

Principales desafíos en la validación de mediciones de aprendizajes en América Latina.

Gilbert A. Valverde, Ph.D.

University at Albany, State University of New York
gvalverde@albany.edu

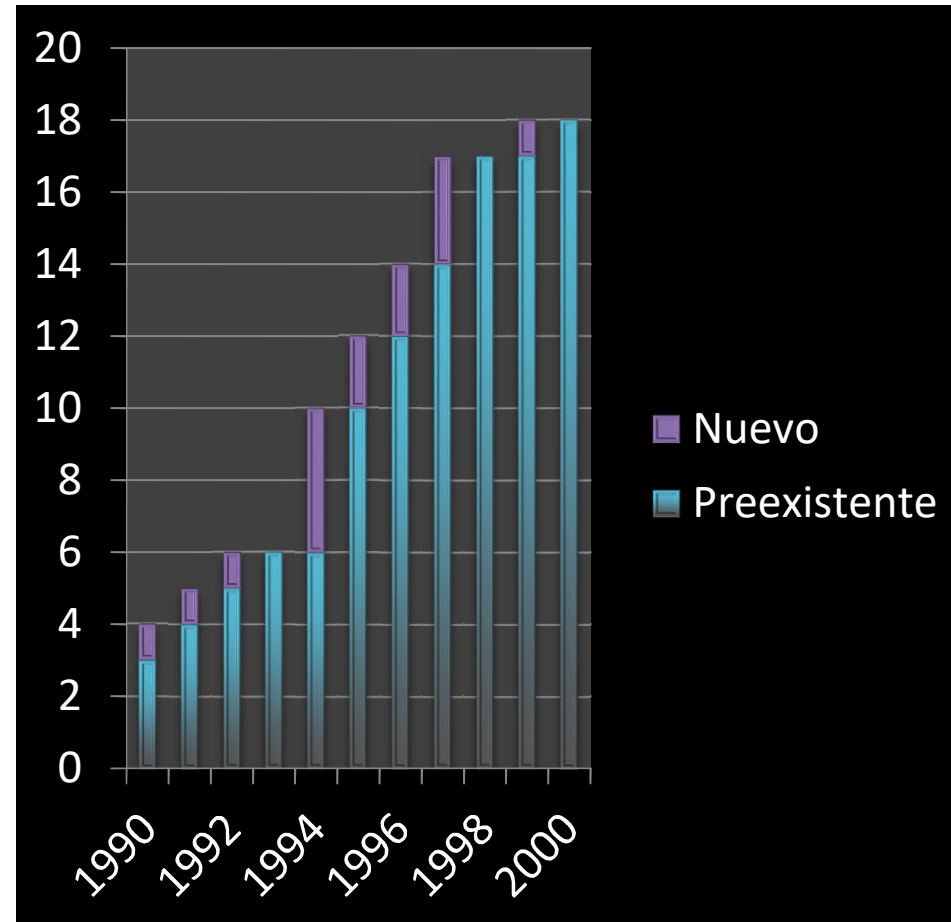
Presentación Inicial
40º Reunión de Coordinadores Nacionales del LLECE
Lima, Perú
24 Setiembre 2018



Institute
for Global Education
Policy Studies

UNIVERSITY AT ALBANY State University of New York

Aumento notable de sistemas de evaluación en América Latina en los años 90.





- Los programas nacionales de evaluación de aprendizajes se han multiplicado con fines de
 - monitorear el logro de los aprendizajes curriculares
 - fomentar mejoras en dichos aprendizajes.
- Si bien la mayoría de los países introdujo evaluaciones nacionales en la década de los 90s, éstas fueron inestables institucionalmente.
- En contextos institucionales precarios, todos los recursos suelen estar puestos en implementar las evaluaciones.

- El grado actual de madurez que están alcanzado las evaluaciones en la región, *y su influencia en la esfera pública*, hacen imperativo priorizar una agenda de validación.
- Esto es, una agenda que asegure la recolección de *evidencia que sustente las interpretaciones, usos, y consecuencias asociadas a las evaluaciones*.
- La **validación** de instrumentos de medición consiste precisamente en la recolección y análisis de esta evidencia.
- Hay una variedad de tipos de validación, pero dado que el enfoque predominante en AL es el de utilizar las evaluaciones con el propósito de monitorear o verificar el cumplimiento de los logros de aprendizaje establecidos en las políticas curriculares, consideramos prioritaria la **validación curricular**.



