, y que resumen otros retos que comparte con la región.

Un ejemplo de la agenda aún abierta es la evidencia en América Latina acerca de las dificultades en término del dominio de las asignaturas escolares por parte de los docentes. La evidencia de esto en la región es contundente. Páginas arriba únicamente describimos una muestra pequeña de la evidencia acerca del alcance y la severidad del problema. Siendo así, ¿cómo evitar pruebas de conocimientos en matemática, ciencias naturales, lengua española y otras materias para los docentes? La evidencia mundial en esto es clara: el conocimiento del docente es el factor que más claramente restringe o estimula su toma de decisiones pedagógicas. Dos tipos de conocimiento están en juego: el conocimiento de la asignatura

y el conocimiento de cómo enseñarla (Goldschmidt & Phelps, 2010; Kansanen, 2009; Kersting, 2008). Medir esto involucra no solamente importantes desafíos psicométricos, supone también desafíos políticos. La posición que adopte el sistema de evaluación docente en Chile frente a esto marcará una pauta importante para la región.

Para finalizar nuestra pequeña lista de algunas de las preguntas y desafíos que se enfrenta en el ámbito de la evaluación docente, resta abordar otro tema en la agenda abierta de Chile y la región, esto es, ¿qué consideración recibirá el aspecto definitorio de la buena enseñanza: el aprendizaje?

La justificación de cualquier definición de qué es un buen docente o una buena enseñanza, incluye aseverar que éstos afectan positivamente los aprendizajes de los estudiantes. Siendo esto definitorio, quizás es por ello lo más controversial. Cuando se objeta por la simplificación de la evaluación docente en un indicador, el indicador que se tiene en mente es el de los resultados de los estudiantes de los docentes en las pruebas estandarizadas. Sin embargo, es difícil justificar esta objeción si fundamentamos nuestras ideas en la noción de verificabilidad. ¿Qué evidencia de buena enseñanza puede omitir la referencia a verificaciones empíricas de qué y cuánto aprenden los estudiantes, sin suscitar objeciones razonables con respecto a su validez?

De nuevo, aquí hay un desafío importante. Primero, y quizás lo más preocupante, es que en América Latina casi nunca se evalúa el aprendizaje, sino que se evalúa el logro. Lo que podría parecer una esotérica distinción técnica, tiene sin embargo profundas implicancias prácticas para la evaluación docente.

La casi totalidad de las evaluaciones nacionales en América Latina mide aquello que los y las estudiantes saben en un punto determinado de su paso por la escuela, por ejemplo en cuarto grado de básica y quizás en otro grado de la educación media. Esto es medición del **logro**, es decir de aquello que se ha logrado hasta un punto determinado en la escolarización del estudiante. El aprendizaje es muy distinto. **Aprendizaje** es la diferencia entre lo que un estudiante sabe en un punto en el tiempo y aquello que termina sabiendo en otro. Por ejemplo, la diferencia entre lo que una alumna o un alumno sabe en la primera semana de 4° básico, con respecto a la última semana del mismo año escolar.

Difícilmente se puede justificar el asociar el logro demostrado por un estudiante en un punto del tiempo, con las acciones de un docente que solo estuvo en contacto con él durante el último año, porque dicho logro es un resultado acumulado de todas las experiencias del estudiante, desde que comenzó su escolarización hasta el momento de la medición. Muchos docentes tuvieron responsabilidad por el estudiante entre esos dos puntos. Asociar el logro con las oportunidades educativas que éste ha tenido únicamente en el año en que rinde la prueba, no solo es ilógico, sino que carece de validez estadística (Schneider, Carnoy, Kilpatrick, Schmidt & Schavelson, 2007).

El *aprendizaje*, sin embargo, se puede asociar con los esfuerzos de los docentes del año en que se toma la prueba, puesto que se refiere únicamente al progreso (o no) que haya experimentado el grupo de estudiantes en el período relevante. Pero inclusive esto es difícil de hacer bien, puesto que cada grupo de estudiantes es, para propósitos prácticos, una muestra muy variable de los estudiantes del sistema. Como bien lo saben los docentes, es común tener en un año dado grupos de estudiantes que representan desafíos inusuales para sus instructores. De hecho, revisiones de sistemas de evaluación de aprendizaje usados para la evaluación de profesores, indican que se requieren al menos 5 años de datos sobre el aprendizaje de grupos de estudiantes a su cargo, con respecto al docente que se quiera evaluar; y aun así, la robustez de los resultados puede ser problemática (Bock, Wolfe & Fisher, 1996).

La región también estará pendiente de cómo resuelve Chile su posición frente a esta pregunta. Y la discusión será compleja: el SIMCE ha sido históricamente (y por requisito legal) una evaluación del logro y no del aprendizaje. Los llamados actuales por el uso de los resultados SIMCE para la evaluación de

los docentes son, por tanto, técnica y éticamente inapropiados. ¿Interesará contar con evaluaciones del aprendizaje de los alumnos que puedan ser utilizados en la evaluación docente? ¿Cómo se resolverán los detalles logísticos y técnicos, como la necesidad antes mencionada de contar con no menos de 5 años de datos de aprendizaje para hacer inferencias mínimamente válidas y confiables acerca de la calidad del trabajo de un docente?

Estos desafíos son de fundamental importancia, y este libro revela el compromiso con el que se han asumido en Chile. Aquí nos hemos enfocado en detallar el panorama regional y puntualizar algunas de sus implicancias para una agenda que recién comienza a enfrentarse.

La buena enseñanza es importante, tan importante que importa evaluarla bien. Este libro revela cómo Chile ha asumido un liderazgo para la región en este importante esfuerzo.



10

Conclusiones

Jorge Manzi • Roberto González
Yulan Sun

Centro de Medición MIDE UC
Pontificia Universidad Católica de Chile

Conclusiones

En este libro hemos reunido y sistematizado la experiencia del Sistema de Evaluación del Desempeño Profesional Docente (en adelante Evaluación Docente), en Chile. En esta última sección del libro, destacaremos algunos aspectos tratados en los capítulos previos, y plantaremos desafíos futuros para potenciar y perfeccionar la evaluación docente chilena.

•

El primer capítulo de este libro sitúa el surgimiento de la Evaluación Docente en el marco de las políticas educacionales de las últimas dos décadas. Como destaca allí Bonifaz, esta evaluación fue el resultado de un largo proceso de análisis y discusión en torno al fortalecimiento de la profesión docente tema que emergió con fuerza desde el restablecimiento de la democracia en 1990. La conciencia de que el país debe mejorar la calidad y las condiciones laborales de los docentes es una materia en la que nuestra sociedad ha progresado lentamente en las últimas dos décadas. Esta conciencia ha sido estimulada por el continuo flujo de información acerca de los logros insuficientes de nuestros escolares en mediciones nacionales e internacionales, así como por la existencia de informes técnicos acerca del sistema educativo chileno (especialmente, el informe de la Comisión Brunner en los 90, el del Consejo Asesor Presidencial de 2006 y la revisión efectuada por la OECD en 2004).

Si bien nuestra sociedad y nuestras élites reconocen la importancia y valor de los docentes como actores claves del proceso educativo, en la práctica ello no se ha traducido en iniciativas que aborden sistemáticamente la cuestión docente. En concreto, ha existido un conjunto de medidas parcialmente conectadas, entre las que cabe destacar las mejoras salariales, la creación del Sistema Nacional de Evaluación del Desempeño de los Establecimientos Subvencionados (SNED) y el Programa de Fortalecimiento de la Formación Inicial Docente (FFID) en los años 90. Estos fueron seguidos, en la década pasada, por la creación de evaluaciones y estímulos individuales para profesores a través de los programas de Evaluación Docente, de Asignación de Excelencia Pedagógica y la Asignación Variable por Desempeño Individual (AVDI), además de la instalación de evaluaciones de egresados de pedagogía (programa Inicia). Las medidas más recientes son la creación de estímulos para atraer a los mejores egresados de la enseñanza media hacia la formación docente (becas Vocación de Profesor del Ministerio de Educación y campaña Elige Educar), así como la revisión y mejoramiento de estímulos asociados al desempeño docente. Sin embargo, está aún pendiente una materia fundamental como es la creación de una carrera profesional docente.

Visto con la perspectiva del tiempo, podemos decir hoy que las medidas hasta aquí implementadas abordan los tres nudos fundamentales para el fortalecimiento de la profesión docente: estamos buscando atraer a mejores jóvenes a la docencia, estamos preocupados de mejorar la formación que reciben en universidades e institutos profesionales, y hemos instalado mecanismos para reconocer y recompensar el buen desempeño docente. Sin embargo, no hemos llegado al punto en que estas distintas medidas formen parte de una política sistemática, sostenida y articulada.

En este contexto, la creación y existencia de la Evaluación Docente ha sido un hito muy significativo. Por una parte, como lo describe Bonifaz, su creación fue fruto de un extenso y complejo proceso de negociaciones que condujo a un acuerdo tripartito entre el gremio de los docentes, la Asociación de Municipios y el Ministerio de Educación. Estas negociaciones se originaron en el marco de las exigencias de evaluación del desempeño profesional que contemplaba el Estatuto Docente.

Cuando se examina el contexto internacional, se observa que una de las principales trabas para instalar dispositivos de evaluación del desempeño profesional docente es la resistencia del gremio docente. En muchos casos éste se ha negado a participar en negociaciones sobre el tema y en otros se ha opuesto a las iniciativas que han buscado implementarlas. El caso chileno es en este sentido excepcional, pues ha sido posible crear el sistema con la participación activa y constructiva de dicho gremio. Más aún, cuando la instancia tripartita acordó el sistema de evaluación, el Colegio de Profesores sometió la propuesta a una consulta nacional, constatándose una aprobación del mismo por dos terceras partes de los participantes. Es decir, se trata de un sistema que fue creado e instalado con participación del gremio, pese a que la evaluación ha contemplado, desde su origen, disposiciones que pueden conducir al término del contrato de docentes, cuando sus evaluaciones son persistentemente negativas.

