

¿Cómo se relacionan los resultados de la Evaluación Docente con el aprendizaje que alcanzan los estudiantes de profesores evaluados?

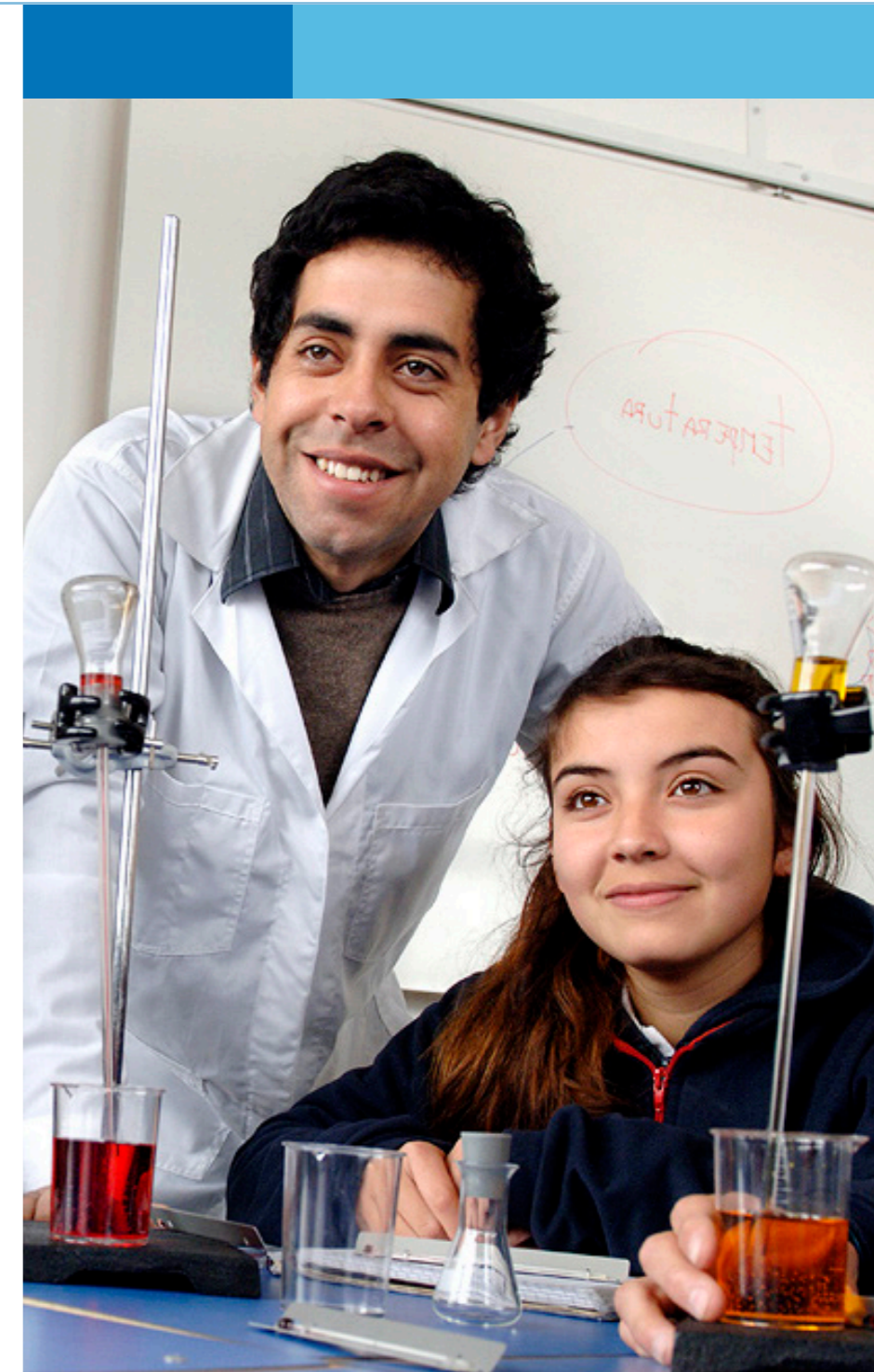


■ Autora: Sandy Taut.