• La evidencia de **validez curricular** contesta preguntas fundamentales:

• *¿En qué medida las evaluaciones permiten efectivamente inferir el logro de los aprendizajes propuestos en las políticas curriculares?*

• *¿En qué medida sus resultados pueden ser interpretados como reflejo de dichos aprendizajes?*

• *¿En qué medida pueden ser utilizados efectivamente para la mejora de los aprendizajes?*



Institute
for Global Education
Policy Studies

UNIVERSITY AT ALBANY State University of New York

- No tener evidencia de validez equivale a no tener fundamentos para juzgar la calidad de la información utilizada para tomar decisiones.
- Es no saber si la información tiene suficiente calidad para sostener las inferencias que hacemos en base a ella
- *Equivale a tomar decisiones sin información alguna.*

- Habitualmente los currículum oficiales no se escriben con el fin de que sean medibles.
- Es poco frecuente que los equipos responsables del currículum cuenten con integrantes especialistas en medición.
- A menudo, los equipos responsables del currículum son suspicaces con respecto al proyecto de operacionalizar el currículum en términos medibles en evaluaciones estandarizadas.
- No es común que los equipos de currículum y evaluación tengan puntos institucionalizados de contacto y/o de trabajo en común.
- *Es frecuente en América Latina que los equipos que deben diseñar evaluaciones tengan la tarea de traducir en términos medibles, un currículum que no se diseñó expresamente para ser medido.*



Estudios realizados en 2016-2017 por G. Valverde y M.J. Ramírez

Estándares de validación curricular de
"libro de texto"

Documentación de procedimientos de
validación en 50 países participantes en
PISA y 12 casos en profundidad
incluyendo entrevistas

*Australia, Canadá (Provincia de Ontario),
España, Inglaterra, Escocia, Holanda,
Nueva Zelanda, EEUU (NAEP), EEUU
(Estados de New York y Virginia)*

Notas e informes de consultorías
realizadas en América Latina 1994 - 2017
(R. Wolfe, G. Valverde, M. J. Ramírez)

10 criterios dispuestos en 3 dimensiones de validación

Estudio comparado con ejemplos en toda América Latina y en 3 casos
en profundidad : Chile, México y Perú

10 Criterios de Validez Prioritarios

Alineamiento de las pruebas con el currículum oficial

- 1 Se justifica el diseño de la evaluación en función del currículum.
- 2 Se operacionaliza el dominio de evaluación tomando en consideración los aprendizajes reales de los estudiantes.
- 3 Monitoreo con precisión y sin sesgos del logro de los aprendizajes curriculares en el tiempo.

Validez de los niveles de desempeño

- 4 Se elaboran niveles de desempeño alineados con el currículum.
- 5 Se elaboran niveles de desempeño tomando en consideración los aprendizajes reales de los estudiantes.
- 6 Los niveles de desempeño dan cuenta de etapas cualitativamente distintas del aprendizaje.
- 7 Los niveles de desempeño equilibran estabilidad y cambio, en el contexto de una política curricular dinámica.

Validez consecucional de las evaluaciones

- 8 Los resultados de las evaluaciones se comunican en forma efectiva.
- 9 Hay mecanismos formales que favorecen el uso de las evaluaciones para la mejora de los aprendizajes.
- 10 Hay mecanismos formales para monitorear las consecuencias de las evaluaciones en el sistema educativo

