

INVESTIGACIONES

Importancia de las características del profesor y de la escuela en la calidad docente: Una aproximación desde la Teoría de Respuesta del Ítem^{*}

The role of personal teacher characteristics and school variables in teacher quality:
A study using Item Response Theory

Importância das características do professor e da escola na qualidade docente: Uma aproximação da Teoria de Resposta ao Item

Santelices Etchegaray, María Verónica^a; Valenzuela Rojas, Francisco^b

^aFacultad de Educación, Pontificia Universidad Católica, Santiago, Chile. Teléfono: 56-2-354-5323.
Correo electrónico: vsanteli@uc.cl

^bFacultad Matemáticas, Pontificia Universidad Católica, Santiago, Chile. Teléfono: 56-2-2354 4511.
Correo electrónico: francisco.valenzuela.rojas@gmail.com

RESUMEN

Esta investigación explora la relación entre la calidad docente, características del docente y características del establecimiento en el que este ejerce. La investigación se vincula con la discusión acerca de la efectividad docente, mediciones de calidad docente, y variables de política educacional efectivas para mejorar dicho desempeño. El estudio se distingue por la utilización de métodos que consideran la habilidad docente como una variable no observable estimada, por considerar el hecho de que los docentes se encuentran anidados al nivel de la escuela, y por utilizar una medida de calidad docente basada en estándares profesionales. Los resultados indican que si bien el desempeño docente medido por estándares profesionales está definido principalmente por características propias del docente, el establecimiento en el que se desempeña juega un rol importante. Las variables incluidas en el modelo, tanto a nivel escuela como a nivel del docente, explican solo un limitado porcentaje de la varianza total (16%).

Palabras clave: calidad docente, estándares profesionales, teoría de respuesta al ítem.

ABSTRACT

This study explores the relationship between teacher quality, teacher attributes and school characteristics. The research presented is related to the discussion regarding teacher quality, how to best measure teacher quality and the most effective policies to improve teacher performance. The study, unlike previous studies, uses methods that consider teacher ability as an unobservable characteristic that is estimated independent from the sample of items and students who took the test. It also considers the nested structure of teachers within schools and uses a measure of teacher quality based on professional standards. Results show that although teacher quality is primarily a consequence of individual teacher characteristics, the school level plays an important role. However, the variables included in the model, both at the teacher and at the school level, explain only a limited proportion of the total variance (16%).

Key words: teacher quality, professional standards, item response theory.

RESUMO

Explora a relação entre a qualidade docente, características próprias do docente e características do estabelecimento educacional em que ele atua. A pesquisa vincula-se com a discussão sobre efetividade docente, medições de qualidade docente e variáveis de política educacional mais efetivas para melhorar tal qualidade. O estudo distingue-se de outros pela utilização de métodos que consideram a habilidade docente como uma variável não

* Esta investigación ha sido posible gracias al financiamiento de FONDECYT de Iniciación mediante el proyecto N° 11110262.

observável, por levar em conta o fato de que os docentes se encontram aninhados nas escolas em que atuam e por utilizar uma medida de qualidade docente baseada em padrões profissionais. Resultados indicam que, ainda que o desempenho docente seja medido por padrões profissionais, ele está definido, principalmente, por características inerentes ao docente e pela importância dada ao estabelecimento em que ele desempenha suas funções. As variáveis incluídas no modelo, tanto para a instituição escolar como para a pessoa, explicam apenas uma limitada porcentagem da variação total (16%).

Palavras chave: Qualidade docente, padrões profissionais, teoria da resposta ao item.

1. INTRODUCCIÓN

El análisis de la literatura respecto de calidad docente, tanto a nivel nacional como internacional, evidencia la tensión en torno a la pregunta de qué se entiende por calidad docente, de qué forma puede o debe medirse y qué variables afectan dichas medidas de calidad. Esta tensión es relevante no solo en cuanto a discusión teórica, sino que también en términos prácticos, en la medida en que (i) existen consecuencias individuales asociadas a dichas mediciones de calidad docente y, (ii) se diseñan políticas y se destinan recursos públicos para la modificación de variables, que se cree tendrán un impacto en la calidad docente. Todo esto con el objetivo final de incrementar la calidad y cantidad de los aprendizajes del profesor.

La discusión de política educacional relativa al mejoramiento de la efectividad docente hoy en Chile está cruzada por al menos dos conceptualizaciones de calidad docente: una basada en la investigación que define calidad a partir de las características propias del profesor y otra basada en la evaluación de competencias profesionales.

Un ejemplo reciente del primer del primer tipo de políticas es el otorgamiento de becas Vocación de Profesor (<http://www.becavocaciondeprofesor.cl/>) para el estudio de pedagogía a alumnos que demostraron cierto nivel de desempeño en la prueba estandarizada de admisión a las universidades chilenas (Prueba de Selección Universitaria o PSU). Esta medida, cuya implementación obligó al gobierno a suplementar el presupuesto del Ministerio de Educación en aproximadamente \$ 2.000 millones de pesos para el 2011¹, presupone que el rendimiento en la PSU es predictivo de un mejor desempeño docente.

Existe un cúmulo de literatura examinando el rol de las variables personales del profesor en el desempeño de sus alumnos, especialmente en Estados Unidos. Dicha investigación comienza a desarrollarse en Chile y sus resultados subrayan la importancia del género del profesor, su certificación como docente y sus años de experiencia (Lara, Mizala y Repetto, 2010). Sin embargo, hasta el momento dicha investigación solo ha analizado el rol de dichas variables relativas al nivel de desempeño de los alumnos. No se conocen estudios que estimen el aporte del profesor, o de sus características personales, al proceso educativo utilizando medidas longitudinales de desempeño del alumno.

Los estudios concernientes a las características personales y del contexto (establecimiento y municipio) que influyen en medidas de calidad basadas en competencias profesionales están menos desarrollados tanto a nivel internacional como nacional. La

¹ <http://www.eligeeducar.cl/ministerio-debera-suplementar-presupuesto-de-becas-para-pedagogia-en-2-000-millones/>

literatura nacional revisada indica que las metodologías utilizadas han considerado la calificación obtenida por el docente en las evaluaciones de su desempeño como una medición directa y verdadera de su habilidad docente, sin tomar en cuenta el hecho de que dicha habilidad es en realidad una variable no observable, medida con cierto grado de error (error de medición) y determinada, en parte, por el grupo que se somete a la evaluación y por la selección de preguntas que el docente debe contestar.

La presente investigación pretende contribuir a la comprensión de la importancia relativa de las características propias del profesor, en relación con aquellas propias del establecimiento, en una estimación de habilidad docente que es independiente del grupo de evaluados y de la muestra de preguntas, o ítems, que deben completar los docentes. Para este estudio se utilizó una evaluación de calidad docente basada en competencias profesionales.

2. ANTECEDENTES SOBRE LAS MEDIDAS DE CALIDAD DOCENTE

A continuación se presentan los antecedentes bibliográficos que han examinado el rol del profesor en el aprendizaje de los alumnos, las distintas aproximaciones a la medición de calidad docente² discutidas en la literatura, las características del profesor que la literatura internacional ha identificado como predictores de calidad docente, el rol menos explorado de las características propias del contexto en el desempeño de los docentes, y el estado de la literatura nacional respecto del tema. La sección concluye discutiendo cómo la investigación propuesta se relaciona con estudios anteriores.

2.1. LA IMPORTANCIA DEL ROL DEL PROFESOR

Existe consenso internacional acerca de la importancia de los profesores en la calidad del proceso educativo (Gordon, Kane & Steiger, 2006; Nye, Konstantopoulos & Hedges, 2004). Así lo evidencian estudios de muy distinta índole: desde el informe McKinsey que estudió las prácticas educacionales en los países con mayores logros educacionales en el mundo (McKinsey & Company, 2007) hasta estudios de tipo más bien estadístico en que la efectividad de los profesores es medida de modo indirecto basándose en los resultados de sus alumnos en pruebas estandarizadas (ver por ejemplo McCaffrey, Lockwood, Koretz, Louis & Hamilton, 2004). En un estudio resumiendo el estado del conocimiento en este campo la National Academy of Education (2009: en línea) señaló que: “La efectividad del profesor es tan preponderante que alumnos de niveles socioeconómicos bajos que tienen la posibilidad de tener un buen profesor durante tres años consecutivos en educación primaria obtienen puntajes estandarizados que, en promedio, son equivalentes a los de un niño de clase media. Por el contrario, la mayor parte de los niños verán su aprendizaje perjudicado si tienen un profesor de poca calidad durante tres años consecutivos”. La importancia de los profesores en el proceso educativo se ve reforzada por el hecho de que el pago de los docentes representa el porcentaje más alto de los ítems presupuestarios de un establecimiento educacional (Hanushek & Rivkin, 2006).

