



VNIVERSITAT
DE VALÈNCIA



Técnicas de construcción de niveles de logro para la presentación de resultados de Prova Brasil

Dr. Jesús M. Jornet Meliá

Departamento de Métodos de Investigación y
Diagnóstico en Educación (MIDE)
UNIVERSITAT DE VALÈNCIA



VNIVERSITAT
ID VALÈNCIA

Hablaremos acerca de...

- REVISIÓN DEL CONCEPTO DE ESTÁNDAR.**

- Métodos para la determinación de estándares y la identificación de puntuaciones de corte.**

- El concepto de validez de los estándares.**



Revisión del concepto de estándar (1).

Precisiones terminológicas

Acepciones de Estándar en la Evaluación Educativa:

- a) Normas de calidad de los planes de evaluación (JCSEE, 1981, 1994)
- b) Criterios o normas de interpretación de las puntuaciones de tests y pruebas

Acepciones de Estándar en Medición:

Estándar y Puntuación de corte se usan como sinónimos

*Propuesta de Kane (1994) *:*

- a) Estándar como “versión conceptual del nivel deseado de competencia”
- b) Estándar como puntuación de corte

**Van der Linden, 1981; Jornet, 1987; Jornet y Suárez, 1989; Kane, 1994
Cizek, 2001; Cizek, Bunch y Koons, 2004*



VNIVERSITAT
ID VALÈNCIA

Revisión del concepto de estándar (2). *Nuestra posición*

Estándares de una prueba:

Sistema de interpretación de sus puntuaciones ...

... y está al servicio de la **toma de decisiones** y de la **comunicación de resultados**.

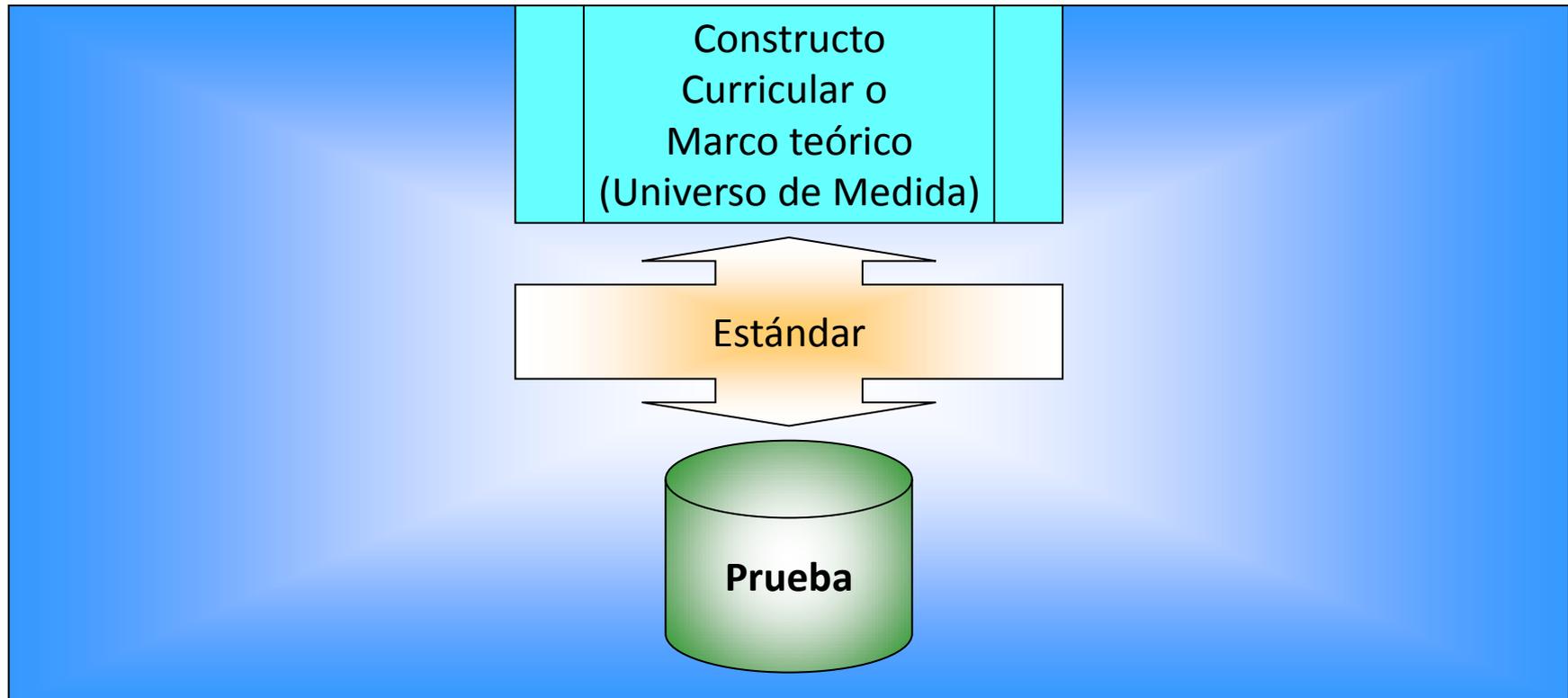
Estándar: Descripción del nivel de competencia