	Criterio	Explicacion	Evidencia para la validacion
Dimensión de alineamiento de las pruebas con el currículum oficial			
1	Se justifica el diseño de la evaluación en función del currículum.	Se establece y justifica para qué, a quién, qué, cómo, y cuando se evalúa, haciendo referencia explícita al currículum. Se identifican las interpretaciones posibles y usos esperados de los resultados de las evaluaciones. Se especifica el dominio de evaluación de las pruebas (ej., matemáticas, lenguaje) en función las competencias, objetivos, contenidos o habilidades del currículum, y de los niveles de desempeño.	<p>* PRODUCTOS: (a) Diseño general de la evaluación [el para qué, a quién, qué, cómo y cuando se evalúa] en función del currículum y (b) Especificaciones de las pruebas; (a) Versión final de las pruebas.</p> <p>* PROCESOS: Métodos y procedimientos para (a) Determinar el diseño general de la evaluación, (b) Hacer las especificaciones de las pruebas, y (c) Llegar a la versión final de las pruebas.</p>
2	Se operacionaliza el dominio de evaluación tomando en consideración los aprendizajes reales de los estudiantes.	Se reconoce la necesidad de que la evaluación se ajuste a lo que saben y pueden hacer todos los estudiantes, desde los más avanzados hasta los más rezagados. El dominio de evaluación y las especificaciones de las pruebas cubren un espectro amplio de aprendizajes. Los resultados de la evaluación permiten dar cuenta de lo que saben y pueden hacer todos los estudiantes, incluso los más rezagados.	<p>* PRODUCTOS: (a) Versión final de las pruebas; (b) Dominio de evaluación y especificaciones de las pruebas; (c) Resultados de las pruebas; (d) Información sobre los aprendizajes reales de los estudiantes [ej., resultados de evaluación de aula, o de evaluaciones internacionales].</p> <p>* PROCESOS: Métodos y procedimientos para (a) Operacionalizar el dominio de evaluación y hacer las especificaciones de las pruebas; (b) Recabar información sobre los aprendizajes reales de los estudiantes [ej., observación de aula, revisión de resultados de evaluaciones de aula y de evaluaciones internacionales].</p>
3	Los resultados de las pruebas permiten monitorear con precisión y sin sesgos el logro de los aprendizajes curriculares en el tiempo.	Los resultados de las pruebas pueden ser interpretados como logro de los objetivos curriculares. Los resultados se publican indicando el margen de error o incertidumbre asociado a los puntajes, porcentajes, diferencias, u otros estadísticos. El margen de error está dentro de lo aceptable, dadas las consecuencias asociadas a los resultados. Los resultados están libres de sesgos (ej. de género, culturales/ lingüísticos/ étnicos). Los resultados de distintas pruebas han sido equiparados y puestos en una misma escala de puntajes.	<p>* PRODUCTOS: Resultados de aprendizaje y errores de medición, resultados del equiparamiento de escala (equating), y de los análisis psicométricos (ej., dificultad, discriminación, comportamiento diferencial de items).</p> <p>* PROCESOS: Métodos y procedimientos para calcular resultados de aprendizaje y errores de medición, para hacer el equiparamiento de escala (equating), y para calcular características psicométricas de las pruebas.</p>

	Criterio	Explicación	Evidencia para la validación
Dimensión de validez curricular de los niveles de desempeño			
4	Se elaboran niveles de desempeño alineados con el currículum.	Se describen y justifican los niveles de desempeño haciendo referencia a distintos aspectos del currículum (ej. competencias, complejidad, contextos). Los niveles de desempeño constituyen hitos de aprendizaje necesarios para lograr los objetivos curriculares. Hay un nivel de desempeño que corresponde al logro de los objetivos curriculares.	<ul style="list-style-type: none"> * PRODUCTOS: Versión final de los niveles de desempeño. * PROCESOS: Métodos y procedimientos para desarrollar niveles de desempeño alineados con el currículum.
5	Se elaboran niveles de desempeño tomando en consideración los aprendizajes reales de los estudiantes.	Los niveles de desempeño se diseñan para asegurar que la evaluación pueda describir lo que todos los estudiantes saben y pueden hacer. Cubren un espectro amplio de aprendizajes, desde los más avanzados hasta los más básicos. Hay un nivel de desempeño que describe lo que saben y pueden hacer los estudiantes con mayor retraso en sus aprendizajes.	<ul style="list-style-type: none"> * PRODUCTOS: (a) Resultados por niveles de desempeño; (b) Versión final de los niveles de desempeño. * PROCESOS: Métodos y procedimientos para (a) Recabar información sobre los aprendizajes reales de los estudiantes (ej., observación de aula, evaluaciones de aula y evaluaciones internacionales), y (b) Utilizar dicha información en la elaboración de los niveles de desempeño.
6	Los niveles de desempeño dan cuenta de etapas cualitativamente distintas del aprendizaje.	Los niveles de desempeño dan cuenta de hitos dentro de una trayectoria de aprendizajes. Lo que saben y pueden hacer los estudiantes en un nivel es claramente distinto de lo que saben y pueden hacer en el siguiente. Los estudiantes son clasificados en los niveles de desempeño en función de si logran o no los aprendizajes asociados a los distintos puntos de corte. Hay un nivel de desempeño que se define por defecto, y que corresponde a los estudiantes que no alcanzan el punto de corte más bajo en la escala de puntajes.	<ul style="list-style-type: none"> * PRODUCTOS: (a) Versión final de los niveles de desempeño, con puntos de corte, descripciones de lo que saben y pueden hacer los estudiantes en cada nivel, y nombre de cada nivel; (b) Resultados por niveles de desempeño. * PROCESOS: Métodos y procedimientos para (a) Elaborar los niveles de desempeño, (b) Clasificar a los estudiantes en los distintos niveles de desempeño.
7	Los niveles de desempeño equilibran estabilidad y cambio, en el contexto de una política curricular dinámica.	Se utilizan dos versiones de niveles de desempeño: <ul style="list-style-type: none"> 1. Niveles de desempeño "actualizados" para monitorear cambios en el aprendizaje en el corto plazo (ej. de un año a otro). Son revisados y actualizados de acuerdo a las reformas y ajustes curriculares. 2. Niveles de desempeño "invariantes" para monitorear cambios en el aprendizajes en el largo plazo (ej. tendencias de 10 o 20 años). Describen aprendizajes nucleares que suelen ser menos afectados por ajustes curriculares. 	<ul style="list-style-type: none"> * PRODUCTOS: (a) Niveles de desempeño "actualizados"; (b) Resultados de aprendizaje asociados a los Niveles de desempeño "actualizados"; (c) Niveles de desempeño "invariantes"; (d) Resultados de aprendizaje asociados a niveles de desempeño "invariantes". * PROCESOS: Métodos y procedimientos para (a) Revisar los niveles de desempeño "actualizados"; (b) Evaluar a los estudiantes utilizando los niveles de desempeño "actualizados"; (b) Desarrollar niveles de desempeño "invariantes"; (b) Evaluar a los estudiantes utilizando los niveles de desempeño "invariantes".

