

EVALUACIÓN DE LA FORMACIÓN INICIAL DOCENTE MEDIANTE ESTIMACIONES DE VALOR- AGREGADO

VERÓNICA SANTELICES, PHD¹
EDGAR VALENCIA, PHD²


¹ FACULTAD DE EDUCACIÓN,
P . UNIVERSIDAD CATÓLICA DE CHILE

² CENTRO DE MEDICIÓN MIDE UC, ESCUELA DE PSICOLOGÍA,
P . UNIVERSIDAD CATÓLICA DE CHILE



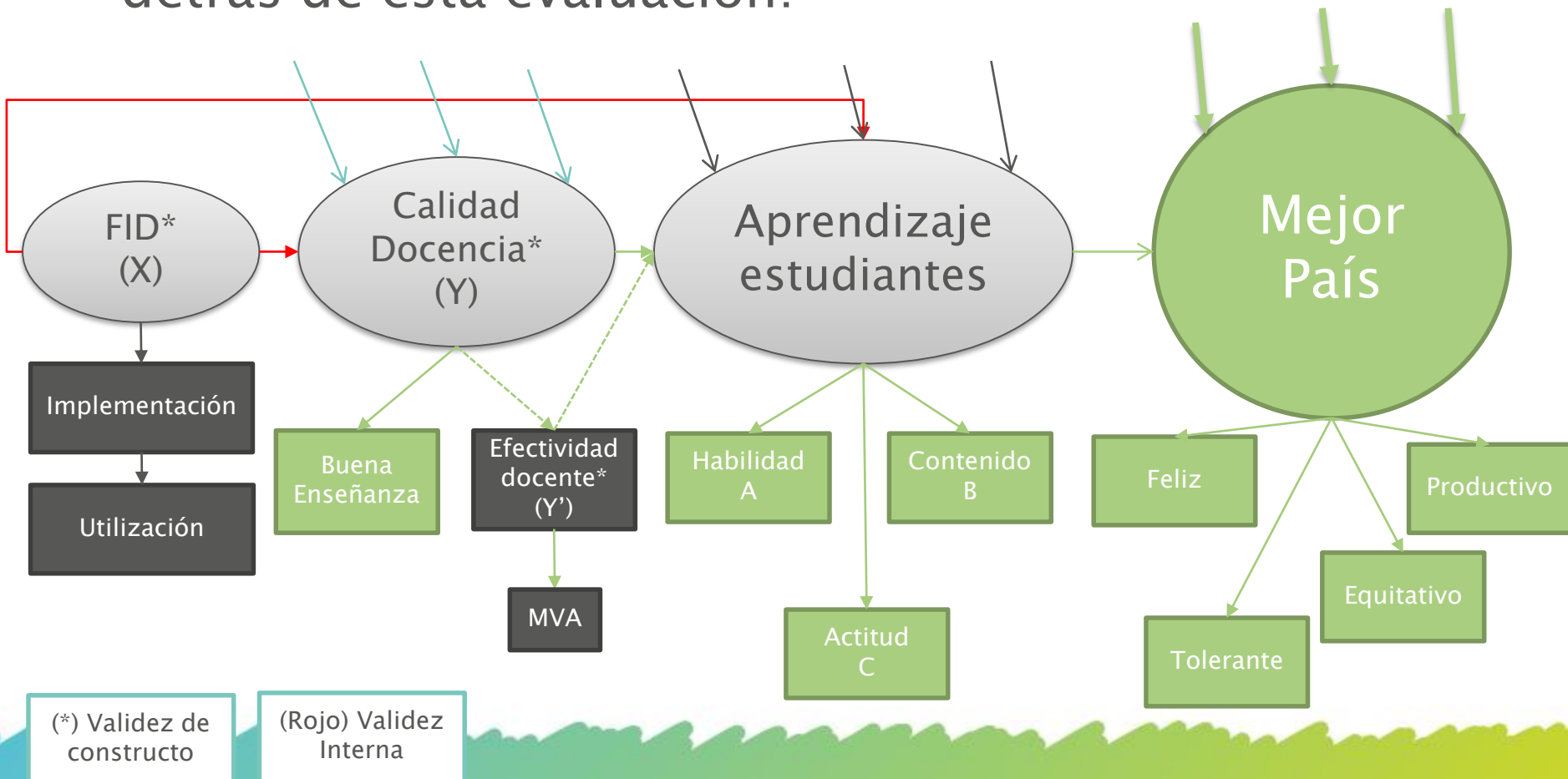
Centro UC
Medición - MIDE

¿CÓMO EVALUAR UN PROGRAMA FID?