Puntuación de corte: puntuación en la prueba que diferencia entre dos niveles de competencia.



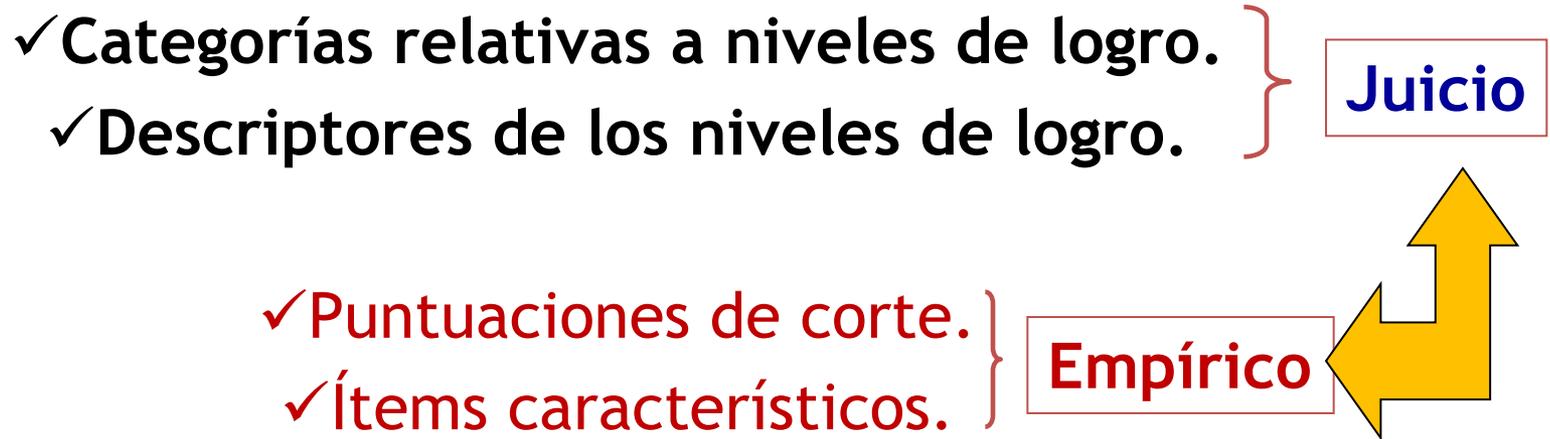
Revisión del concepto de estándar (3).

Nuestra posición





Revisión del concepto de estándar (4). *Elementos en la definición de Estándares*





Revisión del concepto de estándar (5).

Ejemplos de etiquetas de niveles de logro.

Etiquetas	Origen
Básico, Competente, Avanzado	National Assessment of Educational Progress
En camino, Progresando, Cerca de la competencia, Competente, Avanzado	Terranova (2ª. ed.) (CTB / McGraw Hill)
Limitado, Básico, Competente, Acelerado, Avanzado	Pruebas de rendimiento del Estado de Ohio
Muy por debajo del nivel básico, Debajo del nivel básico, Básico, Competente, Avanzado	Estado de California, Tests de California
No llega al nivel usual, Llega al nivel usual, Rendimiento destacado	Estado de Texas, Estándares de valoración de Texas de conocimientos y destrezas
Sin etiquetas, identificación mediante la cualificación de la escala numérica	INCE: Diagnóstico del sistema educativo español, 1998
Cinco niveles, sin etiquetas	Proyecto PISA (2000): Lectura
Escala de habilidad: Máximo, Medio, Mínimo	Proyecto PISA (2000): Matemáticas y Ciencias
Por debajo del Básico, Básico, Medio, Avanzado	INEE-México



Revisión del concepto de estándar (6).

Ejemplo de Descriptor de nivel de logro.