El valor de la Evaluación Docente para la política educacional ha quedado demostrado en el rol que la información generada por esta evaluación ha tenido en la discusión pública acerca de los docentes. Como muestra el capítulo de Sun, Correa, Zapata y Carrasco, la distribución de los docentes en la evaluación ha mostrado que cerca de un tercio de ellos obtienen calificaciones por debajo del estándar de la evaluación (nivel de desempeño competente). Más aún, los resultados del portafolio han permitido identificar con precisión los ámbitos en los que es necesario focalizar las acciones para mejorar el desempeño de los profesores. Así, la Evaluación Docente ha contribuido a situar la discusión sobre la cuestión docente en un nuevo contexto, en el que se dispone de información pública y consistente acerca de las fortalezas y debilidades de estos profesionales, como base empírica para orientar la formación inicial y en servicio.

•

El valor de la Evaluación Docente como parte de la política pública deriva en forma importante de la opción adoptada al momento de crear el programa. Tal como señala el capítulo 2, de Flotts y Abarzúa, los sistemas de evaluación de profesores tienden a basarse en una de dos opciones: se fundan en un modelo basado en los resultados externos de la acción docente (logro de los estudiantes), o bien, lo hacen en medidas directas del desempeño profesional basadas en estándares de desempeño. La experiencia internacional de la última década está mostrando las fortalezas y debilidades relativas de estos enfoques.

En términos generales, se puede decir que los modelos que optan por evaluar a los docentes a través del logro de los estudiantes, tienen la ventaja de aparecer ante la sociedad como una medida más válida del desempeño, en tanto se entiende que el sentido fundamental del quehacer de estos profesionales es promover el aprendizaje de sus estudiantes. Sin embargo, la amplia evidencia acerca de las dificultades de atribuir los resultados de los estudiantes a la acción de un determinado profesor (incluso cuando se emplean los más avanzados modelos de valor agregado), demuestra que no es recomendable reducir la evaluación de docentes a una medida de su "efectividad".

En contraste, los sistemas de evaluación basados en estándares de desempeño, tienen la ventaja de focalizarse directamente en el docente evaluado, por lo que hay menos razones para dudar de la pertinencia de esta medida como reflejo de la calidad de su trabajo. La dificultad, en este caso, radica en dos condiciones: (a) contar con un marco de estándares adecuado, consistente con el conocimiento experto y legitimado en el sistema educativo, y (b) resolver adecuadamente las dificultades técnicas involucradas en la medición directa del desempeño. En el caso chileno, la primera de estas condiciones estuvo dada por la existencia del Marco para la Buena Enseñanza, documento que fue preparado con la participación de expertos nacionales e internacionales, y que contó con la legitimidad de una amplia discusión de los borradores iniciales del mismo por parte de los docentes en ejercicio. La segunda condición ha sido el principal desafío técnico de los primeros años de implementación de la Evaluación Docente, cuya principal preocupación reside en asegurar que los instrumentos de evaluación empleados, especialmente los que demandan un juicio profesional por parte de quienes los utilizan (Portafolio y

Entrevista por un Evaluado Par), sean aplicados tomando todos los resguardos necesarios para asegurar la confiabilidad y validez de las puntuaciones a que dan lugar (el capítulo de Flotts y Abarzúa, detalla el proceso de construcción y corrección de estos instrumentos, mientras que el de Taut, Santelices y Manzi aporta evidencia acerca de la validez de las puntuaciones que a partir de ellos se reportan).

Es importante enfatizar que la elección de un modelo basado en estándares tiene múltiples consecuencias positivas para la evaluación docente, que van más allá de su valor como dispositivo de evaluación. En primer lugar, por basarse en prescripciones explícitas acerca del desempeño, los sistemas de medición de este tipo permiten retroalimentar a los evaluados en forma precisa acerca de sus fortalezas y debilidades relativas, lo que entrega pistas directas a los profesores para identificar los aspectos que deben mejorar o fortalecer. En otras palabras, este tipo de sistema de evaluación tiene la virtud de hacer converger lo que se evalúa con respecto a lo que debe ser mejorado. En contraste, los sistemas de evaluación que solo retroalimentan a los docentes con respecto a su efectividad (por ejemplo, a partir de mediciones de logro de los estudiantes), pueden llegar a identificar docentes efectivos o no efectivos, pero no informan acerca de las razones del resultado reportado, lo que limita el valor de estas evaluaciones para fundamentar decisiones de desarrollo profesional por parte de los docentes evaluados y otros actores del sistema escolar (como los directivos, por ejemplo). En lo fundamental, estos sistemas pueden llegar a identificar cuándo un docente es efectivo, pero no pueden establecer las razones de su efectividad.

Un segundo argumento que respalda a las evaluaciones basadas en estándares, tiene relación con su foco en las prácticas de aula. Como señala Ravela en su comentario, el aspecto más relevante en las políticas de mejoramiento de la calidad docente, es la revisión y mejoramiento de las prácticas de aula. En la medida que esta evaluación aporta antecedentes directos sobre este ámbito, contribuye a identificar los patrones más regulares de la docencia en el aula (ver capítulo 8), lo que a su vez posibilita identificar necesidades de mejoramiento e incluso divulgar ejemplos de buenas prácticas.

•

En el capítulo 3, Sun, Calderón, Valerio y Torres abordan las complejidades asociadas al desarrollo e implementación de la Evaluación Docente a escala nacional. Los desafíos técnicos y logísticos que encierra el proceso son significativos, considerando la cantidad de profesores evaluados cada año, el volumen de material que debe ser movilizado hacia y desde las comunas, y los requerimientos que plantea a quienes cumplen roles claves (en las comunas y en los centros de corrección del portafolio). Una de las consecuencias prácticas de este despliegue anual es que gradualmente se han ido fortaleciendo las capacidades de gestión y evaluación en el nivel local, así como en los centros universitarios responsables de la corrección del portafolio.

Desde sus inicios, el diseño de la Evaluación Docente se propuso promover la transferencia de capacidades hacia los actores locales del proceso (especialmente los Coordinadores Comunales de Evaluación y Evaluadores Pares), lo que se refleja, especialmente, en el gran número de seminarios de capacitación que se realizan cada año. El fortalecimiento de las capacidades locales queda plasmado particularmente en el trabajo de las Comisiones Comunales de Evaluación, las que han asumido un rol crecientemente activo en el análisis de los antecedentes aportados por los instrumentos de la evaluación y en la determinación de la calificación final de cada docente. De igual forma, a medida que se ha consolidado la evaluación, es posible observar que, en forma creciente, las autoridades locales desarrollan autónomamente iniciativas para motivar a los docentes con buen desempeño o apoyar a aquellos que obtienen un mal resultado (algunas de las cuales son mencionadas en el capítulo 3). Por último, no debe olvidarse que, como efecto de la evaluación, cada año cerca de dos mil profesores son capacitados como evaluadores pares y correctores del portafolio.

•

Los resultados de la evaluación, que se sintetizan en el capítulo 4, demuestran con claridad la importancia de contar con una herramienta diagnóstica sobre el desempeño docente. El capítulo es especialmente ilustrativo cuando analiza, fundamenta y ejemplifica los aspectos que son evaluados mediante el portafolio. Este análisis tiene especial relevancia para las instituciones responsables de la formación inicial y en servicio de los docentes, pues allí se identifican con claridad los aspectos mejor desarrollados y los más deficitarios del trabajo cotidiano de los profesores con sus estudiantes. Es también relevante prestar atención a las diferencias entre profesores que se desempeñan en distintos niveles y sectores, pues en varios casos estas diferencias son muy marcadas.

Los resultados asociados a los distintos instrumentos de la Evaluación Docente también deben ser objeto de atención. Como muestra el capítulo de Sun, Correa, Zapata y Carrasco, la autoevaluación ha presentado en todos los años resultados muy elevados, dejando poco espacio para diferenciar el desempeño de los profesores. La evaluación de pares y de personal directivo, por su parte, también presentan promedios relativamente elevados, aunque permiten diferenciar en mayor grado a los docentes según su desempeño. En el marco del consenso existente en torno a fortalecer el rol de los directores como líderes pedagógicos, parece razonable revisar la forma en que directores y jefes de UTP participan en la evaluación, estableciendo procedimientos que favorezcan el ejercicio de la función evaluativa como parte regular de su rol conductor en la gestión pedagógica de las escuelas.

•

El capítulo 5, de Cortés y Lagos presenta una visión panorámica acerca de las consecuencias formativas de la Evaluación Docente, los Planes de Superación Profesional (PSP). El establecimiento de dispositivos de perfeccionamiento docente que derivan de los resultados de la evaluación (y son obligatorios para profesores calificados con el nivel Insatisfactorio y Básico), da cuenta del carácter formativo de ésta. Adicionalmente, a través de ellos se busca promover una mayor sintonía entre los programas de formación en servicio y las debilidades que se han detectado en los docentes evaluados. Como es sabido, en Chile, como en muchos otros países, la conexión existente entre necesidades profesionales y oferta de perfeccionamiento es más bien débil, por lo que el establecimiento de planes de formación que deban fundamentarse en un análisis específico de las debilidades de quienes participarán en ellos, es una muy buena noticia. Sin embargo, la tasa aún insuficiente de participación de los docentes en los PSP (especialmente de los calificados con nivel Básico), así como la baja proporción de planes que se basan en el uso de procedimientos formativos más efectivos, representan señales de alerta que deben ser consideradas para revisar y potenciar esta importante herramienta asociada a la Evaluación Docente. Los recientes cambios a la Evaluación Docente, que aumentan las consecuencias para quienes sean calificados con nivel Básico, seguramente incidirán en una mayor participación de éstos en los PSP, por lo que se hace relevante acelerar el mejoramiento de la oferta formativa asociada a éstos, así como la implementación de estudios que permitan evaluar su efectividad.