² En este documento se entiende “desempeño docente” como sinónimo de “calidad docente”.

2.2. ¿QUÉ SE ENTIENDE POR CALIDAD DOCENTE?

La discusión pertinente a calidad docente y efectividad se ha desarrollado en al menos dos ámbitos académicos: el ámbito de la economía de la educación y el ámbito de la medición de competencias profesionales del docente que provienen de una vertiente educacional y psicométrica. En esta discusión participan activamente académicos de distintas disciplinas utilizando implícitamente alguna de estas dos aproximaciones en sus investigaciones.

2.2.1. *El concepto de calidad docente basado en el desempeño de los alumnos*

En el ámbito de la economía de la educación, la efectividad del profesor se ha estudiado utilizando un enfoque que asimila el proceso educativo a una función de producción y se ha entendido “calidad docente” como las características propias del profesor que influyen el desempeño de sus alumnos. Una línea de trabajo se ha centrado en el estudio de *características observables* propias del profesor, mientras que otra línea ha investigado la efectividad docente *indirectamente*. En ambos casos se ha utilizado el desempeño de alumnos en pruebas estandarizadas como el criterio (o variable dependiente) para evaluar la calidad del docente³.

Dentro de la primera línea de investigación, las características *observables* del profesor más frecuentemente estudiadas han sido la trayectoria académica del docente, sus años de experiencia y su nivel de perfeccionamiento (Hanushek & Rivkin, 2006). Dichas características se han medido en la práctica mediante las siguientes variables: el desempeño en pruebas estandarizadas de admisión a la universidad⁴, el nivel de selectividad de la institución donde se estudió pedagogía, el hecho de poseer certificación docente o no poseerla, la realización de postgrados, el número y la disciplina en la que se cursan dichos postgrado así como el número y tipo de perfeccionamientos realizados (Hanushek & Rivkin, 2006; Millman, 1997). Los resultados de este tipo de estudios entregan algunas luces respecto de las variables directamente observables y propias del profesor que influyen en el mejor aprendizaje de los alumnos, pero los hallazgos son difícilmente concluyentes, ya que numerosos estudios encuentran resultados opuestos para una misma variable y la mayor parte de los estudios no logra explicar un porcentaje alto de los resultados de sus alumnos con las variables estudiadas.

Por ejemplo, el informe elaborado por la Comisión Educacional de los Estados (Allen, 2003) encuentra evidencia que apoya solo parcialmente la importancia de la formación inicial del docente para su efectividad y enfatiza las similitudes entre los programas de formación tradicional y alternativos. Rivkin, Hanushek & Kain (2005) asimismo encuentran que ninguna de las características observables de los docentes logra explicar una proporción importante del desempeño de los alumnos. Rice (2003), en cambio, al revisar la literatura sobre características de calidad docente relevantes para acciones de política

³ El desempeño de los alumnos en pruebas estandarizadas durante dos o más períodos consecutivos ha sido utilizado como base para el desarrollo de mediciones indirectas de la calidad docente llamadas efecto profesor o de valor agregado. Dichas mediciones intentan capturar la contribución específica que cada profesor hace del proceso de aprendizaje de sus alumnos tomando en consideración las variables que están fuera de su control directo, como el nivel socioeconómico de sus alumnos (McCaffrey et al., 2004). Este tipo de datos también han sido utilizados para estimar el “valor agregado” de establecimientos escolares.

⁴ PAA o PSU en Chile y SAT o ACT en Estados Unidos.

educacional, encuentra evidencia a favor de la certificación docente, de la preparación docente en el área enseñada, de la preparación pedagógica y de los años de experiencia. Estos resultados son apoyados por el resultado de los estudios de Akiba, LeTendre & Scriber (2007), Nye et al. (2004), Darling-Hammond, Berry & Thoreson (2001) y Wilson, Floden & Ferrini Mundy (2001). Por último, Hanushek & Rivkin (2006), al resumir los estudios realizados en Estados Unidos con anterioridad a 1994, y considerando la calidad de sus estimaciones, solo encuentran evidencia sólida acerca de la importancia de los años de experiencia del profesor en el desempeño de sus alumnos.

Kennedy (2008) ejemplifica esta falta de consenso citando dos investigaciones que llegan a conclusiones contrapuestas (énfasis agregados):

Más de 200 estudios han encontrado que profesores que tienen *mejor preparación en su disciplina* y tienen *mayor conocimiento del proceso de enseñanza-aprendizaje* son *mejor evaluados* y *más exitosos* con estudiantes en áreas desde Párvulos y Educación Básica hasta Matemáticas, Ciencias y Educación Técnica (Darling-Hammond & McLaughlin, 1999: 377-378).

Si tuviese que decir algo, yo concluiría que *existe poca evidencia* que sustente la *importancia del conocimiento disciplinario del profesor en el aprendizaje* de sus alumnos (Friedman, 2000: 20).

2.2.2. Calidad docente entendida como medición de competencias profesionales

Existen también mediciones alternativas del desempeño docente basadas en la evaluación de la práctica profesional del profesor, ya sea mediante una observación directa del docente en el aula, por medio de la evaluación de una clase filmada o por medio de material producido por el docente. El uso de este tipo de mediciones para la evaluación de profesores es habitualmente considerada más válida⁵ que las estimaciones estadísticas basadas en el desempeño de los alumnos, ya que permiten evaluar dimensiones propias de la labor profesional docente, como la preparación de una clase, el diseño de material, la implementación de una clase, el diseño de evaluaciones, su aplicación y uso en la retroalimentación de sus alumnos, y la reflexión acerca de la propia práctica de la pedagogía. Además, la evaluación de la labor profesional con este tipo de instrumentos se realiza de manera independiente del grupo de alumnos con el que trabaja el profesor y de las características de dichos alumnos que pudiesen dificultar o facilitar su aprendizaje.

Estas mediciones de competencias profesionales del profesor fueron desarrolladas en Estados Unidos como resultado de una inquietud por profesionalizar la labor docente (Milanowski, 2011) y se han plasmado en programas de certificación de excelencia docente como PRAXIS III del Educational Testing Service y el programa de certificación del National Board for Professional Teaching Standards. El primer programa certifica profesores que recién comienzan su labor profesional, mientras que el segundo certifica profesores que ya se encuentran ejerciendo su labor en establecimientos educacionales. El modelo de competencias o estándares profesionales escrito por Charlotte Danielson (1996) fue una pieza clave en el diseño de estos programas y en la adopción de este tipo de medidas de calidad docente por más de 200 distritos educacionales norteamericanos.

⁵ La palabra "validez" se utiliza en su acepción psicométrica y se refiere a qué tan cercana se encuentra la medición al constructo o dimensión que interesa medir.

Su traducción casi textual es conocida en Chile con el nombre de “Marco Para la Buena Enseñanza” (Ministerio de Educación, 2004) y fue la base del diseño del sistema de Evaluación Docente (ED), obligatorio para todos los docentes que trabajan en establecimientos educacionales públicos desde el 2005 (CPEIP, 2005a, 2005b).

El mencionado modelo de competencias profesionales de Danielson define dimensiones de la labor docente que son evaluadas de acuerdo con el nivel de desempeño en escalas del 1 al 4 (Danielson, 2007, 1996). La práctica docente se divide en cuatro ámbitos: planificación y preparación de la clase, ambiente de la clase, instrucción y responsabilidades profesionales. Las categorías finales de la evaluación son insatisfactorio, básico, competente y distinguido.

Si bien el uso de mediciones basadas en competencias profesionales se ha extendido entre distritos de los Estados Unidos durante los últimos años, su implementación ha sido de tipo “local” por lo que cada distrito ha adaptado el marco de estándares profesionales así como los instrumentos y las rúbricas utilizados para su evaluación (Milanowski, 2011). Esto ha redundado en que, hasta la implementación de la ED en Chile, existieran solo experiencias de escala pequeña en que se utilizaran mediciones de calidad docente basadas en estándares profesionales, lo que ha limitado la generalizabilidad y comparabilidad de los resultados de las investigaciones realizadas hasta la fecha (Archibald, 2007).

