



**UNIVERSIDAD
DEL PACÍFICO**

Economía

Facultad de Economía y Finanzas

**LOS EFECTOS DE LA CAPACITACIÓN DOCENTE SOBRE LAS
PRÁCTICAS PEDAGÓGICAS Y LOS APRENDIZAJES DE LOS
ESTUDIANTES**

**Trabajo de Suficiencia Profesional
presentado para optar al Título Profesional de
Licenciado en Economía**

**Presentado por
Marcelo Alejandro Rondoy López**

Lima, marzo 2024



REPORTE DE EVALUACIÓN DEL SISTEMA ANTIPLAGIO
FACULTAD DE ECONOMÍA Y FINANZAS

A través del presente, la Facultad de Economía y Finanzas deja constancia de que el Trabajo de Suficiencia Profesional titulado “Los efectos de la capacitación docentes sobre las prácticas pedagógicas y los aprendizajes de los estudiantes” presentado por MARCELO ALEJANDRO RONDOY LOPEZ, identificado con DNI N° 70392588, para optar al Título Profesional de Licenciado en Economía, fue sometido al análisis del sistema antiplagio Turnitin el 29 de febrero de 2024. El siguiente fue el resultado obtenido:



De acuerdo con la política vigente, el porcentaje obtenido de similitud con otras fuentes se encuentra dentro de los márgenes permitidos.

Se emite el presente documento para los fines estipulados en el Reglamento de Grados y Títulos de la Facultad de Economía y Finanzas.

Lima, 12 de abril de 2024.


Juan Francisco Castro
Decano
Facultad de Economía y Finanzas

RESUMEN

Este trabajo investiga el impacto de la formación docente en las prácticas pedagógicas y los aprendizajes de los estudiantes. A través de una revisión exhaustiva de la literatura y evidencia empírica, identifica características clave de programas de formación efectivos. Los hallazgos sugieren que la formación docente en servicio (*in-service*) puede mejorar significativamente el desarrollo cognitivo y socioemocional de los estudiantes, solo si se mejoran las habilidades pedagógicas en el proceso. Además, hay poca evidencia de que la formación previa al servicio (*pre-service*) mejore el rendimiento docente, sino que estos aprenden a través de mecanismos informales, como el aprendizaje por experiencia o entre colegas. La literatura mide el aprendizaje cognitivo con las calificaciones de exámenes, mientras que las habilidades socioemocionales son poco abordadas. Las prácticas pedagógicas, por otro lado, se miden con observaciones del aula y métricas de percepción de los estudiantes. Finalmente, el enfoque econométrico preferido por los investigadores son los tratamientos controlados aleatorios (RCT), en tanto consideran la asignación selectiva de docentes o estudiantes. Los diseños cuasi experimentales controlan las características observadas y no observadas, y reconocen la presencia restante de heterogeneidad (QEX-I) o asumen que los controles utilizados son suficientes para eliminar cualquier posible sesgo (QEX-II).

ABSTRACT

This study investigates the impact of teacher training on pedagogical practices and student learning outcomes. Through a comprehensive review of literature and empirical analysis, it identifies key characteristics of effective training programs. The findings suggest that in-service teacher training can significantly improve students' cognitive and socio-emotional development, only if they enhance teachers' skills in the process. Also, there is little evidence that pre-service training improves teacher performance, although teachers learn through informal mechanisms, such as learning-by-doing or learning from peers. Researchers often measure students' cognitive learning with test scores, whilst socio-emotional skills are hardly taken into account. Pedagogical practices, on the other hand, are measured with classroom observations and students' perception metrics. Finally, the preferred econometric approach by researchers is randomized controlled trials (RCT) since they account for selective teacher or student assignment. The quasi-experimental designs covered in this paper control for observed and unobserved characteristics, and either recognize the remaining presence of heterogeneity (QEX-I) or assume the controls used are enough to account for any potential bias (QEX-II).

TABLA DE CONTENIDO

RESUMEN.....	ii
ABSTRACT.....	iii
ÍNDICE DE ANEXOS.....	v
INTRODUCCIÓN.....	1
CAPÍTULO I. MARCO TEÓRICO.....	3
1.1 Teorías del aprendizaje.....	3
1.2 Capacitación docente y prácticas pedagógicas.....	5
CAPÍTULO II. EVIDENCIA EMPÍRICA.....	7
2.1 Aprendizaje del estudiante.....	7
2.1.1 Calificaciones.....	7
2.1.2 Aprendizaje no cognitivo.....	8
2.2 Prácticas pedagógicas.....	9
2.2.1 Observaciones del aula.....	9
2.2.2 Valoraciones del estudiante.....	10
2.3 Tipos de capacitación docente.....	10
2.3.1 Capacitación formal <i>pre-service</i>	11
2.3.2 Capacitación formal <i>in-service</i>	12
2.3.2.1 Desarrollo profesional.....	12
2.3.2.2 Estudios de posgrado.....	15
2.3.2.3 Coaching.....	15
2.3.3 Capacitación informal.....	16
2.3.3.1 Aprendizaje por experiencia (<i>learning-by-doing</i>).....	16
2.3.3.2 Aprendizaje por colegas (<i>learning from peers</i>).....	17
2.4 Estrategias econométricas.....	18
CONCLUSIONES.....	20
REFERENCIAS BIBLIOGRÁFICAS.....	21
ANEXOS.....	26

ÍNDICE DE ANEXOS

Anexo 1.	Artículos académicos sobre programas de capacitación docente.....	27
Anexo 2.	Instrumentos de observación utilizados en el proyecto MET.....	28
Anexo 3.	Representación visual del marco teórico.....	30
Anexo 4.	Cuestionario del Proyecto Trípode para mejorar la enseñanza	31
Anexo 5.	Resultados de estudios revisados sobre los efectos de la capacitación docente en las prácticas pedagógicas y/o aprendizajes	32

INTRODUCCIÓN

La calidad de la enseñanza es un factor determinante en el éxito de los estudiantes. A pesar de la abundancia de estudios que documentan la importancia de la calidad del docente, existe un debate sobre los factores que mejoran o incluso predicen esta calidad. La capacitación docente emerge, en consecuencia, como un componente crítico, en tanto potenciar las estrategias pedagógicas mejora los resultados de aprendizaje de los alumnos (Darling-Hammond, 2017; Kraft et. al, 2018; Taylor, 2023). El anexo 1 muestra los principales campos de investigación en los cuales la literatura académica se concentra al momento de analizar la relación entre la capacitación y el desempeño docente. El presente trabajo explora los efectos de la capacitación docente sobre los aprendizajes de los estudiantes, y se basa, de manera similar, en una revisión de los principales estudios que abordan esta relación.

La literatura distingue distintos tipos de entrenamiento docente, desde la educación universitaria previa al servicio, o *pre-service*; hasta el desarrollo profesional en servicio, o *in-service*. Estos programas son considerados capacitaciones formales en tanto sean estructurados o provengan de un diseño previo. Existe, sin embargo, otras oportunidades que tienen los profesores para aprender nuevas habilidades. La formación informal incluye el aprendizaje a través de la experiencia, y de otros profesores. Ambos mejoran consistentemente el desempeño docente. La revisión empírica sobre los programas formales, en cambio, muestra resultados mixtos. Aunque la evidencia causal es limitada, estos programas son esenciales, pero no suficientes por sí solos para garantizar una mejora en el desempeño docente.

Los programas exitosos presentan características similares (Darling-Hammond, 2017), las cuales no se limitan a mejorar el conocimiento teórico del docente. Habilidades como el buen manejo del aula de clases, técnicas de enseñanza basada en cuestionamiento al alumno, y la retroalimentación personalizada hacia el docente son algunas de las características presentes en las capacitaciones exitosas. Por lo tanto, la hipótesis del trabajo plantea que la capacitación docente es efectiva y conduce a mejoras significativas en el aprendizaje de los estudiantes, únicamente cuando los contenidos de dicha capacitación están destinados, entre otras cosas, a mejorar las prácticas pedagógicas del docente. Este enfoque práctico promueve una aplicación directa de las estrategias de enseñanza en el aula, lo que se traduce en una mejora tangible en la calidad de la educación.

El trabajo se organiza de la siguiente manera. El capítulo I compone el marco teórico de la investigación. Resume las principales teorías educativas que explican el proceso de aprendizaje de los individuos, al igual que las prácticas pedagógicas usualmente aceptadas por la literatura. El capítulo II muestra la evidencia empírica que explora la relación entre el docente y los aprendizajes. Primero, se detallan las principales variables utilizadas para medir el desempeño estudiantil: las calificaciones de exámenes y métricas de aprendizaje no cognitivo (conducta, esfuerzo, motivación, etc.). Igualmente, se recopilan las formas de medir las prácticas pedagógicas: observaciones del aula y valoraciones del estudiante. Después, se realiza una breve disertación de los hallazgos sobre el efecto de cada tipo de programa de capacitación (*pre-service*, *in-service*, capacitación informal) en las variables de interés (prácticas pedagógicas y aprendizajes). Asimismo, se discuten las distintas técnicas econométricas empleadas en los trabajos revisados, y se mencionan brevemente los problemas de estimación más comunes: sesgo de selección y deserción (o *attrition*). Finalmente, se presentan las conclusiones del trabajo, las cuales recogen todas las teorías que han sido aprobadas y deja constancia de los problemas pendientes.

CAPÍTULO I. MARCO TEÓRICO

Esta sección compone el marco teórico del trabajo, el cual pretende resaltar la influencia de la formación docente en el rendimiento y aprendizaje de los estudiantes. Primero se sintetizan las principales teorías educativas que explican el proceso de aprendizaje de los individuos. Después, se explica cómo la formación docente impacta en las prácticas pedagógicas y, en consecuencia, en el aprendizaje de los estudiantes.

