



INFORMES TÉCNICOS MIDE UC /TECHNICAL REPORTS MIDE UC

Centro de medición MIDE UC/Measurement Center MIDE UC

**Mejoras en el desempeño docente de los profesores del sistema municipal**

ÁLVARO ZAPATA Y MAXIMILIANO HERRADA

Diciembre del 2013



ESCUELA DE PSICOLOGÍA

# INDICE

<b>1. Resumen .....</b>	<b>3</b>
<b>2. Introducción.....</b>	<b>4</b>
2.1 Evidencia y argumentos sobre el progreso en la evaluación docente .....	5
2.2 Confiabilidad en la ganancia de puntajes .....	8
<b>3. Objetivos del estudio .....</b>	<b>9</b>
<b>4. Metodología .....</b>	<b>10</b>
4.1 Obtención de la muestra e indicadores relevantes .....	10
4.1.1 Elección de los indicadores comparables .....	11
4.1.2 Elección de los correctores con mayor consistencia.....	13
4.2 Análisis de Datos .....	13
<b>5. Resultados .....</b>	<b>14</b>
5.1 Caracterización de la muestra.....	14
5.2 Comparabilidad de los puntajes del portafolio .....	17
5.3 Estimación de la confiabilidad de la ganancia de puntaje y categorización de los docentes que mejoran .....	20
5.4 Variables que pudieran explicar el cambio positivo del desempeño docente. ....	24
<b>6. Discusión y Conclusiones .....</b>	<b>27</b>
<b>7. Referencias .....</b>	<b>30</b>

<b>8. Anexos .....</b>	<b>31</b>
8.1 Indicadores no comparables entre el periodo 2007 y 2011 .....	31
8.2 Contrastes en la comparación de re-evaluados.....	32
8.3 Confiabilidad, correlación e Índice en la ganancia de Puntaje .....	32
8.4 Predictores utilizados para explicar la mejora significativa en los puntajes .....	33

## 1. Resumen

El programa de evaluación docente, en conjunto con otras políticas, tiene como finalidad, apoyar la la mejora continua de los profesores municipales Considerando los 10 años que lleva ejecutándose la Evaluación Docente, existe un creciente interés por reportar progresos en los desempeños de los evaluados como medida de impacto de los programas de mejora profesional docente.

La posibilidad de reportar progresos en los puntajes, enfrenta algunas dificultades técnicas tales como lograr la comparabilidad de los Portafolios entre evaluaciones realizadas en distintos períodos, corregir considerando los índices de confiabilidad de la corrección del Portafolio y estimar la confiabilidad en las diferencias detectadas para determinar la magnitud de las diferencias que podemos identificar como progreso significativo.

Para llevar a cabo este estudio la muestra se consideró en base a los re-evaluados 2007-2011.

Los resultados indican que considerando la muestra de docentes corregidos por correctores más consistentes en los indicadores comparables del Portafolio, existe en promedio una mejora en el puntaje global de portafolio de 0.08 puntos, con una alta varianza en la distribución. Dada la magnitud de la mejora y la heterogeneidad de las trayectorias, se hizo necesario estimar índices de confiabilidad de los cambios. Los resultados de este ejercicio señalan que solo es posible afirmar que 14 de cada 100 reevaluados logra obtener una mejora sustantiva en su desempeño entre una y otra evaluación. El mayor predictor de este cambio es el puntaje previo del docente. Luego pueden considerarse, la participación en cursos del Programa de Perfeccionamiento Fundamental (previo al 2007), la zona geográfica del establecimiento en que trabajan y la edad del docente al evaluarse por primera vez.

Se discute sobre la replicabilidad de este estudio y el establecimiento de una metodología para reportar progresos, además del contexto actual sobre la influencia en la mejora del desempeño docente.

## 2. Introducción

La evaluación docente se enmarca en una política pública más general que articula el diagnóstico sobre las fortalezas y debilidades del trabajo docente en aula con Planes de Superación Profesional (PSP) orientados a los docentes con bajo desempeño que permitan el desarrollo profesional continuo. A 10 años de implementación de la Evaluación Docente se hace cada vez más necesario responder a la pregunta de si es posible reportar mejoras en el desempeño de los docentes, expresadas en una evolución favorable en de los puntajes que obtienen en el Portafolio y si estas, están asociadas con los esfuerzos explícitos para generar desarrollo profesional.

El reglamento de la evaluación docente (*Artículo N°1, Decreto 192, AAAA*), documento legal, establece que la evaluación docente debe apuntar a la mejora continua de la labor del profesor:

***"Evaluación Docente":*** Sistema de evaluación de los profesionales de la educación que se desempeñen en funciones de docencia de aula, de carácter formativo, orientado a mejorar la labor pedagógica de los educadores y a promover su desarrollo profesional continuo según lo establecido en el artículo 70 del Decreto con Fuerza de Ley N° 1 de 1996, del Ministerio de Educación, con niveles de desempeño que corresponden a destacado, competente, básico o insatisfactorio. (p.2)

En consecuencia es de interés que una política pública de esta envergadura sea capaz de señalar mejoras en los docentes. En caso contrario, la evaluación docente pierde sentido dado que los profesores no reportan cambios, ya sea porque las políticas de cambio, como PSP, han sido pobres o porque el instrumento no es sensible a los cambios.

A pesar que no existen consecuencias asociadas a la mejora de los docentes, si es necesario evaluar el progreso de los docentes dado que es la intención final de la política.

## ***2.1 Evidencia y argumentos sobre el progreso en la evaluación docente***

El sistema de evaluación docente recoge evidencias de trabajo docente a través de un Portafolio estandarizado corregido por docentes de aula, entrenados en el uso de rúbricas estandarizadas que puntúan el trabajo docente. En total los correctores, a través de las rúbricas, corrigen 24 competencias profesionales las que se agrupan en 8 dominios relevantes del trabajo docente<sup>1</sup>: Organización de la unidad, Análisis de las clases, Calidad de la evaluación, Reflexión a partir de los resultados de los alumnos, Reflexión pedagógica, Ambiente de la clase, Estructura de la clase e Interacción pedagógica.

Cada docente recibe un reporte con el nivel de desempeño final y una retroalimentación específica correspondiente al nivel de desempeño obtenido en cada uno de los indicadores evaluados en su evidencia. Con esos indicadores se calcula un puntaje promedio en dimensión y un Portafolio global. El puntaje de Portafolio, que es el promedio de las 8 dimensiones, representa un 60% del puntaje final del docente evaluado que determina su nivel de desempeño final<sup>2</sup>. El resultado en estos indicadores también se pone a disposición de los ejecutores de los Planes de Superación Profesional, para que estos implementen los planes de ayuda a los docentes con menor desempeño.

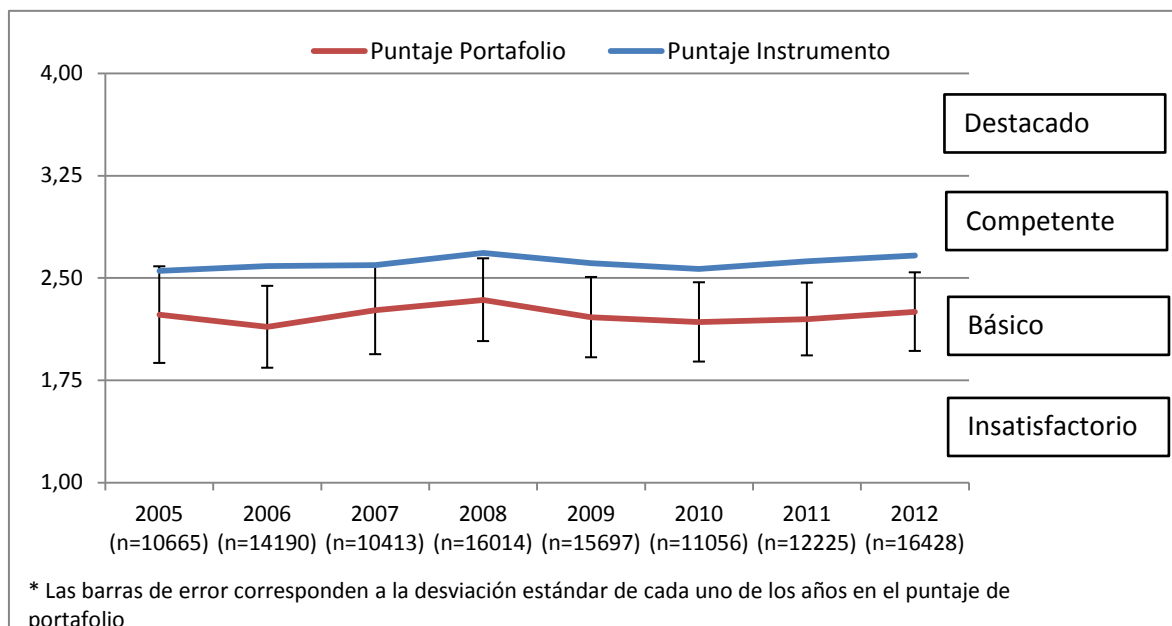
Una primera aproximación para cuantificar cambios en el desempeño de los docentes municipales, consiste en la comparación de los resultados por año (Gráfico 1). Dependiendo de la curva que se describa se puede determinar si cada una de las cohortes evaluadas han mantenido, empeorado o mejorado su resultado en relación a las otras.