Etiqueta de nivel de rendimiento	Descriptor de nivel de rendimiento
Avanzado	<p><i>Los estudiantes de cuarto grado que están en el nivel avanzado deben poder generalizar sobre los tópicos en la selección de lecturas y demostrar un conocimiento suficiente acerca de cómo los escritores componen y usan las estrategias literarias. Cuando leen textos apropiados para cuarto grado, deben poder juzgarlos de forma crítica y, en general, dar respuestas minuciosas que demuestren que han comprendido el texto.</i></p> <p>Por ejemplo, cuando leen textos literarios los estudiantes de nivel avanzado deben poder hacer las generalizaciones sobre lo relevante de la historia y prolongar su significado integrando las experiencias personales y las otras interpretaciones con las ideas indicadas por el texto. Deben poder identificar los recursos literarios como la lengua figurada.</p> <p>Cuando leen textos informativos los alumnos de cuarto grado de nivel avanzado deben poder explicar el propósito del escritor usando material de soporte del texto. Deben poder hacer juicios críticos sobre la forma y el contenido del texto y explicar sus juicios claramente.</p>

Genérico

Ejemplos



VNIVERSITAT
ID VALÈNCIA

Revisión del concepto de estándar (7).

Finalidad del proceso de determinación de estándares.

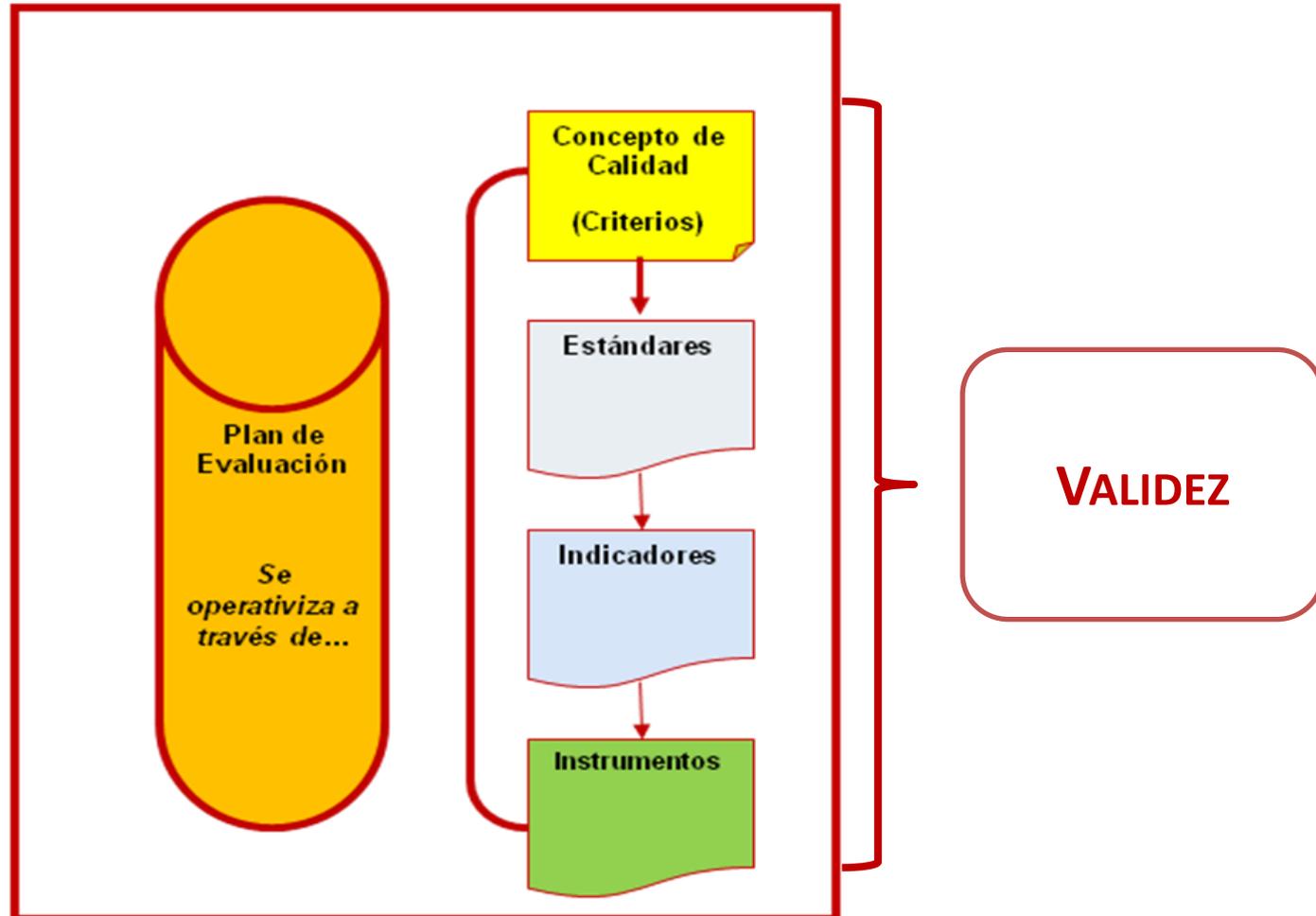
- ✓ Aportar un sistema de interpretación de puntuaciones de las pruebas que esté al servicio de la **toma de decisiones** y de la **comunicación de resultados**.
- ✓ ... De manera que el tipo de estándares a desarrollar debe ser **coherente con la finalidad y objetivos del plan general de evaluación**.