2.3. LA INFLUENCIA DEL CONTEXTO EN EL DESEMPEÑO DOCENTE

Es importante reconocer que el desempeño docente no solo está dado por variables propias del profesor, sino que también está influido por variables del contexto escolar, entendiendo contexto como aquellas características propias del establecimiento escolar y su comunidad, propias del distrito o municipalidad y la interacción de dichas características. Al tratar el tema, algunos autores han hecho la distinción entre el rol de aspectos físicos (seguridad, recursos educacionales, infraestructura) y aspectos organizacionales como el liderazgo al interior del establecimiento, las características del equipo docente, la relación que se da entre sus miembros, la relación con padres y apoderados y la relación con los mismos alumnos (Odden, Borman & Fermanich, 2004). En este sentido Heneman & Milanowski (2004) plantean que los sistemas de administración de recursos humanos a nivel de distrito y de establecimiento son importantes para el mejoramiento del logro académico de los estudiantes. Por su parte, Heck (2006) plantea un modelo de variables latentes que explica los resultados de los alumnos a partir de variables del alumno, del contexto de la escuela, y de la calidad del profesor aunque encuentra poco apoyo empírico para su modelo.

Kennedy (2011) va aún más allá y propone que la investigación no debería quedarse solo en el rol de las variables propias del profesor y en las prácticas profesionales, sino que en el futuro debería focalizarse en las características propias de las situaciones organizacionales en las que trabajan los profesores y que determinan qué y cómo enseñan los profesores durante sus clases. Algunas de las variables a las que hace referencia son: tiempo disponible para planificación de clases, esquema de asignación de clases, disponibilidad especificada y alineamiento entre materiales para utilizar en la sala de clase.

Si bien el énfasis de las investigaciones basadas en el desempeño de los alumnos pareciera puesto en variables propias del docente, principalmente trayectoria académica y nivel de perfeccionamiento o postgrado, algunas investigaciones han incorporado variables

observables o no observables propias del establecimiento y del distrito (Hanushek & Rivkin, 2006).

2.4. LA INVESTIGACIÓN RELATIVA A CALIDAD DOCENTE EN CHILE

En primer lugar se presentan los estudios que investigan los factores asociados al desempeño de los alumnos en pruebas estandarizadas (SIMCE) y en una segunda sección se presentan aquellos que investigan las variables asociadas a medidas de competencia profesional del profesor, como el puntaje en la ED.

2.4.1. *Estudios que investigan variables relacionadas con el desempeño de los alumnos en Chile*

La mayor parte de la investigación concerniente a la calidad docente se ha centrado en estudiar el rol de distintas variables en el desempeño de sus alumnos en la prueba SIMCE (Sistema de Medición de la Calidad) e incluyen como variables explicativas indicadores de competencias profesionales del docente (ver, por ejemplo, Lara et al., 2010; Toledo, 2010; Bravo, Falck, González, Manzi y Peirano, 2008; León, Manzi y Paredes, 2008). Si bien algunos estudios han incluido variables del establecimiento en sus análisis, el foco de la discusión ha estado puesto en la capacidad predictiva de la ED y cómo las correcciones metodológicas que consideran el ordenamiento de profesores entre escuelas cambian o no cambian las variables del profesor que influyen en el desempeño de sus alumnos. En términos comparativos, menor atención se ha dedicado al rol que variables propias del establecimiento escolar y de la municipalidad pudiesen tener en el desempeño de los alumnos. Los estudios presentados, además, han abordado el rol de distintos factores en el nivel de desempeño de los alumnos en el SIMCE y no en su progreso durante el año que trabajan con un determinado profesor. El análisis del desempeño de alumnos en una sola medición estandarizada dificulta la interpretación de los resultados, ya que el nivel de aprendizaje no es atribuible solo al profesor que le enseña en un grupo de alumnos determinado, sino que también es resultado de la labor de los docentes que les enseñaron en años anteriores.

2.4.2. *Estudios que investigan variables relacionadas con las competencias profesionales de los docentes*

Al menos tres estudios han investigado las variables relacionadas con competencias profesionales del docente utilizando modelos multinivel y la categoría final obtenida en el Sistema Nacional de Evaluación del Desempeño Profesional Docente (ED) como variable dependiente.

La ED es una evaluación obligatoria para más de 70.000 docentes de aula que se desempeñan en establecimientos municipales a lo largo de Chile. Su objetivo es fortalecer la profesión docente y así contribuir a mejorar los aprendizajes de los alumnos (Taut, Santelices, Araya & Manzi 2010). La evaluación se realiza por medio de cuatro instrumentos, los que permiten reunir información complementaria del desempeño docente: (i) portafolio (60% de la categoría final), (ii) pauta de autoevaluación (10% de la categoría final), (iii) entrevista por un evaluador par (20%), (iv) informe de referencia a terceros (10%). Todos los instrumentos de la ED se construyen basados en los dominios, criterios y descriptores señalados en el MBE. El MBE establece estándares de una adecuada práctica profesional docente, las responsabilidades que contribuyen significativamente al

aprendizaje de sus alumnos y las responsabilidades que asume en el aula, en la escuela y en la comunidad donde se inserta.

En los modelos multinivel la importancia de las características observables propias del profesor viene dada por la significancia estadística de las variables incluidas en el modelo, mientras que el porcentaje total de la varianza asociada al nivel de profesor indica la importancia relativa de las características del profesor, en relación a características propias del establecimiento o del municipio, para explicar el desempeño del docente en la ED. En este caso, la proporción de varianza asociada al nivel profesor indica la importancia tanto de variables observables como no observables del profesor. Asimismo, la importancia del contexto en estos modelos viene dada por la proporción de la varianza explicada por los niveles superiores de anidación así como la significancia estadística de las variables explicativas incluidas en esos niveles superiores de anidación.

Manzi, Strasser, San Martín & Contreras (2007) encontraron que solo el 10% de la varianza del puntaje del portafolio se puede atribuir al nivel del establecimiento y 27% del puntaje de la evaluación del par se puede atribuir a este nivel de análisis. A nivel del profesor, los investigadores encuentran que las características más importantes a la hora de explicar desempeño docente fueron género, capacitación profesional y el subsector que enseña.

Eisenberg (2008), utilizó un modelo jerárquico de dos niveles para el estudio del desempeño en la ED en 5.164 docentes de primero a cuarto básico y encontró que la mayor parte de la varianza es atribuible al primer nivel de análisis, que en su caso agregaba el nivel del profesor y del establecimiento, mientras que 12% de la varianza es explicada por el segundo nivel de análisis correspondiente al municipio. Dicha investigadora encontró que género, duración de la carrera y modalidad de la formación inicial, años de experiencia en el área disciplinaria que se enseña y realización de postgrado se asocian estadísticamente a desempeño docente, pero las variables observables incluidas en este primer nivel solo explican una proporción pequeña de la varianza del primer nivel de análisis. Es decir, la mayor parte de las características del docente y del establecimiento asociadas al desempeño docente, no son capturadas por las variables observables incluidas en el modelo. El nivel socioeconómico del establecimiento está asociado a un mejor desempeño de los docentes (nivel socioeconómico más alto, mejor desempeño en la ED), mientras que el promedio SIMCE de la escuela se asoció en forma positiva a dicho desempeño. En el segundo nivel de análisis las variables más importantes fueron el tipo de administración del sistema educacional del municipio, el nivel socioeconómico de la municipalidad (mejor desempeño en escuelas de menor nivel socioeconómico) y el puntaje SIMCE promedio.

Valencia (2009), por su parte, analizó la evaluación de docentes evaluados entre 2005 y 2007 (N=23,995), en los tres ciclos de educación y diversos subsectores empleando un modelo jerárquico de tres niveles en que el docente (o sala de clases) corresponde al primer nivel, la escuela al segundo nivel y la municipalidad al tercer nivel de información. Lo mismo se hace en el estudio de Santelices, Valencia, Taut y Manzi (2010), pero con docentes evaluados entre 2006 y 2008, solo de 4° y 8° básico (n=1,422), con el objetivo de incluir variables acerca de aspectos motivacionales y de creencias de los docentes así como nuevos antecedentes de la sala de clases. Ambos estudios coinciden en reportar que la mayor variabilidad de los resultados se encuentra a nivel de los docentes. La variabilidad del contexto (escuela y municipio consideradas en conjunto) corresponde a 20% y

24%. Ambos estudios encuentran que las profesoras mujeres, docentes que han realizado post-título o postgrados, y los docentes que estudiaron programas de modalidad diurna y de mayor duración (normalistas o universitarios en el caso de Santelices et al., 2010) obtienen mejor evaluación en su ED. Sin embargo, difieren en la importancia de participar en desarrollo profesional. Mientras Valencia (2009) encuentra asociación positiva entre desempeño docente y PPF, al igual que Eisenberg 2008, esta relación no es estadísticamente significativa en Santelices et al. (2010). Ambos estudios encuentran que el nivel socioeconómico de la escuela no está significativamente relacionado con calidad docente medido por la ED. Además ambos encuentran que los recursos tecnológicos disponibles en la sala de clases, las expectativas de escolaridad de los apoderados y el número de profesores con certificación AEP en la escuela están positivamente relacionados con el desempeño individual de los docentes. Los resultados de ambos estudios también son consistentes en indicar que el tipo de administración escolar no está significativamente asociado a desempeño docente. Hay algunos indicios que apuntan a que mayor pobreza a nivel municipal mejor es el desempeño docente (% pobreza CASEN en Valencia, 2009 y GSE Bajo en Eisenberg, 2008).