- Indiscutido interés por “incrementar” la “calidad” de la docencia, énfasis razonable en la formación inicial.
 - Los programas de FID “deben” y “están interesados en” evaluarse.
 - Esta presentación es sobre las múltiples complejidades de evaluar el impacto de la FID en la efectividad docente.
- 


¿CÓMO EVALUAR UN PROGRAMA FID?

Diagrama simplificado de los principales supuestos detrás de esta evaluación:



CÓMO EVALUAR UN PROGRAMA FID?


Tres tipos de desafíos (Harris y Sass, 2010):

- Limitaciones prácticas:
 - Carencia de información sobre características de la formación inicial que reciben los docentes.
 - Dificultad para vincular la limitada información existente con resultados de los estudiantes en pruebas estandarizadas.
 - Validez del constructo del outcome (efectividad docente)
 - Inestabilidad de estimaciones efectividad docente.
 - Dificultad de aislar estimaciones de efectividad docente de otros factores.
 - Validez Interna (Atribución Causal)
 - Ausencia de contrafactual: Solo Post Test Grupo Único, sin asignación aleatoria.
 - Existen potenciales hipótesis alternativas a la acción del programa.
- 

¿CÓMO EVALUAR UN PROGRAMA FID?

- Limitaciones prácticas: Identificamos, localizamos, reunimos, integramos y analizamos diversas fuentes de información para conducir una evaluación de impacto de un programa de FID utilizando un diseño cuasi-experimental.
- Validez de constructo del outcome: Examinamos la estabilidad de distintas estimaciones de MVA.
- Atribución causal: complementamos la información del diseño cuasi-experimental con otras fuentes de evidencia disponible y argumentación (Eckert, 2000; Mayne, 2001; Yin, 2000).

¿CÓMO EVALUAR UN PROGRAMA FID?

1. Diseño Cuasi-Experimental Solo Post-Test con Grupo de Comparación (resultado de resolver limitaciones prácticas):
 - Los grupos son **egresados y no egresados** del PFID (X').
 - Obtuvimos una medida de **efectividad docente (Y')** utilizando MVA y resultados de estudiantes en pruebas estandarizadas.
 - Comparamos la efectividad de docentes de egresados y no egresados para obtener una medida de **impacto (efecto) del programa**.
 2. Estabilidad de las medidas de efectividad docente (validez de constructo del outcome):
 - En función de la especificación (matemática) del MVA.
 - En función de diferencias en el efecto del programa por contenido enseñado (profesores de matemática y lenguaje) y cohorte de estudiantes (dos cohortes).
 3. Fortalecer Diseño de Grupo Único con Grupo Comparación sin asignación aleatoria (atribución causal) usando situational enhancement of validity:
 1. Rule-in the construct
 2. Rule-out plausible rival explanations
- 

¿CÓMO EVALUAR UN PROGRAMA FID?

- (1) Solo Post Grupo Comparación No Aleatorio y (2) Estabilidad

Cohorte / Contenido	No Egresados (control)			Egresados (programa)		
	N	Media	SD	N	Media	SD
Lenguaje 8° 2011	1183	.03	11.14	19	-1.96	9.08
Matemática 8° 2011	1286	-.08	10.81	9	11.42	11.85
Lenguaje 8° 2013	1033	-.04	9.96	7	6.4	9.60
Matemática 8° 2013	1113	.02	11.14	13	1.82	13.57

Nuestra estimación de efectividad docente (MVA) es relativamente estable frente a diferencias en la especificación de los modelos.