---

<sup>1</sup> Esta descripción del portafolio obedece a la versión del instrumento para los períodos 2007 y 2011, que son los empleados en este estudio. Desde 2012, se redujo el número de dimensiones a 7 y el número de competencias evaluadas a 22.

<sup>2</sup> Este nivel de desempeño es posteriormente considerado por la Comisión Comunal de Evaluación, la cual, a la luz del reporte y otras consideraciones, puede ratificar o modificar el nivel de desempeño final. Hasta 2012, los docentes ratificados por la comisión comunal asciende a 94,2%, lo que implica que en la gran mayoría de los casos esta comisión ratifica el nivel de desempeño otorgado por la evaluación.

**Gráfico 1. Promedio de Puntaje de Portafolio y de Todos los Instrumentos según Periodo<sup>3</sup>**



En el Gráfico 1 se muestra el promedio de los puntajes de portafolio y el de los puntajes considerando todos los instrumentos, según el año de evaluación. Las curvas que describen estos puntajes se mueven relativamente en paralelo, y sin mostrar una evolución clara. Aunque hay una diferencia importante entre estos dos puntajes. Mientras el puntaje de instrumentos está en el rango de Competente, el de portafolio se encuentra en Básico. Esto revela que la evaluación del portafolio es más exigente que la del resto de los instrumentos. Incluso con la baja ponderación de los otros instrumentos hay un impacto sustantivo en los puntajes (ver nota al pie).

Esta comparación simple es una de las formas en las que se podría evaluar un progreso en las cohortes de los docentes y la conclusión sería que no existe una tendencia positiva hacia la mejora. No obstante, es osado aventurar alguna conclusión definitiva, dado que en estos resultados no existe necesariamente comparación de los mismos sujetos ya que cada

<sup>3</sup> El puntaje de todos los instrumentos considera el puntaje de portafolio (60%), el informe de referencia de terceros (10%), la entrevista del evaluador par (20%) y la autoevaluación (10%), considerando las ponderaciones establecidas en el reglamento de la evaluación docente.

docente debe evaluarse cada cuatro años si tiene los niveles Básico, Competente, Destacado y al año siguiente si tiene resultado Insatisfactorio<sup>4</sup>. Por lo tanto un primer problema que se presenta en la comparación son las diferencias que tienen las cohortes de evaluados. Este aspecto tiene mayor peso, porque en cada período existe la inclusión progresiva de docentes no evaluados anteriormente y de nuevos sectores y modalidades de enseñanza, como Educación Parvularia, Educación Especial y Educación de Adultos

Por lo tanto para reportar mejoras es necesario que se compare la evolución de los mismos sujetos en el tiempo. Los docentes que han sido evaluados en más de una oportunidad ascienden, el año 2012, a un 46.0% de un total de 78.717 docentes<sup>5</sup>. Para esta proporción de profesores ya podrían reportarse cambios, una vez se acuerden los criterios bajo los que esta comparación se establezca.

Un segundo problema consiste en que el Portafolio de la evaluación docente ha experimentado modificaciones en los indicadores que lo conforman año a año. Estos cambios se deben en primer lugar, a la necesidad de adecuar la rúbrica a los sectores nuevos que se van incorporando, de modo tal de asegurar una cobertura de las competencias y un nivel de dificultad similar para todos los evaluados. En segundo lugar, se han realizado modificaciones en algunos indicadores, para prevenir la posibilidad de que un mismo portafolio pueda ser reutilizado en dos instancias distintas y la copia de los mismos, para algunas de las preguntas que lo conforman. Por tanto, para reportar mejoras será necesario también emplear el conjunto de indicadores más estable en el Portafolio, entre las cohortes antes mencionadas.

Finalmente, un tercer problema es que existe un error de medición que debe ser controlado. Investigaciones realizadas por MIDE UC con un estudio de generalizabilidad, concluyó que el factor que más permitía aumentar la confiabilidad en la corrección de portafolios, es aumentar el número de veces que es corregida una evidencia. El sistema de evaluación docente actual, tal y como está planteado, solo contempla la doble corrección de un 20% de

---

<sup>4</sup> A partir del año 2011, los docentes que resulten básicos deberán evaluarse al año subsiguiente.

<sup>5</sup> Se excluye a los docentes que se han inscrito, pero que no se han evaluado por motivos ajenos a la evaluación.



la evidencia. Por tanto, se deberá encontrar una forma de aumentar al máximo la confiabilidad de los puntajes, para la cohorte y considerando el instrumento en cuestión.

## ***2.2 Confiabilidad en la ganancia de puntajes***

Al comparar dos evaluaciones de un conjunto de sujetos, además de considerar los puntajes inicial y final, se puede considerar como una variable la diferencia que existe entre ambas mediciones. De esta manera esta diferencia tiene un parámetro y un error asociado a ese parámetro.

Esta nueva variable que es la diferencia entre una evaluación y otra también es necesario que cumpla ciertos requisitos para estimar la confiabilidad de esta diferencia. En algunas investigaciones se han realizado avances en esta línea (Zimmerman & Williams, 1998) que apuntan específicamente al modo en que se estima la confiabilidad de esta diferencia de puntajes. Entre los errores frecuentes para este ejercicio está el de promediar las confiabilidades de la medición pre y la post. Este cálculo en realidad sobre estima la confiabilidad real de la ganancia. Se deben considerar más índices en este punto. Es importante tener la proporción de las varianzas y las confiabilidades de cada evaluación, además de la correlación que existe entre ambos periodos. Zimmerman & Williams (Op Cit.) plantean que mientras sean desiguales las varianzas y las confiabilidades de las medidas pre y post, la confiabilidad de las diferencias entre estas mediciones puede ser mayor que en un escenario donde estos indicadores son más parecidos.

Este indicador además da pie para determinar cuánto debe mejorar un docente para decir que su cambio es realmente un cambio positivo. En esa línea se puede determinar un intervalo de confianza (Crocker & Algina, 2008) que permita establecer quienes están dentro y fuera del cambio esperable y que razones hay para aquello.

En síntesis, se tienen que reportar los cambios de los docentes que participan en la evaluación docente y que caracteriza las posibles mejoras. Esto dentro de un marco, en que se está trabajando por presentar la evolución de los docentes. Este es un estudio preliminar, dado que se toman solo dos mediciones de los docentes. De hecho, el próximo año se podrán tomar 3 cohortes 2005-2009-2013, en la que podrá mostrarse con mayor evidencia temporal la evolución de los docentes.

Como consideración adicional, se debe tener en cuenta que una de las críticas que ha recibido la evaluación es que no ha sabido llegar a la formación de los profesores. No obstante, cada año se están implementando innovaciones que cumplan con este objetivo. Esta apuesta no va únicamente dirigida al docente sino también al equipo de liderazgo escolar en el establecimiento y a los municipios.

En el año 2013 la Organización para la Cooperación y el Desarrollo Económico (OCDE) realizó un diagnóstico y algunas sugerencias respecto al sistema de evaluación docente implementado en Chile. En este informe se establecen algunas recomendaciones que pudieran colaborar en el desarrollo del desempeño docente. Dichas recomendaciones se relacionan por un lado con ajustes a los instrumentos de evaluación y por otro con la difusión de los resultados de tal manera que se forme dialogo pedagógico formal en el establecimiento. La apuesta que renace con fuerza es que dentro del mismo establecimiento se creen o fomenten las comunidades pedagógicas y que éstas puedan potenciar las habilidades de los docentes, de una manera más directa y sistemática que se refleje finalmente en la evaluación tanto de docentes como de estudiantes.

### **3. Objetivos del estudio**

El objetivo general que guía este trabajo es explorar una metodología de comparación de puntajes de docentes y en base a ella determinar si existe una mejora en el desempeño docente al evaluarse por segunda vez. Los objetivos específicos para esta evidencia son:

- Seleccionar una muestra comparable, corregida por los correctores más consistentes.
- Seleccionar criterios de evaluación comparables entre el año 2007 y 2011.
- Evaluar la magnitud y confiabilidad de la diferencia entre ambos períodos.
- Validar investigando cuáles son las condiciones de los docentes que aumentan la probabilidad de una mejora efectiva.

## 4. Metodología

### 4.1 Obtención de la muestra e indicadores relevantes

Para levantar la información se utilizaron múltiples fuentes de datos. El universo al que se apunta son docentes que hayan sido evaluados en el periodo 2007 y en el período 2011 que hayan sido corregidos por el 80% de los correctores más consistentes en cada año. Por lo tanto, lo primero que se determinó fueron los docentes re-evaluados 2007-2011<sup>6</sup>. Una vez escogidos estos docentes, se realizó otro proceso aparte para determinar cuál es el 80% de los correctores con mayor consistencia en el periodo 2007 y en el periodo 2011.

La mayoría de las bases de datos obtenidas proviene de la Evaluación Docente. Las principales son las bases de resultados y caracterización de los docentes sobre el periodo de evaluación 2007 y 2011. De estas bases fueron extraídos los docentes que fueron evaluados en ambos periodos sus resultados y su caracterización. Un segundo conjunto de datos fue la información sobre los correctores, de las cuales se calculó el porcentaje de consistencia por corrector. Las dos bases que no pertenecen a la evaluación docente son las extraídas del sitio web de Grupo Educativo (Índice de vulnerabilidad) y de la agencia de calidad de la educación (Grupo Socio Económico)<sup>7</sup>.