Las investigaciones descritas en esta sección han utilizado el puntaje obtenido por el docente en la Evaluación Docente como proxy de calidad docente y han supuesto que la categoría final de desempeño observado (en el contexto de pruebas, el equivalente a puntaje) refleja en forma precisa la habilidad verdadera del docente. Sin embargo, dicha categoría final es en realidad un índice compuesto del puntaje en cuatro instrumentos, cada uno de ellos influido por las condiciones en las que el docente es evaluado, el tipo de tareas que el docente debe completar (ítems) y el desempeño del grupo que es evaluado junto a ese docente. La verdadera habilidad del docente es una destreza no observable (Hambleton, Swaminathan & Rogers, 1991).

Esta investigación tiene por objetivo explorar la relación la habilidad docente y los aspectos individuales del profesor y de su contexto (del establecimiento educacional), pero considerando la verdadera naturaleza de la variable dependiente. Para ello, la habilidad estimada del docente se obtiene de la implementación del modelo de Crédito Parcial sobre la base de los resultados de los docentes en el portafolio del ED. El modelo de Crédito Parcial fue desarrollado en el contexto de la Teoría de Respuesta al Ítem (TRI) y permite la modelación de respuestas en las que se entrega crédito parcial por respuestas incompletas. Las estimaciones de habilidad obtenidas mediante modelos TRI (y también de los parámetros de los ítem), en los casos en los que los modelos tienen un buen ajuste, son independientes de la muestra de individuos y de la muestra de ítems utilizados.

Esta investigación considera también la anidación de los resultados de la ED a nivel de la escuela, al ser este el nivel inmediatamente superior al docente, foco de interés de esta investigación. Van den Noortgate, Opendakker & Oghena (2005) han recomendado la consideración del nivel superior, e inmediatamente inferior al nivel de interés. Ignorar la estructura de la información redundante en sesgo de los parámetros estimados de efectos fijos y efectos aleatorios así como en sus errores estándar (Cervini, 2012; Van den Noortgate et al., 2005; Opendakker & Van Damme, 2000). Al utilizar modelos jerárquicos de dos niveles (docentes y escuelas) se reconoce explícitamente la estructura de la información que está influida por las relaciones interpersonales y el mismo contexto laboral. La relación que existe entre los docentes que trabajan en una

misma escuela ha sido incluso documentada cualitativamente por otras investigaciones por medio de entrevistas y *focus groups* en los que los docentes han reconocido colaborar en la realización del portafolio e incluso contar con apoyo del jefe de UTP del establecimiento escolar (Taut, Santelices & Stecher, 2012). Investigaciones anteriores (Santelices et al., 2010 analizando el resultado agregado de la ED y Manzi et al. 2007 analizando el resultado en el portafolio) sugieren que aproximadamente 10% de la varianza total del desempeño docente (en modelos incondicionales o sin covariables) es atribuible al establecimiento educacional.

3. PREGUNTAS DE INVESTIGACIÓN

El presente artículo tiene por objetivo responder las siguientes preguntas de investigación:

- ¿Cuál es la importancia del nivel del docente y del nivel escuela en el desempeño docente, medido como una habilidad no observable estimada mediante el enfoque de Teoría de Respuesta al Ítem?
- ¿Se observa asociación entre las variables propias del docente (como el tipo de institución, duración y modalidad de formación inicial) y desempeño docente cuando se utiliza una estimación de habilidad docente obtenida mediante el enfoque de Teoría de Respuesta al Ítem?
- ¿Existe asociación entre las características de la escuela en que el docente trabaja y desempeño docente cuando se utiliza una estimación de habilidad docente obtenida mediante el enfoque de Teoría de Respuesta al Ítem?

4. METODOLOGÍA

En esta sección se presenta la información utilizada, los modelos de medición y regresión estimados y los análisis estadísticos realizados.

4.1. INFORMACIÓN UTILIZADA

Los datos utilizados para el modelamiento corresponden a los resultados del portafolio de docentes trabajando en escuelas municipales el 2010 de norte a sur de Chile (excluyendo el territorio insular). El portafolio es un instrumento de evaluación en el que el docente debe presentar evidencia que dé cuenta de su mejor práctica pedagógica. Debe desarrollarse en el sector en que se encuentra inscrito el docente para ser evaluado y este cuenta con 12 semanas para su elaboración. El portafolio se estructura en dos módulos. Para la elaboración del módulo 1, Descripción, Evaluación y Reflexión de una Unidad Pedagógica, el docente debe diseñar e implementar una unidad pedagógica de ocho horas, una evaluación de término de esa unidad y completar preguntas referidas a su quehacer docente. Para la elaboración del módulo 2, Grabación de una Clase, el docente debe presentar la grabación de una clase de 40 minutos, sin cortes ni interrupción, con un curso con el que trabaja regularmente y completar una ficha con información relativa a esa clase. De acuerdo con lo evaluado en el portafolio, el docente

es categorizado en uno de cuatro niveles de desempeño: Destacado, Competente, Básico o Insatisfactorio. El portafolio incluye 24 indicadores que miden 8 dimensiones del desempeño docente: (i) Organización de los elementos de la unidad (3 ítems), (ii) Calidad de las actividades de la clase (3 ítems), (iii) Calidad del instrumento de evaluación (3 ítems), (iv) Utilización de los resultados de la evaluación (2 ítems), (v) Reflexión acerca de su quehacer docente (3 ítems), (vi) Ambiente de la clase (3 ítems), (vii) Estructura de la clase (3 ítems) y (viii) Interacción Pedagógica (4 ítems). Se optó por analizar el desempeño en el portafolio de la ED por ser el instrumento con mayor ponderación en el desempeño final del docente y por ajustarse de mejor manera, comparado con el desempeño final del docente que combina el desempeño en cuatro instrumentos, el supuesto de una escala unidimensional.

Las variables propias del profesor se obtuvieron del Cuestionario Complementario que el docente debe completar como parte de la ED. El cuestionario incluye preguntas sobre la edad, educación y perfeccionamiento, entre otras. La información respecto del nivel socioeconómico del establecimiento del docente provino del sitio *web* del SIMCE disponible en www.simce.cl. Los establecimientos escolares están categorizados de la siguiente manera: nivel socioeconómico bajo (A), medio-bajo (B), medio (C), medio-alto (D) y alto (E).

4.2. CARACTERÍSTICAS DE LA MUESTRA

En el 2010 se cuenta con N=8,358 profesoras y profesores sometidos a la (ED), de todos los ciclos y subsectores del aprendizaje. La mayor parte de los profesores evaluados proviene de la Región Metropolitana, seguido de la Región de Los Lagos. El 70% corresponde a mujeres (67,3%) y la mayoría de ellos trabajan en colegios de zonas urbanas (71,1%). El 89,7% de los docentes poseen un título de profesor o educador en el 2010 mientras que 1,6% corresponde a psicopedagogos y profesionales de otras áreas. Hay un porcentaje no menor de profesores para los cuales no se dispone información relativa a su título (8,7%). Los docentes generalistas son los profesores con mayor presencia en la ED (28,5%). En gran parte de las regiones del país los profesores fueron calificados mayoritariamente como Competentes (en un rango desde 52,6% hasta 68,2%), mientras que en la Región de Tarapacá predominan los profesores calificados como Básico (51,4%). En el total de la muestra 2% de los docentes fueron calificados como insatisfactorio, 32% como básico, 59% como competente y 7% como destacado. Del total de docentes incluidos 31% trabajaba en escuelas de nivel socioeconómico bajo y 51% en escuelas de nivel socioeconómico medio-bajo.