1.1 Teorías del aprendizaje

El fundamento teórico del aprendizaje se basa en la premisa de que la calidad de la educación depende significativamente de la calidad de la enseñanza, la cual está intrínsecamente relacionada con la capacitación y preparación de los docentes. La formación docente es vista no solo como un proceso inicial que habilita a una persona para ejercer la profesión, sino también como un proceso continuo que permite a los educadores actualizarse en métodos de enseñanza, tecnologías educativas, enfoques pedagógicos innovadores, entre otros. Existen varias teorías que proporcionan una base para entender cómo la capacitación docente puede influir en las prácticas pedagógicas (aprender a enseñar) y, a su vez, en los aprendizajes.

La teoría del conductismo (*behaviourism*, en inglés) sostiene que el comportamiento de los individuos es influenciado por estímulos externos observables (Skinner, 1953). En el campo pedagógico, los profesores pueden emplear incentivos (reforzamientos), como recompensas o castigos para moldear el comportamiento de sus estudiantes hacia la consecución de objetivos específicos. El conductismo también enfatiza el rol del entorno, lo que sugiere que la creación de un entorno de aprendizaje estructurado y de apoyo constantes estaría relacionado a un proceso de aprendizaje más adecuado. Las aplicaciones del conductismo están enfocadas principalmente en la mejora de prácticas pedagógicas, como el manejo y control del salón de clases (Vance Hall et. al, 1968), la participación de los alumnos (Kane y Staiger, 2012), etc.

El cognitivismo está abordado en Piaget (1952), y se enfoca en los procesos mentales internos que ocurren durante el aprendizaje. Piaget propone que los infantes experimentan cuatro etapas de desarrollo cognitivo: la etapa sensoriomotriz (hasta los 2 años), donde aprenden a través de sus sentidos y acciones; la etapa preoperacional (de 2 a 7 años), donde comienzan a utilizar símbolos y lenguaje, aunque aún carecen de razonamiento lógico; la etapa operativa (7 a 11 años), a partir de la cual piensan lógicamente sobre eventos concretos; y la etapa operacional formal (a partir de los 12 años), donde empiezan a pensar de manera abstracta y razonar sobre problemas hipotéticos. En este marco, las estrategias de enseñanza deberían variar según la

etapa de aprendizaje en la que el alumno se encuentra. En la práctica, la literatura muestra que los programas de capacitación tradicionales fallan en generar efectos significativos en el aprendizaje, principalmente porque estos suelen incluir contenidos genéricos y estandarizados (Darling-Hammond et. al, 2017; Taylor, 2023). Es más, la mayoría de los estudios muestra efectos diferenciados según el grado de instrucción del alumnado: inicial, primaria, secundaria, universitaria (Harris y Sass, 2011; Coenen et. al, 2018). Esto sugiere que cada estudiante, según su edad y, por lo tanto, según su etapa de aprendizaje, requiere técnicas de enseñanza diferentes, tal como la teoría del cognitivismo plantea.

El enfoque constructivista, en cambio, sostiene que las personas construyen activamente su conocimiento a través de experiencias y reflexiones (Vygotsky y Cole, 1978). Esta teoría difiere del enfoque de Piaget, al postular que el desarrollo de las funciones psicológicas superiores de un niño está fundamentalmente condicionado por las interacciones sociales y el contexto cultural dentro del cual ocurren estas interacciones. Las prácticas educativas contemporáneas que siguen este enfoque abogan por estrategias de instrucción que prioricen la colaboración, el diálogo y una pedagogía culturalmente receptiva.

El conectivismo surge como una teoría para la era digital, enfatizando la importancia de las redes digitales y otras herramientas tecnológicas en el proceso de aprendizaje (Siemens, 2005). Se enfatiza la importancia de las interconexiones que existen entre distintas fuentes de información, individuos y organizaciones. Siemens argumenta que los enfoques tradicionales del conductismo, cognitivismo y constructivismo no son adecuados para explicar por completo el proceso de aprendizaje que ocurre en un entorno digital de constante cambio. Bajo el enfoque conectivista, el conocimiento está esparcido a lo largo de distintas “redes de conexiones” y el aprendizaje consiste en la habilidad de navegar a través de estas redes para obtener información (Siemens, 2005). Naturalmente, el conectivismo es un aporte significativo para el campo de la educación a distancia (e-learning). Es más, la mayoría de los programas de capacitación exitosos incluyen al menos un tipo de material audiovisual, u otra herramienta tecnológica que permita al docente permanecer actualizado con la información más reciente para mejorar su productividad (Darling-Hammond et. al, 2017).

Finalmente, el enfoque humanista está profundamente arraigado a las teorías del desarrollo personal y psicológico planteadas por Rogers (1975). Este enfoque sostiene que para que se produzca un aprendizaje significativo, la educación debe ser personalizada, es decir, estar centrada en los estudiantes y no en el plan de estudios o en el docente. Se enfatiza también la

importancia de crear un entorno de aprendizaje democrático que permita a los estudiantes explorar, descubrir y aprender de acuerdo con sus intereses y a su propio ritmo. En consecuencia, el papel del docente es más el de un facilitador del aprendizaje, y no una autoridad del conocimiento (Rogers, 1975). Este enfoque holístico ha incidido en el desarrollo de métodos de enseñanza que priorizan el bienestar emocional y psicológico de los estudiantes, en lugar de únicamente el desarrollo intelectual. Los programas de coaching (Kraft et. al, 2018) y mentoría (Boyd et. al, 2009) son de particular interés, en tanto hacen un seguimiento personalizado hacia el estudiante.

En la práctica, el resultado de los programas de capacitación docente puede ser explicado al considerar aspectos de una o más de las teorías antes mencionadas. Por ejemplo, la evidencia empírica señala efectos diferenciados entre docentes de inicial, primaria, y secundaria (Taylor, 2023), lo cual indicaría la necesidad de diseñar capacitaciones específicas según la etapa cognitivo de los estudiantes de cada docente. También, pretender mejorar las prácticas pedagógicas, concepto muchas veces relacionado al desarrollo de habilidades blandas (Vance Hall et. al, 1968; Kraft, 2019), requiere emplear aspectos del enfoque conductista y/o humanista.

1.2 Capacitación docente y prácticas pedagógicas

Los programas de capacitación docente pueden clasificarse en varias categorías, según la etapa de la carrera del docente en la que se aplican (antes del servicio, o *pre-service*; y durante el servicio, o *in-service*); o de acuerdo con el nivel de estructura y diseño del entrenamiento (formales e informales). La literatura académica acepta ampliamente que los docentes tienen un impacto significativo en los resultados de sus estudiantes (Darling-Hammond, Hyler, y Gardner, 2017; Kraft, Blazar, y Hogan, 2018; Taylor, 2023). Por consecuencia, resulta de principal interés identificar las características relevantes en la productividad de los docentes, y los mecanismos por medio de las cuales estas características pueden potenciarse.

La sección anterior abordó las principales teorías del aprendizaje. Estas no solo son fundamentales para comprender los procesos mediante los cuales los estudiantes asimilan y construyen conocimientos, sino también para comprender las dinámicas de aprendizaje de los docentes y optimizar el diseño de los programas de capacitación (enseñar a enseñar). Al adentrarnos en estas teorías, obtenemos una guía esencial para desarrollar metodologías y estrategias pedagógicas efectivas, lo que se traduce en una formación docente más sólida y enfocada.

El proyecto *Measures of Effective Teaching* (MET) reúne a académicos, profesores y organizaciones educativas comprometidas con la investigación de mejores formas de identificar y desarrollar una enseñanza eficaz. Kane y Staiger (2012) lideran el proyecto y evalúan el efecto de las prácticas pedagógicas en distintas variables de interés. Estas incluyen, pero no se limitan a mejoras en las calificaciones, nivel de esfuerzo y satisfacción de los estudiantes. El MET analiza cinco esquemas diferentes de rúbricas de evaluación de prácticas pedagógicas, los cuales están presentados en el anexo 2. Por ejemplo, el *Framework for Teaching* (FFT) desarrollado por Danielson (1996) toma en consideración si el docente plantea preguntas durante la clase, y si estas son preguntas que requieren que los estudiantes revelen su razonamiento. Otro ejemplo es *Classroom Assessment Scoring System* (CLASS) de Pianta et. al, 2008, que evalúa la eficacia con la que el docente interactúa y motiva a su salón de clases. En general, estas rúbricas incluyen lo que en la literatura se acepta como prácticas pedagógicas efectivas. Suelen incluirse la innovación que el docente pueda hacer en el currículo, la motivación y compromiso del estudiante, la buena gestión y conducta dentro del aula, adecuadas estrategias de evaluación, etc. Kane y Staiger (2012) muestran que los profesores con puntuaciones más altas en estas rúbricas tienen estudiantes que aprendieron más y se desempeñaron mejor en otros aspectos fuera del aula.

Finalmente, el anexo 3 representa visualmente el marco teórico del presente trabajo. Este sostiene que las capacitaciones docentes, en tanto cuenten con las características documentadas por Darling-Hammond (2017) y Kraft et. al (2018), tendrán un impacto en las prácticas pedagógicas. Estas prácticas se reflejan en habilidades como el buen manejo del aula de clase, la claridad de las explicaciones, la capacidad de crear un ambiente de confianza entre maestro y alumno, entre otras detalladas en las rúbricas de evaluación de las observaciones del aula (Kane y Staiger, 2012). Esta mejora en las prácticas pedagógicas impactará, a su vez, en los aprendizajes de los estudiantes, medidos ya sea por calificaciones u otras métricas de aprendizaje no cognitivo (conducta, esfuerzo, motivación, etc.).

CAPÍTULO II. EVIDENCIA EMPÍRICA

Las escuelas utilizan distintas métricas para medir el desempeño de sus docentes. Algunas se basan en resultados cuantificables, como calificaciones de los estudiantes (); mientras que otras, en aspectos cualitativos, como observaciones de las prácticas pedagógicas dentro del aula (). Independientemente del método de evaluación utilizado, las características específicas de los docentes son importantes. La literatura distingue entre características adquiridas y sociodemográficas (Coenen et. al, 2018). Las primeras se relacionan con la educación (grado académico alcanzado, calidad universitaria, certificados de enseñanza, puntajes de exámenes docentes y años de experiencia)¹; mientras que las segundas, en elementos como el género y origen étnico del docente.