Cizek (2001b): *"El establecimiento de estándares es quizás la rama de la psicometría que mezcla -más que cualquier otra- ingredientes artísticos, políticos y culturales en la preparación de sus productos "* (p.5).



Revisión del concepto de estándar (8).

Validez de la Evaluación





VNIVERSITAT
ID VALÈNCIA

Hablaremos acerca de...

- Revisión del concepto de estándar.

- MÉTODOS PARA LA DETERMINACIÓN DE ESTÁNDARES Y LA IDENTIFICACIÓN DE PUNTUACIONES DE CORTE.**

- El concepto de validez de los estándares.



Una cuestión previa

Tipología de Estándares (Linn, 1994).

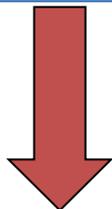
- ✓ **Exhortación**, cuando representan metas deseables de logro a las que debe tender la mejora de un sistema educativo o de los estudiantes,
- ✓ **Ejemplificación o muestra de rendimiento**, cuando representan las habilidades o competencias características de diversos niveles de ejecución,
- ✓ **Rendición de cuentas para educadores**, cuando representan metas curriculares precisas, orientando la evaluación hacia el contraste entre del currículum diseñado, el implementado y los logros educativos, y
- ✓ **Certificación del logro del estudiante**, cuando identifican un nivel mínimo de competencia del logro de estudiantes.



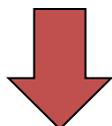
Métodos para la determinación de estándares y puntuaciones de corte (1).

Evolución metodológica

**Pruebas Referidas
al Criterio**



**Pruebas Referidas
a Estándares**



Pruebas a Gran Escala

- Años 60 (Glaser)
- Polémica juicio vs. empírico (Glass vs. Popham, Block, Hambleton, Shepard, Berk...)
- Orientado a pruebas de admisión/certificación
- Esencialmente decisiones dicotómicas