4.3. MODELO DE MEDICIÓN Y DE REGRESIÓN

El análisis se realizó en dos etapas. La primera etapa consistió en obtener estimaciones de la habilidad docente mediante el modelo de Crédito Parcial (CP) para luego, en una segunda etapa, relacionar las estimaciones obtenidas con variables propias del profesor y de la escuela utilizando un modelo lineal jerárquico de dos niveles. El modelo de CP estimó el parámetro de habilidad de los docentes utilizando como base los 24 indicadores completados por cada docente en el portafolio.

4.3.1. Teoría de Respuesta al Ítem y Modelo de Crédito Parcial

La Teoría de Respuesta al Ítem es una técnica confirmatoria que utiliza información a nivel de ítem para modelar datos de variables dicotómicas y polítomos. La esencia de la Teoría de Respuesta al Ítem es que la probabilidad de responder una pregunta correctamente, o de alcanzar cierto nivel de desempeño, se modela como función de la habilidad individual o variable latente (Embretson & Reise, 2000). En la medida en que el nivel de habilidad individual aumenta, la probabilidad de responder el ítem correctamente, o alcanzar cierto nivel de desempeño, también aumenta. La representación más conocida de esta relación es la curva característica del ítem, que entrega la probabilidad de la respuesta correcta de un ítem en función de parámetros de la persona y de los ítems. En el contexto de pruebas educacionales, el nivel de habilidad de un examinado no es una transformación directa del número de ítems o preguntas respondidas correctamente, sino que se estima a partir del patrón individual de respuestas de cada examinado considerando también el parámetro de dificultad de cada ítem (Embretson & Reise, 2000). Este modelo general también se caracteriza por la estimación de parámetros de personas e ítems y por tener estadísticos mínimos suficientes para la estimación de parámetros de personas e ítems (número de eventos observables) (Wright & Masters, 1982).

El modelo de Crédito Parcial es el modelo TRI más adecuado para analizar respuestas de tipo ordinal como las de la ED. Se usó el modelo de CP porque pese a que todos los 24 ítems del portafolio tienen las mismas categorías de respuestas, no se esperaba encontrar distancias similares entre pasos para todos los ítems. Esto último es una restricción del modelo de escala de clasificación (*rating scale model*), alternativa para analizar datos polítomos dentro de la TRI.

El modelo de CP, descrito por Wright & Masters (1982) como un caso especial del modelo de Rasch, comparte la forma algebraica con otros modelos Rasch. Esta forma algebraica describe curvas logísticas con pendientes comunes para todos los ítems, por lo que los parámetros de los ítems y de las personas son expresados en la misma escala de medida.

El modelo de CP entrega crédito parcial a niveles intermedios de desempeño en un ítem. Las curvas características de un modelo de CP están dadas por la siguiente expresión (Wright & Masters, 1982):

$$\phi_{nix} = \frac{\pi_{nix}}{\pi_{nix-1} + \pi_{nix}} = \frac{\exp(\beta_n - \delta_{ix})}{1 + \exp(\beta_n - \delta_{ix})} \quad x = 1, 2, \dots, m_i \quad (0)$$

Donde ϕ_{nix} denota la probabilidad de que el profesor n tenga puntaje parcial de x en el ítem i , es decir, x corresponde a la categoría de desempeño en cada ítem, siendo $x=0,1,2,3$. Además i identifica al ítem ($i=1, \dots, 24$) que tiene m_i categorías de desempeño, β_n se refiere a la habilidad del profesor n , δ_{ix} es el parámetro de dificultad de la categoría de desempeño x del ítem i , x es el número de categorías de desempeño completadas de forma exitosa. La ecuación (0) describe la probabilidad de que un profesor n obtenga una calificación de x , en vez de $x-1$, en un ítem como función del parámetro de personas β_n y de un parámetro de ítem δ_{ix} . Ambos parámetros definen la transición desde el nivel de desempeño $x-1$ al nivel de desempeño x en el ítem i . Las categorías de desempeño se encuentran ordenadas.

Análisis estadísticos. Para medir la calidad del ajuste del modelo CP se calcularon los *infits* (media de los residuos estudentizados cuadráticos ponderados por su varianza) y *outfits* (media de los residuos estudentizados cuadráticos no ponderados) de cada ítem (Cecilio, 2011). Los valores de estos indicadores revelan si hay una mayor o menor variación entre el patrón observado de respuestas de los sujetos y el patrón esperado de respuestas. Si las diferencias son pequeñas se puede concluir que el modelo se ajusta a los valores observados; en caso contrario no se puede suponer que el modelo, es una simplificación apropiada de los datos. Linacre (2002) considera que los valores que oscilan entre 0,5 y 1,5 revelan diferencias despreciables. El modelo de CP se estimó utilizando el *software* R.

4.3.2. Modelos jerárquicos de 2 niveles

Una vez estimados los parámetros de dificultad de los ítems (24 indicadores) y los parámetros de habilidad, para cada docente, se estimaron dos modelos. El primero fue un modelos de regresión lineal con variables propias del profesor que incluyó además un intercepto aleatorio (Modelo 1) con el fin de examinar el rol de la escuela en la definición de calidad docente.

$$\begin{aligned}\theta_{ij} &= \alpha_j + \beta_k \cdot X_{ij} + \varepsilon_{ij} && \text{(Modelo 1)} \\ \alpha_j &= \gamma_1 + u_j\end{aligned}$$

Donde las variables propias del docente están incluidas en X_i . La calidad de los docentes está también relacionada con factores aleatorios representados por ξ_{ij} . Los parámetros (β_1) miden la relación entre las variables explicativas y la calidad docente (θ_{ij}). Los supuestos de este modelo son los siguientes:

$$\begin{aligned}E(\varepsilon_{ij}) &= 0; \text{Var}(\varepsilon_{ij}) = \varphi_\varepsilon; \text{Cov}(\varepsilon_{ij}, \varepsilon_{kl}) = 0, \forall i \neq k, j \neq l \\ E(u_j) &= 0; \text{Var}(u_j) = \varphi_u; \text{Cov}(u_j, u_{kl}) = 0, \forall j \neq k \\ \text{Cov}(\varepsilon_{ij}, u_1) &= 0, \forall i, j \neq 1\end{aligned}$$

Es decir, la varianza de cada uno de los profesores dentro de los colegios es φ_ε y la varianza de los colegios es φ_u .

El segundo modelo incluyó, además, el nivel socioeconómico de la escuela como efecto fijo (modelo 2). Los supuestos del modelo 2 son iguales a los supuestos del modelo 1.

$$\begin{aligned}\theta_{ij} &= \alpha_j + \beta_k X_{ij} + \varepsilon_{ij} && \text{(Modelo 2)} \\ \alpha_j &= \gamma_1 + \gamma_2 z_{2j} + u_j\end{aligned}$$

En el modelo 2 z_{2j} es el nivel socioeconómico del colegio j , codificado de la siguiente manera:

$$z_{2j} = \begin{cases} 1, & \text{si el colegio } j \text{ es del grupo económico A} \\ 2, & \text{si el colegio } j \text{ es del grupo económico B} \\ 3, & \text{si el colegio } j \text{ es del grupo económico C} \\ 4, & \text{si el colegio } j \text{ es del grupo económico D} \end{cases}$$

También se ajustan los modelos 1 y 2 sin la variable explicativa relativa a las expectativas laborales de los docentes como efecto fijo con el objetivo de explorar posibles cambios en los resultados. Dicha variable explicativa contaba con valores perdidos, por lo que el tamaño de la muestra se redujo a 4.655 observaciones.

Análisis estadísticos. Los modelos multinivel 1 y 2 se estimaron paralelamente usando los *software* R y SPSS y el ajuste de ambos modelos se comparó utilizando los estadísticos de $-2 \log$ de la verosimilitud restringida, el criterio de información de Akaike (AIC) y el criterio bayesiano de Schwarz (BIC). El AIC está dado por la siguiente fórmula, en la que k es el número de parámetros incluidos en el modelo:

$$AIC = -2\log(L) + 2(k + 1)$$

El BIC está dado por:

$$BIC = -2\log(L) + \ln(n)(k + 1)$$

En el que k se refiere al número de parámetros en el modelo y n es el número de observaciones en la muestra. El AIC y BIC entregan dos formas de ajustar por el número de términos en el modelo y por el número de observaciones utilizadas. Los tres estadísticos descritos son recomendados para comparar modelos anidados utilizando la misma muestra; valores más bajos de los estadísticos indican modelos con mayor ajuste (Long, 1997). Además se computaron los coeficientes de regresión para los efectos fijos y aleatorios y los valores p para el modelo incondicional y los modelos condicionales 1 y 2.