Esta sección resume la evidencia empírica sobre el efecto de los programas de capacitación docente en i) las prácticas pedagógicas y ii) los aprendizajes de los estudiantes. Primero, se analiza las formas en las que la literatura suele cuantificar ambas variables de interés. Los aprendizajes pueden medirse a través de las calificaciones, aprendizaje no cognitivo, y tasa de competencia de los alumnos. Segundo, se explora los distintos tipos de capacitación docente generalmente aplicados. Los programas se categorizan según la etapa de la carrera del docente en la que se aplican (*pre-service* e *in-service*); o de acuerdo con el nivel de estructura y diseño del entrenamiento (formales e informales). Finalmente, se discuten los principales problemas de estimación y especificaciones econométricas empleadas en la literatura.

2.1 Aprendizaje del estudiante

La forma más intuitiva para medir el desempeño de un docente es estimar su contribución al aprendizaje de sus alumnos. La aproximación más usada para son las calificaciones obtenidas en los exámenes (Hanushek, 1971; Harris y Sass, 2011; Coenen et. al, 2018; Taylor, 2023). Si bien estos datos son estadísticamente convenientes de usar, no capturan todo el espectro del conocimiento. Para complementar las notas, la literatura emplea métricas de aprendizaje no cognitivo (Jackson, 2018; Kraft, 2019).

2.1.1 Calificaciones

La contribución de un docente individual al rendimiento académico de sus alumnos se denomina comúnmente “medida de valor agregado” (*Value Added Measures*, o VAM en inglés). El uso de estas métricas se remonta a Hanushek (1971), a partir del cual empezaron a

¹ Estas características se adquieren a través de programas de capacitación docente, los cuales serán abordados en esta sección.

ganar popularidad hasta 2019, donde dos tercios de los estados de EE. UU. exigían el uso de alguna métrica de valor agregado (Taylor, 2023).

Taylor (2023) encuentra que una desviación estándar (SD) en la distribución del desempeño docente equivale alrededor de 0.10 y 0.20 SD de las calificaciones de sus estudiantes. Es más, una clase promedio con un maestro dentro del percentil 75 de la distribución obtendrá una puntuación entre 0.13 y 0.27 SD más alta que la misma clase con un maestro del percentil 25. Asimismo, existe evidencia alentadora que muestra que los docentes con VAM más altas afectan positivamente los resultados a largo plazo de sus estudiantes, como su admisión a la universidad (Jackson, 2018) o su nivel de ingresos (Chetty et. al, 2014).

Uno de los principales problemas de estos modelos es la asignación de estudiantes y docentes. Uno de los supuestos es que esta asignación es aleatoria, y las características del estudiante no determinan su emparejamiento con un profesor en particular, ni viceversa (Taylor, 2023). En la práctica, sin embargo, es posible, por ejemplo, que los maestros más productivos sean asignados a los alumnos con menores calificaciones. En general, estudios cuasiexperimentales apoyan el supuesto de asignación aleatoria (Chetty et. al, 2014; Bacher-Hicks et. al, 2019); aunque la evidencia no es unánime (Harris y Sass, 2011; Rothstein, 2017).

Un enfoque alternativo al VAM es considerar simplemente el porcentaje de estudiantes que aprueban el año escolar. La literatura referencia a esta medida como “tasa de competencia”, la cual se usa mayormente para evaluar docentes de manera grupal, y no individual (Taylor, 2023). Otra diferencia es que no se evalúa la contribución del maestro al aprendizaje, sino simplemente si es que el estudiante cuenta con el conocimiento necesario para aprobar graduarse o pasar de año. (Taylor, 2023) precisa que los puntajes de los exámenes generalmente se dividen en niveles de logro, como avanzado, competente y básico. La evidencia muestra que lo que un estudiante debe saber para ser considerado “competente” varía considerablemente, a pesar de que las puntuaciones límites son establecidas por un panel de expertos (Koretz, 2008).

2.1.2 Aprendizaje no cognitivo

Naturalmente, las notas no capturan todos los aspectos del conocimiento. Del mismo modo en que se puede inferir que un estudiante con notas más altas tiene habilidades cognitivas superiores a las de un estudiante con notas bajas; se puede inferir que un estudiante irrespetuoso, indisciplinado, que falta a clases y no entrega las tareas a tiempo probablemente tenga habilidades no cognitivas inferiores.

Los resultados de Jackson (2018) sugieren que las contribuciones de los docentes se reflejan también en las habilidades conductuales de sus estudiantes. El autor crea un índice de conductas para representar las habilidades no cognitivas y encuentra que los maestros de noveno grado impactan en la deserción escolar, graduación de la escuela secundaria, las intenciones de asistir a una universidad, entre otros.

Kraft (2019) evalúa el impacto de los docentes en las habilidades socioemocionales. En el experimento, los estudiantes completaron breves cuestionarios para medir su determinación y su mentalidad de mejora. El índice de determinación captura la tendencia de los estudiantes a mantener el interés y esfuerzo en sus clases; mientras que la mentalidad de mejora mide el grado de confianza de los estudiantes en su propio potencial. En ambos casos, la contribución del docente es positiva y significativa (Kraft, 2019).

2.2 Prácticas pedagógicas

Las escuelas también evalúan el desempeño del docente observando su trabajo en el aula de clases (Barr, 1928). Actualmente, los evaluadores se basan en rúbricas de puntuación detalladas que toman en cuenta prácticas pedagógicas deseables en un maestro (Kane y Staiger, 2012).

2.2.1 Observaciones del aula

Las observaciones del aula son conducidas por un evaluador especializado, ya sea externo (auditor) o interno (otro docente). A diferencia del primer grupo de variables, estas evalúan las cualidades del docente (habilidades comunicativas, interacción con los estudiantes, etc.), mas no sus resultados (calificaciones del estudiante). Por ejemplo, muchas rúbricas piden al observador que califique si el docente hace preguntas a sus alumnos, pero no si es que estas generaron aprendizaje en los estudiantes (Taylor, 2023). Kane y Staiger (2012) comparan las métricas de evaluación más empleadas por los observadores, las cuales están detalladas en el anexo 2. En base a las características evaluadas por estas rúbricas, resulta intuitivo pensar que los docentes que cuenten con estas habilidades desempeñarán mejor su rol como educadores. Bacher-Hicks et. al (2019) realizan un experimento aleatorio en salones de matemática primaria, y concluyen que el puntaje obtenido por los docentes en observaciones del aula (CLASS, MQI, Trípode) es un predictor de su desempeño.

Una crítica común a este tipo de métricas es el hecho de que son pocos los docentes que reciben calificaciones bajas (Bell et. al, 2022; Taylor, 2023). El sesgo de “indulgencia” pone en discusión la objetividad de los evaluadores, en especial si estos no son observadores externos, sino colegas del profesor evaluado. En estos casos, la resiliencia a registrar una baja

calificación últimamente sobrestima el desempeño docente. Taylor (2023) encuentra que apenas el 5-10% de los docentes obtienen una puntuación de 1 o 2, en una escala de 1-5.

2.2.2 Valoraciones del estudiante

Algunas escuelas brindan la posibilidad de que los estudiantes califiquen a sus maestros. El anexo 4 contiene la encuesta de percepción de los estudiantes del MET se basa en el Proyecto Trípode para la Mejora Escolar (Kane y Cantrell, 2010). Las preguntas del Trípode se agrupan en siete títulos o conceptos, llamados las siete Cs: *Care, Control, Clarify, Challenge, Captivate, Confer, Consolidate*. Lo que se le pregunta al estudiante es muy similar a lo que se evalúa en las rúbricas durante las observaciones de aula. Un hallazgo importante fue que, en las escuelas secundarias con menor rendimiento, menos de la mitad de los alumnos respondió de manera afirmativa; mientras que, en las escuelas de rendimiento superior, más del 80% estuvo de acuerdo (Kane y Cantrell, 2010).

2.3 Tipos de capacitación docente

Independientemente del método de evaluación utilizado, existe variabilidad entre el desempeño entre docentes. Estas diferencias sugieren la presencia de habilidades pedagógicas distintas, lo que motiva el interés para investigar cómo los docentes adquieren esas habilidades. Los programas de capacitación docente se pueden clasificar según distintas dimensiones, dependiendo de si estos que ocurren antes o durante el ejercicio de la docencia (*pre-service* e *in-service*), o si se rigen bajo un esquema estructurado o no (formales e informales).

Por un lado, los programas *pre-service* están diseñados para aquellos maestros en preparación, y tienen como objetivo equiparlos con conocimientos y habilidades necesarios para ingresar a la profesión docente. Los programas de formación continua, en cambio, están destinados a mejorar el desarrollo profesional de los docentes en ejercicio (*in-service*). Si bien la intuición detrás de cada tipo de programa puede sugerir un efecto positivo en la enseñanza, la evidencia empírica favorece a los programas *in-service* sobre los *pre-service* (Harris y Sass, 2011; Koedel et. al, 2015; Taylor, 2023).

Por otro lado, los programas formales están estructurados bajo un plan de estudios específico que conducen a una certificación reconocida. Los programas informales, por el contrario, son más flexibles. Algunos ejemplos son talleres, seminarios o cursos en línea en los que los docentes participan voluntariamente. En el sector educativo, la capacitación formal predomina, a pesar de la escasa evidencia de que estos programas tradicionales mejoren el desempeño docente (Taylor, 2023). Sin embargo, la literatura apoya la idea que los docentes aprenden de

manera informal, ya sea por experiencia (Harris y Sass, 2011), o de sus colegas (Papay et. al, 2020).