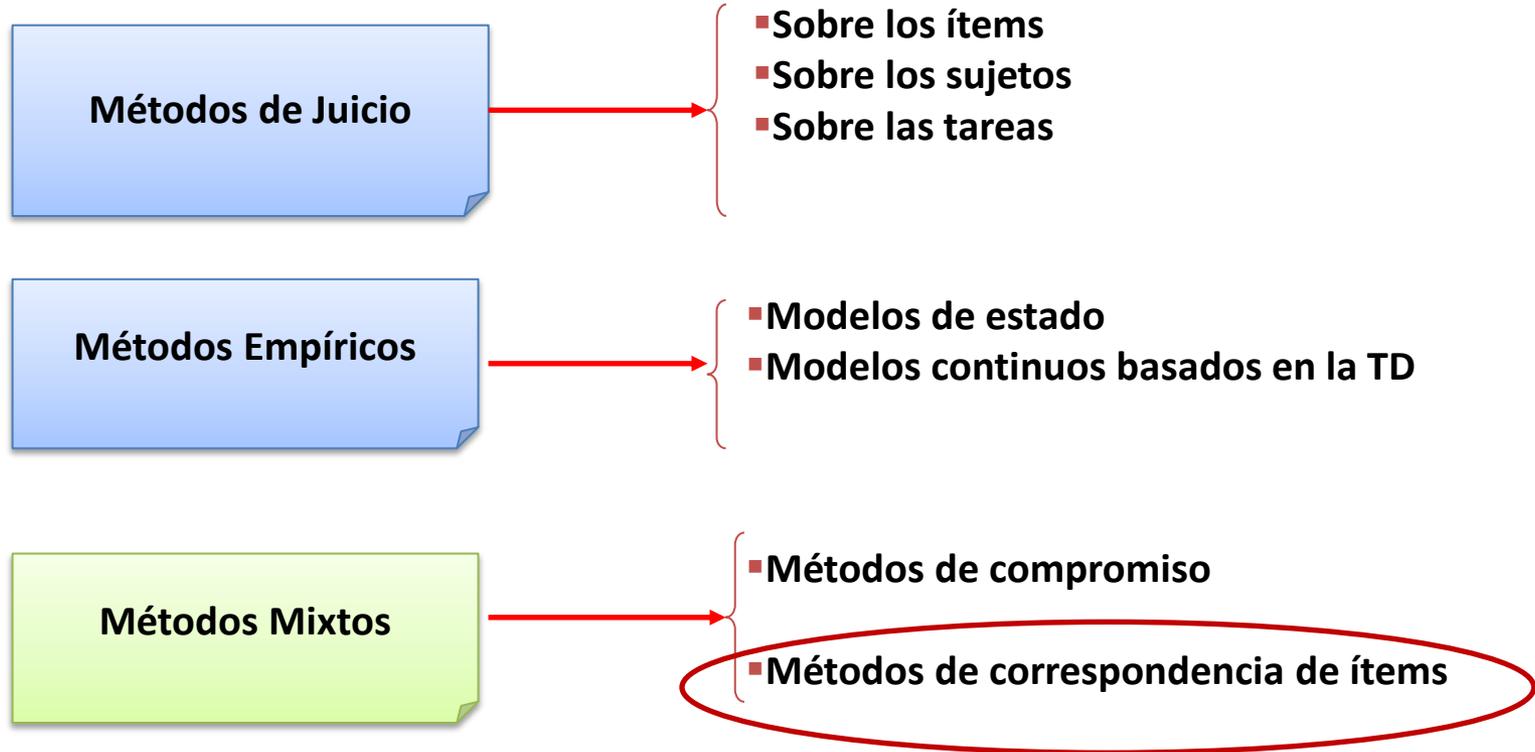
- Años 90
- Generalización de procesos mixtos juicio-empíricos
- Reconocimiento en Estándares legales y profesionales
- Incorporación a pruebas a gran escala
- Decisiones politómicas

- Años 2000
- Programas Internacionales y nacionales
- Generalización de procesos mixtos juicio-empíricos



Métodos para la determinación de estándares y puntuaciones de corte (2).

Clasificación de métodos.



Métodos para la determinación de estándares y puntuaciones de corte (3).

Etapas para la determinación de estándares

(Hambleton, 1998,2001)



VNIVERSITAT
DE VALÈNCIA

1. Seleccionar un comité de expertos grande y representativo, como base de la validez y fiabilidad de los estándares.
2. Elegir el método de determinación de estándares; preparar materiales de formación y el programa de reuniones para la determinación de estándares.
3. Preparar las descripciones de las categorías de rendimiento.
4. Formar a los participantes en el uso del método de determinación de estándares.
5. Recopilar clasificaciones de ítems y otras valoraciones de los participantes y producir información descriptiva /resumen u otra realimentación para los participantes.
6. Facilitar la discusión entre participantes de la información descriptiva /resumen inicial.
7. Realizar una segunda sesión de clasificaciones/valoraciones; compilar la información y facilitar la discusión como en los pasos 5 y 6.
8. Dar una oportunidad final a los participantes de examinar la información y llegar a los estándares finales de rendimiento recomendados.
9. Llevar a cabo una **EVALUACIÓN DEL PROCESO** de determinación de estándares, recogiendo información sobre la confianza de los participantes en el proceso y los estándares de rendimiento resultantes.
10. Reunir la documentación del proceso de determinación de estándares y cualquier otra **EVIDENCIA DE LA VALIDEZ** de los estándares de rendimiento resultantes.



VNIVERSITAT
DE VALÈNCIA

Métodos para la determinación de estándares
y puntuaciones de corte (4).