De todas maneras, a pesar de haber elegido esta sub-muestra, como se mostrará en los resultados, se realizaron también contrastes con el grupo completo de re-evaluados, con el fin de tener un punto de comparación y poder ponderar el cambio en la elección de este subgrupo.

Cabe aclarar que el cambio de puntaje entre una evaluación y otra siempre se observó mediante el puntaje de portafolio. Este instrumento apunta directamente al desempeño que el docente tiene en su ejercicio profesional. En ese sentido, suena más sensato decir que una

---

<sup>6</sup> Dado que el nivel de desempeño está asociado al periodo de evaluación, existe una restricción insorteable respecto a la cantidad de casos con resultado final Insatisfactorio. Esto se corrige durante los análisis.

<sup>7</sup> [www.grupoeducativo.cl/download/IVE\\_2011.xls](http://www.grupoeducativo.cl/download/IVE_2011.xls) y [www.agenciaeducacion.cl/simce/bases-de-datos-nacionales](http://www.agenciaeducacion.cl/simce/bases-de-datos-nacionales) respectivamente.

evolución de su desempeño se debe a su trabajo en el aula, más que fuera de ella o por razones que se alejan de este espacio.

#### 4.1.1 Elección de los indicadores comparables

Como ya se mencionó, el portafolio es modificado año a año dentro de la evaluación. Las razones tienen que ver principalmente con la inclusión progresiva de nuevos sectores, por los cuales los criterios de evaluación deben ser re-adaptados, del tal manera que la evaluación siga siendo lo más equitativa posible.

En este escenario, existe una dificultad en comparar directamente los puntajes de portafolio obtenidos por los docentes en los distintos años de evaluación. Es por eso que el equipo de instrumentos del proyecto Docentemás ha registrado año a año cómo se van creando, reformulando, reubicando y eliminando ítems. Con esta información se determinó que de las 24 competencias que el Portafolio evalúa, aproximadamente un 50% de los indicadores es comparable para los períodos 2007 a 2011. (Tabla 1).

**Tabla 1. Indicadores Comparables 2007-2011<sup>8</sup>**

2007		2011	
Indicador	Nombre 2007	Indicador	Nombre 2011
A1	Formulación de Objetivos	A1	Formulación y coherencia entre los objetivos
B1	Relación entre actividades y objetivos	A2	Relación entre actividades y objetivos
A2	Secuencia de clases	A3	Calidad de la secuencia entre las clases de la unidad
D2	Variedad de estrategias didácticas	E1	Variedad de estrategias didácticas
C3	Calidad de las preguntas o instrucciones de la evaluación	C1	Calidad del instrumento de evaluación
C2	Adecuación de la evaluación a los objetivos de aprendizaje de la unidad	C2	Calidad de la pauta de corrección de la evaluación
		C3	Adecuación de la evaluación a los objetivos de aprendizaje de la unidad realizada
D1	Atribución de resultados de aprendizaje	D2	Atribución de los resultados de aprendizaje
E3	Detección y satisfacción de las propias necesidades de aprendizaje	E3	Detección y mejora de las propias necesidades de aprendizaje
F1	Focalización de los alumnos en las actividades propuestas	F1	Promoción de la focalización de los alumnos en las actividades
F3	Promoción de la participación de todos los alumnos	F2	Promoción de la participación de todos los alumnos
G3	Contribución de las actividades al logro de los objetivos y contenidos de la clase	G3	Contribución de las actividades al logro de los objetivos de la clase
H1	Calidad de las explicaciones	H1	Calidad de la explicación

Particularmente en este estudio, la comparabilidad de indicadores se utilizó para construir un nuevo puntaje de portafolio. Este puntaje corresponde al promedio simple de todos los indicadores en cada uno de los años. Al igual que con la selección de correctores se conserva el puntaje de portafolio oficial para realizar la comparación de ambos puntajes.

---

<sup>8</sup> En Anexo (8.1) se encuentran la lista de indicadores que no tienen comparabilidad entre estos dos años.

#### 4.1.2 Elección de los correctores con mayor consistencia

El grado de consistencia del corrector fue determinado en base a un proceso denominado "salto a triple corrección". Si en una evidencia que es corregida dos veces existe un alto grado de desacuerdo<sup>9</sup>, entonces esta evidencia es corregida una tercera vez por el Supervisor de la Corrección, quien funciona como un árbitro entre ambos correctores. La cantidad de este tipo de inconsistencias en el modulo en que el corrector era el más alejado del supervisor se considera como una inconsistencia. Luego, para cada corrector, se obtuvo un porcentaje de evidencia corregido consistentemente. Teniendo la distribución de esos porcentajes, se selecciona el 80% de los correctores<sup>10</sup>. Esto se realiza para el periodo 2011 y para el 2007.

Una vez determinado los correctores consistentes de cada periodo, en la base de docentes se seleccionaron los docentes que habían sido corregidos por estos correctores consistentes.

### **4.2 *Análisis de Datos***

Luego de la consolidación de las distintas bases de datos, se utilizaron técnicas de estadística descriptiva para varios propósitos. El primero para ver las distribuciones de las variables de caracterización que fueran más relevantes, por ejemplo variables socio-demográficas como género, edad o región, o también variables pedagógicas como sector y nivel de enseñanza. También se utilizaron descriptivos para calcular los promedios en cada uno de los años y de los grupos de re-evaluados, ver las variaciones que estos tenían y como se asociaban las mediciones de ambos años.

Una vez teniendo estos datos el siguiente paso realizado fue el cálculo de la confiabilidad en la ganancia de puntaje. Para esto se requería de las desviaciones estándar y las

---

<sup>9</sup> El desacuerdo está determinado por la suma de las distancias en cada uno de los indicadores de la dimensión

<sup>10</sup> En promedio, los correctores elegidos como consistentes, doble-corrigieron un 87.3% ( $\pm 5\%$ ) de la evidencia consistentemente, es decir no saltaron a triple corrección.

confiabilidades de los puntajes de ambos periodos de evaluación<sup>11</sup>, además del índice de correlación Pearson.

Con este indicador de la confiabilidad en la ganancia de puntaje se establece un intervalo de confianza del 95% (bilateral) para determinar los docentes que han tenido un incremento sobresaliente de puntaje entre la primera y segunda evaluación.

Por último se utilizó regresión logística para analizar que variables observadas explican el cambio significativo de puntaje. Particularmente la presentación de los odds ratio.

## 5. Resultados

### 5.1 Caracterización de la muestra

Se muestra la cantidad total de docentes que participaron en el estudio señalando las muestras a las que pertenecen (Tabla 2):

**Tabla 2. Universo y muestra**

	<b>N</b>	<b>Porcentaje</b>
Universo (Dotación 2007)	10.413	100,0%
Docentes Re-evaluados en 2011	4.808	46,2%
M1 corregido por un corrector consistente en 2007 y en 2011	2.718	26,1%

Dentro de la muestra final (2.718) los docentes se agrupan en 1.641 establecimientos y el promedio de docentes evaluados por cada establecimiento asciende a 1.65 docentes por establecimiento (mínimo 1; moda 1; Máximo 8). Un 23.2% de los docentes pertenecientes a la muestra cambiaron su establecimiento en una segunda evaluación. Se seleccionaron las variables principales que describen los docentes que fueron seleccionados en la muestra final.

---

<sup>11</sup> Se utilizó como medida de confiabilidad, la confiabilidad inter corrector (Hallgren, 2012). Esta confiabilidad estima si existe variación en la corrección que sea explicada por el corrector, lo que es poco deseable. Mientras más alto sea este indicador, menor es la influencia del corrector en la corrección.

La caracterización socio-demográfica indica que la muestra en su mayoría está compuesta por mujeres, en edad se concentran entre los 40 y 59 años. Los establecimientos donde trabajan se ubican principalmente en zonas urbanas y pertenecen al grupo socio-económico Medio bajo o Bajo (Tabla 3).

**Tabla 3. Caracterización socio demográfica de la muestra**

<b>Variable</b>	<b>Condición</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Género	Masculino	825	30.4%
	Femenino	1893	69.7%
Edad	Perdidos	11	0.4%
	29 o menos	164	6.0%
	30 a 34	188	6.9%
	35 a 39	267	9.8%
	40 a 44	425	15.6%
	45 a 49	546	20.1%
	50 a 54	692	25.5%
	55 a 59	400	14.7%
	60 o más	25	0.9%
	Zona Geográfica del Establecimiento	Perdidos	57
Rural		707	26.0%
Urbano		1954	71.9%
Grupo Socio Económico del Establecimiento	Perdidos	57	2.1%
	Bajo	643	23.7%
	Medio Bajo	1447	53.2%
	Medio	502	18.5%
	Medio Alto	69	2.5%

A continuación se registran las variables relativas a la profesión (Tabla 4). En la tabla se observa que en el nivel de enseñanza la proporción de docentes es decreciente a medida que el nivel de enseñanza es mayor. Dado que los profesores generalistas corresponden a los docentes de primer ciclo, este porcentaje se mantiene en el subsector del docente. A esta cifra le sigue la de docentes de Matemática (17.7%) no así los de lenguaje con un porcentaje más bajo.