2.3.1 Capacitación formal *pre-service*

Gran parte de la formación de los docentes ocurre antes de empezar a enseñar. Ya sea mediante la obtención de un título universitario o cursos especializados en educación, los maestros prospectos reciben algún tipo de entrenamiento previo. Según la teoría del capital humano, se podría esperar que exista una relación positiva entre el nivel de educación del docente y el aprendizaje de sus estudiantes. No obstante, la literatura sobre este tipo de programas no solo es escasa, sino que no proporciona evidencia consistente de que los docentes con mayor nivel de educación se desempeñen mejor en el aula (Aaronson et. al, 2007; Harris y Sass, 2011; Coenen et. al, 2018; Goldhaber et. al, 2020).

La mayor parte de la literatura se centra en el efecto diferenciado de tener una maestría o un doctorado, en comparación con una licenciatura (Coenen et. al, 2018). En general, el efecto es no significativo, a excepción de docentes de ciencias o matemáticas con maestría. De manera similar, Koedel et. al (2015) comparan el desempeño de docentes que participaron en distintos programas de preparación (TPP, en inglés). Sus resultados sugieren que las variaciones en la eficacia pedagógica provienen de diferencias en las características del docente, mas no de diferencias en los programas de preparación.

Si bien estos estudios no hallan diferencias significativas entre las capacitaciones *pre-service*, esto no implica que haber participado en algún programa sea irrelevante en el desempeño futuro del docente. Taylor (2023) comenta que un grupo de control más efectivo son los profesores que enseñan sin certificación, o con algún entrenamiento informal de corta duración, como programas de verano. Kane, Rockoff, y Staiger (2008) toman como muestra las escuelas públicas de Nueva York, y no hallan diferencias en el promedio ni en la dispersión de las notas de estudiantes que tuvieron un maestro certificado versus uno no certificado. Por el contrario, los primeros dos años de experiencia del docente son un mejor predictor de la eficacia futura. Constantine et. al (2009) reportaron resultados similares en un experimento aleatorio en múltiples estados, sin encontrar diferencias significativas en las calificaciones.

En general, la eficiencia del docente parece no estar relacionada ni con la carrera universitaria elegida (Aaronson et. al, 2007), ni con haber llevado cursos de pregrado vinculados educación (Harris y Sass, 2011). Solamente Clotfelter et. al (2010) encuentran una relación positiva significativa entre el nivel de prestigio de la escuela de pregrado y la productividad de los

docentes de educación secundaria. Para cursos de comprensión lectora, los docentes que tienen un título en educación desempeñan significativamente mejor que aquellos que no (Harris y Sass, 2011). Otra excepción es para cursos de matemáticas de nivel secundaria, donde las calificaciones mejoran si un profesor está certificado y/o tiene un título o cursos completos relacionados con matemáticas (Wayne y Youngs, 2003; Harris y Sass, 2011). Asimismo, Wayne y Youngs (2003) encuentran que los estudiantes generalmente aprenden más de profesores con puntajes más altos en sus pruebas docentes y en su historial universitario (GPA). Sin embargo, esto parece estar más relacionado a las habilidades intelectuales del docente, mas no al efecto de meramente asistir a la universidad.

La aparente ineficacia de la capacitación *pre-service* puede deberse a que la productividad docente es específica del contexto (por ejemplo, el tipo de estudiante y docente) o que los programas de capacitación puedan no estar centrados en desarrollar el tipo de prácticas pedagógicas que mejoren el aprendizaje. Esto también es consistente con la evidencia de que la experiencia, que es menos estandarizada, tiene efectos significativos en la productividad docente (Harris y Sass, 2011; Taylor, 2023).

Otro ejemplo de capacitaciones previas, aunque menos estudiado, es el programa de asistentes de enseñanza (*teaching assistant* o TA en inglés). Los participantes suelen ser estudiantes universitarios o recién graduados, los cuales colaboran bajo la supervisión de un docente en el diseño, elaboración y conducción de las clases. Goldhaber, Krieg, y Theobald (2020) encuentran que los docentes desempeñan mejor cuando los cursos que dictan son similares a los que dictaron cuando fueron TA. También, los programas que permiten a los estudiantes participar como TA son de los mayores predictores (Boyd et. al, 2009).

2.3.2 Capacitación formal *in-service*

Bajo el esquema en el que los docentes trabajan para las escuelas, estas tienen un incentivo de capacitar a su personal en ejercicio. Estos programas están comúnmente estipulados en los contratos; aunque los profesores pueden tener incentivos para capacitarse por su cuenta. La literatura sobre las capacitaciones *in-service* es más amplia, aunque la evidencia empírica no conduce a una respuesta clara sobre el impacto en el aprendizaje (Taylor, 2023).

2.3.2.1 Desarrollo profesional

Los programas de desarrollo profesional (PD en inglés) son iniciativas estructuradas destinadas a mejorar las habilidades, conocimientos y eficacia de los docentes en el aula. Yoon et. al (2007) revisaron alrededor de 1300 estudios de programas PD. Uno de los problemas de

estimación más comunes es el sesgo de selección², a raíz del cual solo 9 de esos estudios tenían un diseño cuasiexperimental suficiente para encontrar un efecto causal de estos programas. Asimismo, Hill et. al (2013) encuentran que los especialistas suelen estar mayormente de acuerdo en los contenidos que deben incluirse en estos programas, a pesar de que, al implementarse, estos no producen mejoras en la enseñanza.

Por un lado, Jacob y Lefgren (2004) estudian escuelas de alta pobreza en Chicago, sin encontrar efectos significativos del entrenamiento docente sobre el desempeño en matemáticas ni lectura de alumnos de primaria. Los autores emplean una regresión discontinua utilizando como punto de quiebre la introducción de una reforma educativa que ponía a las escuelas con menor desempeño en un periodo de prueba (*probation*). Estas escuelas recibían apoyo financiero dedicado específicamente a mejorar la instrucción en el aula, principalmente a través de la capacitación docente y el desarrollo del personal. Por otro lado, Angrist y Lavy (2001) realizan un estudio similar para las escuelas no religiosas de Jerusalén, aunque encontrando un efecto positivo en el entrenamiento *in-service*. Los docentes capacitados recibían 10.5 horas más de entrenamiento semanal en matemáticas que el grupo de control (y 7.7 horas más de lectura). Mediante una estimación de diferencias en diferencias, los alumnos de las escuelas de tratamiento incrementaron sus calificaciones en 0.46 desviaciones estándar en matemáticas; y 0.68 en lectura (Angrist y Lavy, 2001). Una diferencia importante respecto al experimento de Chicago es la inclusión de componentes de retroalimentación que recibieron los docentes de Jerusalén luego de terminar el programa, lo que sugeriría que la mera capacitación no es suficiente para mejorar resultados.

Esta inconsistencia en los resultados también está presente en estudios más recientes. En general, los autores encuentran impactos significativos cuando segmentan por grado de instrucción (Harris y Sass, 2011), materia (Abeberese et. al, 2014), tipo de entrenamiento (Kerwin y Thornton, 2021), y variable de interés (Loyalka et. al, 2019). Asimismo, los efectos no son permanentes, sino que se difuminan en el corto plazo (Lombardi, 2019).

Primero, Abeberese et. al (2014) encuentran efectos a corto plazo en la comprensión lectora de estudiantes filipinos, aunque sin mejoras en otros cursos. Las aulas de los docentes del grupo de tratamiento incrementaron sus calificaciones en 0.13 SD, aunque el efecto se redujo a 0.06 en los siguientes tres meses. Además, Abeberese et. al (2014) muestran mejoras los hábitos de lectura de los estudiantes, incluso luego de acabado el programa. En otro estudio, programas

² Abordado en más detalle en la sección 2.4

de capacitación destinados a docentes de bajo rendimiento en Chile mejoraron sus prácticas pedagógicas y las calificaciones de sus alumnos (Lombardi, 2019). A pesar de esto, los resultados no son robustos y el efecto desaparece después de un año.

Para Uganda del Norte, Kerwin y Thornton (2021) comparan dos versiones de capacitación de alfabetización estudiantil, encontrando diferencias significativas en las calificaciones de lectura y redacción. La versión completa logró mejorar las notas en 0.64 y 0.45, respectivamente; mientras que la versión reducida no tuvo efecto en la comprensión lectora y, en cambio, redujo las notas de redacción en -0.33 SD. La estructura de la capacitación siguió un esquema *train-the-trainer*, en el que los formadores transmiten habilidades y competencias a tutores de centros coordinadores (CCT), quienes luego capacitan a los docentes, y realizan visitas periódicas para monitorear sus resultados (Kerwin y Thornton, 2021). La versión reducida, sin embargo, no tuvo este esquema, sino que proveyó a los docentes directamente de guías y recursos audiovisuales para que se capaciten por su cuenta. Asimismo, la ausencia de monitoreo en la versión reducida sugiere algo similar a Angrist y Lavy (2001), sobre la importancia del componente de retroalimentación.

Loyalka et. al (2019) realizaron un experimento para asignar de manera aleatoria la participación de maestros en programas in-service en China. El programa se destinó a docentes de matemáticas, los cuales recibían un entrenamiento de dos semanas en modalidad presencial, seguido de material suplementario online. Al igual que Kerwin y Thornton (2021), se comparó distintas versiones del programa. Ninguna de estas versiones conduce a mejoras significativas en las variables de interés planteadas por Loyalka et. al (2019), respecto al grupo de control: calificaciones de matemáticas (0.023 SD, 0.036 SE), tasa de deserción de los alumnos (-0.002 SD, 0.009 SE), conocimiento del docente (0.153 SD, 0.138 SE), ni en habilidades pedagógicas (0.008 SD, 0.048 SE), entre otros.

En esa misma línea, Harris y Sass (2011) encuentran que el entrenamiento in-service no tiene efecto en la productividad e incluso puede reducirla por diversas razones. Primero, participar en programas PD quita tiempo a los maestros para preparar y dictar sus clases. Si se contrata sustitutos menos eficientes durante el tiempo que dura el programa, esto puede reducir los resultados promedio del aula del docente capacitado. Por último, debido a que generalmente ha pasado una parte del año escolar una vez concluido el programa PD, los maestros tienen pocas oportunidades de incorporar inmediatamente lo que aprenden. En consecuencia, Harris y Sass (2011) encuentran que el efecto inicial de la capacitación es negativo, aunque esto

revierte después de dos años. Estos resultados son significativos únicamente para profesores de matemáticas de secundaria y preparatoria.