Esta edición de Midevidencias presenta una investigación que analiza mediante el uso de modelos de valor agregado la relación entre los resultados de la Evaluación Docente, en conjunto y para cada uno de sus instrumentos, con el aprendizaje de los estudiantes de los profesores evaluados, medido a través del SIMCE de Lenguaje y Matemática. Los resultados del estudio muestran que el portafolio es el instrumento de la Evaluación Docente que más se relaciona con el aprendizaje de los estudiantes, siendo la Interacción pedagógica, el Manejo del aula y la Reflexión sobre la propia práctica, los aspectos que tienen relaciones más significativas en ambas disciplinas. Los resultados del estudio apoyan la recomendación de la investigación internacional sobre el uso del portafolio con fines de certificación y formativos.

Esta investigación analiza la relación entre los resultados de la Evaluación Docente en Chile y el aprendizaje de los estudiantes de profesores evaluados. Este tema es relevante para determinar si la Evaluación Docente es un sistema de medición que recoge información válida. ¿Por qué es relevante examinar la validez de un sistema de medición como la Evaluación Docente? En el campo de la medición educacional la validez es un concepto clave que implica un juicio respecto de si decisiones basadas en los resultados de una medición son defendibles o inadecuadas (AERA, APA & NCME,

2014; Kane, 2006). Una validación es especialmente crucial en contextos de decisiones que tienen importantes consecuencias (high stakes) para los evaluados, como es el caso de la Evaluación Docente en Chile. Ya existe un importante cuerpo de evidencia respecto de la validez de la Evaluación Docente (Taut, Santelices & Stecher, 2012) pero el estudio descrito aquí entrega un aspecto fundamental que estaba pendiente: ¿Cómo se relacionan los resultados de la Evaluación Docente con el aprendizaje que alcanzan los estudiantes de profesores evaluados?



Relación entre Evaluación Docente y logros de aprendizaje: La evidencia chilena

En Chile existen varios estudios que asocian los resultados que obtienen los profesores en la Evaluación Docente con el aprendizaje de los estudiantes. Estos estudios muestran una relación consistente entre los resultados de la evaluación y los logros de aprendizaje de los estudiantes, medidos a través de la prueba SIMCE (Sistema de Medición de la Calidad de la Educación). En particular, el portafolio es el instrumento que se asocia más fuertemente con el aprendizaje.

Sin embargo, estos estudios utilizan datos de la prueba SIMCE en solo un momento del tiempo y correlacionan estos resultados de los estudiantes con los resultados de los docentes, en su gran mayoría estableciendo el vínculo entre estudiantes y docentes a nivel escuela (Manzi, Strasser, San Martín, & Contreras, 2008; Bravo, Falck, González, Manzi & Peirano, 2008; Eisenberg, 2008; León, 2008; Alvarado, Cabezas, Falck & Ortega, 2012; Ministerio de Educación, 2012; Loyola, 2013). El estudio presentado aquí es particularmente relevante porque es el primero en establecer un vínculo directo entre el resultado de profesores en la Evaluación Docente y el progreso en el aprendizaje de sus estudiantes, medidos en dos momentos del tiempo, mientras que trabajaban juntos.

Aspectos metodológicos del estudio

Para responder la pregunta que plantea este número de Midevidencias, el presente estudio utiliza una metodología especialmente robusta. Se combinaron datos de los profesores en la Evaluación Docente con datos de sus estudiantes en la prueba SIMCE. Específicamente, se usaron datos de una cohorte de estudiantes que rindió la prueba SIMCE dos veces en el tiempo, primero en 8° básico (en 2004) y luego en II medio (en 2006). Para el análisis sólo se incluyeron docentes que se evaluaron en 2006 y que enseñaron al mismo grupo de estudiantes durante I y II medio (2005-2006). Además, tenían que tener datos completos para un mínimo de 10 estudiantes. Así, obtuvimos una muestra de 150 profesores de matemática (y 5.445 estudiantes) y 162 profesores de Lenguaje (y 5.998 estudiantes). Aplicamos la metodología de valor agregado (MVA) para obtener una estimación de la efectividad de los docentes¹. En un segundo paso se correlacionaron estos índices de VA de los profesores con los resultados que ellos habían obtenido en la Evaluación Docente.