El nivel de desempeño tiene una distribución un tanto diferente a la que podría tener la muestra total, dado el curso de las evaluaciones dependiendo de este factor. La muestra seleccionada considera los docentes que se evalúan cada 4 años, por lo tanto excluye a la



gran mayoría de los insatisfactorios que deben evaluarse necesariamente al año siguiente<sup>12</sup>. Este desajuste en la proporción de nivel final tiene un impacto en la distribución de los puntajes de portafolio. Más adelante se relata la solución a este problema.

**Tabla 4. Caracterización profesional de la muestra**

Variable	Condición	Frecuencia	Porcentaje
Nivel de Enseñanza	Primer Ciclo	1379	50.7%
	Segundo Ciclo	900	33.1%
	Educación Media	439	16.2%
Sector de Enseñanza	Biología	9	0.3%
	Ciencias Naturales	158	5.8%
	Ciencias Sociales	178	6.6%
	Educación Física	236	8.7%
	Generalista	1379	50.7%
	Historia	12	0.4%
	Lenguaje	238	8.8%
	Matemática	482	17.7%
	Química	26	1.0%
Nivel de Desempeño Final (2007)	Insatisfactorio	2	0.1%
	Básico	1013	37.3%
	Competente	1520	55.9%
	Destacado	183	6.7%
Participación en PSP	Sí	839	30.9%
	No	1879	69.1%
Participación en Cursos del Programa Perfeccionamiento Fundamental	Sí	1537	56.6%
	No	1181	43.4%

Particularmente en la línea del desarrollo profesional se incluyen también la participación en planes de superación profesional (PSP) y en cursos del Programa de Perfeccionamiento Fundamental (PPF). Un 30.9% de los docentes declara haber participado en actividades de perfeccionamiento luego de la evaluación 2007. Un 99.6% del porcentaje anterior

<sup>12</sup> Los 2 docentes insatisfactorios encontrados lograron, de distintas maneras, evaluarse cuatro años más tarde, porque suspendieron su evaluación hasta ese tiempo. Esta asociación entre el periodo de evaluación y el nivel de desempeño final también afecta a los básicos, a partir del año 2011, producto del cambio en la ley 20.501. Ahora los docentes con este nivel de desempeño se evaluarán cada 2 años.

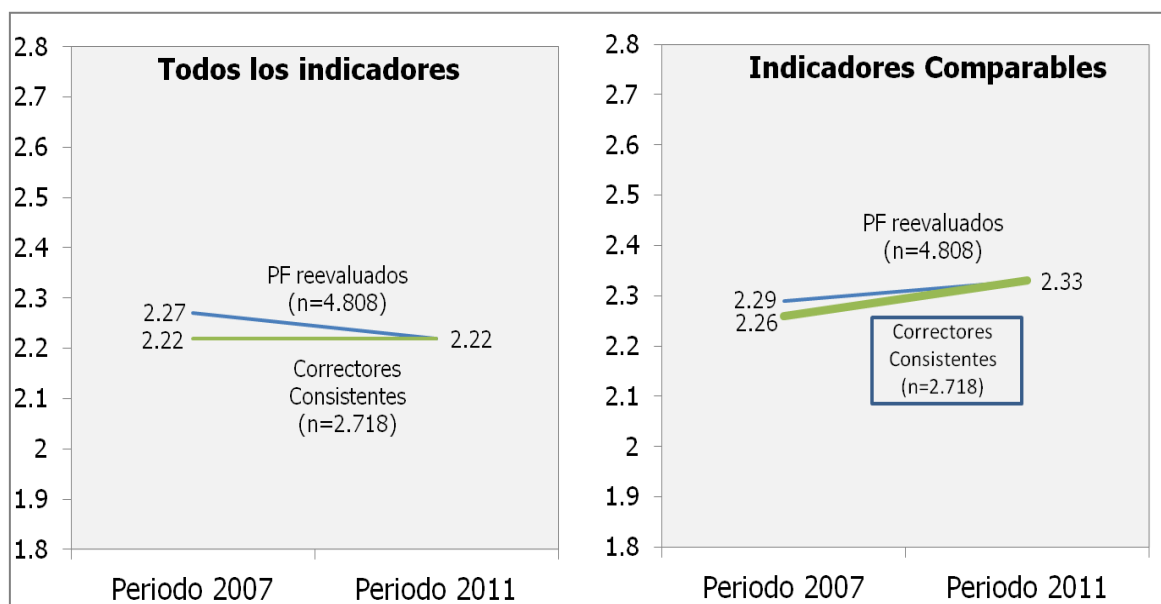
corresponden a docentes que resultaron básicos previamente. No obstante, dentro de este nivel de desempeño previo la participación en PSP asciende al 82.5%. En el caso de la participación en el Programa de Perfeccionamiento Fundamental, más de la mitad de los docentes de la muestra (56.6%) declara haber participado.

## ***5.2 Comparabilidad de los puntajes del portafolio***

Una vez obtenida la muestra requerida, se calcularon las medias obtenidas en el pre y en el post (

Gráfico 2). El gráfico muestra cuánto cambia la pendiente entre la evaluación 2007 y la evaluación 2011 considerando cuatro condiciones. La primera es seleccionando a todos los docentes que tienen evaluación en ambos períodos y considerando el puntaje de portafolio oficial, o sea con indicadores comparables y no comparables. La segunda es seleccionando solo a los docentes corregidos, en ambos periodos, por correctores consistentes y considerando los puntajes de portafolio oficial. La tercera condición (lado derecho del gráfico) selecciona a todos los re-evaluados y considera los indicadores comparables. Y por último la condición seleccionada considera ambos cambios, esto es docentes corregidos, en ambos periodos, por correctores consistentes y puntaje de portafolio solo con indicadores comparables.

**Gráfico 2. Medias para cada uno de los grupos contruidos.**



Al considerar indicadores comparables los puntajes de la evaluación 2011, independiente si los docentes fueron o no corregidos por correctores consistentes, tiene una leve mejora respecto a la evaluación 2007. Por otro lado, si se considera la comparación oficial de re-evaluados, es decir todos los docentes re-evaluados con el puntaje de portafolio oficial, entonces la inclinación de la pendiente tiene una leve tendencia negativa. Dicha tendencia se regula cuando se consideran los correctores consistentes.

Se realiza en esa línea un contraste t para muestras dependientes en cada una de las cuatro condiciones. Adicionalmente, dado la cantidad de docentes a comparar, se agrega el tamaño del efecto en cada una de las comparaciones (Anexo 8.2).

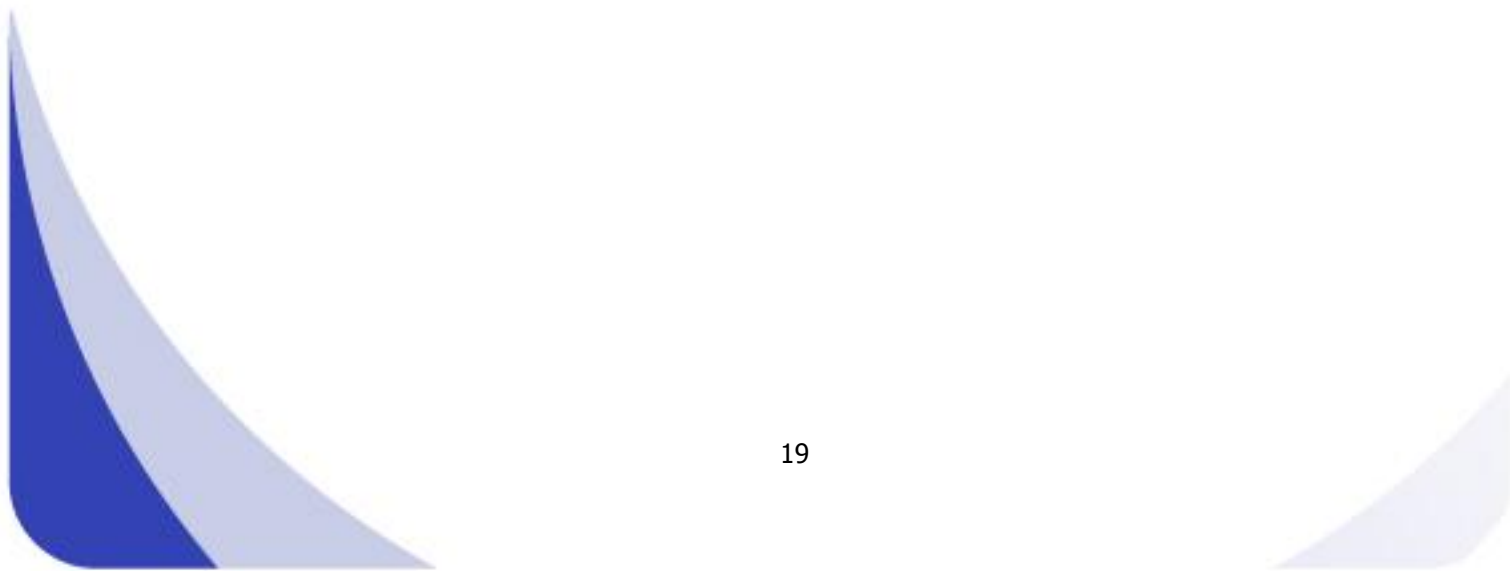
Llama la atención que al considerar todos los indicadores y solamente limpiar el efecto que el corrector pudiera tener sobre la corrección, la diferencia negativa, se anule y deje de ser significativa (diferencia pasa de  $-.05$  a  $.00$  y la significancia cambia de  $<.0001$  a  $.54$ , donde el alfa teórico es de  $0.01$ ). Por otro lado, la mayor diferencia se encuentra en la condición 4, con ambos tratamientos, de corrector consistente e indicadores comparables ( $t:9.84$ ;  $gl: 2.717$ ;  $p<.0001$ ;  $TE:.19$ ). El tamaño del efecto de dicha comparación es modesto para decir que en esta condición los profesores en el segundo periodo presentan mejores resultados que en el primero. No obstante, no se debe dejar de considerar que la tendencia es positiva.