•

Taut, Santelices y Manzi dan cuenta en el capítulo 6 del importante esfuerzo de validación que ha acompañado al desarrollo de la Evaluación Docente. A partir de una revisión actualizada del importante concepto de validez en la medición educacional, los autores resumen los diferentes estudios hasta aquí desarrollados, los que abarcan desde aspectos que apuntan a la calidad de las puntuaciones reportadas (estudios acerca de la confiabilidad, generalizabilidad y análisis factorial de los instrumentos), hasta otros que aportan evidencia respecto a la validez de constructo y consecencial de la evaluación. Esta amplia agenda de validación revela la preocupación que ha existido por asegurar que este esfuerzo pionero cuente con revisiones periódicas de sus instrumentos, de manera de mejorarlos continuamente, y poder así respaldar los usos y consecuencias de la Evaluación Docente.

Los resultados reportados por Taut, Santelices y Manzi dan cuenta de que, en lo sustantivo, los instrumentos de la evaluación son internamente consistentes y que se emplean procedimientos adecuados para monitorear y garantizar la calidad de las puntuaciones que asignan diversos correctores en el portafolio. Sin embargo, las complejidades asociadas a la asignación de puntuaciones en instrumentos que requieren jueces entrenados, sugieren que se deben ampliar los esfuerzos por incrementar la consistencia entre distintos correctores, lo que idealmente debiera conducir a un aumento gradual en la tasa de doble corrección de los portafolios evaluados, hasta llevarla al 100% de doble corrección (hasta la fecha ese porcentaje solo alcanza al 20%).

•

En el capítulo 7, Valencia y Manzi analizan los factores de carácter personal y contextual que se asocian a mejores y peores desempeños docentes. Desde el inicio de la evaluación se han recabado antecedentes formativos y sociodemográficos de los profesores que participan en ella, lo que ha permitido enriquecer el análisis e interpretación de los resultados. Este capítulo presenta, primero de manera descriptiva y luego mediante el empleo de técnicas estadísticas avanzadas, una visión panorámica de los factores que más se asocian al desempeño profesional. En su conjunto, estos análisis revelan que los factores individuales tienen un peso relevante, destacando entre ellos factores sociodemográficos como el sexo (mejor rendimiento de mujeres) y la edad (mejor rendimiento en edades intermedias), así como los antecedentes formativos (mejores resultados en quienes reportan haber obtenido su título en carreras con programas de estudio más extensos y quienes han participado en programas de formación en servicio). En el caso de los factores contextuales, es relevante mencionar que la condición socioeconómica del establecimiento, así como de la comuna en que se localiza la escuela o liceo, no tienen mayor relación con el desempeño docente; esto indica, al menos parcialmente, que no habría una distribución especialmente sesgada de docentes según las características socioeconómicas de los estudiantes, aunque es importante considerar que la varianza socioeconómica en los establecimientos municipales es restringida.

•

La Evaluación Docente no solo ha servido para identificar las fortalezas y debilidades individuales o colectivas de los profesores evaluados, sino que también ha contribuido significativamente impulsando la investigación educacional. La existencia de varios miles de videos de clases o de documentos que muestran la forma en que los docentes planifican, evalúan y reflexionan acerca de su profesión, ha estimulado a distintos grupos de investigadores para indagar acerca de la realidad pedagógica en nuestro país. Cabe señalar que el acceso a los datos de la evaluación para fines académicos está mediado por procedimientos que garantizan el respeto a la confidencialidad de la información y se basan en un consentimiento expreso que los profesores voluntariamente firman para autorizar el uso de la evidencia que proporcionan con fines investigativos. Hasta el presente, cerca de un 80% de los profesores evaluados ha otorgado la autorización para el uso de sus datos, lo que ha hecho posible que varios grupos de investigadores puedan contar con un gran volumen de datos así como de evidencias directas de desempeño, para estudiar las prácticas docentes. La naturaleza nacional de estos datos hace posible la obtención de muestras representativas que de otra manera sería muy costoso lograr. Los resultados ya conocidos de varios estudios emprendidos en este contexto, entregan información muy valiosa acerca de los patrones instruccionales que predominan en las aulas chilenas, tal como lo muestra el capítulo 8, al reportar los estudios de Cornejo, Silva y Olivares; Galdames, Medina, San Martín, Gaete y Valdivia; Martinic; y Preiss. Estos estudios complementan y enriquecen lo ya observado en la evaluación, permitiéndonos conocer algunas de las razones que subyacen a las debilidades o fortalezas que ésta señala en el desempeño. Por ejemplo, a través del Portafolio, la Evaluación Docente ha mostrado que la interacción pedagógica entre profesores y estudiantes presenta importantes debilidades (ver capítulo de Sun, Correa, Zapata y Carrasco). Este resultado ha sido complementado y

enriquecido por la evidencia aportada por la investigación académica que se reporta en el capítulo 8, por ejemplo, acerca del tipo de habla que predomina en los docentes, el tipo de intervenciones que éstos hacen en su interacción con los estudiantes, o la forma en que los profesores manejan el tiempo durante sus clases. En varios casos, estos estudios realizan análisis complementarios cuyos resultados se correlacionan con aquellos de la Evaluación.

Cabe mencionar que las oportunidades de investigación educacional que se han abierto con los datos de la Evaluación Docente, se han reflejado en todos los instrumentos de apoyo a la investigación científica existentes en nuestro país. Es así como, hasta el presente, hay investigaciones ya finalizadas o en curso empleando datos de la evaluación con apoyo de FONDECYT, FONIDE, FONDEF y de los centros de investigación educacional establecidos por CONICYT. Una mención especial cabe para un proyecto FONDEF liderado por investigadores de la Escuela de Psicología y de la Facultad de Educación de la Universidad Católica de Chile. Este proyecto propone crear una videoteca de buenas prácticas docentes a partir de la revisión y análisis de videos de la Evaluación Docente, generando material pedagógico que pueda ser aprovechado por las entidades responsables de la formación inicial o en servicio de los docentes, así como por los profesores que busquen enriquecer o analizar sus prácticas. Este proyecto demuestra que la evaluación puede contribuir de manera directa al mejoramiento de las prácticas de aula a través de la identificación y divulgación de buenas prácticas.

Desafíos futuros para la Evaluación Docente

Aunque la Evaluación Docente chilena ha alcanzado un importante grado de consolidación, la revisión realizada en este libro también revela algunos desafíos que quisiéramos consignar al finalizar esta última sección de conclusiones. Algunos de esos desafíos aluden al sistema mismo de evaluación, mientras que otros se refieren a sus consecuencias, así como al contexto o entorno del programa.

- a. El amplio debate que se ha generado en años recientes acerca de la posibilidad de considerar los logros de los estudiantes como parte de los sistemas de evaluación docente, debiera llevar a plantearnos la misma pregunta en nuestro medio. Hemos advertido en este libro acerca de las limitaciones y complejidades que conlleva el uso de resultados de los estudiantes para evaluar a docentes, incluso cuando se recurre a los modelos más completos y sofisticados para analizar los datos de mediciones educacionales (modelos de valor agregado). Esta materia también fue abordada en los comentarios de Martínez, Ravella y Valverde. Las reservas que surgen del análisis de la experiencia internacional en este campo, sin embargo, no deben entenderse como un obstáculo para analizar la posibilidad de incorporar antecedentes de ese tipo en una versión futura de la evaluación. Nuestro planteamiento, que concuerda con lo expresado por la mayoría de los expertos que se han pronunciado sobre esta materia en el contexto del debate norteamericano, es que los resultados de los estudiantes deben ser solo una parte de dicha evaluación. Más aún, atendiendo a las dificultades o a la virtual imposibilidad de contar con información acerca del logro escolar en todos los ámbitos curriculares, se debe analizar la conveniencia de contemplar los logros de los estudiantes como un factor colectivo de la evaluación (asociado a la escuela o liceo en su conjunto), que pudiera combinarse con la información individual acerca del desempeño de cada docente.
- b. Con respecto a los instrumentos que actualmente forman parte de la evaluación, se debe analizar la conveniencia de introducir algunas revisiones o ajustes. En el caso del Portafolio, la recomendación más importante es la de avanzar hacia un mayor porcentaje de doble corrección, pues de esta manera se reduciría sustantivamente el riesgo de sesgo en las puntuaciones asociado a efectos de los correctores. En el caso del informe que emiten los directores, se debiera revisar el rol que éstos cumplen en la evaluación, potenciando su involucramiento en ella. Además de los cambios en las pautas que los directores deben completar, lo que se requiere es estimular que

éstos incorporen la evaluación como parte de sus prácticas cotidianas de liderazgo pedagógico, lo que debiera traducirse entre otras acciones, en la existencia de entrevistas de inicio y cierre del proceso evaluativo, así como del establecimiento de oportunidades para observar directamente el desempeño de los docentes. Existe evidencia que ello ya está ocurriendo, sin embargo, es necesario potenciarlo como práctica permanente de parte de los directores. Finalmente, es importante indicar que a lo largo de los años de implementación del programa de evaluación docente, se ha constatado la limitada capacidad discriminativa que tiene el instrumento de autoevaluación. Esta limitación, sin embargo, no se deriva necesariamente de la naturaleza del instrumento. De hecho, la autoevaluación constituye en sí misma un insumo muy relevante de considerar al momento de evaluar el conjunto de evidencias que da cuenta del desempeño docente, sin embargo, cuando ésta se vincula a consecuencias (como ocurre en esta evaluación), se observa que los docentes tienden en su mayoría autoevaluarse en términos muy favorables. Dado este comportamiento poco discriminativo en el tiempo y los efectos que introduce en los resultados de la misma, es recomendable que dicho instrumento siga siendo utilizado para proveer información valiosa del docente a la Comisión Comunal de Evaluación, pero no como insumo al momento de estimar la categoría final de desempeño de los docentes. Dado el carácter formativo de la evaluación docente, la información proveniente de la autoevaluación debiera contribuir principalmente a identificar los aspectos que para cada docente constituyen fortalezas o debilidades profesionales, y muy especialmente, para orientar recomendaciones y acciones de desarrollo profesional.