Existen grandes variaciones en el efecto del programa dependiendo del contenido y cohorte, el *tamaño del efecto* del programa varía entre no efecto y un efecto largo (n.s.).

¿CÓMO EVALUAR UN PROGRAMA FID?

- (3) Atribución Causal
- Validez de constructo del programa:
 - Nuestro análisis de documentos internos y externos, sugieren que aspectos como diseño del programa, currículum, planta docente, infraestructura, elementos de soporte están debidamente implementados (implementación).
 - La utilización del programa (horas de dedicación a cursos y la secuencia de cursos reportado por los egresados cuando eran estudiantes) varía significativamente entre egresados y se desvía con respecto a lo intencionado por los planificadores, debilitando la validez de constructo del programa. Sin embargo, creemos que esto no necesariamente afecta la relación entre el programa de FID y efectividad de la docencia sino más bien implica la redefinición del programa.

¿CÓMO EVALUAR UN PROGRAMA FID?

- (3) Atribución Causal
- Plausibilidad de explicaciones alternativas (Yin, 2000; Eckart, 2002; Shadish, Cook & Campbell, 2002):
 - *Selección*: el programa de FID selecciona candidatos debido a su alto rendimiento académico (prueba de selección universitaria) y socioeconómico (capacidad de pago), por lo que dichas características de los participantes podrían ser causa del efecto observado en efectividad de la docencia en lugar del programa. Esta es la amenaza más seria que podría llevar a una incorrecta inferencia causal entre FID y efectividad de la docencia.
 - *Mortalidad-Deserción*: en nuestro caso, la deserción de los participantes durante los años de duración del programa es un problema que puede estar relacionado con el programa y causar una incorrecta inferencia sobre el efecto de FID sobre la efectividad de la docencia. Queda por determinar la medida en que dicha mortalidad obedece a razones espontáneas o aleatorias u obedece a razones que están conectadas el programa.
 - *Efecto aditivo e interacción*: La suma de los efectos de selección y deserción y la interacción entre ambas también es una explicación alternativa que no es posible descartar. El efecto aditivo implica que ambas explicaciones alternativas son ciertas, y la interacción podría ocurrir, por ejemplo, cuando los participantes privilegiados con mayor rendimiento académico previo son los que logran terminar y egresar del programa.

¿CÓMO EVALUAR UN PROGRAMA FID?RIO

Presentamos:

1. Desafíos para evaluar FID usando efectividad docente como criterio.
2. Resultados de resolver limitaciones prácticas, en atención a problemas de validez de constructo del outcome y utilizando estrategias para fortalecer la atribución causal.

Algunas reflexiones:

- Los resultados son consistentes con literatura previa que indica que las estimaciones de valor-agregado del profesor pueden variar por cohorte y contenido de las pruebas (ej. Papay, 2010). Evaluaciones de impacto de FID utilizando MVA debieran considerar múltiples cohortes, años, y una variedad de contenidos.
- Las mediciones de logro educativo nacionales no están diseñadas para evaluar progreso de los estudiantes salvo por unas pocas excepciones fortuitas, que fueron aquellas cohortes con pre y posttest utilizadas en esta evaluación.
- Las mediciones de logro educativo nacionales no recogen información relevante sobre las características de los programas de formación inicial docente.
- La información administrativa que permite vincular egresados de programas de formación y docentes en ejercicio es de difícil acceso y están distribuida burocráticamente entre diferentes instituciones.
- La evaluación de la FID debiera estar incorporada en el diseño del programa.

¿CÓMO EVALUAR UN PROGRAMA FID?

- Generamos varias hipótesis respecto al problema de selección y mortalidad:
 - Examinar relación entre utilización del programa (X') y efectividad docente (Y') (variación conjunta entre causa y efecto).
 - Junto a examinar amenazas a validez interna (explicaciones alternativas), y validez de constructo del programa (implementación, utilización), examinar amenazas a la validez de conclusión estadística.
 - Incluir otras medidas de calidad de la docencia, por ejemplo, observaciones de aula, y la percepción de los egresados (¿cuánto contribuyó el programa a sus habilidades como docente en esta escuela?
- 