	Critero	Explicacion	Evidencia para la validacion
Dimensión de validez consecucional de las evaluaciones			
8	Los resultados de las evaluaciones se comunican en forma efectiva.	Los resultados se publican en forma clara y correcta, en distintos medios y formatos (ej., informes, talleres, folletos, videos) adecuados para distintas audiencias, y se difunden en tiempo oportuno.	* PRODUCTOS: (a) Productos de comunicación utilizados para difundir los resultados e información de las evaluaciones; (b) Estrategia de comunicación utilizada (ej. productos comunicacionales, audiencias, tiempos); (c) Niveles de acceso, valoración, comprensión de los resultados por parte de diferentes actores educativos. * PROCESOS: Métodos y procedimientos para diseñar la estrategia de comunicación (ej. consulta a audiencias o actores claves).
9	Hay mecanismos formales que favorecen el uso de las evaluaciones para la mejora de los aprendizajes.	Existen mecanismos formales/ institucionales que fomentan el uso de las evaluaciones para la mejora (ej., planes de mejora para las escuelas, orientaciones para la supervisión de escuelas, orientaciones para la formación docente, recursos pedagógicos coherentes con el currículum y los niveles reales de aprendizaje de los estudiantes).	* PRODUCTOS: (a) Catastro de mecanismos formales que utilizan insumos de las evaluaciones. * PROCESOS: Metodología y procedimientos utilizados para recabar información sobre (a) Usos de los resultados de las evaluaciones por parte de diferentes actores educativos.
10	Hay mecanismos formales para monitorear las consecuencias de las evaluaciones en el sistema educativo	Se hacen regularmente estudios sobre las consecuencias, esperadas y no esperadas, de las evaluaciones en el sistema educativo. Por ejemplo, sobre su impacto en el aprendizaje de los estudiantes y las prácticas pedagógicas de los docentes, sobre el involucramiento de los padres en la educación de sus hijos, sobre su uso para decisiones de política.	* PRODUCTOS: Existencia de estudios sobre las consecuencias de las evaluaciones * PROCESOS: Existencia de mecanismos formales para realizar dichos estudios de validez consecucional: comisiones o auditorias, financiamiento interno o externo para investigaciones.

Fortalezas y debilidades en la evidencia de validez curricular en América Latina en una selección de criterios en cada dimensión.

ALGUNOS EJEMPLOS



Institute
for Global Education
Policy Studies

UNIVERSITY AT ALBANY State University of New York

	Criterio	Explicacion	Evidencia para la validacion
Dimensión de alineamiento de las pruebas con el currículum oficial			
1	Se justifica el diseño de la evaluación en función del currículum.	Se establece y justifica para qué, a quién, qué, cómo, y cuando se evalúa, haciendo referencia explícita al currículum. Se identifican las interpretaciones posibles y usos esperados de los resultados de las evaluaciones. Se especifica el dominio de evaluación de las pruebas (ej., matemáticas, lenguaje) en función las competencias, objetivos, contenidos o habilidades del currículum, y de los niveles de desempeño.	<p>* PRODUCTOS: (a) Diseño general de la evaluación [el para qué, a quién, qué, cómo y cuando se evalúa] en función del currículum y (b) Especificaciones de las pruebas; (a) Versión final de las pruebas.</p> <p>* PROCESOS: Métodos y procedimientos para (a) Determinar el diseño general de la evaluación, (b) Hacer las especificaciones de las pruebas, y (c) Llegar a la versión final de las pruebas.</p>