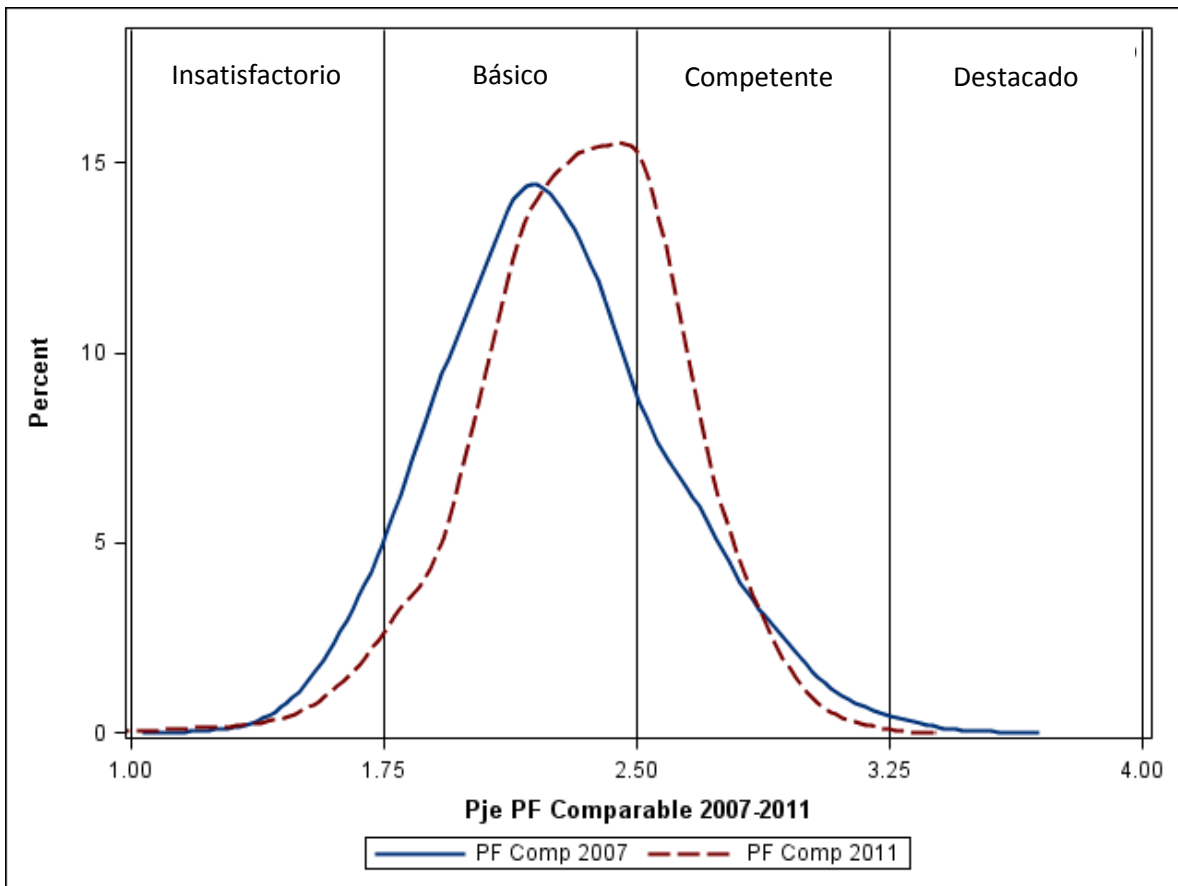
Completada la primera etapa de análisis se optó por pasar a la siguiente solamente con una de las condiciones a fin de simplificar los cálculos a realizar. La condición seleccionada fue la que incorporaba ambos criterios de filtro, el de correctores consistentes y el de indicadores comparables, dado que se observa un cambio en los puntajes de la evaluación al limpiar estos dos efectos. De todas maneras cabe aclarar que, entre estos dos ajustes, sigue siendo el de indicadores comparables el que pareciera tener mayor impacto. Con estos criterios los docentes en promedio, mejoraron entre 2007 y 2011, 0.08 puntos. La desviación estándar de la diferencia entre la primera y la segunda evaluación es de 0.41, lo que da cuenta de una alta variabilidad en los cambios de puntaje en portafolio entre la primera y la segunda evaluación.

Con el fin de presentar la distribución de los puntajes, y no solo sus promedios, el

Gráfico 3 muestra la densidad de los puntajes para ambos periodos 2007 y 2011. Consistente con la comparación de medias realizadas anteriormente se observa una leve inclinación hacia la derecha del periodo 2011.

**Gráfico 3. Distribuciones de la medida Pre – Post (Kernel)**





### ***5.3 Estimación de la confiabilidad de la ganancia de puntaje y categorización de los docentes que mejoran***

Cómo se presentó más arriba, la comparación de puntajes entre una evaluación y otra requiere de la estimación de una confiabilidad en sí mismo y no el mero promedio de las confiabilidades de cada periodo de medición (Zimmerman & Williams, 1998). Además es necesario considerar cuál es la razón de las variaciones entre una medición y otra, junto con el grado de asociación que éstas tienen. Este último indicador fue modificado por un método que se llama asociación corregida por restricción de rango (Tran, 2011). En los puntajes de portafolio iniciales hay un sesgo o restricción de rango, producto que la muestra de re-evaluados considera un periodo de cuatro años fijo entre cada evaluación, excluyendo inmediatamente a los insatisfactorios, quienes se evalúan al año siguiente. De este modo la

asociación entre las re-evaluaciones puede verse disminuida innecesariamente. Es por eso que se aplica dicha corrección.

En la Tabla 5 están los indicadores mencionados, para la condición seleccionada de correctores consistentes e indicadores comparables<sup>13</sup>. La confiabilidad en la ganancia de puntajes es de .61, considerando los demás resultados mostrados en la tabla. Esta cifra indica que el cambio de puntaje es confiable entre una medición y otra. En otras palabras es esperable que la mayoría de los docentes evaluados tenga una leve mejora en una segunda evaluación.

**Tabla 5. Confiabilidad en la Ganancia de Puntaje entre la medición Pre y Post**

Estadísticos	Pre	Post
Media	<b>2.26</b>	<b>2.33</b>
Sd	<b>0.35</b>	<b>0.30</b>
Confiabilidad (IRR)	<b>0.71</b>	<b>0.70</b>
Correlación pre-post (r de Pearson)	<b>0.26</b>	
<i>Confiabilidad en la ganancia</i>	<b>0.61</b>	

Para tener mayor claridad sobre la diferencia entre las dos evaluaciones, se presenta la densidad de las diferencias de los docentes evaluados en ambos periodos ((2)

Algo muy importante de considerar es que el parámetro o punto de referencia desde el cual se agrega y se resta el intervalo de confianza es el punto 0. Concretamente, esto significa determinar un intervalo de confianza al no cambio, que permita determinar desde que punto de corte se puede considerar una mejora (o un empeoramiento), a partir de la diferencia conseguida entre evaluaciones. En este caso, por ejemplo si un docentes obtiene medio punto más en su segunda evaluación, respecto dela primera, entonces se puede decir que este docente ha presentado una mejora sustantiva.

Una vez determinado el punto de corte de la diferencia positiva (.51), se encuentra que cerca del 14% de la muestra está sobre este punto de corte. Estos docentes se convierten por lo tanto en el grupo de interés a comparar.

---

<sup>13</sup> Los mismos indicadores se presentan para las otras condiciones en el Anexo 8.3.



Gráfico 4). Los docentes modificaron sus puntajes entre evaluación entre -1.8 y 1.8 puntos. Más de la mitad de los docentes de la muestra, mejora sus puntajes en una segunda evaluación, aunque esa diferencia para la mayoría no es una mejora sustancial.