4.4. LIMITACIONES

Es importante reconocer explícitamente que el modelo de medición utilizado en el análisis de la información supone que el portafolio es unidimensional. Este supuesto se funda en el hecho de que en la actualidad el puntaje final del portafolio se reporta como un solo resultado, sin distinguir entre dimensiones. En este supuesto puede constituirse en una limitación en la medida en que el portafolio no corresponda a una escala unidimensional. Además se utilizaron modelos jerárquicos de dos niveles bajo el supuesto de que estos eran suficientemente complejos como para permitir modelar la variabilidad de la información a nivel del docente y a nivel de su contexto escolar, pero no se controla por la posible anidación de la información a nivel de la comuna. Por último, este estudio no intenta controlar por la posible asignación no aleatoria de profesores a salas de clases o escuelas, ni de niños a escuelas. Los resultados deben interpretarse solo en forma descriptiva, como indicación de asociación entre variables, y no desde una perspectiva causal.

5. RESULTADOS

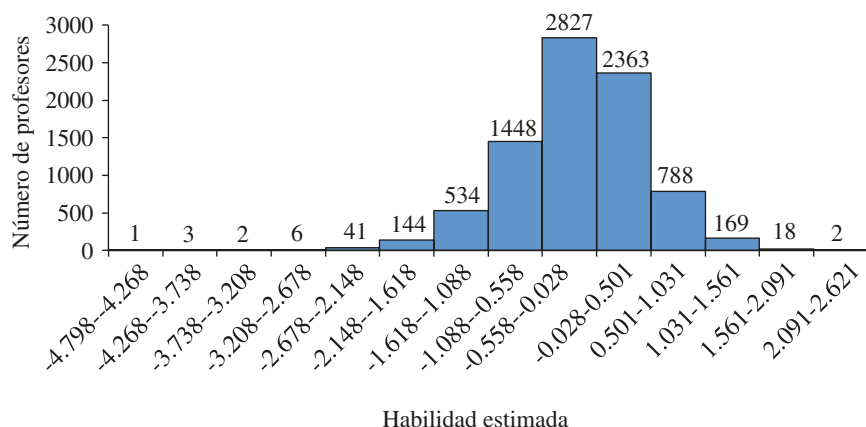
Esta sección presenta los resultados de los análisis realizados para estimar la habilidad de los docentes utilizando el modelo de CP, y aquellos implementados para explorar la relación entre dicha habilidad y las variables del docente y de la escuela.

5.1. ESTIMACIÓN DE HABILIDAD DOCENTE

La distribución de las estimaciones de la habilidad docente (θ) obtenidas al implementar el modelo de CP se muestran en la Figura 1. La media de la habilidad estimada

es -0,198 logits, la mediana es -0,146 logits, la desviación estándar es 0,66 logits. Los coeficientes de asimetría y curtosis corresponden a -0,48 y 1,34.

Figura 1. Histograma de la distribución de habilidad docente, estimada usando el modelo de CP



Para el modelo de CP se observa un comportamiento escalonado de los pasos de progresión de acuerdo con lo esperado en aproximadamente la mitad de los ítems (13 de 24 ítems) y todos los ítems tienen infits y outfits en los rangos 0.5 y 1.5, lo que señala un buen ajuste del modelo utilizado.

Al observar las medias de la habilidad estimada de los profesores, por nivel socioeconómico al cual pertenece cada colegio (ver Tabla 1) se observa que conforme el nivel socioeconómico del colegio es más alto, la habilidad media estimada resulta más alta. Además, la desviación estándar es menor conforme el grupo socioeconómico de pertenencia del colegio es más alto.

Tabla 1. Habilidad media estimada por nivel socioeconómico de los establecimientos educacionales usando el modelo de CP

Nivel socioeconómico	Habilidad estimada media	Desviación estándar	N	Valor mínimo	Valor máximo
A	-0,255	0,658	2,491	-4,083	2,12
B	-0,200	0,658	4,142	-4,083	1,88
C	-0,118	0,645	1,418	-4,798	2,62
D	0,073	0,615	135	-1,347	1,66

5.2. AJUSTE DE LOS MODELOS

La Tabla 2 muestra los criterios de información para los modelos incondicional, 1 y 2. Los tres modelos incluyen un intercepto aleatorio para considerar el hecho de que los profesores se encuentran anidados a nivel escuela, por lo que los errores de sus estimaciones estarán correlacionados. Los modelos difieren en el número de variables explicativas incluidas: el modelo incondicional no incluye variables explicativas adicionales al intercepto aleatorio, el modelo 1 incluye además variables explicativas propias del docente, mientras que el modelo 2 agrega a estas el nivel socioeconómico de la escuela. Se puede observar que el modelo 2 ofrece un mejor ajuste.

Tabla 2. Criterios de información modelos 1 y 2 usando modelo de CP

Criterios de información, Modelo Incondicional, respuesta estimada vía PCM	
2 log de la verosimilitud restringida	16.620,4
Criterio de información de Akaike (AIC)	16.624,4
Criterio bayesiano de Schwarz (BIC)	16.636,5
Criterios de información, Modelo 1, respuesta estimada vía PCM	
-2 log de la verosimilitud restringida	10.517,1
Criterio de información de Akaike (AIC)	10.521,1
Criterio bayesiano de Schwarz (BIC)	10.534,4
Criterios de información, Modelo 2, respuesta estimada vía PCM	
-2 log de la verosimilitud restringida	10.251,3
Criterio de información de Akaike (AIC)	10.255,3
Criterio bayesiano de Schwarz (BIC)	10.268,6

5.3. VARIABLES EXPLICATIVAS

La Tabla 3 muestra las estimaciones de varianzas en los modelos incondicional, 1 y 2. Ambas varianzas resultan significativas (varianza del desempeño de los profesores dentro de los colegios, y varianza del desempeño docente entre los colegios), lo que significa que tanto el nivel del docente como del establecimiento son importantes a la hora de explicar la variabilidad de la habilidad docente estimada, esto es, se producirían diferencias de habilidad entre los profesores dentro de los colegios y diferencias de habilidad entre colegios. La mayor variabilidad, sin embargo, reside en el nivel de los docentes. Tanto en el modelo incondicional como en los dos modelos condicionales el nivel escuela explica aproximadamente 10% de la varianza total.

Las variables incluidas en el modelo 2, el que presenta el mejor ajuste global del modelo, explican el 15% de la varianza observada a nivel docentes y el 20% de la varianza a nivel escuela. La Tabla 4 presenta las estimaciones de efectos fijos del modelo 2. Los resultados a nivel del profesor indican que ser mujer, haber realizado un postítulo y proyectarse en el sector educacional se relacionan positivamente con un mejor desempeño docente. Tener mayor edad estaría relacionado con menores niveles de

Tabla 3. Estimaciones de parámetros de varianza, modelos incondicional 1 y 2, estimación de habilidad usando modelo de CP

Parámetro	Estimación	Error típico	Wald Z	Sig.
Modelo Incondicional				
θ_e	0,388	0,007	54,55	0,000
θ_u	0,048	0,005	9,16	0,000
Modelo 1				
θ_e	0,330	0,008	42,186	0,000
θ_u	0,043	0,006	7,006	0,000
Modelo 2				
θ_e	0,330	0,008	42,217	0,000
θ_u	0,038	0,006	6,538	0,000

desempeño docente, aunque la magnitud de la relación es pequeña. Es interesante notar que no resultan significativas la institución (universidad), ni la modalidad de formación inicial (presencial), ni el trabajar en más de un establecimiento escolar al mismo tiempo o haber participado en programas de formación profesional. Se observan, sin embargo, diferencias en la calidad docente entre profesores que enseñan en distintos subsectores: el desempeño medido por el portafolio es especialmente débil en educadoras de párvulos y profesores de inglés (2° ciclo y educación media) cuando se compara con el desempeño de los docentes generalistas (categoría de referencia).