- c.** Con respecto a las consecuencias de la evaluación, se deben establecer mecanismos sistemáticos para monitorear la calidad de la implementación de las consecuencias previstas (especialmente las de tipo formativo), pero se deben contemplar también las consecuencias no previstas o negativas. La reciente investigación acerca de validez consecucional de la Evaluación Docente (ver capítulo 6, de Taut, Santelices y Manzi), arroja pistas importantes acerca del tipo de monitoreo que se debiera establecer. Derivado de esta investigación, se sugiere que se explicita el modelo de acción de la evaluación docente (lo que supone identificar y divulgar las consecuencias previstas de la evaluación), y que se establezca un programa de investigaciones que permita indagar en muestras de establecimientos, profesores, directores y sostenedores, el grado en que las consecuencias previstas se están materializando. Este tipo de estudios también debiera entregar pistas acerca de consecuencias no previstas, tanto positivas como negativas.
- d.** También se deben ampliar las acciones destinadas a lograr un mejor aprovechamiento de la información asociada a la Evaluación Docente. Aunque ya existen algunos buenos ejemplos de uso de la evaluación para retroalimentar y potenciar la formación pedagógica, estimular la investigación científica o alimentar la discusión pública en torno a los temas docentes y política educativa, en todos estos campos los avances hasta el presente son solo parciales. Se requiere un análisis para cada uno de estos contextos de manera que la riqueza de información disponible no se desaproveche. Es particularmente importante concentrar la atención en las instituciones formadoras de profesores, puesto que hasta el presente muy pocas de ellas han contemplado a la Evaluación Docente como parte de los antecedentes para fundamentar revisiones de sus planes de estudio.
- e.** Por último, no es posible dejar de mencionar el rol que la evaluación pueda tener en la futura carrera profesional docente. Hay un amplio consenso acerca de la necesidad de establecer una nueva carrera docente que sea atractiva y estimulante. En este contexto, se ha considerado que la Evaluación Docente puede tener un rol clave en la promoción dentro de esa carrera o para asignar incentivos. No obstante, es importante ser muy prudentes a la hora de asociar la evaluación a este tipo de consecuencias de alto impacto, pues es conocido que cuando los sistemas de evaluación elevan sus consecuencias, se incrementa el riesgo de malas prácticas que pueden, en el peor de los casos, conducir a una desacreditación del sistema de evaluación, tal como lo advierte el comentario de Martínez.

Al concluir este libro, es importante retornar al punto que originó la decisión de implementar una evaluación docente en Chile. Como se ha dicho, ello ocurrió en el marco de políticas educacionales que lenta, pero sostenidamente han conducido a una revalorización de la profesión docente como el eje central del mejoramiento de la calidad de la educación. Una vez establecida, la Evaluación Docente enriqueció la discusión en torno a las políticas docentes, aportando información precisa y verificable acerca de las virtudes y falencias de nuestros profesores, facilitando la priorización de recursos y opciones de política educacional. Más allá de los ajustes que pueda experimentar la actual evaluación, su rol debiera seguir siendo fundamental para monitorear y retroalimentar la efectividad de las políticas actualmente en curso y para orientar las futuras decisiones en este campo.

El trabajo presentado en este libro reafirma el valor fundamental de los profesores en las políticas de mejoramiento educativo en nuestro país. Luego de décadas en que la profesión docente ha sido relegada a lugares secundarios en términos de su reconocimiento material y simbólico, nuestra sociedad enfrenta ahora al desafío de asumir las implicancias que conlleva la elevación del valor social de la docencia. Queremos que nuestros mejores jóvenes vean en la docencia una carrera desafiante y atractiva, lo que a su vez requiere que sus padres y la sociedad en su conjunto estimulen, en lugar de desalentar, esta vocación. Necesitamos a su vez, un nuevo discurso en las elites y medios de comunicación, que supere la visión pesimista y desesperanzada acerca de esta profesión, que muchas veces lleva a la búsqueda de soluciones ilusorias y cortoplacistas. Necesitamos también que quienes tienen directa relación con la formación y desarrollo de la profesión docente asuman oportuna y eficazmente las exigencias que se producen en esta nueva fase de la política educativa, marcada por la certeza de que la calidad de nuestra educación está indisolublemente asociada a la calidad de nuestros docentes.

Referencias

- Agudelo–Valderrama, C., Clarke, B., & Bishop, A. J. (2007). Explanations of attitudes to change: Colombian mathematics teachers' conceptions of the crucial determinants of their teaching practices of beginning algebra. *Journal of Mathematics Teacher Education*, 10(2), 69–93.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Anastasi, A. (1986). Evolving concepts of test validation. *Annual Review of Psychology*, 37(1), 1–16. doi:10.1146/annurev.ps.37.020186.000245
- Anderson, L. W. (2004). *Increasing teacher effectiveness* (2ª Ed.). Paris, Francia: UNESCO International Institute for Educational Planning.
- Angrist, J. D., & Guryan, J. (2008). Does teacher testing raise teacher quality? Evidence from state certification requirements. *Economics of Education Review*, 27(5), 483–503. doi:10.1016/j.econedurev.2007.03.002
- Aronson, J., Zimmerman, J., & Carlos, L. (1999). Improving student achievement by extending school: Is it just a matter of time? WestEd, San Francisco, CA. Recuperado a partir de <http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED435127>
- Assael, J. (2009, Noviembre). *La profesión docente en Chile y su estatuto: Marco jurídico en disputa*. Panel la profesión docente: Análisis de casos latinoamericanos presentado en Congreso Internacional Presente y Futuro de la Profesión Docente en Colombia, Bogotá, Colombia.
- Assael, J., & Pavez, J. (2008). La construcción e implementación del sistema de evaluación del desempeño docente chileno: Principales tensiones y desafíos. *Revista Iberoamericana de Evaluación Educativa*, 1(2), 2008.
- Avalos, B. (2002). *Profesores para Chile: Historia de un proyecto*. Santiago, Chile: Ministerio de Educación.
- Baker, E. L., Barton, P. E., Darling–hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., Ravitch, D., et al. (2010). *Problems with the use of student test scores to evaluate teachers* (Briefing Paper No. 278). EPI Briefing Paper. Washington, DC: Economic Policy Institute. Recuperado a partir de <http://www.epi.org/publications/entry/bp278>
- Bandura, A. (1997). *Self–efficacy: The exercise of control*. New York, NY: W. H. Freeman.
- Barber, M., & Mourshed, M. (2007). *How the world's best–performing schools systems come out on top*. McKinsey, & Company.
- Bellei, C. (2009). Does lengthening the school day increase students' academic achievement? Results from a natural experiment in Chile. *Economics of Education Review*, 28(5), 629–640. doi:10.1016/j.econedurev.2009.01.008
- Berliner, D. C. (2005). The near impossibility of testing for teacher quality. *Journal of Teacher Education*, 56(3), 205–213. doi:10.1177/0022487105275904
- Bock, R. D., Wolfe, R. G., & Fisher, T. H. (1996). *A review and analysis of the Tennessee Value–Added Assessment System*. Nashville, TN: State of Tennessee Comptroller of the Treasury.
- Boyd, D., Goldhaber, D., Lankford, H., & Wyckoff, J. (2007). The effect of certification and preparation on teacher quality. *Future of Children*, 17(1), 45–68.

- Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2008). *Who Leaves? Teacher attrition and student achievement* (NBER Working Paper No. 14022). Recuperado a partir de <http://www.nber.org/papers/w14022>
- Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). Explaining the short careers of high-achieving teachers in schools with low-performing students. *American Economic Review*, *American Economic Review*, 95(2), 166–171.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academies Press.
- Braun, H. I. (2005). *Using student progress to evaluate teachers: A primer on Value-Added Models* (Policy Information Perspective No. PIC-VAM). Princeton, NJ: Educational Testing Service.
- Bravo, D., Falck, D., González, R., Manzi, J., & Peirano, C. (2008). *La relación entre la evaluación docente y el rendimiento de los alumnos: Evidencia para el caso de Chile*. Santiago, Chile: Centro de Microdatos, Departamento de Economía, Universidad de Chile. Recuperado a partir de http://www.microdatos.cl/docto_publicaciones/Evaluacion%20docentes_rendimiento%20escolar.pdf
- Brennan, R. L. (2001). *Generalizability theory*. New York, NY: Springer.
- Briggs, D., & Domingue, B. (2011). *Due diligence and the evaluation of teachers: A review of the value added analysis underlying the effectiveness rankings of Los Angeles unified school district teachers* by the Los Angeles Times. Boulder, CO: National Education Policy Center.
- Bryk, A. S., & Raudenbush, S. W. (2001). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: SAGE Publications, Inc.
- Charlesworth, R., Hart, C. H., Burts, D. C., & Hernández, S. (1991). Kindergarten teachers beliefs and practices. *Early Child Development and Care*, *70*(1), 17–35. doi:10.1080/0300443910700103
- Clark, I. (2008). Assessment is for learning: Formative assessment and positive learning interactions. *Florida Journal of Educational Administration, & Policy*, *2*(1), 1–16.
- Coggs, J. G., Ott, A., & Lasagna, M. (2010). *Retaining teacher talent: Convergence and contradictions in teachers* (pp 24). Washington, DC: Learning Point Associates, & Public Agenda.
- Coll, C.. (2003). Constructivismo y educación: La concepción constructivista de la enseñanza y el aprendizaje. En C. Coll, J. Palacios, & A. Marchesi (Eds.), *Desarrollo psicológico y educación: Psicología de la educación* (Vol. 2). Madrid, España: Alianza.
- Coll, C., & Edwards, D. (Eds.). (1997). *Teaching, learning and classroom discourse: Approaches to the study of educational discourse*. Madrid, España: Fundación Infancia y Aprendizaje.
- Collins, A. (2006). Cognitive apprenticeship. En R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pps 47–60). New York, NY: Cambridge University Press.
- de Corte, E. (2010). Historical developments in the understanding of learning. En H. Dumont, D. Istance, & F. Benavides (Eds.), *The nature of learning: Using research to inspire practice, educational research and innovation* (pp 35–67). Paris, Francia: OECD Publishing.
- Cortés, F., Taut, S., Santelices, M. V., & Lagos, M. J. (2011, Enero). *Formación continua en profesores y la experiencia de los Planes de Superación Profesional (PSP) en Chile: Fortalezas y debilidades a la luz de la evidencia internacional*. Paper presentado en Segundo Encuentro Anual de la Asociación Chilena de Políticas Públicas, Santiago, Chile.
- Cotton, K. (1989). *Educational time factors* (Close-Up No. 8). School Improvement Research Series (SIRS). Portland, OR: Northwest Regional Educational Laboratory. Recuperado a partir de http://notendur.hi.is/joner/eaps/wh_etim8.htm

- Cronbach, L. (1989). Construct validation after thirty years. En R. L. Linn (Ed.), *Intelligence: Measurement, theory, and public policy* (pps 147–171). Urbana, OH: University of Illinois Press.
- Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58(4), 438–481. doi:10.3102/00346543058004438
- Cross, K. P. (1998). Classroom research: Implementing the scholarship of teaching. *New Directions for Teaching and Learning*, 1998(75), 5–12. doi:10.1002/tl.7501
- Cudeck, R. (2000). Exploratory factor analysis. En H. Tinsley, & S. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pps 265–296). San Diego, CA: Academic Press.
- Cueto, S., Ramirez, C., & Leon, J. (2006). Opportunities to learn and achievement in mathematics in a sample of sixth grade students in Lima, Peru. *Educational Studies in Mathematics*, 62(1), 25–55.
- Danielson, C. (1996). *Enhancing professional practice: A framework for teaching* (1st ed.). Alexandria, VA: Association for Supervision and Curriculum Development.
- Darling–Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education policy analysis*, 8(1). Recuperado a partir de <http://epaa.asu.edu/ojs/article/view/392>
- Darling–Hammond, L., Berry, B., & Thoreson, A. (2001). Does teacher certification matter? Evaluating the evidence. *Educational Evaluation and Policy Analysis*, 23(1), 57–77. doi:10.3102/01623737023001057
- Darling–Hammond, L., Wei, R. C., & Johnson, C. M. (2009). Teacher preparation and teacher learning. A changing policy landscape. En G. Sykes, B. L. Schneider, & D. N. Plank (Eds.), *Handbook of education policy research* (pps 596–612). New York, NY: Taylor, & Francis.
- Darling–Hammond, L., & Youngs, P. (2002). Defining «highly qualified teachers»: What does «scientifically-based research» actually tell us? *Educational Researcher*, 31(9), 13–25.
- Dee, T. S., & Keys, B. J. (2004). Does merit pay reward good teachers? Evidence from a randomized experiment. *Journal of Policy Analysis and Management*, 23(3), 471–488. doi:10.1002/pam.20022
- DeFord, D. E. (1985). Validating the construct of theoretical orientation in reading instruction. *Reading Research Quarterly*, 20(3), 351–367. doi:10.2307/748023
- Delhaxhe, A. (1997). Le temps comme unité d'analyse dans la recherche sur l'enseignement. *Revue Française de Pédagogie*, 118(1), 106–125. doi:10.3406/rfp.1997.1180
- Díaz, H., & Saavedra, J. (2000). *La carrera de maestros: Factores institucionales, incentivos económicos y desempeño* (Research Network working paper No. R–410). Washington, DC: Banco Interamericano de Desarrollo. Recuperado a partir de <http://ideas.repec.org/p/idb/wpaper/3109.html>
- Dillon, S. (2010a, Diciembre 3). Video eye aimed at teachers in 7 school systems. *The New York Times*. Recuperado a partir de <http://www.nytimes.com/2010/12/04/education/04teacherside.html>
- Dillon, S. (2010b, Diciembre 3). Teacher evaluations may get a video assist. *The New York Times*. Recuperado a partir de <http://www.nytimes.com/2010/12/04/education/04teacher.html>
- Ding, C., & Sherman, H. (2006). Teaching effectiveness and student achievement: Examining the relationship. *Educational Research Quarterly*, 29(4), 39–49.
- Edwards, D., & Mercer, N. (1987). *Common knowledge: The development of understanding in the classroom*. London: Routledge.
- Eisenberg, N. V. (2008). *The performance of teacher in Chilean public elementary schools: Exploring its relationship with teacher backgrounds and student achievement, and its distributions across schools and municipalities* (Disertación doctoral no publicada). University of California Los Angeles.
- Fenstermacher, G. D., & Richardson, V. (2005). On making determinations of quality in teaching. *The Teachers College Record*, 107(1), 186–213.

- Fenstermacher, G. D. (1994). The Knower and the Known: The nature of knowledge in research on teaching. *Review of Research in Education*, 20(1), 3–56. doi:10.3102/0091732X020001003
- Fernández, C., & Yoshida, M. (2004). *Lesson study: A Japanese approach to improving mathematics teaching and learning*. Mahwah, NY: Lawrence Erlbaum Associates.
- Fullan, M., & Stiegelbauer, S. M. (1997). *El cambio educativo: Guía de planeación para maestros*. Ciudad de México: Trillas.
- Givvin, K. B., Hiebert, J., Jacobs, J. K., Hollingsworth, H., & Gallimore, R. (2005). Are there national patterns of teaching? Evidence from the TIMSS 1999 video study. *Comparative Education Review*, 49(3), 311–343.
- Glazerman, S., Loeb, S., Goldhaber, D., Staiger, D., & Raudenbush, S. (2010). *Evaluating teachers: The important role of value-added*. Washington, DC: Brown Center on Education Policy.
- Goldhaber, D. D., & Brewer, D. J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis*, 22(2), 129–145. doi:10.3102/01623737022002129
- Goldhaber, D., Perry, D., & Anthony, E. (2004). The National Board for Professional Teaching Standards (NBPTS) process: Who applies and what factors are associated with NBPTS certification? *Educational Evaluation and Policy Analysis*, 26(4), 259–280.
- Goldhaber, D., & Anthony, E. (2011). Can teacher quality be effectively assessed? National Board certification as a signal of effective teaching. *Review of Economics and Statistics*, 89(1), 134–150. doi:10.1162/rest.89.1.134
- Goldhaber, D., & Hansen, M. (2008). *Is it just a bad class? Assessing the stability of measured teacher performance* (CRPE Working Paper No. 2008–5). Seattle, WA: Center on Reinventing Public Education, University of Washington.
- Goldschmidt, P., & Phelps, G. (2010). Does teacher professional development affect content and pedagogical knowledge: How much and for how long? *Economics of Education Review*, 29(3), 432–439. doi:10.1016/j.econedurev.2009.10.002
- Goodrich, H. (1997). Understanding rubrics. *Educational Leadership*, 54(4), 14–17.
- Grupo de Trabajo sobre Estándares y Evaluación PREAL. (2009). Boletín N°9 Observatorio Regional de Políticas De Evaluación Educativa. Recuperado a partir de http://www.preal.org/Grupo2.asp?Id_Grupo=3
- Hakel, M. D., Koenig, J. A., & Elliott, S. W. (Eds.). (2008). *Assessing accomplished teaching: Advanced-level certification programs*. Washington, DC: National Academies Press.
- Hanushek, E. A., Kain, J. F., O'Brien, D. M., & Rivkin, S. G. (2005). *The market for teacher quality* (NBER Working Paper No. 11154). Cambridge, MA: National Bureau of Economic Research. Recuperado a partir de <http://www.nber.org/papers/w11154>
- Harris, D., & McCaffrey, C. (2010). *Value-added: Assessing teachers' contributions to student achievement*. En M. M. Kennedy (Ed.), *Handbook of teacher assessment and teacher quality*. San Francisco: Jossey Bass.
- Harris, D. N., & Sass, T. R. (2007). What makes for a good teacher and who can tell? *Unpublished manuscript, Florida State*. Recuperado a partir de http://www.caldercenter.org/upload/harris_sass.pdf
- Hart, D. (1994). *Authentic assessment: A handbook for educators*. White Plains, NY: Dale Seymour Publications.
- Heneman III, H. G., Milanowski, A., Kimball, S. M., & Odden, A. (2006). *Standards-based teacher evaluation as a foundation for knowledge-and skill-based pay* (CPRE Policy Brief No. RB–45). Philadelphia, PA: University of Pennsylvania, Graduate School of Education. Recuperado a partir de http://www.cpre.org/index.php?option=com_content&task=view&id=68&Itemid=73