Justamente, con el objetivo de determinar un punto de corte que permitiera definir los docentes que efectivamente habían mejorado, se establece un intervalo de confianza en base al error estándar de la medición (Crocker & Algina, 2008) que permita determinar dicho punto. El error estándar de medición a su vez está compuesto por la desviación estándar de la diferencia entre los puntajes 2007 y 2011 y la confiabilidad en la ganancia de esta diferencia. A continuación se presentan las fórmulas de cómo fue calculado el punto de corte:

$$(1) \sigma_{E_{DD'}} = \sigma_{DD'} \sqrt{1 - \rho_{DD'}} \Rightarrow \sigma_{E_{DD'}} = .41 \sqrt{1 - .61} = .26$$

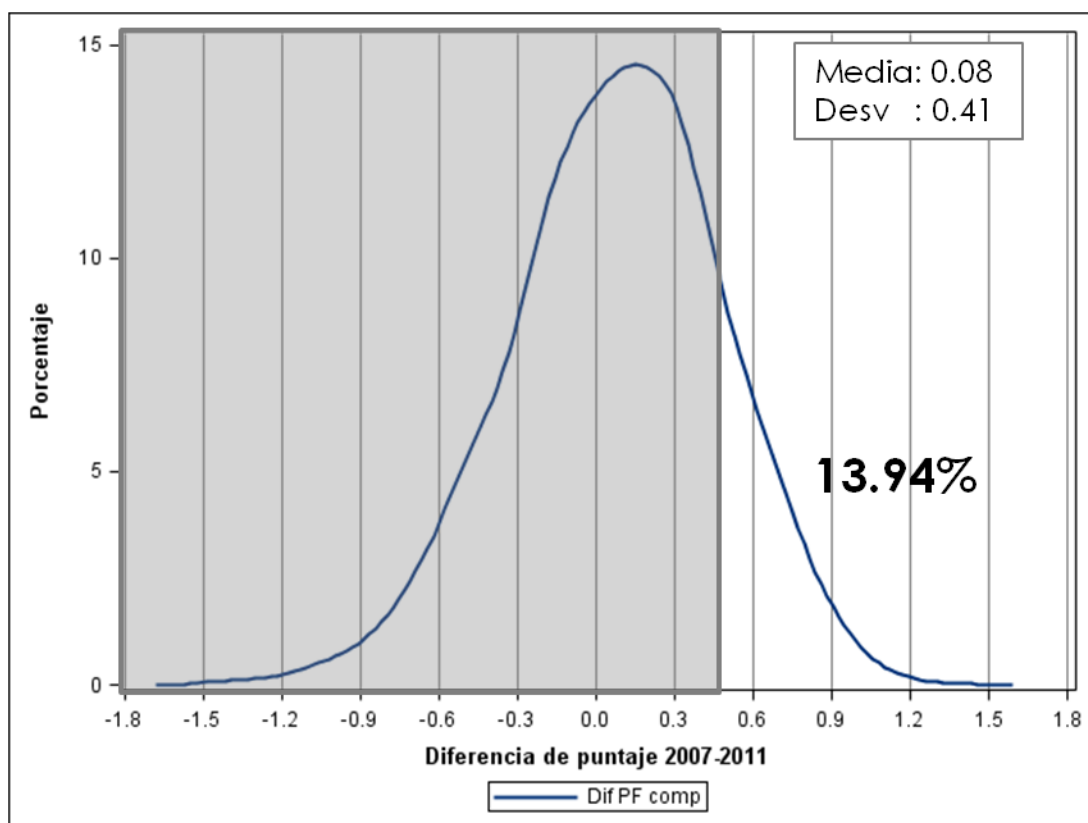
$$(2) {}_{95\%}IC_0 = 0 \pm 1.96\sigma_E \Rightarrow {}_{95\%}IC_0 = 0 \pm 1.96 * 0.26 = \pm 0.51$$

Algo muy importante de considerar es que el parámetro o punto de referencia desde el cual se agrega y se resta el intervalo de confianza es el punto 0. Concretamente, esto significa determinar un intervalo de confianza al no cambio, que permita determinar desde que punto de corte se puede considerar una mejora (o un empeoramiento), a partir de la diferencia conseguida entre evaluaciones. En este caso, por ejemplo si un docentes obtiene medio punto más en su segunda evaluación, respecto de la primera, entonces se puede decir que este docente ha presentado una mejora sustantiva.

Una vez determinado el punto de corte de la diferencia positiva (.51), se encuentra que cerca del 14% de la muestra está sobre este punto de corte. Estos docentes se convierten por lo tanto en el grupo de interés a comparar.



Gráfico 4. Densidad de las diferencias entre pre y post (Kernel)



#### ***5.4 Variables que pudieran explicar el cambio positivo del desempeño docente.***

Se espera que aquellos docentes que mejoraron tengan un perfil relativamente diferente de quienes no lo hicieron. Más aun, idealmente este perfil diferente debiese tener alguna relación con la realización de algún perfeccionamiento que explique este puntaje. Si bien el cambio de puntaje, tanto negativo como positivo, puede deberse a múltiples factores, lo deseable es que los factores seleccionados tengan alguna relación con una intervención directa en el desempeño del docente. De hecho, los pasos anteriores de selección de la muestra e indicadores comparables, limpiaron los ruidos que pudieran realizar las características del proceso de evaluación, no obstante quedan muchos otros aspectos que controlar.

Para ver que variables estaban asociadas a los docentes que obtuvieron una mejora efectiva, se realizó un análisis de regresión logística, con el método stepwise, considerando como predictores todas las variables del Cuestionario Complementario de la evaluación docente comparables en ambos períodos, más el Índice de Vulnerabilidad Escolar y la Zona Geográfica (urbano/rural) del Establecimiento, ambas variables según datos del SIMCE. Se tomó como variable dependiente la pertenencia del docente al grupo que había mejorado su puntaje en una cantidad mayor o igual al umbral considerado de mejora, es decir:

$$\ln\left(\frac{P_i}{1 - P_i}\right) = \alpha + b_1X_{1i} + b_2X_{2i} \dots + b_nX_{ni}$$

Donde P es la probabilidad del sujeto i, de pertenecer al grupo de docentes que mejora (mayor o igual que 0.51 puntos) entre una evaluación y otra en el Portafolio.  $\alpha$  es el intercepto y cada uno de los b es el coeficiente de regresión logística para la variable independiente X considerada en el modelo.

El modelo que más ajustó se presenta en la Tabla 6, mostrando solo las variables seleccionadas por el método stepwise.

**Tabla 6. Variables predictoras de la mejora efectiva**

Mejora vs. No mejora	b (SE)	95% IC para Razón de Verosimilitud (Odds Ratio)		
		Mínimo	Razón de Verosimilitud	Máximo
Intercepto	11.10 (0.93) ***			
Pje Portafolio 2007	-5.65 (0.38) ***	0.00	0.00	0.01
Edad	-0.03 (0.01) **	0.95	0.97	0.99
Participación en Cursos del Programa de Perfeccionamiento Fundamental 2007	0.49 (0.18)**	1.14	1.63	2.32
Zona Rural-Urbano (Ref:Rural)	-0.37 (0.10) ***	0.32	0.48	0.71

Nota:  $R^2=0.21$  (Cox & Snell),  $R^2=0.38$  (Nagelkerke), Model  $X^2(56)=51.634$ ,  $p=.641$ ,  
\*  $p<.05$ , \*\*  $p<.01$ , \*\*\*  $p<.001$

Se observa que el mayor predictor de logro del docente es el puntaje que obtuvo en la primera medición. Ambas variables están inversamente relacionadas. La probabilidad de ser un docente que mejora es mucho mayor para los docentes con menores rendimientos en la primera evaluación, que aquellos con mayores rendimientos. La edad y la zona donde se ubica el docente son también variables que influyen en la proporción de docentes que mejoran. Aunque el impacto no es tan fuerte como el del puntaje de portafolio 2007, si se observa una leve tendencia que los jóvenes tengan una mayor probabilidad de mejorar en su próxima evaluación. También se observa que los docentes que se desempeñan en establecimientos en zonas urbanas tienen mayor probabilidad de mejorar en su siguiente evaluación en comparación con docentes que se desempeñan en establecimientos en zonas rurales.

En cuanto al desarrollo profesional, existen diversos resultados en relación a la mejora entre evaluaciones. La única que en este análisis demostró tener un impacto fue la participación en PPF. Los docentes que declaran haber participado en cursos del Programa de Perfeccionamiento Fundamental, previo o durante el año 2007 tienen mayor probabilidad de mejora que quienes no participaron. En el caso la participación en PSP no resultó aumentar la probabilidad de mejora de los docentes que participaron en ella. Esta variable fue introducida en un primer ejercicio de manera dicotómica (asiste/no asiste) y luego de manera continua atribuyendo en ambos el resultado fue negativo. Luego se realizó un contraste filtrando solo los docentes que resultaron básicos en la primera evaluación, pero aun así la probabilidad de obtener una mejora en la segunda evaluación era similar. Por último, la participación en post grados no fue una variable importante en la mejora de los docentes en este periodo. Aunque pudiera ser que al replicar este análisis con otras cohortes, la realización de magister pueda considerarse como una variable que haga una diferencia.

En el Anexo 8.4 se muestran el resto de las variables ingresadas a la regresión.

## 6. Discusión y Conclusiones

Al considerar el puntaje de Portafolio elaborado con los indicadores comparables y restringiendo la muestra a una cohorte reevaluada al cabo de cuatro años, corregida por los correctores más consistentes, puede concluirse que el progreso de los docentes, como grupo es modesto aún.

Al incorporar factores tales como el error de medición y obtener un punto de corte que considere la confiabilidad del progreso en puntajes, podemos decir que 14 de cada 100 docentes reevaluados ha experimentado una mejora significativa en sus prácticas docentes. Además, en base a esas características métricas se estableció que una diferencia de 0.51 puntos entre una evaluación y otra es una mejora considerable.