Tabla 4. Estimaciones de efectos fijos, modelo 2, estimación de habilidad usando modelo CP

Parámetro	Estimación	Error típico	t	Sig.	Intervalo de confianza, 95%	
					Lím. inferior	Lím. superior
Edad*	-0,008	0,001	-12,32	0,000	-0,009	-0,007
Sexo*	0,093	0,0184	5,08	0,000	0,057	0,129
Modalidad Estudios Presencial	0,037	0,036	1,04	0,297	-0,033	0,108
Postítulo (Postítulo, Magíster, Doctorado)*	0,087	0,017	5,20	0,000	0,054	0,120
Perfeccionamiento Docente (PPF)	0,024	0,087	0,27	0,786	-0,148	0,195
Múltiples Colegios	-0,044	0,024	-1,82	0,069	-0,090	0,003
Asistió Universidad	-0,022	0,024	-0,89	0,373	-0,069	0,026
Proyección Educación*	0,090	0,018	5,14	0,000	0,056	0,125
2° Ciclo Lenguaje*	-0,100	0,029	-3,44	0,001	-0,158	-0,043
2° Ciclo Matemática	0,013	0,028	0,45	0,652	-0,043	0,068

Parámetro	Estimación	Error típico	t	Sig.	Intervalo de confianza, 95%	
					Lím. inferior	Lím. superior
2º Ciclo Cs. Naturales*	-0,129	0,034	-3,77	0,000	-0,196	-0,062
2º Ciclo Cs. Sociales*	-0,159	0,035	-4,54	0,000	-0,229	-0,091
2º Ciclo Inglés*	-0,698	0,056	-12,39	0,000	-0,809	-0,588
2º Ciclo Ed. Física*	-0,384	0,050	-7,62	0,000	-0,482	-0,285
2º Ciclo Otro*	-0,347	0,040	-8,56	0,000	-0,426	-0,267
Ed. Media Castellano*	-0,171	0,043	-4,03	0,000	-0,255	-0,088
Ed. Media Matemáticas	0,126	0,048	2,61	0,009	0,031	0,221
Ed. Media Historia	-0,082	0,058	-1,42	0,156	-0,195	0,031
Ed. Media Inglés*	-0,641	0,062	-10,31	0,000	-0,762	-0,519
Ed. Media Biología/Física/Química*	-0,427	0,072	-5,92	0,000	-0,568	-0,285
Ed. Media Ed Física*	-0,271	0,077	-3,50	0,000	-0,422	-0,119
Ed. Media Otro*	-0,414	0,052	-8,00	0,000	-0,515	-0,312
Ed. Especial	0,028	0,036	0,79	0,429	-0,042	0,098
Párvulo*	-0,829	0,053	-15,67	0,000	-0,933	-0,725
Nivel Socioeconómico Establecimiento*	0,067	0,012	5,65	0,000	0,044	0,091

A nivel escuela observamos que el nivel socioeconómico resulta estadísticamente significativo y positivamente asociado a desempeño docente, es decir, se observa mejor desempeño en profesores que trabajan en escuelas de mayor nivel socioeconómico⁶.

6. DISCUSIÓN

El estudio presentado explora la relación entre el desempeño docente y variables tanto del docente como de la escuela. Como variable dependiente se utiliza una estimación de habilidad docente basada en el desempeño de los docentes en el portafolio de la ED. La estimación de habilidad fue obtenida utilizando el modelo de Crédito Parcial, que supone la habilidad docente como una variable no observable. Estas estimaciones de habilidad son independientes del grupo de docentes evaluados y del tipo de ítems completados y considera además el hecho de que los docentes reciben crédito parcial cuando entregan

⁶ La estimación del modelo 2 sin la variable "Proyección en Educación", que permite utilizar una muestra de mayor tamaño, arroja resultados similares. Hay dos diferencias importantes, sin embargo: (i) la formación inicial en una universidad, versus instituciones técnico-profesionales, en este caso sí resulta significativo y positivamente asociado a un mejor desempeño docente. Lo mismo ocurre con (ii) formación presencial. Ambas variables muestran correlaciones importantes con el desempeño docente. Además se observa que la edad en este caso se asocia positivamente al desempeño docente pero la magnitud del coeficiente es pequeño. Los resultados se encuentran disponibles para aquellos lectores interesados.

respuestas parciales. Los resultados del modelo de medición muestran una distribución simétrica de la habilidad docente, centrada cerca del cero y un comportamiento escalonado de los pasos de progresión de acuerdo con lo esperado en aproximadamente la mitad de los ítems.

La investigación implementó además un modelo jerárquico de dos niveles para estudiar la relación entre variables explicativas y la variable dependiente con el fin de reconocer explícitamente la estructura anidada de los datos (docentes en escuelas). El modelo de regresión de mejor ajuste es el que incluye tanto variables propias del docente como controles aleatorios y fijos (nivel socioeconómico) a nivel de la escuela donde trabaja el profesor.

Los resultados coinciden con estudios anteriores (Santelices et al., 2010; Valencia, 2009; Eisenberg, 2008) en mostrar que, aun cuando el nivel escuela es importante, la mayor parte de la variabilidad del desempeño se observa a nivel del docente. El nivel escuela representa aproximadamente 10% de la varianza total del desempeño docente. Es interesante notar que investigaciones con datos del SERCE y del LLECE en los que se usaron modelos jerárquicos de tres y cuatro niveles para modelar desempeño de los alumnos, se estiman que la escuela también representa aproximadamente 10% de la varianza (Cervini, 2012), y de Argentina (Cervini, 2009). Asimismo, los resultados coinciden con las tres investigaciones anteriores en indicar que las docentes mujeres y docentes con postgrado o postítulo obtienen mejores evaluaciones. Al excluir una de las variables explicativas, para la que hay numerosos casos perdidos, se observa que, al igual que en los tres estudios anteriores, resultan significativas la formación del profesor en modalidad presencial y la formación del profesor en universidades (versus instituciones técnicas).

Algunos resultados difieren de lo reportado en estudios anteriores. A diferencia de Santelices et al. (2010), en este estudio se encuentra que proyectarse en el sector educacional se asocia positivamente al desempeño docente. Además los resultados relativos a la importancia de los programas de formación profesional son bastante inestables (significativo en Valencia (2009), no significativo en Santelices et al. (2010) y en esta investigación). Por último, los resultados de esta investigación difieren de Valencia (2009) y Santelices et al. (2010) al encontrar una asociación positiva entre el nivel socioeconómico de la escuela y el desempeño docente. Estos resultados coinciden, sin embargo, con lo reportado por Eisenberg (2008).

Los resultados de esta investigación, junto a estudios anteriores en esta misma línea, subrayan la importancia de la formación inicial y de postítulos y posgrados en el desempeño docente. Esto coincide con los resultados de investigaciones norteamericanas como Akiba et al. (2007), Nye et al. (2004), Rice (2003), Wilson et al. (2001) y Darling-Hammond et al. (2001), y no apoya los hallazgos de Hanushek & Rivkin (2006) y Rivkin, Hanushek & Kain (2005). La importancia de la formación profesional es menos clara.

Del mismo modo, esta línea de investigación indica que los resultados de la ED, y más específicamente los resultados del portafolio, están efectivamente influidos por el contexto, y no solo por las características o atributos personales del profesor. En este sentido, el estudio evidencia que las medidas de prácticas profesionales basadas en estándares profesionales, al igual que medidas de calidad docente basadas en el desempeño longitudinal de los estudiantes, no están aisladas del contexto.

7. CONCLUSIÓN

Los resultados de esta investigación muestran que al utilizar estimaciones de habilidad docente basadas en el modelo de Crédito Parcial el nivel del docente corresponde al porcentaje más importante en la varianza total del desempeño docente medido por el portafolio. Sin embargo, el contexto, conceptualizado como un segundo nivel correspondiente a la escuela, corresponde aproximadamente a 10% de la varianza total del desempeño docente. Estos resultados coinciden con resultados de investigaciones anteriores (Santelices et al., 2010; Manzi et al., 2007). Además, los resultados coinciden con investigaciones anteriores (Santelices et al., 2010; Valencia, 2009; Eisenberg, 2008) respecto de la importancia de la formación inicial, los postítulos y posgrados en la calidad de su práctica profesional se mantienen al utilizar una estimación de habilidad docente basada en la Teoría de Respuesta al Ítem. Al utilizar la estimación de habilidad docente basada en el modelo de CP no se encuentra evidencia respecto de la importancia de los programas de formación profesional, pero sí resulta importante el nivel socioeconómico de la escuela en que trabaja el docente.

Es importante confirmar en futuros estudios si estos hallazgos se mantienen al utilizar modelos de tres niveles en combinación con estimaciones de habilidad basados en la Teoría de Respuesta al ítem como los empleados en esta investigación. Además, próximas investigaciones deberían explorar la dimensionalidad del portafolio de manera de confirmar si el supuesto de unidimensionalidad utilizado en esta investigación es apropiado. Conceptualmente, además, el hecho de que las distintas medidas de calidad docente desarrolladas hasta ahora estén influenciadas por el contexto plantea desafíos tanto a la medición de las prácticas como a la formación de profesores.