Chile	Perú	México
<p>La ley establece que la evaluación nacional debe verificar el cumplimiento del currículum usando estándares de desempeño.</p> <p>La justificación del diseño de la evaluación se basa en los resultados de análisis curriculares usados como bases de las especificaciones técnicas de la evaluación. Estas son refinadas a partir de realimentación de equipos de expertos del Ministerio, incluyendo responsables del currículum.</p>	<p>El Currículum Nacional de Educación Básica se estructura en torno a competencias acompañadas de estándares de aprendizaje que definen el nivel que se espera puedan alcanzar todos los estudiantes al finalizar los ciclos de la Educación Básica.</p> <p>La justificación del diseño de la evaluación (en este caso, la Tabla de Especificaciones) se hace en base a los estándares de aprendizaje.</p>	<p>No existe un solo documento que especifica el currículum nacional, los referentes de las evaluaciones son: planes y programas de estudios, libros de texto, fichas de trabajo y otros materiales.</p> <p>Las evidencias de validez curricular son los resultados de los análisis curriculares de este diverso material, usados en la justificación de los dominios medidos en las pruebas.</p>

Una salida al dilema del currículum no diseñado para ser medido

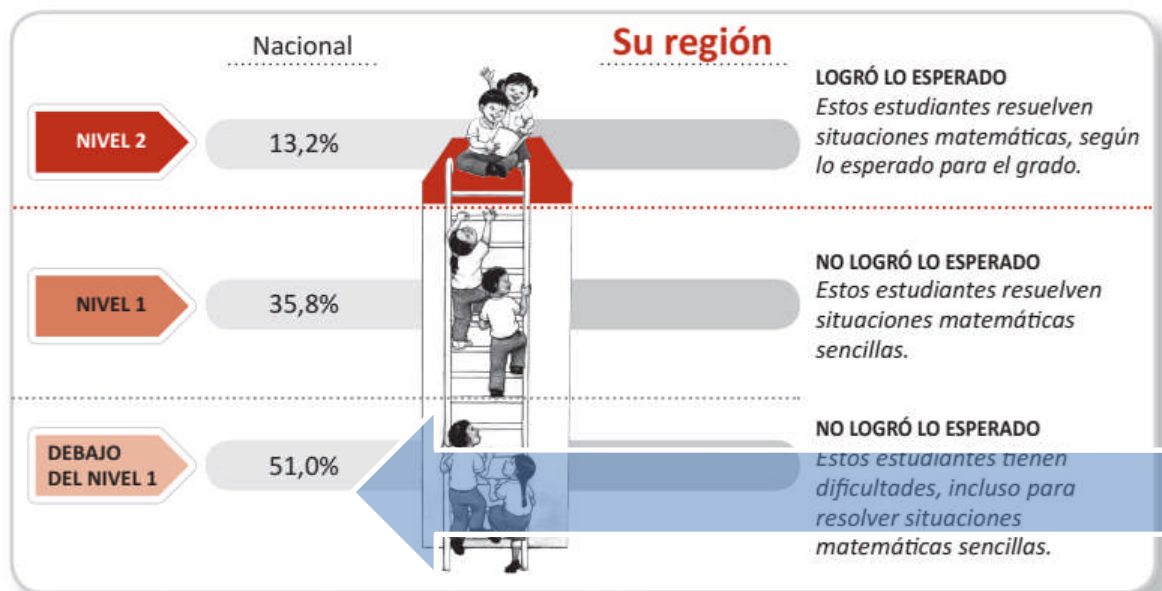
- Podemos considerar a México como el ejemplo más instructivo para la situación más común en AL.
- México es excepcional en no contar con un documento único que encarna el currículum oficial, sino que cuenta con un conjunto de documentos que juntos representan ese currículum.
- Lo que es instructivo, es la manera en la cual el INEE en México resuelve trabajar con ese conjunto de documentos:
 - mediante un procedimiento de riguroso y disciplinado análisis curricular.
- Su metodología es interesante por el potencial que tiene en abordar el dilema que afrontan muchos países en la región: la necesidad de traducir, en términos medibles, un currículum oficial que no fue diseñado apropiadamente para ser medido.
- La documentación de los *procedimientos* es extensiva, lo que no hay aun es evidencia acerca de validez de los *productos* del procedimiento.