- Heneman, H., & Milanowski, A. (2004). Alignment of human resource practices and teacher performance competency. *Peabody Journal of Education*, 79(4), 108–125. doi:10.1207/s15327930pje7904_6
- Herman, J. L., Aschbacher, P. R., & Winters, L. (1992). *A practical guide to alternative assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Hernan, J., & Baker, E. (2009). Assessment policy: Making sense of the babel. *Handbook of education policy research* (pps 176–190). New York, NY: Routledge.
- Hiebert, J., Gallimore, R., Garnier, H., Givvin, K. B., Hollingsworth, H., Jacobs, J., Chui, A. M.–Y., et al. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video study*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Hindman, A. H., & Wasik, B. A. (2008). Head Start teachers' beliefs about language and literacy instruction. *Early Childhood Research Quarterly*, 23(4), 479–492. doi:10.1016/j.ecresq.2008.06.002
- Hoyle, R. H., & Duvall, J. L. (2004). Determining the number of factor in exploratory and confirmatory factor analysis. En D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the Social Sciences* (pps 301–315). Thousand Oaks, CA: SAGE.
- Ingersoll, R. M. (2007). *A comparative study of teacher preparation and qualifications in six nations*. Philadelphia, PA: Consortium for Policy Research in Education. Recuperado a partir de <http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED498318>
- Ingvarson, L., & Rowe, K. (2008). Conceptualising and evaluating teacher quality: Substantive and methodological issues. *Australian Journal of Education*, 52(1), 5–35. doi:10.1016/j.apmr.2010.02.005
- Isoré, M. (2009). *Teacher evaluation: Current practices in OECD countries and a literature review* (OECD Education Working Papers No. 23). Paris, Francia: OECD Publishing.
- Jacob, B., & Lefgren, L. (2005). *Principals as agents: Subjective performance measurement in education* (No. W11463). NBER Working Paper. Cambridge, MA: National Bureau of Economic Research. Recuperado a partir de <http://ssrn.com/abstract=755699>
- Jacob, B., & Lefgren, L. (2006). When principals rate teachers. *Education Next*, 26(2), 59–69.
- Jacob, B., & Lefgren, L. (2008). Can principals identify effective teachers? Evidence on subjective performance evaluation in education. *Journal of Labor Economics*, 26, 101–136. doi:10.1086/522974
- Joint Committee on Standards for Educational Evaluation. (2009). *The personnel evaluation standards: How to assess systems for evaluating educators* (2ª Ed.). Thousand Oaks, CA: Corwin Press.
- Kagan, D. M. (1992). Professional growth among preservice and beginning teachers. *Review of Educational Research*, 62(2), 129–169. doi:10.3102/00346543062002129
- Kane, M. T. (2001). Current concerns in validity theory. *Journal of Educational Measurement*, 38(4), 319–342. doi:10.1111/j.1745–3984.2001.tb01130.x
- Kane, M. T. (2006). Validation. En R. L. Brennan (Ed.), *Educational measurement* (4th ed., pps 17–64). Westport, CT: Praeger Publishers.
- Kansanen, P. (2009). Subject–matter didactics as a central knowledge base for teachers, or should it be called pedagogical content knowledge? *Pedagogy, Culture, & Society*, 17(1), 29. doi:10.1080/14681360902742845
- Kersting, N. (2008). Using video clips of mathematics classroom instruction as item prompts to measure teachers' knowledge of teaching mathematics. *Educational and Psychological Measurement*, 68(5), 845–861. doi:10.1177/0013164407313369
- Koretz, D. M. (2008). *Measuring up: What educational testing really tells us*. Cambridge, MA: Harvard University Press.

- Linn, R. L. (1997). Evaluating the validity of assessments: The consequences of use. *Educational Measurement: Issues and Practice*, 16(2), 14–16.
- Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher*, 20(8), 15–21. doi:10.3102/0013189X020008015
- LLECE. (2010). *SERCE: Factores asociados al logro cognitivo de los estudiantes de América Latina y el Caribe*. Santiago, Chile: UNESCO / OREALC.
- Lockwood, J. R., McCaffrey, D. F., Hamilton, L. S., Stecher, B., Le, Vi-Nhuan, L., & 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.
- Luke, D. A. (2004). *Multilevel modeling*. Quantitative Application in the Social Sciences. Thousand Oaks, CA: SAGE Publications, Inc.
- Luna, E. A., Valverde, G., & Roncagliolo, R. (2009). *Informe de evaluación externa. Reporte TEF-Matemáticas 2009*. Santiago de los Caballeros, República Dominicana: Pontificia Universidad Católica Madre y Maestra.
- Mahony, P., & Hextall, I. (2001). «Modernizing» the teacher. *International Journal of Inclusive Education*, 5(2/3), 133.
- Manzi, J. (2007a, Octubre). *Evaluación Docente: Antecedentes, resultados y proyecciones*. Presentado en Seminario Evaluación Docente en Chile: Fundamentos, experiencias y resultados, Santiago, Chile. Recuperado a partir de <http://mideuc.cl/seminario/presentaciones.html>
- Manzi, J. (2007b). Gestión del conocimiento generado por la evaluación docente: desafíos para las universidades, los sostenedores, y diseñadores de políticas. En D. Doren (Ed.), *Nuevos desafíos para el desarrollo profesional docente* (pps 41–48). Santiago, Chile: Unidad de Comunicaciones del Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas, Ministerio de Educación.
- Manzi, J., Preiss, D. D., González, R., & Sun, Y. (2008, Marzo). *Design and implementation of a national project of teaching assessment*. Paper presentado en Annual Meeting of the American Educational Research Association, New York, NY.
- Manzi, J., Strasser, K., San Martin, E., & Contreras, D. (2007). *Quality of education in Chile: Final report of the Interamerican Development Bank project*. Washington, DC: BID. Recuperado a partir de <http://www.iadb.org/res/laresnetwork/files/pr300finaldraft.pdf>
- Marchesi, A., & Martín, E. (1998). *Calidad de la enseñanza en tiempos de cambio*. Madrid, España: Alianza.
- Mares, G., Guevara, Y., Rueda, E., Rivas, O., & Rocha, H. (2004). Análisis de las interacciones maestra-alumnos durante la enseñanza de las ciencias naturales en primaria. *Revista Mexicana de Investigación Educativa*, 9(22), 721–745.
- Martinic, S., & Vergara, C. (2007). Gestión del tiempo e interacción del profesor-alumno en la sala de clases de establecimientos de Jornada Escolar Completa en Chile. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 5(5), 3–20.
- Marzano, R. J., D’Arcangelo, M., & Sheets, D. (2004). *Classroom management that works DVD and facilitator’s guide*. Association for Supervision and Curriculum Development.
- Marzano, R. J., Marzano, J. S., & Pickering, D. J. (2003). *Classroom management that works: Research-based strategies for every teacher*. Alexandria, VA: Association for Supervision and Curriculum Development.
- McKenzie, P., & Santiago, P. (2005). *Teachers matter: Attracting, developing and retaining effective teachers*. Paris, Francia: OECD Publishing.

- McMillan, J. H. (2005). Understanding and improving teachers' classroom assessment decision making: Implications for theory and practice. *Educational Measurement: Issues and Practice*, 22(4), 34–43. doi:10.1111/j.1745–3992.2003.tb00142.x
- Mehrens, W. A. (1989). Combining evaluation data from multiple sources. En J. Millman, & L. Darling–Hammond (Eds.), *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers* (pps 322–336). Newbury Park, CA: Sage.
- Mehrens, W. A. (2005). Using performance assessment for accountability purposes. *Educational Measurement: Issues and Practice*, 11(1), 3–9. doi:10.1111/j.1745–3992.1992.tb00220.x
- Mercer, N. (2010). The analysis of classroom talk: Methods and methodologies. *The British Journal of Educational Psychology*, 80(Pt 1), 1–14. doi:10.1348/000709909X479853
- Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking: A sociocultural approach*. London: Routledge.
- Messick, S. (1989). Validity. En R. L. Linn (Ed.), *Educational Measurement* (3ª Ed., pps 13–110). New York, NY: MacMillan.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23(2), 13–23. doi:10.3102/0013189X023002013
- Messick, S. (1995). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice*, 14(4), 5–8.
- MET Project. (2010a). *Working with teachers to develop fair and reliable measures of effective teaching*. Bill, & MeL. Gates Foundation. Recuperado a partir de <http://www.metproject.org/downloads/met-framing-paper.pdf>
- MET Project. (2010b). *Validation engine for observational protocols*. Bill, & MeL. Gates Foundation. Recuperado a partir de http://www.metproject.org/downloads/Validation_Engine_concept_paper_092410.pdf
- MIDE UC. (2010). *Informe final de Planes de Superación Profesional: Resultados del estudio PSP 2009* (Documento Interno MIDE UC). Santiago, Chile: Pontificia Universidad Católica de Chile, Escuela de Psicología, Centro de Medición MIDE UC.
- Milanowski, A. T., Steven M. K., & Allan O. (2005). Teacher accountability measures and links to learning. En L. Stiefel, A. E. Schwartz, R. Rubenstein, J. Zabel (Eds.), *Measuring school performance and efficiency: Implications for practice and research* (Yearbook of the American Education Finance Association) (pp. 137-161).
- Ministerio de Educación. (2000). *Estándares de desempeño para la formación inicial de docentes*. División de Educación Superior, Programa de Fortalecimiento de la Formación Inicial Docente.
- Ministerio de Educación. (2003). *Marco para la buena enseñanza*. Santiago, Chile: Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas.
- Ministerio de Educación. (2008). *Informe nacional de resultados SIMCE 2007*. Santiago, Chile: Unidad de Curriculum y Evaluación.
- Ministerio de Educación. (2009). *Informe nacional de resultados SIMCE 2008*. Santiago, Chile: Unidad de Curriculum y Evaluación.
- Ministerio de Educación. (2010). *Documento estadísticas de la educación en Chile 2009*. Departamento de Estudios y Desarrollo. Recuperado a partir de http://w3app.mineduc.cl/DedPublico/anuarios_estadisticos
- Myford, C. M., & Engelhard, G. (2001). Examining the psychometric quality of the National Board for Professional Teaching Standards Early Childhood/Generalist assessment system. *Journal of Personnel Evaluation in Education*, 15(4), 253–85.