2.3.2.2 Estudios de posgrado

Aparte de la capacitación organizada por las escuelas, los docentes cuentan con incentivos económicos (salarios más altos) para completar cursos de posgrado³. La literatura se concentra mayormente en el efecto de la obtención de una maestría o doctorado en la productividad del docente (Rivkin et. al, 2005; Clotfelter et. al, 2007). En general, la evidencia apunta a que no hay impacto en la obtención de dichos grados, probablemente porque el contenido está enfocado, salvo excepciones, en reforzar las habilidades técnicas y no pedagógicas, cuando son estas últimas las que tienen efectos en el aprendizaje. Rivkin et. al (2005) no proporcionan evidencia de que poseer una maestría aumente la productividad de los docentes, independientemente de la materia o grado que dicten. Harris y Sass (2011) encuentran una mejora solo para docentes de matemáticas de primaria y secundaria; mientras que, para comprensión lectora en secundaria, el efecto es negativo.

Del mismo modo, los hallazgos de Clotfelter et. al (2007) sugieren que contar con un posgrado no es predictor de mayor productividad entre docentes. Por el contrario, encuentran un efecto negativo de contar con un profesor con doctorado. Los autores argumentan que esto puede estar más relacionado a las características de los docentes con Ph.D. que a la eficacia de obtener el posgrado (Clotfelter et. al, 2007). En cambio, los efectos de la maestría dependen del nivel de enseñanza y el momento en la carrera del docente en que se cursó. Para secundaria, al igual que Harris y Sass (2011), la maestría es un buen predictor de la productividad docente. Para otros grados, en cambio, los docentes que empezaron a enseñar luego de obtener la maestría son marginalmente más efectivos. Por el contrario, estudiar un posgrado con cinco o más años de experiencia docente tiene efectos negativos en el aprendizaje de los estudiantes. Esto se interpreta como si los profesores menos eficaces son los que deciden por obtener títulos de posgrado en etapas más tardías de sus carreras (Clotfelter et. al, 2007).

2.3.2.3 Coaching

Estos programas han ganado popularidad respecto a métodos de capacitación tradicionales. La introducción del coaching se remonta a los años ochenta, conceptualizado inicialmente como herramienta que facilitaba la habilidad del docente para transmitir sus conocimientos hacia sus alumnos, mediante la observación y retroalimentación continua (Joyce y Showers, 1982). Si

³ No necesariamente vinculados a educación, sino a la carrera que estudiaron o a la materia que dictan.

bien esta definición puede calzar con la de programas PD, el coaching pretende ser individualizado, intensivo, prolongado, especializado y centrado en habilidades pedagógicas (Kraft et. al, 2018). De hecho, el aparente fracaso del enfoque tradicional de PD para mejorar la instrucción docente se justifica, en parte, por su estructura de corta duración y contenidos genéricos (Darling-Hammond, 2017).

Si bien suele aplicarse en docentes con poca experiencia previa, otros incluyen también a docentes veteranos (Blazar y Kraft, 2017). Asimismo, la mayor parte de la evidencia apunta a beneficios en la educación inicial y primaria. Kraft et. al (2018) recopilan datos de 60 estudios donde se aplica algún programa de coaching, y evalúan sus efectos en las prácticas pedagógicas del docente (*instruction*) y en los aprendizajes (*student achievements*). En ambos casos, el efecto es positivo, con un impacto de 0.49 SD y 0.18 SD, respectivamente. Asimismo, los autores observan efectos diferenciados en programas de coaching específicos de alguna materia (lectura, matemáticas, ciencias, etc.) y programas más genéricos. El primer tipo tiene un efecto mayor, tanto en las observaciones del aula (0.51 SD vs 0.47 SD), como en las calificaciones de los alumnos (0.20 SD vs 0.07 SD). Finalmente, el efecto se desagrega en distintos grados de instrucción (inicial, primaria y secundaria), aunque no se encuentran diferencias significativas (Kraft et. al, 2018).

En cuanto al efecto diferenciado del coaching respecto a otros programas, Cilliers et. al (2020) realizan un experimento de asignación aleatoria en escuelas de Sudáfrica. Las escuelas fueron divididas tres categorías (coaching, capacitación grupal, y control). Por un lado, las capacitaciones grupales fueron de dos días al comienzo de cada semestre, con un capacitador asignado a cada siete maestros. El tratamiento de coaching, en cambio, consistía en visitas mensuales a cada docente, donde el capacitados observaba sus clases, brindaba retroalimentación y dictaba una clase modelo. Cilliers et. al (2020) encuentran que ambos programas tuvieron un impacto positivo respecto al grupo de control, aunque las calificaciones mejoraron más con el coaching que con la capacitación grupal (0.24 SD vs 0.12 SD).

2.3.3 Capacitación informal

2.3.3.1 Aprendizaje por experiencia (*learning-by-doing*)

Si bien existe discrepancia en la literatura respecto al efecto de las capacitaciones formales, un aparente hecho estilizado es que la experiencia impacta tanto en la calidad docente como en los aprendizajes (Kraft, Papay, y Chi, 2020; Taylor, 2023). Asimismo, el efecto es positivo independientemente de qué métrica se use, ya sea calificaciones de alumnos (Harris y Sass,

2011) u observaciones del aula (Bell et. al, 2022). Por otro lado, Ost (2014) sugiere que el desempeño al inicio de la carrera depende de la experiencia específica en un determinado grado de instrucción (los docentes que cambian de grado recurrentemente mejoran menos).

Existen varios problemas al momento de realizar inferencias sobre los efectos de la experiencia docente. Primero, los mecanismos por medio de los cuales mejora la productividad no están claramente definidos, razón por la cual, en la práctica, no se la considera como un programa de capacitación (Taylor, 2023). Después, Bell et. al (2022) destacan la posibilidad de que los evaluadores sean más indulgentes en sus calificaciones con profesores experimentados. Finalmente, existe un sesgo de selección, en el sentido que los docentes con menor desempeño son más propensos a dejar la profesión y, por lo tanto, acumular menos años de experiencia (Taylor, 2023).

2.3.3.2 Aprendizaje por colegas (*learning from peers*)

Resulta viable pensar que los docentes aprenden nuevas habilidades de sus colegas. Aprender de sus pares también puede contribuir a la experiencia entre los docentes que recién inician su carrera (Taylor, 2023). Sin embargo, el aprendizaje entre pares es de interés en sí mismo y podría tener beneficios más allá de los primeros años de carrera.

Por un lado, Papay et. al (2020) muestran evidencia de que los docentes aprenden de sus compañeros. En las escuelas de tratamiento, se formó parejas de profesores y se les pidió que trabajaran juntos para mejorar su enseñanza. Cada pareja incluía un maestro “objetivo”, de bajo desempeño; y un maestro “compañero”, de alto desempeño. El nivel de desempeño se midió en base a habilidades pedagógicas, como hacer preguntas en el aula (Papay et. al, 2020). El rendimiento de los estudiantes de las escuelas de tratamiento mejoró 0.12 SD en las aulas de los profesores objetivo. Por otro lado, Murphy, Weinhardt, y Wyness (2021) no encuentran efectos significativos de colaborar entre maestros. Al igual que el ejemplo anterior, los docentes fueron divididos en parejas, las cuales se reunían para diseñar estrategias pedagógicas y proveer retroalimentación de sus clases.

Un aspecto importante por considerar es el potencial costo de este tipo de programas: el tiempo y esfuerzo invertido por el mentor. Por ejemplo, un mentor que deja su propio salón para pasar tiempo observando las clases de un compañero y, por lo tanto, los propios estudiantes del mentor se ven perjudicados (Taylor, 2023). Finalmente, los programas de asistentes de enseñanza (Boyd et. al, 2009; Goldhaber et. al, 2020) podrían ser considerados dentro de este grupo. Si bien estos tienen una naturaleza más jerárquica y suelen ocurrir *pre-service*, el

estudiante asistente (mentorizado) es emparejado con un docente experimentado (mentor), por medio del cual aprende y mejora sus habilidades pedagógicas.

2.4 Estrategias econométricas

El anexo 5 sintetiza los resultados de los estudios previamente mencionados, sobre los efectos de la formación docente en las prácticas pedagógicas y el rendimiento estudiantil. La mayoría emplea datos de escuelas estadounidenses y emplea las calificaciones como métrica de aprendizaje. Los resultados se dividen por tipo de capacitación (*pre-service*, *in-service*, capacitación informal), materia (matemáticas, comprensión lectora, multidisciplinarios) y grado de instrucción (primaria, secundaria), lo que permite resaltar efectos diferenciados.

Existen algunos consensos en la literatura sobre las estrategias de estimación necesarias para obtener resultados robustos. Primero, se reconoce que los niveles de logro estudiantil son el resultado de un proceso acumulativo, influenciado por docentes y rendimientos previos; y que las posibles contribuciones del docente de tratamiento dependen en gran medida de las características socioeconómicas del entorno estudiantil (Coenen et. al, 2018). Por lo tanto, es crucial que los diseños de estudio controlen por estas diferencias. Segundo, existe también un consenso en que únicamente los diseños (cuasi-)experimentales pueden derivar en resultados robustos (Coenen et. al, 2018). Sin embargo, son escasos los estudios donde tanto los estudiantes como docentes sean asignados aleatoriamente a las clases. Aun así, se ha demostrado que la mayor parte del sesgo resultante de esta asignación no aleatoria puede compensarse controlando por los resultados previos en pruebas de los estudiantes (Chetty, Friedman, & Rockoff, 2014).