Dimensión de validez curricular de los niveles de desempeño

4	Se elaboran niveles de desempeño alineados con el currículum.	Se describen y justifican los niveles de desempeño haciendo referencia a distintos aspectos del currículum (ej. competencias, complejidad, contextos). Los niveles de desempeño constituyen hitos de aprendizaje necesarios para lograr los objetivos curriculares. Hay un nivel de desempeño que corresponde al logro de los objetivos curriculares.	<p>* PRODUCTOS: Versión final de los niveles de desempeño.</p> <p>* PROCESOS: Métodos y procedimientos para desarrollar niveles de desempeño alineados con el currículum.</p>
5	Se elaboran niveles de desempeño tomando en consideración los aprendizajes reales de los estudiantes.	Los niveles de desempeño se diseñan para asegurar que la evaluación pueda describir lo que todos los estudiantes saben y pueden hacer. Cubren un espectro amplio de aprendizajes, desde los más avanzados hasta los más básicos. Hay un nivel de desempeño que describe lo que saben y pueden hacer los estudiantes con mayor retraso en sus aprendizajes.	<p>* PRODUCTOS: (a) Resultados por niveles de desempeño; (b) Versión final de los niveles de desempeño.</p> <p>* PROCESOS: Métodos y procedimientos para (a) Recabar información sobre los aprendizajes reales de los estudiantes (ej., observación de aula, evaluaciones de aula y evaluaciones internacionales), y (b) Utilizar dicha información en la elaboración de los niveles de desempeño.</p>

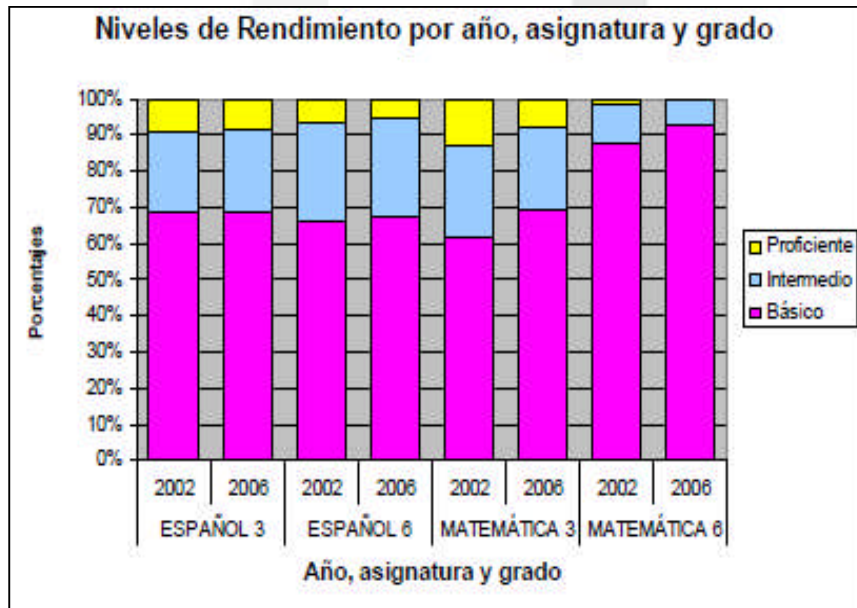
- Observamos que muchas evaluaciones en la región tienen poca o ninguna información acerca de los logros de las poblaciones mas vulnerables de estudiantes
- Estas pruebas no describen sus logros ni se representan estos en los informes nacionales.