¿Qué son los modelos de valor agregado para profesores, y para qué sirven?

Los modelos de valor agregado intentan capturar la contribución específica de un profesor al aprendizaje de sus alumnos usando dos o más mediciones de desempeño de los mismos alumnos en el tiempo (McCaffrey et al., 2003; Lissitz, 2005). Todos los modelos de valor agregado se basan en controlar el punto de partida de los alumnos, y en muchos casos, controlan además variables adicionales de características de los alumnos, de sus pares, y del contexto escolar (por ejemplo: género, nivel socioeconómico, varianza en el rendimiento del grupo de pares), todos los cuales describen aspectos que influyen en la labor docente pero que están fuera del control del profesor (Braun, 2005). Así se puede estimar la contribución del docente al aprendizaje de los estudiantes de forma más precisa, la cual además es relativa al grupo de profesores incluidos en el análisis. Es importante destacar que nosotros estamos usando el aprendizaje de los estudiantes solamente para examinar la validez de la Evaluación Docente y sus instrumentos y no proponemos el uso del aprendizaje de los estudiantes para evaluar a sus profesores dadas las complejidades que esto tiene (Baker et al., 2010; OCDE, 2013; Papay, 2011; Amrein-Bearnsley, 2014).

1. Obtuvimos estas estimaciones de valor agregado para cada profesor a través de regresiones multinivel, contrastando diferentes especificaciones para los modelos.

En la Evaluación Docente, el portafolio es el instrumento que más se asocia al aprendizaje de los estudiantes

Antes de responder la pregunta sobre si los profesores con mejor desempeño en la Evaluación Docente, logran mejores resultados en el aprendizaje de sus estudiantes -lo que se expresaría en que tuvieran un mayor valor agregado-, es necesario aclarar si existen diferencias de aprendizaje entre aulas. Si existen diferencias entre aulas, entonces podemos suponer que una variable importante que incide en ello es el profesor. En cambio, si no la hay, entonces estudiar la efectividad del profesor para promover el aprendizaje perdería sentido. El estudio encuentra que en Matemática, un 42% de la variación en los resultados SIMCE de los estudiantes de II medio corresponde a diferencias entre aulas, y un 58% corresponde a diferencias entre estudiantes dentro de estas aulas².

En términos de la pregunta central, el estudio encuentra una relación positiva entre los resultados en la Evaluación Docente (ED) de los profesores y el aprendizaje de sus estudiantes. Estos resultados son evidencia que apoya la validez de la Evaluación Docente. Las relaciones son más fuertes entre aprendizaje de los alumnos y resultados del profesor en el portafolio de la ED y con el resultado final (ver Tabla 1). Encontramos relaciones positivas pero más débiles con la evaluación que entrega el director, y cercanas a cero con la entrevista del evaluador par y la autoevaluación. Las relaciones se observan más fuertemente en Matemática que en Lenguaje. La Tabla 1 muestra las relaciones (correlaciones)³ para

Tabla 1. Correlaciones entre puntajes ED y resultados de aprendizaje (Modelo de Valor Agregado, MVA⁴) en Matemática (N=150) y Lenguaje (N=162).

| | Valor agregado Lenguaje | Valor agregado Matemática |
|---|-------------------------|---------------------------|
| Auto-evaluación | -0.08 | 0.07 |
| Evaluación Par | 0.07 | 0.08 |
| Evaluación Directivos | 0.05 | 0.14 |
| Portafolio | 0.19* | 0.33** |
| Resultado final | 0.16* | 0.28** |
| ** Los asteriscos indican que estas correlaciones son estadísticamente significativas (diferentes de cero). | | |

Matemática y Lenguaje, respectivamente. A partir de estudios internacionales previos se ha establecido que es evidencia clara de validez cuando se encuentran correlaciones de 0.3 a 0.5 (Milanowski, 2004; Hill, Kapitula & Umland, 2011; Kane, Taylor, Tyler & Wooten, 2010; Grossman, Loeb, Cohen, Hammerness, Wyckoff, Boyd & Lankford, 2010). Correlaciones por debajo de 0.3 pero significativamente diferente a cero muestran una relación, aunque más débil, y no debieran ignorarse.