En línea con lo anterior, la tercera columna del anexo 5 identifica los métodos de estimación empleado por los estudios revisados, los cuales emplean datos de panel. Por un lado, en los tratamientos controlados aleatorios (RCT, en inglés), los estudiantes son asignados aleatoriamente a las clases, de modo que los efectos estimados del docente no pueden ser resultado de un sesgo selectivo de estudiantes a clases específicas. No obstante, en algunos casos no considera la asignación de docentes, que pueden elegir en qué escuela trabajar y, hasta cierto punto, a qué tipo de estudiantes enseñar. Después, los estudios cuasiexperimentales de tipo I (QEX-I) reconocen que las características heterogéneas no observadas pueden sesgar los resultados. Finalmente, los diseños tipo QEX-II asumen que las variables de control incluidas son suficientes para solucionar el problema de heterogeneidad.

Una de las principales deficiencias de los primeros estudios que intentan explorar estas relaciones es su dependencia de las características observadas de los estudiantes (covariables) para explicar la heterogeneidad. Existe evidencia, en cambio, de que los docentes más capacitados y con mayor experiencia tienden a ser asignados a estudiantes más capaces y disciplinados (Clotfelter et. al, 2007; Harris y Sass, 2011). No controlar por estas características no observadas de los estudiantes sobrestima las estimaciones el efecto de las capacitaciones.

Asimismo, los efectos también pueden estar sesgados si no se considera el problema de deserción o desgaste (*attrition*). Por ejemplo, los docentes menos eficaces podrían tener más probabilidades de abandonar el programa y esto puede dar la apariencia de que el tratamiento aumenta el valor añadido docente cuando, en realidad, los docentes menos eficaces simplemente están saliendo de la muestra (Rockoff, 2004; Kane et. al 2006). Por otro lado, los docentes más capaces y con mayores costos de oportunidad pueden tener más probabilidades de migrar de profesión (y, en consecuencia, abandonar también el tratamiento), lo que lleva a una correlación negativa entre la capacitación docente y el rendimiento de los estudiantes (Harris y Sass, 2011). En general, los trabajos de Kraft et. al (2018), Coenen et. al (2018) Taylor (2023) realizan una revisión extensa de la relación entre la capacitación docente, las prácticas pedagógicas y los aprendizajes, donde contemplan las herramientas con las que se tratan estos problemas de estimación. Estas incluyen la implementación de variables para controlar la heterogeneidad, como efectos fijos del estudiante, docente y/o colegio.

CONCLUSIONES

La eficacia en la enseñanza es crucial para el éxito académico de los estudiantes. Pese a la gran cantidad de investigaciones que resaltan la relevancia de un docente competente, persiste un debate en torno a los elementos que contribuyen o predicen eficazmente dicha competencia. En este contexto, la formación docente se presenta como un elemento esencial, dado que el fortalecimiento de las técnicas pedagógicas favorece el progreso académico de los estudiantes (Darling-Hammond, 2017; Kraft et. al, 2018; Taylor, 2023). El enfoque econométrico preferido por la literatura son los ensayos controlados aleatorizados (RCT), ya que permiten controlar la asignación selectiva de docentes o estudiantes. Los diseños cuasiexperimentales discutidos en este documento toman en cuenta las características observadas y no observadas, y bien reconocen la presencia de heterogeneidad restante (QEX-I) o suponen que los controles aplicados son suficientes para compensar cualquier sesgo potencial (QEX-II).

En general, existe poca evidencia que demuestre que el nivel de educación previo del docente (pregrado, maestría, u otras certificaciones *pre-service*) tenga efectos significativos en el aprendizaje de los estudiantes. Algunos estudios, por el contrario, encuentran una relación negativa entre docentes con maestría y desempeño estudiantil (Clotfelter et. al, 2010). Un hallazgo importante es que los profesores que fueron asistentes de enseñanza (TA) en su época estudiantil se desempeñan mejor, en especial si dictan el mismo curso (Goldhaber, 2020).

Por otro lado, los programas de desarrollo profesional (PD) muestran resultados mixtos a lo largo de los años y en distintos países, ya sea por la estrategia econométrica empleada (Coenen et. al, 2018) o por los contenidos abordados en los programas (Darling-Hammond, 2017). En cuanto a lo primero, los diseños de tratamientos controlados aleatorios (RCT) son los más preferidos, en tanto eliminan la posibilidad de asignación selectiva de estudiantes o maestros al aula. Sobre lo segundos, programas de PD que tengan un diseño personalizado y cuenten con algún tipo de retroalimentación hacia el docente muestran efectos positivos significativos en las calificaciones y otras métricas de aprendizaje no cognitivo (Kraft et. al, 2018, Kerwin y Thorton, 2021). Finalmente, la evidencia muestra que los docentes aprenden por experiencia (Harris y Sass, 2011) o por colegas (Papay et. al, 2020).

REFERENCIAS BIBLIOGRÁFICAS

- Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics*, 25(1), 95-135.
- Abeberese, A. B., Kumler, T. J., & Linden, L. L. (2014). Improving reading skills by encouraging children to read in school: A randomized evaluation of the Sa Aklat Sisikat reading program in the Philippines. *Journal of Human Resources*, 49(3), 611-633.
- Angrist, J. D., & Lavy, V. (2001). Does teacher training affect pupil learning? Evidence from matched comparisons in Jerusalem public schools. *Journal of Labor Economics*, 19(2), 343-369.
- Bacher-Hicks, A., Chin, M. J., Kane, T. J., & Staiger, D. O. (2019). An experimental evaluation of three teacher quality measures: Value-added, classroom observations, and student surveys. *Economics of Education Review*, 73, 101919.
- Barr, A. S. (1928). An evaluation of items to observe in classroom supervision. *The Journal of Educational Research*, 18(1), 53-65.
- Bell, C., James, J., Taylor, E. S., & Wyckoff, J. (2022). *Measuring returns to experience using observations of teaching*. EdWorkingPaper No. 22-526.
- Blazar, D., & Kraft, M. (2017). Teacher and teaching effects on students' attitudes and behavior. *Educational Evaluation and Policy Analysis*, 39(1), 146-170.
- Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis*, 31(4), 416-40.
- Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Reasuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review*, 104(9), 2633-2679.
- Cilliers, J., Fleisch, B., Prinsloo, C., & Taylor, S. (2020). How to improve teaching practice? An experimental comparison of centralized training and in-classroom coaching. *Journal of Human Resources*, 55(3), 926-962.

- Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review, 26*(6), 673-682.
- Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2010). Teacher credentials and student achievement in high school: a cross-subject analysis with student fixed effects. *Journal of Human Resources, 45*(3), 655–681.
- Coenen, J., Cornelisz, I., Groot, W., Maassen van den Brink, H., & Van Klaveren, C. (2018). Teacher characteristics and their effects on student test scores: A systematic review. *Journal of Economic Surveys, 32*(3), 848-877. doi:10.1111/joes.12210
- Constantine, J., Player, D., Silva, T., Hallgren, K., Grider, M., & Deke, J. (2009). *An evaluation of teachers trained through different routes to certification*. NCEE 2009-4043. U.S. Department of Education.
- Danielson, C. (1996). *Enhancing professional practice: A framework for*. Alexandria, VA: Association of Supervision and Curriculum Development.
- Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective Teacher Professional Development*. Palo Alto, CA: Learning Policy Institute.
- Goldhaber, D., Krieg, J., & Theobald, R. (2020). Exploring the impact of student teaching apprenticeships on student achievement and mentor teachers. *Journal of Research on Educational Effectiveness, 13*(2), 213-234.
- Grossman, P. (2009). *Protocol for language arts teaching observations*. Stanford, CA: Stanford University.
- Hanushek, E. A. (1971). Teacher characteristics and gains in student achievement: Estimation using micro data. *American Economic Review, 61*(2), 280-288.
- Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of Public Economics, 95*, 798–812.
- Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and instruction, 26*(4), 430-511.

- Hill, H., Beisiegel, M., & Jacob, R. (2013). Professional development research: Consensus, crossroads, and challenges. *Educational Researcher*, 42(9), 476-487.
- Jackson, C. K. (2018). What do test scores miss? The importance of teacher effects on non-test score outcomes. *Journal of Political Economy*, 126(5), 2072-2107.
- Jacob, B. A., & Lefgren, L. (2004). The impact of teacher training on student achievement quasi-experimental evidence from school reform efforts in Chicago. *Journal of Human Resources*, 39(1), 50-79.
- Jacob, B. A., Rockoff, J. E., Taylor, E. S., Lindy, B., & Rosen, R. (2018). Teacher applicant hiring and teacher performance: Evidence from DC public schools. *Journal of Public Economics*, 166, 81-97.
- Joyce, B., & Showers, B. (1982). The coaching of teaching. *Educational leadership*, 40(1), 4-10.
- Kane, T. J., & Cantrell, S. (2010). *Learning about teaching: Initial findings from the measures of effective teaching project*. Seattle, WA: Bill & Melinda Gates Foundation.
- Kane, T. J., & Staiger, D. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains*. Seattle, WA: Bill & Melinda Gates Foundation.
- Kane, T., Rockoff, J., & Staiger, D. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review*, 27(6), 615-631.
- Kerwin, J. T., & Thornton, R. L. (2021). Making the grade: The sensitivity of education program effectiveness to input choices and outcome measures. *Review of Economics and Statistics*, 103(2), 251-264.
- Koedel, C., Parsons, E., Podgursky, M., & Ehlert, M. (2015). Teacher preparation programs and teacher quality: Are there real differences across programs? *Education Finance and Policy*, 10(4), 508-534.
- Koretz, D. (2008). *Measuring up: What educational testing really tells us*. Cambridge, MA: Harvard University Press.