Tabla 2. Resultados nacionales y de su región en **Matemática**

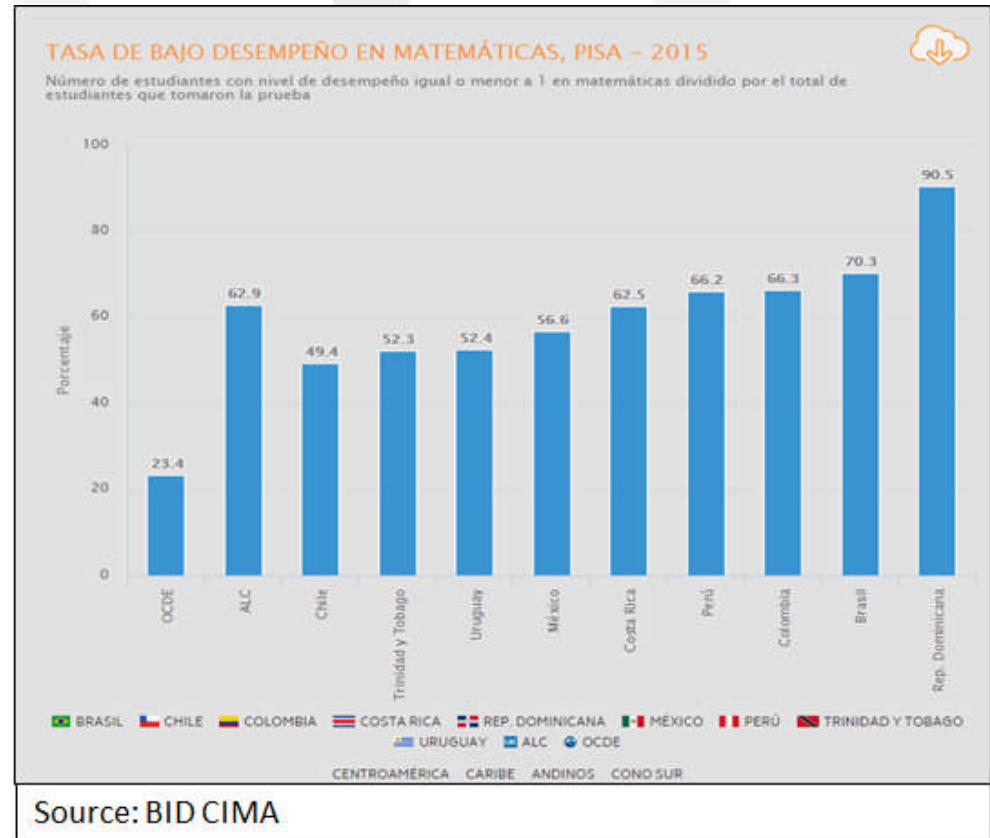


Los logros en matemática de aproximadamente la mitad de los estudiantes en Perú estuvieron debajo de los dos niveles definidos de desempeño en la prueba ECE de 2011.

Esta ausencia de evidencia acerca de los logros de proporciones significativas de estudiantes de poblaciones vulnerables, es común en pruebas en la región.



Nicaragua: Evaluación Nacional del Rendimiento Académico



Source: BID CIMA

	Criterio	Explicación	Evidencia para la validación
Validez consecucional de las evaluaciones			
8	Los resultados de las evaluaciones se comunican en forma efectiva.	Los resultados se publican en forma clara y correcta, en distintos medios y formatos (ej., informes, talleres, folletos, videos) adecuados para distintas audiencias, y se difunden en tiempo oportuno.	<p>* PRODUCTOS: (a) Productos de comunicación utilizados para difundir los resultados e información de las evaluaciones; (b) Estrategia de comunicación utilizada (ej. productos comunicacionales, audiencias, tiempos); (c) Niveles de acceso, valoración, comprensión de los resultados por parte de diferentes actores educativos.</p> <p>* PROCESOS: Métodos y procedimientos para diseñar la estrategia de comunicación (ej. consulta a audiencias o actores claves).</p>
Chile		Perú	
<p>Los resultados de las evaluaciones se difunden a través de diversos medios para distintas audiencias. Incluye difusión a través de sitio de internet oficial, informe nacional e informes para las instituciones educativas, entre otros.</p> <p>Por ley, la evaluación tiene consecuencias para las escuelas, que son clasificadas en cuatro categorías de desempeño y esta categorización está asociada a un sistema de incentivos y sanciones fuertes. No tiene consecuencias para los estudiantes.</p> <p>Estudio de valoración, acceso, comprensión y usos de la información muestra que docentes y directivos por lo general tienen acceso a los resultados SIMCE, pero aún tienen dificultades para comprenderlos y utilizarlos para la mejora.</p>		<p>Se entregan los resultados de las pruebas nacionales censales a través de informes de resultados a nivel nacional, regional, para directivos (institución educativa), docentes y padres o apoderados.</p> <p>Por decreto ministerial, se difunden resultados en un taller de reflexión nacional sobre los resultados de la evaluación nacional.</p> <p>Los resultados de las evaluaciones no tienen consecuencias para estudiantes, docentes o escuelas.</p> <p>Durante el año 2014 se realizó una encuesta sobre la recepción y uso de los informes de resultados de la Evaluación Censal (ECE). Existe también una evaluación realizada en 2017 que indaga sobre el uso de los informes de evaluación en el sistema educativo.</p>	
México		<p>Los resultados se difunden en el sitio de internet del INEE, en un Banco de Indicadores Educativos (un sistema nacional de indicadores educativos disponible desde el ciclo 2004-2005) y el informe nacional anual, denominado “Panorama Educativo de México”.</p> <p>Los resultados de las evaluaciones no tienen consecuencias para estudiantes, docentes o escuelas.</p> <p>No hay evidencias formales acerca de la eficacia de las comunicaciones</p>	

Consideraciones para afrontar nuevos desafíos.