2. En este estudio no incluimos el nivel escuela en nuestros análisis, por lo que no sabemos en qué medida esta variación corresponde a variación entre escuelas. 3. Las correlaciones (denominado con la letra "r") se expresan en una escala donde -1 significa una relación inversa perfecta, 0 representa ninguna relación entre ED y aprendizaje, y 1 significa una relación directa perfecta. 4. Modelo de valor agregado que controla el rendimiento previo a nivel individual y grupal (del aula).

Los aspectos del portafolio que más se asocian al aprendizaje de los estudiantes

El estudio también examina la relación entre seis aspectos específicos de la calidad del proceso de enseñanza-aprendizaje que se evalúan en el portafolio de la ED, por un lado, y el aprendizaje de los estudiantes, por otro lado. La Tabla 2 muestra que la relación más fuerte ($r=0.21$) y consistente tanto para Matemática como Lenguaje existe con la calidad de la Interacción pedagógica (la cual incluye, por ejemplo, la calidad de las explicaciones que entregan los docentes y el tipo de preguntas que hacen para promover el pensamiento⁵). Otros aspectos evaluados por el portafolio que resultan con relaciones significativas para ambas asignaturas corresponden al Manejo del aula y a la Reflexión sobre la propia práctica.

La validez del portafolio respalda su uso para evaluar docentes

El portafolio es el principal instrumento de la Evaluación Docente, y es uno de los instrumentos considerados para determinar avance en la carrera docente. Este estudio aporta evidencia relevante sobre la validez de este instrumento. Los resultados muestran que los profesores más efectivos, es decir, aquellos que promueven mayores progresos en los aprendizajes de sus estudiantes, son quienes poseen mejores resultados en el portafolio, y por el contrario, los docentes menos efectivos obtienen más bajos resultados en este instrumento.

Tabla 2.

Correlaciones entre aspectos específicos que evalúa el portafolio y resultados de aprendizaje (Modelo de Valor Agregado, MVA⁶) en Matemática y Lenguaje.

| | Valor agregado Lenguaje | Valor agregado Matemática |
|---|-------------------------|---------------------------|
| Módulo 1 del portafolio (evidencia escrita) | | |
| Planificación pedagógica | 0.00 | 0.14 |
| Evaluación de aprendizajes | 0.09 | 0.16* |
| Reflexión sobre la propia práctica | 0.15* | 0.21* |
| Módulo 2 del portafolio (clase grabada) | | |
| Interacción pedagógica | 0.21* | 0.21** |
| Manejo del aula | 0.17 | 0.21** |
| Estructura de la clase | 0.14 | 0.19 |
| ** Los asteriscos indican que estas correlaciones son estadísticamente significativas (diferentes de cero). | | |

Los otros instrumentos de la ED muestran relaciones más bajas, por lo que no son tan robustos para identificar profesores efectivos en términos del aprendizaje que generan en sus estudiantes –al menos en el contexto actual de altas consecuencias-. Es probable que la falta de relaciones para estos instrumentos reflejen un sesgo al alza que esconde diferencias reales entre los profesores evaluados. Sin embargo, aun cuando estos instrumentos no parecen resultar útiles para discriminar docentes según su efectividad, sí podrían ser pertinentes para la evaluación formativa en contextos locales.

5. Esta información se basa en el video que forma parte de la Evaluación Docente de cada profesor evaluado. 6. Modelo de valor agregado que controla el rendimiento previo a nivel individual y grupal (del aula).



Dado el tamaño de las relaciones que se encontraron, es importante recordar que hay muchos otros factores no incluidos en los modelos estadísticos que inciden en el aprendizaje de los estudiantes, reflejando la complejidad de evaluar la calidad del proceso de enseñanza-aprendizaje. Considerando esta complejidad, a la hora de tener que evaluar la calidad docente en un sistema nacional de calificaciones para el cual se debe comprobar su validez, se sugiere mantener los instrumentos que ya cuentan con esta evidencia de validez, como es el caso del portafolio.