- Kraft, M. (2019). Teacher effects on complex cognitive skills and social-emotional competencies. *Journal of Human Resources*, 54(1), 1-36.
- Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, 88(4), 547-588.
- Kraft, M. A., Papay, J. P., & Chi, O. L. (2020). Teacher skill development: Evidence from performance ratings by principals. *Journal of Policy Analysis and Management*, 39(2), 315-347.
- Lombardi, M. (2019). Is the remedy worse than the disease? The impact of teacher remediation on teacher and student performance in Chile. *Economics of Education Review*, 73, 101928.
- Loyalka, P., Popova, A., Li, G., & Shi, Z. (2019). Does teacher training actually work? Evidence from a large-scale randomized evaluation of a national teacher training program. *American Economic Journal: Applied Economics*, 11(3), 128-54.
- Murphy, R., Weinhardt, F., & Wyness, G. (2021). Who teaches the teachers? A RCT of peer-to-peer observation and feedback in 181 schools. *Economics of Education Review*, 82, 102091.
- Ost, B. (2014). How do teachers improve? The relative importance of specific and general human capital. *American Economic Journal: Applied Economics*, 6, 127–51.
- Papay, J. P., Taylor, E. S., Tyler, J. H., & Laski, M. E. (2020). Learning job skills from colleagues at work: Evidence from a field experiment using teacher performance data. *American Economic Journal: Economic Policy*, 12(1), 359-88.
- Piaget, J. (1952). *The origins of intelligence in children*. W W Norton & Co. doi:<https://doi.org/10.1037/11494-000>
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment Scoring System (CLASS): Dimensions overview*. Baltimore, MD: Paul H Brookes Publishing.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2), 417-458.
- Rogers, C. R. (1975). *Freedom to learn: A view of what education might become*. CE Merrill.

- Rothstein, J. (2017). Measuring the impacts of teachers: Comment. *American Economic Review*, 107(6), 1656-84.
- Siemens, G. (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning*, 2(1).
- Skinner, B. F. (1953). *Science and human behavior*. Macmillan.
- Taylor, E. S. (2023). Teacher Evaluation and Training. (E. A. Hanushek, S. Machin, & L. Woessmann, Edits.) *Handbook of the Economics of Education*, 7, 61-141.
- Vance Hall, R., Panyan, M., Rabon, D., & Broden, M. (1968). Instructing beginning teachers in reinforcement procedures which improve classroom control. *Journal of Applied Behavior Analysis*, 1(4), 315-322.
- Vygotsky, L. S., & Cole, M. (1978). *Mind in society: Development of higher psychological processes*. Harvard University Press.
- Walkington, C., Arora, P., Ihorn, S., Gordon, J., Walker, M., Abraham, L., & Marder, M. (2012). *Development of the UTeach observation protocol: A classroom observation instrument to evaluate mathematics and science teachers from the UTeach preparation program*. Austin, TX: University of Texas, UTeach Natural Sciences.
- Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, 73(1), 89-122.
- Yoon, K. S., Duncan, T., Lee, S. W., Scarloss, B., & Shapley, K. L. (2007). *Reviewing the evidence on how teacher professional development affects student achievement*. REL 2007-33. U.S. Department of Education.

ANEXOS

Anexo 1: Artículos académicos sobre programas de capacitación docente

		Rechazados	Aceptados
Criterios de búsqueda			
<i>Portal bibliográfico</i>	Science Direct, Sage Journals		
<i>Fecha de búsqueda</i>	22 de febrero de 2024		
<i>Términos</i>	("teacher training") AND ("teacher effectiveness" OR "teacher quality" OR "student performance" OR "student test score" OR "student achievement") AND ("pre-service" OR "in-service" OR "formal training" OR "informal training")		3851
Selección de artículos			
<i>Áreas temáticas</i>	Ciencias sociales, psicología, educación	1202	2649
<i>Calidad</i>	Revistas de los cuartiles Q1 y Q2, según Scimago Journal Rank (SJR) 2022	144	2505
<i>Relevancia</i>	Términos relevantes a la investigación ("teacher" AND "training") incluidos en "Abstract, Título o Palabras clave"	931	1574
		Aceptados	Porcentaje
<i>Tipo de capacitación</i>	In-service		
	<i>Desarrollo profesional</i> ("professional development" OR "PD")	464	69.88%
	<i>Coaching</i> ("coaching")	69	10.39%
	<i>Mentorías</i> ("mentoring")	131	19.73%
	Pre-service		
	<i>Educación superior</i> ("undergrad" OR "master" OR "phd" OR "doctorate")	146	92.99%
	<i>Asistentes de enseñanza</i> ("teaching assistant" OR "TA")	11	7.01%
	Informal		
	<i>Aprendizaje por experiencia</i> ("learning-by-doing" OR "experience")	484	77.69%
	<i>Aprendizaje por pares</i> ("peer learning" OR "peers")	139	22.31%
<i>Variable de interés</i>	Aprendizajes de los estudiantes		
	("student achievement" OR "scores" OR "grades")	641	65.74%
	Prácticas pedagógicas		
	("classroom observation" OR "control" OR "teaching skills")	334	34.26%
<i>Otras especificaciones</i>	Grado de instrucción		
	Inicial ("kindergarten")	145	25.09%
	Primaria ("elementary school" OR "middle school")	170	29.41%
	Secundaria ("high school")	138	23.88%
	Universitaria ("college" OR "undergrad")	125	21.63%
	Materia dictada		
	Matemáticas ("math")	116	17.82%
	Comprensión lectora ("reading")	229	35.18%
	Habilidades de redacción ("writing")	75	11.52%
	Materias de ciencias ("science")	231	35.48%

Nota: Elaboración propia.

Anexo 2: Instrumentos de observación utilizados en el proyecto MET

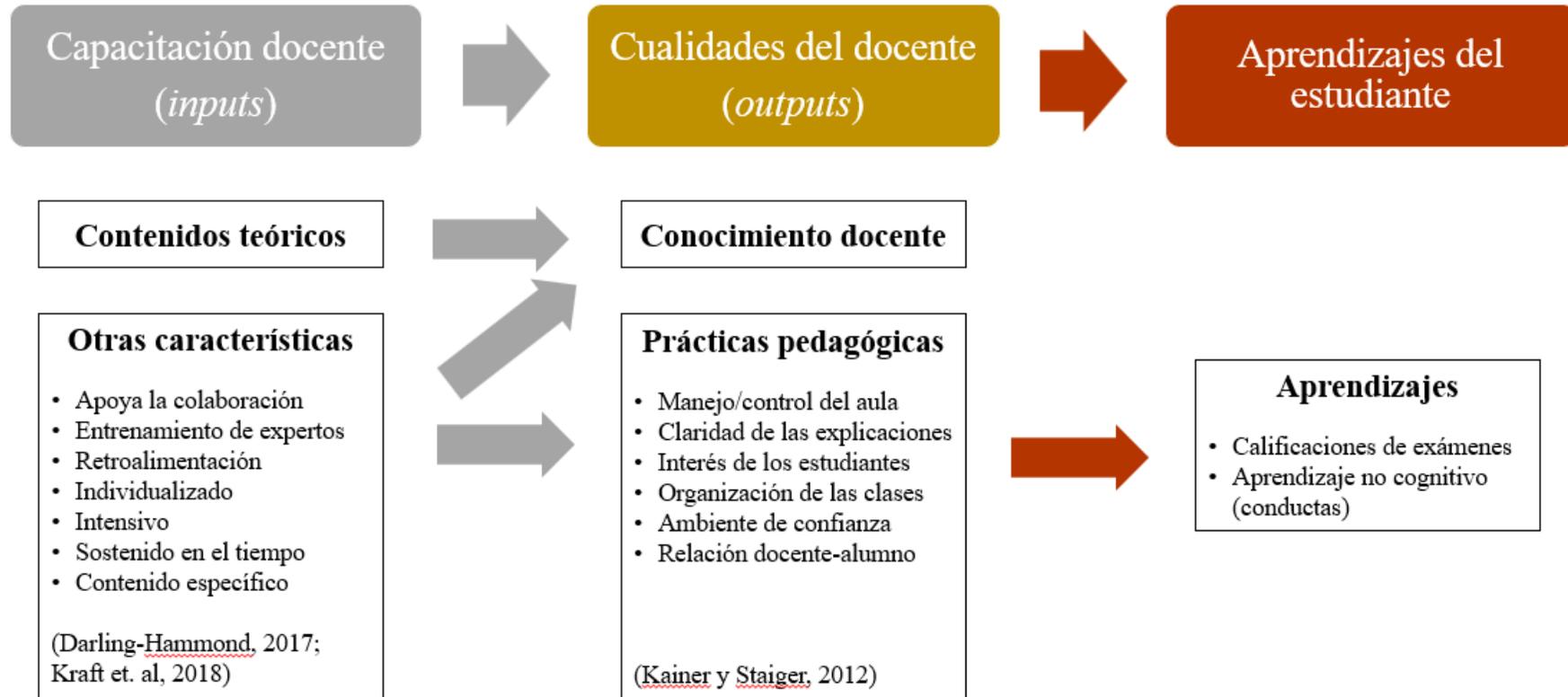
Rúbrica	Descripción	Grado¹	Materia²	Contenidos evaluados sobre el docente
<i>Framework for Teaching</i> (FFT; Danielson, 1996)	Basado en el aprendizaje y el compromiso intelectual del estudiante. Enfoque: Conductismo, constructivismo	K, ES, MS, HS	MD	I. Ambiente de clase: a) ambiente de respeto y simpatía; b) cultura del aprendizaje; c) gestión del aula; d) conducta del alumnado. II. Instrucción: a) comunicación con el estudiante; b) uso de preguntas y técnicas de discusión; c) involucrar a los estudiantes.
<i>Classroom Assesment Scoring System</i> (CLASS; Pianta, LaParo, & Hamre, 2008)	Evalúa la calidad y eficacia de las interacciones entre profesores y alumnos. Enfoque: Conductismo, humanismo	K, ES, MS, HS	MD	I. Soporte emocional: a) ambiente dentro del aula; b) sensibilidad del docente; c) interés por los aportes del estudiante. II. Organización: a) manejo conductual III. Soporte del docente: a) calidad del <i>feedback</i> ; b) entendimiento del currículo; c) resolución de problemas; d) dialogo que facilite el aprendizaje. IV. Compromiso del estudiante
<i>Protocol for Language Arts Teaching Observations</i> (PLATO; Grossman, 2009)	Enfatiza la importancia del discurso, <i>feedback</i> y el rol del docente como modelo. Enfoque: Conductismo, constructivismo, humanismo	ES	CL, R	I. Manejo conductual II. Manejo del tiempo III. Aplicación de estrategias pedagógicas IV. Discusiones en clase V. Motivación intelectual