En base a los resultados del modelo logístico se observó de forma exploratoria, que el mejor predictor de la mejora es el puntaje de portafolio en la evaluación previa. En menor medida resultan ser predictores la edad, la zona geográfica y la participación en cursos del PPF.

Este resultado indica que si bien la mejora es modesta, para los docentes que tienen menor desempeño esta mejoría se hace más probable. Ahora bien, la mejoría de los niveles más bajos no está determinado necesariamente por programas de formación, entre periodos de evaluación necesariamente.

Sólo la participación en cursos del Programa de Perfeccionamiento Fundamental (PPF) resulta aumentar la probabilidad de mejora. Este programa implementado entre los años 1999 y 2002, tenía la finalidad de "... contribuir al mejoramiento de la calidad y equidad de la educación a través de la implementación de la Reforma Curricular<sup>14</sup>." (Dirección de Presupuesto (DIPRES), 2002). La población objetivo de este programa fueron todos los docentes y directivos del sistema subvecionado. La cobertura alcanzó cerca de un 40% de

---

<sup>14</sup> Esta Reforma Curricular corresponde a la implementación de una reforma educacional iniciada en el gobierno de Eduardo Frei, en 1996 y que continuó en los siguientes gobiernos. Esta reforma abarcó varios aspectos, además del currículo, entre ellos: textos escolares, masificación de las TICs, la Jornada Escolar Completa (JEC) (Elaqua).

dicha población en la fase de instalación, pero disminuye a menos del 10% en la de finalización, siendo estas las únicas dos fases (op cit).

Las razones para señalar que este programa revela un efecto en docentes que se evalúan mucho después, son un tanto enigmáticas aun. Probablemente, esta diferencia esté mediada por características similares entre los profesores que participaron de estos programas, en comparación con quienes no lo hicieron y que no fueron observadas esta vez.

En relación a las variables sociodemográficas consideradas, la edad resultó un factor explicativo, con una relación negativa. A mayor edad, menor la probabilidad de mejorar en una siguiente evaluación. Este efecto fue persistente al evaluar todos los modelos, sin embargo la diferencia es pequeña.

Finalmente está zona geográfica, la diferencia es un poco mayor y, probablemente tiene que ver con el acceso de recursos que se pueden dificultar más en las zonas rurales. Si bien una gran ayuda pueden ser los recursos tecnológicos, también sería necesario pensar en que otros recursos se dificultan en este tipo de zonas. Sobre todo porque el uso del computador por parte de los docentes es una característica que se puede considerar prácticamente como característica universal.

Llama bastante la atención que ni la realización de un PSP ni su valoración sean variables explicativas de la mejora de los docentes, ni tampoco la valoración de los Informes de Evaluación Individual. En base a varios estudios se sabe que los docentes hacen poco uso de los resultados de la evaluación docente. Más aún las posibilidades de retroalimentación que reciben sobre su desempeño docente son, en la mayoría de los casos escasas y de carácter informal. Por otro lado, se sabe que el sistema ha estado vinculado a una evaluación sumativa, más que formativa. La mejora de los profesores en su conjunto es un objetivo en el que se debiera y se está interviniendo actualmente por Docentemás. En ese sentido, también ha colaborado enormemente el informe de la OECD (Santiago et al, 2013). Uno de los aspectos que se está trabajando por mejorar es la labor formativa de la evaluación docente. Se podrían ver un aumento en los puntajes en general, dado las recomendaciones del informe de la OECD. Específicamente, se podría ver cambios en la medida que se implementen las modificaciones definitivas al manual del portafolio y a la entrega de los resultados, además del mayor involucramiento del docente, el establecimiento y los

municipios con esta evaluación. En el fondo, como señaló Barbara Bruns (2013) la evaluación docente en Chile ha servido de ejemplo a nivel internacional, pero ahora es tiempo de implementar un conjunto de innovaciones que apunten a que esta permita y apoye el desarrollo profesional continuo de los docentes.

En cuanto a lo metodológico, se abren un conjunto de preguntas sobre el modo en que debieran realizarse los contrastes de progreso y como éstos debieran ser reportados. Los aspectos que quizás son más importantes de considerar se relacionan con la comparabilidad del instrumento entre los años, dado los pequeños cambios que tiene cada año, y a la confiabilidad de la corrección. Al considerar los puntajes con todos los indicadores la pregunta que se abre es sobre la atribución que se puede hacer sobre ese cambio, mejor dicho, cómo se puede diferenciar si el cambio del docente es producto de un cambio en su desempeño o es producto de cambios en el portafolio. Es una pregunta difícil de resolver, pero que se puede resolver en la medida que los cambios en los indicadores sigan haciendo referencia a un mismo constructo. Por otro lado, el aumentar la confiabilidad y validez de la corrección, ya sea, aumentando la proporción de doble correcciones o mejorando el entrenamiento de los correctores, favorecen el que estas comparaciones puedan tener una mayor validez.

En este estudio se resolvió este problema de la comparación mediante los indicadores comparables. Es posible que en el futuro se implementen nuevas formas de comparación. Como se menciona al principio, este trabajo no es sólo la presentación de un resultado particular, sino también un antecedente para establecer una metodología de comparabilidad. Probablemente, sea replicada en otras cohortes de docentes, para examinar si se mantienen estos resultados. O también abra la posibilidad a otro tipo de estudios que permitan responder, finalmente, la pregunta por el progreso, de la forma más limpia posible.

## 7. Referencias

- Bruns, B.(2013).La importancia de la evaluación docente. *Seminario: Evaluación Docente en Chile: Un Enfoque Internacional*. [Presentación Power Point].
- Crocker, L. y Algina, J. (2008). Chapter 6. Reliability and the classical true score model. *Introduction to classical and modern test theory*. Mason, Ohio, Cengage Learning.
- Cronbach LJ (1951). Coefficient alpha and the internal structure of tests. *Psychometrika* 16 (3): 297–334.
- Elacqua, G.(s. a.).Breve historia de las reformas educacionales en Chile (1813 - presente):Coberturas, condiciones, calidad y equidad. Recuperado desde: <http://bicentenario.camara.cl/seminario/pdf/mesa%204%20-%20Gregory%20Elacqua.pdf>
- Hallgren, K. (2012). Computing Inter-Rater Reliability for observational Data: An Overview and Tutorial. *Tutor Quant Methods Psychol.* 2012 ; 8(1): 23–34.
- Decreto No. 192 [Ministerio de Educación] Reglamento sobre Evaluación Docente. Diario Oficial de la República de Chile. 11 de Junio, 2005.
- Ministerio de Educación.(2002). Informe Final de Evaluación: Programa Perfeccionamiento Fundamental de Docentes (MINEDUC). Recuperado de: [http://www.dipres.gob.cl/574/articles-14857\\_doc\\_pdf.pdf](http://www.dipres.gob.cl/574/articles-14857_doc_pdf.pdf)
- Santiago, P, Benavides, F, Charlotte, D, Goe, L & Nusche D.(2013). Teacher Evaluaction in Chile. *Reviews of Evaluation and Assessment in Education*. Organización para la Cooperación y el Desarrollo Económico, OCDE.
- Tran, M. (2011). Robustness Of Two Formulas To Correct Pearson Correlation For Restriction Of Range. *Educational Policy Studies Dissertations*. Paper 84.
- Zimmerman, D. W. and Williams, R. H. (1998). Reliability of gain scores under realistic assumptions about properties of pre-test and post-test scores. *British Journal of Mathematical and Statistical Psychology*, 51: 343–351.

## 8. Anexos

### ***8.1 Indicadores no comparables entre el periodo 2007 y 2011***

<b>INDICADORES 2011</b>	
B1	Características de los alumnos consideradas en la unidad
B2	Análisis de las fortalezas de la unidad realizada
B3	Análisis de los aspectos por mejorar de la unidad realizada
D1	Calidad de la retroalimentación
E2	Adecuación de las acciones para motivar a los alumnos
F3	Supervisión de las actividades
G1	Calidad del inicio de la clase
G2	Calidad del cierre de la clase
H2	Calidad de las preguntas
H3	Calidad de la retroalimentación a los alumnos
H4	Indicador por subsector o área
<b>INDICADORES 2007</b>	
A3	Análisis de la unidad implementada
B2	Análisis de una clase
B3	Clima para el aprendizaje
C1	Adecuación de los ítemes a los aprendizajes declarados
D3	Análisis de la evaluación
E1	Desarrollo de las fortalezas de los alumnos
E2	Promoción de la formación de valores
F2	Manejo del grupo curso
G1	Calidad de la estructura de la clase
G2	Utilización del tiempo instruccional
H2	Calidad de la interacción que promueve el profesor
H3	Acompañamiento de las actividades en función del aprendizaje