Este estudio también subraya la importancia de evaluar lo que sucede en el aula: el aspecto del portafolio que más se asocia con el aprendizaje de los estudiantes es la calidad de la Interacción pedagógica, la cual se puede evaluar únicamente al observar lo que pasa en el aula. En esta dimensión de la

calidad del proceso de enseñanza-aprendizaje se evalúa que el docente dé cuenta a lo largo de la clase grabada que es capaz de explicar los contenidos de la asignatura, empleando estrategias que promuevan el establecimiento de relaciones significativas entre los contenidos y el aprendizaje previo, que sepa retroalimentar a los alumnos para que puedan identificar la razón que está detrás de sus aciertos y errores, y que sepan formular preguntas que favorezcan el desarrollo de habilidades superiores. El manejo de la conducta en el aula y la capacidad auto-reflexiva de los docentes sobre su propia práctica, son otros aspectos evaluados por el portafolio que se asocian con el aprendizaje de los estudiantes.

A esto se suma que a nivel internacional el portafolio, y especialmente la observación de aula (directa o a través de filmaciones), es el método más respaldado para evaluar el desempeño docente en aula,

sobre todo porque este instrumento también tiene un alto potencial para entregar retroalimentación específica a los docentes sobre los aspectos más y menos logrados de su práctica (OCDE, 2013; Bill & Melinda Gates Foundation, 2013).

A partir de los resultados presentados en este Midevidencias, se puede concluir que el portafolio de la Evaluación Docente constituye una base sólida para evaluar la práctica docente en Chile. Este instrumento debiera ser una herramienta fundamental en este contexto, ya que entrega un diagnóstico de fortalezas y debilidades que puede orientar el desarrollo profesional docente, además de demostrar una relación clara con el aprendizaje de estudiantes de profesores evaluados.