CONCLUSIONES Y RECOMENDACIONES



Institute
for Global Education
Policy Studies

UNIVERSITY AT ALBANY State University of New York



Institute
for Global Education
Policy Studies

UNIVERSITY AT ALBANY State University of New York

- Un principio fundamental de la medición educativa es que aseveraciones acerca de la calidad en educación, deben ser aseveraciones auditables – susceptibles de verificación mediante evidencia empírica.
- Esto requiere un enfoque sobre la validez y la validación que recién comienza a ser discutido en América Latina
- Vemos en la región avances en herramientas y procedimientos para la validación curricular de instrumentos.
- Vemos desafíos también.

Avances

- Equiparación y calibración de pruebas
- Reportar niveles de desempeño
- Procedimientos estandarizados, documentados y públicos acerca del alineamiento de las pruebas con el currículum oficial.
- Esfuerzos generales por documentar y poner a disposición del público los procedimientos de diseño de pruebas

Desafíos

- Niveles de desempeño que dan poca o ninguna información acerca de los desempeños de grandes proporciones de estudiantes vulnerables
 - Presión por alinear niveles con PISA/TIMSS o las escalas UNESCO correspondientes a los ODS.
- Escasa atención a la validación con respecto al currículum ejecutado (validez instruccional)
- Ausencia de procedimientos para asegurar la comparabilidad en contextos dinámicos de reforma y reajuste curricular.
- Necesidad de ir más allá de la documentación de la fidelidad de procedimientos y llevar a cabo auditorías internas y externas de los productos de dichos procedimientos.

Desafíos para la evaluación externa de evidencia de validez

- Es necesario pasar de un esquema de validación de pruebas basada en verificar si se han seguido con fidelidad los procedimientos fijados para desarrollarlas, a un esquema de auditorías internas y externas de los resultados.
 - La formación de “Comisiones SIMCE” en Chile,
 - estudios de uso de informes de evaluación en Perú,
 - paneles de validación en México
- Los esfuerzos recién comienzan y muchas unidades de evaluación en LA son reticentes a este tipo de auditoría por muchas razones.
 - se objeta la cantidad de recursos monetarios y de tiempo que tomaría crear y organizar el tipo de documentación que se requeriría para una auditoría externa de este tipo.
 - Existe temor que las auditorías resultarían en juicios sumativos peligrosos para sistemas de evaluación que están aún en desarrollo, a menudo vulnerables políticamente.

- Administradores y técnicos tienen una genuina preocupación por costos y por los posibles efectos negativos de prácticas de validación que van más allá de asegurar que se siguen con fidelidad los procesos.
- Se objeta que mayores esfuerzos pueden ser muy costosos.
- Todo procedimiento de validación tiene costos.
- Curiosamente, ninguno de nuestros informantes menciona otro costo importante:
 - el costo de tomar una decisión de política equivocada debido a inferencias inválidas acerca de los problemas de enseñanza y aprendizaje que tienen los estudiantes.
 - Este costo, ampliamente estudiado en economía y análisis de políticas, es también llamado el costo del “error tipo III” (Mitroff & Featheringham, 1974),
 - es el error que se comete al tratar de resolver el problema equivocado.
- El sustento del valor de las evaluaciones para la mejora de la educación es la suposición que sus resultados arrojan informaciones válidas acerca de cuan bien dominan los estudiantes los ámbitos curriculares del currículum oficial.
- Si tenemos escasa evidencias de la validez de esas inferencias, el riesgo de caer en Error tipo III, es más alto.
- Este riesgo históricamente no ha recibido atención en AL.



Institute
for Global Education
Policy Studies

UNIVERSITY AT ALBANY State University of New York

+ información

Ramírez, M. J. y G. A. Valverde. *¿Cómo velar por la validez de las evaluaciones de aprendizajes? Criterios prioritarios para América Latina y el Caribe.*

Valverde, G. A. y M. J. Ramírez. *Prácticas contemporáneas en la validación curricular de evaluaciones nacionales de aprendizajes en América Latina: Estudio comparado de casos de Chile, México y Perú.*

En: *Validez de los Sistemas de Evaluación en Educación.* (Editores: J. Manzi, S. Taut, M.R. García). Pontificia Universidad Católica de Chile: MIDE-UC (Publicación en preparación – proyectada para marzo 2019)

gvalverde@albany.edu

¡MUCHAS GRACIAS!



Institute
for Global Education
Policy Studies

UNIVERSITY AT ALBANY State University of New York