- Navarro, J. C. (2002). *Quiénes son los maestros? Carreras e incentivos docentes en América Latina*. Washington, DC: Banco Interamericano de Desarrollo.
- Nichols, S. L., & Berliner, D. C. (2005). *The inevitable corruption of indicators and educators through High-Stakes testing*. Executive summary. Tempe, AZ: Education Policy Studies Laboratory, Arizona State University. Recuperado a partir de <http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED508510>
- Núñez, I. (2000). Los docentes y la política de desarrollo profesional. Ministerio de Educación.
- Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26(3), 237–257. doi:10.3102/01623737026003237
- Nystrand, M., & Gamoran, A. (1991). Instructional discourse, student engagement, and literature achievement. *Research in the Teaching of English*, 25(3), 261–290.
- Ochs, K. (2007). Implementation of the Commonwealth Teacher Recruitment Protocol: Considering the education systems and context. *Perspectives in Education*, 25(2), 15–24.
- Odden, A. (2004). Lessons learned about standards-based teacher evaluation systems. *Peabody Journal of Education*, 79(4), 126–137. doi:10.1207/s15327930pje7904_7
- OECD. (2009). *Creating effective teaching and learning environments: First results from TALIS*. Paris, Francia: OECD Publishing.
- Papay, J. P. (2010). Different tests, different answers: The stability of teacher Value-Added estimates across outcome measures. *American Educational Research Journal*. doi:10.3102/0002831210362589
- Popham, W. J. (1997). Consequential validity: Right concern—wrong concept. *Educational Measurement: Issues and Practice*, 16(2), 9–13. doi:10.1111/j.1745–3992.1997.tb00586.x
- Preiss, D. D. (2005). *The task of teaching: Folk pedagogy and instructional practice in Chile's elementary education* (Disertación doctoral no publicada). Yale University, New Haven, CT.
- Preiss, D. D. (2009). The Chilean instructional pattern for the teaching of language: A video-survey study based on a national program for the assessment of teaching. *Learning and Individual Differences*, 19(1), 1–11. doi:10.1016/j.lindif.2008.08.004
- Preiss, D. D. (2010). Folk pedagogy and cultural markers in teaching: Three illustrations from Chile. En R. J. Sternberg, & D. D. Preiss (Eds.), *Innovations in educational psychology: Perspectives on learning, teaching, and human development* (pp 325–356). New York, NY: Springer Publishing Company.
- Preiss, D. D., San Martín, E., Alegria, I., Espinoza, A. M., Nuñez, M., & Ponce, L. (2011). *Enseñar matemáticas fomentando la apropiación privada de términos y procedimientos: Estructura y discurso en el aula matemática de primer ciclo*. Manuscrito en preparación.
- Prince, C. D., Koppich, J., Morse, T., Bhatt, M., & Witham, P. J. (s.d.). *How well do principals' evaluations of teachers predict student achievement outcomes?* (Research Synthesis). Center for Educator Compensation Reform. Recuperado a partir de http://cecr.ed.gov/researchSyntheses/Research_Synthesis_Q_D19.pdf
- Radovic, D., & Preiss, D. D. (2010). Patrones de discurso observados en el aula de matemática de segundo ciclo básico en Chile. *Psyche (Santiago)*, 19(2), 65–79. doi:10.4067/S0718–22282010000200007
- Ramírez, M. J. (2006). Understanding the low mathematics achievement of Chilean students: A cross-national analysis using TIMSS data. *International Journal of Educational Research*, 45(3), 102–116. doi:10.1016/j.ijer.2006.11.005
- Raudenbush, S. W. (2004). What are value-added models estimating and what does this imply for statistical practice? *Journal of Educational and Behavioral Statistics*, 29(1), 121–129.

- Ravela, P. (2009). Consignas, devoluciones y calificaciones: Los problemas de la evaluación en las aulas de educación primaria en América Latina. *Ppinas de Educación*, 2, 49–89.
- Rice, J. K. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes*. Washington, DC: Economic Policy Institute.
- Richardson, V., Anders, P., Tidwell, D., & Lloyd, C. (1991). The relationship between teachers' Beliefs and practices in reading comprehension instruction. *American Educational Research Journal*, 28(3), 559–586. doi:10.3102/00028312028003559
- Robalino, M., & Körner, A. (2005). *Condiciones de trabajo y salud docente*. Santiago, Chile: ORELAC/ UNESCO.
- Rockoff, J. E., & Speroni, C. (2010). Subjective and objective evaluations of teacher effectiveness. *American Economic Review*, 100(2), 261–266. doi:10.1257/aer.100.2.261
- Rosenthal, V. (2004). Microgenesis, immediate experience and visual processes in reading. En A. Carsetti (Ed.), *Seeing, thinking and knowing: Meaning and self-organization in visual cognition and thought* (pps 221–243). Dordrecht, The Netherlands: Kluwer Academic.
- Rubin, D. B., Stuart, E. A., & Zanutto, E. L. (2004). A potential outcomes view of Value-Added assessment in education. *Journal of Educational and Behavioral Statistics*, 29(1), 103–116. doi:10.3102/10769986029001103
- San Martin, E., Gonzalez, J., & Tuerlinckx, F. (2009). Identified parameters, parameters of interest and their relationships. *Measurement: Interdisciplinary Research and Perspectives*, 7(2), 97–105.
- Sanders, W. L., & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center. Recuperado a partir de [http://www.mccsc.edu/~\sim\\$curriculum/cumulative and residual effects of teachers.pdf](http://www.mccsc.edu/~\sim$curriculum/cumulative%20and%20residual%20effects%20of%20teachers.pdf)
- Santelices, M. V., & Taut, S. (2011). Convergent validity evidence regarding the validity of the Chilean standards-based teacher evaluation system. *Assessment in Education: Principles, Policy, & Practice*, 18(1), 73. doi:10.1080/0969594X.2011.534948
- Santelices, M. V., Taut, S., & Valencia, E. (2009). *Relación entre los resultados de la Evaluación Docente y los Planes de Superación Profesional: Estudio descriptivo*. Internal document MIDE UC. Santiago, Chile: Pontificia Universidad Católica de Chile, Escuela de Psicología, Centro de Medición MIDE UC.
- Santelices, M. V., Valencia, E., Taut, S., & Manzi, J. (2010, Abril 3). *The importance of contextual and personal variables in explaining teacher quality: Evidence using a standards-based, standardized measure of teacher performance*. Poster session presentado en Annual Meeting of the American Educational Research Association, Denver, CO.
- Santelices, M. V., Valencia, E., Taut, S., & Manzi, J. (2011, Enero 19). *Importancia de las características personales y contextuales en el desempeño docente: Evidencia usando una medida de desempeño docente estandarizada y basada en estándares de desempeño profesional*. Paper presentado en 2º Encuentro de la Sociedad Chilena de Políticas Públicas, Santiago.
- Santiago, P., & Benavides, F. (2009). *Teacher evaluation: A conceptual framework and examples of country practices*. Paper presentado en OECD-México Workshop Towards a Teacher Evaluation Framework in México: International Practices, Criteria and Mechanisms, Ciudad de México.
- Santibáñez, L., & Martínez, J. F. (2010). Políticas de incentivos para maestros: Carrera magisterial y opciones de reforma. En A. Arnaut, & S. Giorgiuli (Eds.), *Los grandes problemas de México: 7. Educación*. Ciudad de México: El Colegio de México.