Referencias

- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (AERA, APA & NCME) (2014). *Standards for educational and psychological testing*. Washington, D.C.: American Educational Research Association.
- Alvarado, M., Cabezas, G., Falck, D. & Ortega, M. E. (2012). *La Evaluación Docente y sus instrumentos: Discriminación del desempeño docente y asociación con los resultados de los estudiantes*. Centro de Estudios MINEDUC & PNUD.
- Amrein-Beardsley, A. (2014). *Rethinking value-added models in education. Critical perspectives on tests and assessment-based accountability*. New York: Routledge.
- Baker, E., Barton, P., Darling-Hammond, L., Haertel, E., Ladd, H., Linn, R., Ravitch, D., Rothstein, R., Shavelson, R., Shepard, L. (2010). *Problems with the use of student test scores to evaluate teachers*. Washington, DC: Economic Policy Institute.
- Bill and Melinda Gates Foundation (2012). *Gathering feedback for teaching. Combining high-quality observations with student surveys and achievement gains*. Descargado de: http://www.metproject.org/downloads/MET_Gathering_Feedback_Practitioner_Brief.pdf
- Braun, H. (2005). *Using Student Progress to Evaluate Teachers: A Primer on Value-Added Models*. Educational Testing Services (ETS). Descargado de: <http://www.ets.org/Media/Research/pdf/PICVAM.pdf>
- 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*. Descargado de: http://www.microdatos.cl/docto_publicaciones/Evaluacion%20docentes_rendimiento%20escolar.pdf
- 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*. Doctoral thesis, University of California Los Angeles, USA.
- Grossman, P., Loeb, S., Cohen, J., Hammerness, K., Wyckoff, J., Boyd, D. & Lankford, H. (2010). *Measure for measure: The relationships between measures of instructional practice in middle school English language arts and teachers' value-added scores*. (Working Paper N° 16015). Descargado de: National Bureau of Economic Research Working Paper website: <http://www.nber.org/papers/w16015>
- Hill, H. C., Kapitula, L. & Umland, K. (2011). *A validity argument approach to evaluating teacher value-added scores*. *American Educational Research Journal*, 48(3), 794–831. DOI:10.3102/0002831210387916
- Kane, M. T. (2006). *Validation*. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17–64). Westport, CT: Praeger Publishers.
- Kane, T. J., Taylor, E. S., Tyler, J. H. & Wooten, A. L. (2010). *Identifying effective classroom practices using student achievement data* (NBER Working Paper No. 15803). Cambridge, MA: National Bureau of Economic Research.
- León, M. G. (2008). *Calidad docente y rendimiento escolar en Chile*. Tesis para optar al grado de Magíster en Ciencias de la Ingeniería. Pontificia Universidad Católica de Chile, Santiago, Chile. Descargado de: <http://repositorio.uc.cl/xmlui/bitstream/handle/123456789/1447/507135.pdf?sequence=1>
- Lissitz, R. (Ed.) (2005). *Value Added Models in Education: Theory and Applications*. Maple Grove: JAM Press.
- Loyola, V. (2013). *Efectos Fijos Profesor y la Evaluación Docente: El Caso de Chile*. Tesis de Magíster en Economía. Pontificia Universidad Católica de Chile, Santiago, Chile. Descargado de: <http://www.economia.puc.cl/docs/tesis-vloyola-2013.pdf>
- Manzi, J., Strasser, K., San Martín, E., & Contreras, D. (2008). *Quality of education in Chile*. Descargado de: <http://www.iadb.org/res/laresnetwork/files/pr-300finaldraft.pdf>
- McCaffrey, D. F., Lockwood, J. R., Hamilton, L. & Koretz, D. (2003). *Evaluating Value-Added Models for Teacher Accountability*. Santa Monica, CA: RAND.
- Milanowski, A. (2004). *The relationship between teacher performance evaluation scores and student achievement: Evidence from Cincinnati*. *Peabody Journal of Education*, 79(4), 33-53. DOI: 10.1207/s15327930pje7904_3
- Ministerio de Educación, (2012). *Evaluación Docente y resultados de aprendizaje: ¿Qué nos dice la evidencia? Centro de Estudios MINEDUC*, 1(6), 1-6.

- Organización para la Cooperación y el Desarrollo Económico [OECD]. (2013). *Teachers for the 21st century: Using evaluation to improve teaching*. OECD Publishing.
- Papay, J. (2011). Different tests, different answers: The stability of teacher value-added estimates across outcome measures. *American Education Research Journal*, 48(1), p.163-193. DOI:10.3102/0002831210362589.
- Taut, S., Santelices, V. & Stecher, B. (2012). Validation of a national teacher assessment and improvement system. *Educational Assessment Journal*, 17(4), 163-199. DOI:10.1080/10627197.2012.735913

Más información en:

- Taut, S., Valencia, E., Palacios, D., Santelices, V., Jimenez, D. & Manzi, J. (2014). Teacher performance and student learning: Linking evidence from two national assessment programs. *Assessment in Education: Principles, Policy & Practice*, DOI: 10.1080/0969594X.2014.961406
- Taut, S., Valencia, E. & Escobar, J. (2012). *La validez de la Evaluación Docente en Chile usando como criterio estimaciones de Valor Agregado de profesores de Enseñanza Media*. Descargado de:

<http://mideuc.cl/wp-content/uploads/2013/01/IT1202.pdf>

Acerca de la autora

Sandy Taut es psicóloga de la Universität zu Köln, Alemania y Ph.D. en Educación de la University of California Los Angeles, EE.UU. Actualmente es profesora asistente de la Escuela de Psicología e investigadora del Centro de Medición MIDE UC, Pontificia Universidad Católica de Chile.

[\(staut@uc.cl\)](mailto:staut@uc.cl)

Cómo citar esta publicación:

Taut, S. (2015). ¿Cómo se relacionan los resultados de la Evaluación Docente con el aprendizaje que alcanzan los estudiantes de profesores evaluados? *Mid evidencias 2*, 1-7. Extraído de:

<http://www.mideuc.cl/wp-content/uploads/2015/06/MidEvidencias-N2.pdf>