REFERENCIAS BIBLIOGRÁFICAS

- Akiba, M., LeTendre, G. K. & Scribner, J. P. (2007). Teacher quality, opportunity gap, and national achievement in 46 countries. *Educational Administration Quarterly*, vol. 43 (4), 399-432.
- Allen, M. (2003). *Eight questions on teacher preparation: What does the research say? A summary of the findings*. Denver, CO: Education Commission of the States.
- Archibald, S. (2007). How well do standards-based teacher evaluation scores identify high quality teachers? A multilevel, longitudinal analysis of one district. *Annual Meeting of the American Educational Research Association*. Chicago, United States.
- Bravo, D., Falck, D., González, R., Manzi, J. y Peirano, C. (2008). La relación entre la evaluación docente y el rendimiento de los alumnos: Evidencia para el caso de Chile. *Encuentro Anual de la Sociedad de Economía de Chile*. Viña del Mar, Chile. Recuperado el 8 de noviembre de 2012 desde http://www.microdatos.cl/docto_publicaciones/Evaluacion%20docentes_rendimiento%20escolar.pdf
- Cecilio, D. (2011). *Construcción de un test informatizado transmodal de memoria de reconocimiento*. Salamanca: Universidad de Salamanca.
- Cervini, R. A. (2012). El “efecto escuela” en países de América Latina: Reanalizando los datos del SERCE. *Archivos Analíticos de Políticas Educativas*, vol. 20 (39), 1-25.
- . (2009). Class, school, municipal and state effects on mathematic achievement in Argentina: A multilevel analysis. *School Effectiveness and School Improvement*, vol. 20 (2), 319-340.
- CPEIP (2005a). *Fortalecimiento de la profesión docente*. Recuperado el 9 de octubre de 2007 desde <http://www.rmm.cl/usuarios/equiposite/doc/200501121331010.politica%20fortprof%20docente.doc>

———. (2005b). *Políticas educacionales durante los gobiernos democráticos*. Recuperado el 9 de octubre de 2007 desde <http://www.rmm.cl/usuarios/equiposite/doc/200501121331270.polt%20educ.doc>

Danielson, C. (2007). *Enhancing professional practice: A framework for teaching* (2nd Ed.). Alexandria, V.A.: Association for Supervision and Curriculum Development.

———. (1996). *Enhancing professional practice: A framework for teaching*. Alexandria, V.A.: Association for Supervision and Curriculum Development.

Darling-Hammond, L., Berry, L. B. & Thoreson, A. (2001). Does teacher certification matter? Evaluating the evidence. *Educational Evaluation and Policy Analysis*, vol. 23 (1), 57-77.

Eisenberg, N. (2008). *The performance of teachers in chilean public elementary schools: Exploring its relationship with teacher backgrounds and student achievement, and its distribution across schools and municipalities*. Ph.D. thesis. Los Angeles: University of California.

Embretson, S. & Reise, S. (2000). *Item response theory for psychologists*. Hillsdale, NJ: Lawrence Erlbaum.

Gordon, R., Kane, T. & Steiger, D. (2006). *Identifying effective teachers using performance on the job*. Washington, D.C.: The Brookings Institution.

Hambleton, R., Swaminathan, H. & Rogers, H. (1991). *Fundamentals of item response theory*. New York: Sage.

Hanushek, E. & Rivkin, F. (2006). Teacher quality. In E. Hanushek & F. Rivkin (Eds.), *Handbook of the economics of education* (Vol. 2) (pp. 1052-1075). Amsterdam: Elsevier.

Heck, R. H. (2006). Examining the relationship between teacher quality as an organizational property of schools and students' achievement and growth rates. *Journal of Special Education*, vol. 40 (2), 115-127.

Heneman, H. G. & Milanowski, A. T. (2004). Alignment of human resource practices and teacher performance competency. *Peabody Journal of Education*, vol. 79 (4), 108-125.

Kennedy, M. (2011). Attribution error and the quest for teacher quality. *Educational Researcher*, vol. 39 (8), 591-598.

———. (2008). Contributions of qualitative research to research on teacher qualification. *Educational Evaluation and Policy Analysis*, vol. 30 (4), 344-367.

Lara, B., Mizala, A. y Repetto, A. (2010). Una mirada a la efectividad de los profesores en Chile. *Estudios Públicos*, (120), 147-182.

León, M. G., Manzi, J. y Paredes, R. D. (2008). Calidad docente y rendimiento escolar en Chile: Evaluando la evaluación. *Encuentro Anual de la Sociedad de Economía de Chile*. Viña del Mar, Chile. Recuperado el 7 de octubre de 2009 desde <https://repositorio.uc.cl/handle/11534/1447>

Linacre, J. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transaction*, vol. 16 (2). Retrieved November 8, 2012 from <http://www.rasch.org/rmt/rmt162f.htm>

Long, J. (1997). *Regression models for categorical and limited dependent variables*. Thousand Oaks, C.A.: Sage.

Manzi, J., Strasser, K., San Martín, E. & Contreras, D. (2007). *Quality of education in Chile*. Washington, D.C.: Interamerican Development Bank. Retrieved November 8, 2012 from: <http://www.iadb.org/en/research-and-data/institution-details,3191.html?idD144>

McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., Louis, T. A. & Hamilton, L. S. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics*, vol. 29 (1), 67-101.

McKinsey & Company (2007). *How the world's best-performing school systems come out on top*. Retrieved October 9, 2007 from http://mckinseysociety.com/downloads/reports/Education/Worlds_School_Systems_Final.pdf

Milanowski, A. (2011). Validity research on teacher evaluation systems based on the framework for teaching. *Annual Meeting of the American Educational Research Association*. New Orleans, United States.

Millman, J. (Ed.). (1997). *Grading teachers, grading schools: Is student achievement a valid evaluation measure?* Thousand Oaks, C.A.: Corwin Press.

Ministerio de Educación. (2004). *Marco para la buena enseñanza*. Santiago de Chile: MINEDUC.

National Academy of Education (2009). *Teacher quality education policy white paper*. Retrieved November 9, 2012 from http://www.naeducation.org/cs/groups/naedsite/documents/webpage/naed_080867.pdf

Nye, B., Konstantopoulos, S. & Hedges, L. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, vol. 26 (3), 237-257.

Odden, A., Borman, G. & Fermanich, M. (2004). Assessing teacher, classroom and school effect including fiscal effect. *Peabody Journal of Education*, vol. 79 (4), 4-32.

Opendakker, M. & Van Damme, J. (2000). The importance of identifying in multilevel analysis: An illustration of the effects of ignoring the top or intermediate levels in school effectiveness research. *School Effectiveness and School Improvement*, vol. 11 (1), 103-130.

Rice, J. K. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes*. Washington, D.C.: Economic Policy Institute.

Rivkin, S. G., Hanushek, E. A. & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, vol. 73 (2), 417-458.

Santelices, M. V., Valencia, E., Taut, S. & Manzi, J. (2010). The importance of contextual and personal variables in explaining teacher quality: Evidence using standard-based, standardized measure of teacher performance. *Annual Meeting of the American Educational Research Association*. Denver, United States.

Taut, S., Santelices, M. V. & Stecher, B. (2012). Validation of a national teacher assessment and improvement system. *Educational Assessment*, vol. 17 (4), 163-199.

_____, Araya, C. & Manzi, J. (2010). Theory underlying a national teacher evaluation program. *Evaluation and Program Planning*, vol. 33 (4), 477-489.

Toledo, G. (2010). *Calidad docente y logro escolar: Enfrentando el problema de ordenamiento no aleatorio entre características de profesores y alumnos*. Tesis para optar al grado de Magíster en Economía. Santiago: Universidad de Chile.

Valencia, E. (2009). *La relación entre los antecedentes individuales, las condiciones de trabajo y el contexto local con la calidad del docente del sistema educacional público chileno*. Tesis para optar al grado de Magíster en Psicología Educacional. Santiago: Pontificia Universidad Católica de Chile.

Van den Noortgate, W., Opendakker, M. & Oghena, P. (2005). The effects of ignoring a level in multilevel analysis. *School Effectiveness and School Improvement*, vol. 16 (3), 281-303.

Wilson, S., Floden, R. E. & Ferrini-Mundy, J. (2001). *Teacher preparation research: Current knowledge, gaps, and recommendations*. Seattle, W.A.: Center for the Study of Teaching and Policy.

Wright, B. & Masters, G. (1982). *Rating scale analysis*. Chicago: MESA Press.