<p><i>Mathematical Quality of Instruction</i> (MQI; Hill et al., 2008)</p>	<p>Basado en el conocimiento intelectual, precisión y exactitud del docente. Enfoque: Conectivismo</p>	<p>K, ES</p>	<p>M</p>	<p>I. Conocimientos de matemáticas II. Errores e imprecisiones III. Participación de los estudiantes IV. Precisión y exactitud en las explicaciones</p>
<p><i>U Teach Teacher Observation Protocol</i> (UTOP; Walkington et al. 2012)</p>	<p>Enfatiza la instrucción basada en el planteamiento de preguntas, cultura del aula y diseño instruccional. Enfoque: Conductismo, constructivismo, conectivismo</p>	<p>K, ES, MS, HS, CG</p>	<p>M, C, CC</p>	<p>I. Ambiente de clase: a) generación de ideas y preguntas por parte de los estudiantes; b) accesibilidad, igualdad y diversidad II. Estructura de la clase: a) organización y uso de recursos apropiados; b) motivación de los estudiantes; c) enfoque de investigación y resolución de problemas. III. Contenido matemático: a) relevancia y precisión; b) conocimiento del docente; c) interconexión entre diferentes disciplinas IV. Implementación: a) preguntas a los estudiantes; b) manejo del tiempo; c) involucramiento de los estudiantes</p>

1. K: Kindergarten (inicial); ES: Elementary School (primaria); MS: Middle School (primaria); HS: High School (secundaria); CG: College (pregrado). 2. MD (multidisciplinario), CL (comprensión lectora), R (redacción), M (matemáticas), C (ciencias), CC (ciencias de la computación).

Fuente: Kane & Staiger (2012)

Anexo 3: Representación visual del marco teórico



Nota: Elaboración propia

Anexo 4: Cuestionario del Proyecto Trípode para mejorar la enseñanza

CARE
<ul style="list-style-type: none">▪ Me gusta cómo me trata mi maestro/a cuando necesito ayuda.▪ Mi maestro es amable conmigo cuando hago preguntas.▪ Mi maestro en esta clase me hace sentir que realmente le importo.▪ Si estoy triste o enojado, mi maestro/a me ayuda a sentirme mejor.▪ El maestro en esta clase me anima a dar lo mejor de mí.▪ Mi maestro parece saber si algo me molesta.▪ Mi maestro nos da tiempo para explicar nuestras ideas.
CONTROL
<ul style="list-style-type: none">▪ Mis compañeros se comportan como mi maestro quiere.▪ Nuestra clase se mantiene ocupada y no pierde el tiempo.▪ Los estudiantes se comportan tan mal que ralentiza nuestro aprendizaje.▪ Todos saben lo que deben hacer y aprender en esta clase.
CLARITY
<ul style="list-style-type: none">▪ Mi maestro explica las cosas de manera muy ordenada.▪ En esta clase, aprendemos a corregir nuestros errores.▪ Mi maestro explica claramente las cosas difíciles.▪ Mi maestro tiene varias buenas maneras de explicar cada tema▪ Entiendo lo que se supone que debo aprender en esta clase.▪ Mi maestro sabe cuándo la clase entiende y cuándo no.▪ Si no entiendes algo, mi maestro lo explica de otra manera.
CHALLENGE
<ul style="list-style-type: none">▪ Mi maestro nos incita a pensar mucho sobre las cosas que leemos.▪ Mi maestro empuja a todos a trabajar duro.▪ En esta clase tenemos que pensar mucho sobre la escritura que hacemos.▪ En esta clase, mi maestro no acepta menos que nuestro máximo esfuerzo.
CAPTIVATE
<ul style="list-style-type: none">▪ El trabajo escolar es interesante.▪ Tenemos tareas interesantes.▪ Las tareas me ayudan a aprender.▪ El trabajo escolar no es muy divertido. (¿Estás de acuerdo?)
CONFER
<ul style="list-style-type: none">▪ Cuando él/ella nos está enseñando, mi maestro nos pregunta si entendemos.▪ Mi maestro hace preguntas para asegurarse que lo seguimos mientras enseña.▪ Mi maestro nos dice qué estamos aprendiendo y por qué.▪ Mi maestro quiere que compartamos nuestros pensamientos.▪ Los estudiantes hablan y comparten sus ideas sobre el trabajo en clase.▪ Mi maestro pide que explique mis respuestas—por qué pienso lo que pienso.
CONSOLIDATE
<ul style="list-style-type: none">▪ Mi maestro se toma el tiempo para resumir lo que aprendemos cada día.▪ Cuando mi maestro califica, escribe en mis papeles para ayudarme a entender.

Nota: Kane y Staiger (2010)

Anexo 5: Resultados de estudios revisados sobre los efectos de la capacitación docente en las prácticas pedagógicas y/o aprendizajes

Panel A: Capacitación pre-service

Estudio	País	Método	Programa¹	Métrica²	Grado³	Materia⁴	Resultados⁵
Aaronson et. al (2007)	EE. UU.	QEX-II	Ugrd, Ms, PhD	TS	HS	M	(0), (-*), (0)
Clotfelter et. al (2010)	EE. UU.	QEX-I	AD	TS	HS	MD	(0)
Constantine et. al (2009)	EE. UU.	RCT	AC	TS CO	EL	M/CL	M (0) / CL (0) (+*)
Harris y Sass (2011)	EE. UU.	QEX-I	AD	TS	EL/MI	M/CL	M-EL (0) / M-MI (+*) / CL-EL (0) / CL-MI (-*)
Kane et. al (2008)	EE. UU.	QEX-II	CF, TA, UC	TS	EL/HS	M/CL	(0), (0), (0)

Panel B: Capacitación in-service

Estudio	País	Método	Programa¹	Métrica²	Grado³	Materia⁴	Resultados⁵
Abeberese et. al (2014)	Filipinas	RCT	PD	TS	EL	M/CL	(0) / (-*)
Angrist y Lavy (2001)	Israel	QEX-I (<i>Diff-in-diff</i>)	PD	TS CO	K-12	M/CL	(+*) (+*)
Cilliers et. al (2020)	Sudáfrica	RCT	PD, CH	TS NC	EL	CL	PD (+*), CH (+*) CH (+*)
Clotfelter et. al (2007)	EE. UU.	QEX-I	Ms, AD, PhD	TS	EL	M/CL	(-), (0), (-*)
Clotfelter et. al (2010)	EE. UU.	QEX-I	PD (contenido específico)	TS	HS	MD	(+*)
Jacob y Lefgren (2004)	EE. UU.	QEX-I (<i>Reg. disc</i>)	PD	TS	EL	M/CL	(0) / (0)
Kerwin y Thorton (2021)	Uganda	RCT	PD " <i>Premium</i> " PD " <i>Low-cost</i> "	TS	EL	CL/R	(+*) / (+*) (0) / (-)
Kraft et. al (2018)	CA, Chile, EE. UU.	RCT	CH	TS, NC, CO	K-12	MD	(+*), (+*), (+*)

Lombardi (2019)	Chile	QEX-I (Reg. disc)	PD (Remedial)	TS, CO	EL/MI/HS	MD	Corto plazo (+*), (+*) / Largo plazo (0), (0)
Loyalka et. al (2019)	China	RCT	PD	TS, NC, CO	K-12	M	(0), (0), (0)
Yoon et. al (2007)	AU, CA, UK, US	RCT	PD	TS	K-12	M/CL	(+*)

Panel C: Capacitación informal

Estudio	País	Método	Programa ¹	Métrica ²	Grado ³	Materia ⁴	Resultados ⁵
Aaronson et. al (2007)	EE. UU.	QEX-II	Experiencia	TS	HS	M	(0)
Boyd et. al, 2009).	EE. UU.	QEX-I	TA	TS	K-12	M/CL	(+*)
Clotfelter et. al (2007)	EE. UU.	QEX-I	Experiencia	TS	EL	M/CL	(+*)
Clotfelter et. al (2010)	EE. UU.	QEX-I	Experiencia	TS	HS	MD	(+*)
Goldhaber et. al, 2020	EE. UU.	QEX-I	TA	TS	EL	M/CL	ST (0), LT (+*)
Harris y Sass (2011)	EE. UU.	QEX-I	Experiencia	TS	EL/HS	M/CL	M-HS (0) / CL-HS (-*) / Otros (+*)
Rockoff (2004)	EE. UU.	QEX-I	Experiencia	TS	EL	M/CL	M (+*) / CL (-*)

1. Ugrd (pregrado), Ms (maestría), PhD (doctorado), AD (estudios de posgrado), CF (certificación), AC (certificación alternativa), TA (asistente de enseñanza), UC (no certificado), PD (desarrollo profesional), CH (coaching). 2. TS (calificaciones estudiantiles), NC (métricas de aprendizaje no cognitivo), CO (observaciones de clase). 3. EL: *Elementary school* (primaria), MI: *Middle school* (primaria), HS: *High school* (secundaria), K-12 (K, EL, MI, HS). 4. M (matemáticas), CL (comprensión lectora), R (redacción), MD (multidisciplinario). 5. Los resultados reflejan efectos diferenciados según programa, materia y grado. Por ejemplo: PD-M-HS muestra el efecto de un programa de desarrollo profesional en las calificaciones de matemáticas de nivel secundaria. El efecto se encuentra entre paréntesis (0 = Nulo, "+" = Positivo, "-" = Negativo). El "*" indica si el resultado es significativo o no.

Nota: Elaboración propia.