- Santibáñez, L., Martínez, J. F., Datar, A., McEwan, P. J., Setodji, C. M., & Basurto-Davila, R. (2007). *Breaking ground: Analysis of the assessment system and impact of México's teacher incentive program «Carrera Magisterial»* (No. TR-458-SEP). Technical Reports. Santa Monica, CA: RAND Corporation. Recuperado a partir de <http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED496982>
- Schmidt, W. H., McKnight, C. C., Houang, R. T., Wang, H., Wiley, D. E., Cogan, L. S., & Wolfe, R. G. (2001). *Why schools matter: A cross-national comparison of curriculum and learning*. The Jossey-Bass Education Series. San Francisco, CA: Jossey-Bass Publishers.
- Schmidt, W. H., McKnight, C. C., Wiley, D. E., Houang, R. T., & Valverde, G. A. (1997). *Many visions, many aims: A cross-national investigation of curricular intentions in school mathematics* (Vol. 1). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Schmidt, W. H., Raizen, S., Wolfe, Richard G., Bianchi, L. J., & Britton, E. D. (1997). *Many visions, many aims: A cross-national investigation of curricular intentions in school mathematics* (Vol. 2). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Schneider, B., Carnoy, M., Kilpatrick, J., Schmidt, W. H., & Shavelson, R. J. (2007). *Estimating causal effects using experimental and observational designs*. Washington, DC: American Educational Research Association.
- Schneider, M., & Stern, E. (2010). The cognitive perspective on learning: Ten cornerstone findings. En H. Dumont, D. Istance, & F. Benavides (Eds.), *The nature of learning: Using research to inspire practice, educational research and innovation* (pps 69–90). Paris, Francia: OECD Publishing.
- Schon, D. A. (1990). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions* (1st ed.). San Francisco, CA: Jossey-Bass.
- Shavelson, R. J., & Webb, N. M. (1991). *Generalizability theory: A primer*. Newbury Park, CA: Sage Publications.
- Shepard, L. A. (1997). The centrality of test use and consequences for test validity. *Educational Measurement: Issues and Practice*, 16(2), 5–24.
- Shepard, L. A. (2001). The role of classroom assessment in teaching and learning. En Virginia Richardson (Ed.), *Handbook of research on teaching* (4ª Ed., pps 1066–1101). Washington, DC: American Educational Research Association.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14.
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1–22.
- Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*, 23(4), 323–355.
- Slavin, R. E. (1996). *Salas de clase efectivas, escuelas efectivas: Plataforma de investigación para una reforma educativa en América Latina*. Santiago, Chile: Programa de Promoción de la Reforma Educativa en América Latina.
- Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: SAGE publications Ltd.
- Sorto, M. A., Marshall, J. H., Luschei, T. F., & Martin, C. (2009). Teacher knowledge and teaching in Panama and Costa Rica: A comparative study in Primary and Secondary education. *Revista Latinoamericana de Investigación en Matemática Educativa*, 12(2), 251–290.

- Springer, M. G., Hamilton, L., McCaffrey, D. F., Ballou, D., Le, V.-N., Pepper, M., Lockwood, J. R., et al. (2010). *Teacher pay for performance: Experimental evidence from the project on incentives in teaching*. Nashville, TN: National Center on Performance Incentives, Peabody College of Vanderbilt University. Recuperado a partir de <http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED513347>
- Stage, S. A., & Quiroz, D. R. (1997). A meta-analysis of interventions to decrease disruptive classroom behavior in public education settings. *School Psychology Review*, 26(3), 333–68.
- Stigler, J. W., Gallimore, R., & Hiebert, J. (2000). Using video surveys to compare classrooms and teaching across cultures: Examples and lessons from the TIMSS – Video studies. *Educational Psychologist*, 35(2), 87–100.
- Stigler, J. W., Gonzales, P., Kawanaka, T., Knoll, S., & Serrano, A. (1999). The TIMSS videotape classroom study: Methods and findings from an exploratory research project on eighth-grade mathematics instruction in Germany, Japan, and the United States. *Education Statistics Quarterly*, 1(2), 109–12.
- Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. New York, NY: Simon and Schuster.
- Strauss, A. L., & Corbin, J. M. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage Publications.
- Stronge, J. (1997). *Evaluating teaching: A guide to current thinking and best practice*. Thousand Oaks, CA: Corwin Press.
- Stronge, J. H., Ward, T. J., Tucker, P. D., & Hindman, J. L. (2007). What is the relationship between teacher quality and student achievement? An exploratory study. *Journal of personnel evaluation in education*, 20(3–4), 165–184. doi:10.1007/s11092-008-9053-z
- Stronge, J. H. (2007). *Qualities of effective teachers* (2ª Ed.). Alexandria, VA: Association for Supervision and Curriculum Development.
- Stronge, J. H., & Tucker, P. D. (2003). *Handbook on teacher evaluation: Assessing and improving performance*. Larchmont, NY: Eye on Education.
- Sturman, M. (2003). Searching for the inverted U-shaped relationship between time and performance: Meta-Analyses of the experience/performance, tenure/performance, and age/performance relationships. *Journal of Management*, 29(5), 609–640.
- Taut, S., Santelices, M. V., Araya, C., & Manzi, J. (2010). Theory underlying a national teacher evaluation program. *Evaluation and Program Planning*, 33(4), 477–486. doi:10.1016/j.evalprogplan.2010.01.002
- Taut, S., Santelices, M. V., Araya, C., & Manzi, J. (2011, Abril). *Effects and uses of the national teacher evaluation system in Chilean elementary schools*. Paper Presentado en la Conferencia Anual de la American Education Research Association, New Orleans, LA.
- Taut, S., Santelices, M. V., & Valencia, E. (2009). *Situación laboral y desempeño en evaluaciones posteriores de los docentes calificados como insatisfactorio en el Sistema Nacional de Evaluación de Desempeño Docente* (Documento Interno MIDE UC). Santiago, Chile: Pontificia Universidad Católica de Chile, Escuela de Psicología, Centro de Medición MIDE UC.
- Taut, S., Santelices, M. V., & Valencia, E. (2010). *Resultado de re-evaluaciones y situación laboral de los docentes evaluados por el Sistema de Evaluación de Desempeño Docente entre 2003 y 2008*. Documento Interno MIDE UC. Santiago, Chile: Pontificia Universidad Católica de Chile, Escuela de Psicología, Centro de Medición MIDE UC.
- Taut, S., & Valencia, E. (2011, Enero 11). *Match bases de datos PSP-Docentemás 200912*. Manuscrito en preparación, Santiago, Chile.

- Tombari, M. L., & Borich, G. D. (1999). *Authentic assessment in the classroom: Applications and practice*. Upper Saddle River, NJ: Prentice-Hall.
- Tornero, B. (2009). *Sistema de evaluación del desempeño docente en Chile: Significados y percepciones de profesores rebeldes* (Tesis de Magister). Pontificia Universidad Católica de Chile, Santiago, Chile.
- US Department of Education. (2010). *Race to the Top Program guidance and frequently asked questions*. Washington, DC.
- Vaillant, D., & Rossel, C. (Eds.). (2006). *Maestros de escuelas básicas en América Latina: Hacia una radiografía de la profesión*. Santiago, Chile: PREAL.
- Valencia, E. (2009). *La relación entre los antecedentes individuales, las condiciones de trabajo y el contexto local con la calidad docente del sistema educacional público chileno* (Tesis de Magister). Pontificia Universidad Católica de Chile, Santiago, Chile.
- Valencia, E., & Taut, S. (2007). *Estudio de la dimensionalidad del portafolio de la Evaluación Docente 2007* (Internal document MIDE UC). Santiago, Chile: Pontificia Universidad Católica de Chile, Escuela de Psicología, Centro de Medición MIDE UC.
- Valverde, G. A. (2009). Estándares y evaluación. En S. Schwartzman, & C. Cox (Eds.), *Políticas educativas y cohesión social en América Latina*. Santiago, Chile: Uqbar Editores.
- Valverde, Gilbert, & Näslund-Hadley, E. (2010). *La condición de la educación en matemáticas y ciencias naturales en América Latina y el Caribe* (Notas Técnicas No. IDB-TN-211). Washington, DC: Banco Interamericano de Desarrollo.
- Walberg, H. J. (1986). Syntheses of research on teaching. En M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pps 214-229). New York, NY: MacMillan.
- Wang, M. C., Haertel, G. D., & Walberg, H. J. (1997). What helps students learn? Spotlight on student success. Laboratory for Student Success, Philadelphia, PA. Recuperado a partir de <http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED461694>
- Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, 73(1), 89-122. doi:10.3102/00346543073001089
- Wei, R. C., Darling-Hammond, L., Andree, A., Richardson, N., & Orphanos, S. (2009). Professional learning in the learning profession: A status report on teacher development in the U.S. and abroad. Technical Report. Oxford, OH: National Staff Development Council. Recuperado a partir de <http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED504168>
- Wideen, M., Andrews, I., & Fullam, M. (1987). Implementing the implementation plan. *Staff development for school improvement: A focus on the teacher*. New York, NY: The Falmer Press.
- William, D. (2010). The role of formative assessment in effective learning environments. En H. Dumont, D. Istance, & F. Benavides (Eds.), *The nature of learning: Using research to inspire practice, educational research and innovation* (pp 35-67). Paris, Francia: OECD Publishing.
- Wilson, S. M., Floden, R. E., & Ferrini-Mundy, J. (2001). *Teacher preparation research: Current knowledge, gaps, and recommendations*. Washington, DC: Center for the Study of Teaching and Policy, University of Washington. Recuperado a partir de <http://depts.washington.edu/ctpmail/about/study/14.shtml>
- Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, 11(1), 57-67. doi:10.1023/A:1007999204543

De esta edición:

© Pontificia Universidad Católica de Chile

Coordinación y Edición

Jorge Manzi, Roberto González y Yulan Sun

Publicado por

MIDE UC, Centro de Medición
Pontificia Universidad Católica de Chile
Vicuña Mackenna 4860, Macul
Santiago de Chile
Tel. (56-2) 3547454
mideuc@uc.cl – www.mideuc.cl

ISBN: 978-956-14-1203-3

Impreso en Chile/Printed in Chile

Inscripción N° 203.690

Primera Edición de 5.000 ejemplares, mayo de 2011.

Diseño y Diagramación

www.iunta.cl

Fotografías portada y portadillas

Alan Warren

Impresión

Alfabetas Artes Gráficas Ltda.

Esta publicación está disponible en www.mideuc.cl

“Las opiniones que se presentan en esta publicación, así como los análisis e interpretaciones, son de responsabilidad exclusiva de los autores de cada capítulo o sección de este y no reflejan necesariamente los puntos de vista de MIDE UC.”

Las informaciones contenidas en el presente documento pueden ser utilizadas y traducidas total o parcialmente mientras se cite la fuente.