## 8.2 Contrastes en la comparación de re-evaluados

Muestra	Indicadores Portafolio	Estadístico	2007	2011	Diferencia	t	sig	TE
Todos los re- evaluados (n=4.808)	Todos los indicadores	<b>Promedio</b>	<b>2.27</b>	<b>2.22</b>	<b>-0.05</b>	<b>-9.49</b>	<b>&lt;.0001</b>	<b>-0.14</b>
		DE	0.31	0.26	0.36			
	Indicadores comparables	<b>Promedio</b>	<b>2.29</b>	<b>2.33</b>	<b>0.03</b>	<b>5.43</b>	<b>&lt;.0001</b>	<b>0.07</b>
		DE	0.37	0.31	0.43			
Muestra Corr Consistentes (n=2.718)	Todos los indicadores	<b>Promedio</b>	<b>2.22</b>	<b>2.22</b>	<b>0.00</b>	<b>0.61</b>	<b>0.54</b>	<b>0.01</b>
		DE	0.29	0.25	0.33			
	Indicadores comparables	<b>Promedio</b>	<b>2.26</b>	<b>2.33</b>	<b>0.08</b>	<b>9.84</b>	<b>&lt;.0001</b>	<b>0.19</b>
		DE	0.35	0.30	0.41			

## 8.3 Confiabilidad, correlación e Índice en la ganancia de Puntaje

Muestra	Indicadores Portafolio	Confiabilidad		Correlación (Con Corrección Restricción de Rango)	Índice de confiabilidad ganancia de puntaje
		2007	2011		
Todos los re- evaluados	Todos los indicadores	0.76	0.77	0.21	0.70
	Indicadores comparables	0.71	0.74	0.24	0.64
Muestra Corr Consistentes	Todos los indicadores	0.76	0.77	0.28	0.68
	Indicadores comparables	0.71	0.70	0.26	0.61

## **8.4 Predictores utilizados para explicar la mejora significativa en los puntajes**

<b>Año</b>		<b>DF</b>	<b>Estimador</b>	<b>Error estándar</b>	<b>Valor t</b>	<b>Aprox Pr &gt;  t </b>
	Intercepto	1	8.03	2.30	3.49	0.00
	Pje PF Comp 2007	1	-5.72	0.42	-13.72	<.0001
	Sexo	1	0.41	0.23	1.75	0.08
	Edad	1	-0.03	0.01	-2.62	0.01
2007	Nivel-Educación Media	1	-0.07	0.34	-0.21	0.84
2007	Nivel-Primer Ciclo	1	0.04	0.23	0.18	0.86
2007	Nivel-Segundo Ciclo	0	0.00			
2007	¿Tiene título profesional?	1	1.74	0.88	1.99	0.05
2007	Ha Participado en PPF del SS evaluado	1	0.45	0.20	2.22	0.03
2007	Otro Perfeccionamiento de 60 horas relacionado con el subsector	1	-0.04	0.20	-0.19	0.85
2007	Diplomado	1	-0.41	0.43	-0.96	0.34
2007	Postítulo	1	0.40	0.24	1.70	0.09
2007	Magíster	1	0.66	0.54	1.21	0.23
2007	Doctorado	0	0.00			
2007	Aproximadamente, ¿con qué frecuencia utiliza el computador para su trabajo docente?	1	-0.02	0.12	-0.16	0.87
2007	¿En qué tipo de establecimiento(s) trabaja?	1	-0.02	0.38	-0.06	0.96
2007	¿Qué nivel de escolaridad cree usted que alcanzará la mayoría de sus alumnos?	1	-0.01	0.11	-0.09	0.92
2007	Expectativas laborales próximos dos años	1	0.48	0.25	1.93	0.05
2007	Trabajar como profesor es muy cercano a mi trabajo ideal	1	0.19	0.16	1.15	0.25
2007	Ser profesor me hace sentir muy orgulloso	1	-0.46	0.24	-1.91	0.06
2007	Estoy muy satisfecho de ser profesor	1	0.37	0.26	1.42	0.16
2007	Me he realizado profesionalmente siendo profesor	1	-0.72	0.32	-2.24	0.02
2007	Si tuviese que decidir una carrera nuevamente, volvería a elegir pedagogía	1	0.07	0.15	0.46	0.64
2007	¿Cómo cree que será la situación de la profesión docente en 5 años más en nuestro país	1	0.14	0.13	1.07	0.29
2007	¿Cómo cree que será su situación personal como docente en 5 años más?	1	-0.07	0.14	-0.54	0.59
2007	[Aspectos que interfieren o dificultan la elaboración del Portafolio] - Insuficiente información sobre el proceso	1	0.15	0.27	0.53	0.60
2007	[Aspectos que interfieren o dificultan la elaboración del Portafolio] - Falta de tiempo	1	0.07	0.25	0.27	0.79
2007	Los productos solicitados en el Portafolio son semejantes a las labores que realizo cotidianamente como docente	1	-0.12	0.09	-1.27	0.21
2007	El Portafolio es un buen instrumento para evidenciar mi desempeño profesional como docente	1	0.02	0.08	0.27	0.79
2007	¿Ha visitado alguna vez el sitio web <a href="http://www.docentemas.cl">www.docentemas.cl</a> ?	1	0.43	0.29	1.47	0.14
2007	¿Conoce el Centro de Llamados de la Evaluación Docente (fono 600-222-0011)?	1	0.43	0.22	1.93	0.05
2007	¿Considera que dispuso de suficiente información para llevar a cabo su evaluación?	1	0.17	0.23	0.76	0.45
2011	¿Tiene título profesional?	1	-0.18	0.36	-0.49	0.62
2011	Doctorado	1	0.17	36.70	0.00	1.00
2011	Magíster	1	0.70	0.40	1.74	0.08
2011	Postítulo	1	-0.19	0.23	-0.80	0.43
2011	Diplomado	1	-0.26	0.32	-0.83	0.40
2011	Otro Perfeccionamiento de 60 horas relacionado con el subsector	1	0.43	0.20	2.18	0.03

Año		DF	Estimador	Error estándar	Valor t	Aprox Pr >  t
2011	[¿En qué tipo de establecimiento(s) trabaja?] - Particular Subvencionado	1	-0.29	0.56	-0.51	0.61
2011	[¿En qué tipo de establecimiento(s) trabaja?] - Particular pagado	1	-1.59	3.40	-0.47	0.64
2011	Aproximadamente, ¿con qué frecuencia utiliza el computador para su trabajo docente?	1	-0.03	0.17	-0.20	0.84
2011	Cuando he enfrentado una dificultad en mi trabajo, he recibido el apoyo del Director para resolverla	1	0.04	0.12	0.30	0.77
2011	El equipo directivo utiliza métodos efectivos para informarse sobre el desempeño de los profesores	1	0.15	0.15	1.03	0.30
2011	Los objetivos que busca alcanzar el establecimiento son conocidos por todos los profesores	1	0.02	0.14	0.13	0.90
2011	En este establecimiento, las personas tenemos claras nuestras responsabilidades	1	0.07	0.14	0.51	0.61
2011	La cantidad de tareas que debo realizar en esta escuela superan el tiempo que tengo disponible	1	0.00	0.08	-0.01	0.99
2011	Los profesores de esta escuela son respetados por la comunidad educativa	1	0.02	0.13	0.16	0.87
2011	Profesores y alumnos se llevan bien en esta escuela	1	-0.07	0.15	-0.46	0.65
2011	El director me retroalimenta respecto a mi desempeño	1	-0.19	0.13	-1.49	0.14
2011	En mi establecimiento se reflexiona acerca de los resultados obtenidos en la Evaluación Docente	1	0.09	0.09	1.00	0.32
2011	En el establecimiento donde trabajo, los profesores trabajan colaborativamente (se planifica, desarrolla actividades y comparte recursos, entre otras actividades, junto con otros docentes)	1	-0.08	0.10	-0.80	0.42
2011	[Durante el presente año, ¿cuántas veces ha participado en las siguientes instancias?] - Ha asistido a entrevistas con el equipo directivo, con el fin de recibir retroalimentación sobre su desempeño	1	-0.03	0.03	-1.03	0.30
2011	[Durante el presente año, ¿cuántas veces ha participado en las siguientes instancias?] - El equipo directivo ha observado sus clases	1	-0.01	0.03	-0.41	0.68
2011	[Durante el presente año, ¿cuántas veces ha participado en las siguientes instancias?] - El equipo directivo ha revisado sus planificaciones	1	0.02	0.02	1.55	0.12
2011	[Aspectos que interfieren o dificultan la elaboracion del Portafolio] - Insuficiente información sobre el proceso	1	0.72	0.40	1.81	0.07
2011	[Aspectos que interfieren o dificultan la elaboracion del Portafolio] - Falta de tiempo	1	-0.21	0.22	-0.93	0.35
2011	Los productos solicitados en el Portafolio son semejantes a las labores que realizo cotidianamente como docente	1	0.02	0.11	0.19	0.85
2011	El Portafolio es un buen instrumento para evidenciar mi desempeño profesional como docente	1	-0.04	0.08	-0.54	0.59
2011	¿Conoce el Centro de Llamados de la Evaluación Docente (fono 600-222-0011)?	1	-0.04	0.10	-0.37	0.72
2011	¿Considera que dispuso de suficiente información para llevar a cabo su evaluación?	1	-0.63	0.49	-1.28	0.20
2008-2011	Porcentaje de asistencia PSP	1	0.05	0.24	0.20	0.85
2011	Grupo Socio Economico del Establecimiento	1	0.12	0.14	0.83	0.41
2011	Zona Establecimiento-Rural	1	-0.71	0.24	-3.00	0.00
2011	Zona Establecimiento-Urbano	0	0.00	.	.	.