Un ejemplo: EXCALE-INEE México

Modelo para la determinación de Niveles de Logro y Puntos de Corte de los Exámenes de la Calidad y el Logro Educativos (Excale)

Jesús M. Jornet Meliá* y Eduardo Backhoff Escudero**

CUADERNO No. 30



VNIVERSITAT
DE VALÈNCIA

Métodos para la determinación de estándares y puntuaciones de corte (5). *Un ejemplo: EXCALE-INEE México*

Adaptación del Método Bookmark (o del Marcador)
de Lewis, Mitzel y Green (1996) y Lewis, Mitzel, Green
y Patz (1999)

Formato de trabajo:

La tarea de los jueces es identificar el/los ítems que actúan como punto de inflexión entre dos niveles de logro a partir de la revisión de los ítems ordenados por dificultad



VNIVERSITAT
DE VALÈNCIA

Métodos para la determinación de estándares y puntuaciones de corte (6). *Un ejemplo: EXCALE-INEE México*

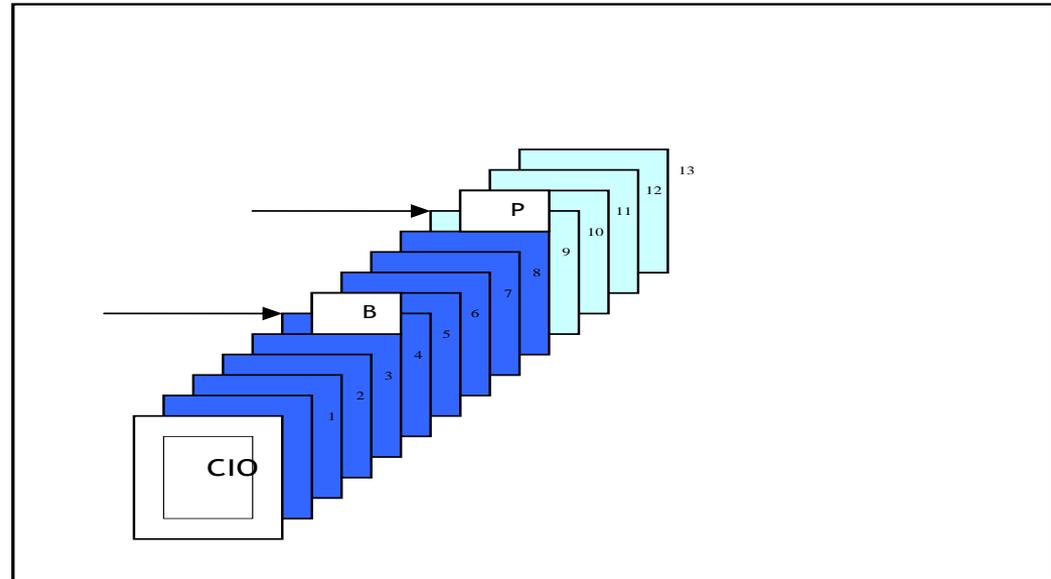


Ilustración de Cuadernillo de Ítems ordenados e Ítems marcadores.



VNIVERSITAT
ID VALÈNCIA

Métodos para la determinación de estándares y puntuaciones de corte (7). *Un ejemplo: EXCALE-INEE México*

Problemas:

- ✓ Pocos...
- ✓ Dudas respecto al límite de probabilidad que se asigna al ítem (67 ó 50%), aunque los estudios realizados avalan la primera opción
- ✓ Es tanto mejor el método cuanto más se simplifique el formato de juicio

Ventajas:

- ✓ Realista y simple
- ✓ Se puede aplicar con ítems de diferente ponderación, incluso con ítems de respuesta abierta
- ✓ Los estudios al respecto avalan en buena medida las propuestas del método, si bien no son concluyentes



Métodos para la determinación de estándares y puntuaciones de corte (8). *Un ejemplo: EXCALE-INEE México*

Aproximación empírica

Componente de trabajo	Finalidad
Revisión teórica	Aportar información a participantes para la toma de decisiones acerca del Modelo
Propuesta del Modelo del INEE	Propuesta de un Modelo de determinación de estándares y el ajuste del diseño específico para el INEE
Estudios de Validación	Recabar evidencias de validez de los estándares
Estudio I	Analizar el proceso de determinación de estándares e identificación de puntos de corte
Estudio II	Analizar los estándares como producto
Conclusiones y recomendaciones de mejora	<ol style="list-style-type: none">1. Aportar una síntesis de la calidad del trabajo desarrollado para determinar estándares y PCs2. Extraer conclusiones de mejora del Modelo de trabajo

Métodos para la determinación de estándares y puntuaciones de corte (9).

Un ejemplo: EXCALE-INEE México



VNIVERSITAT
DE VALÈNCIA

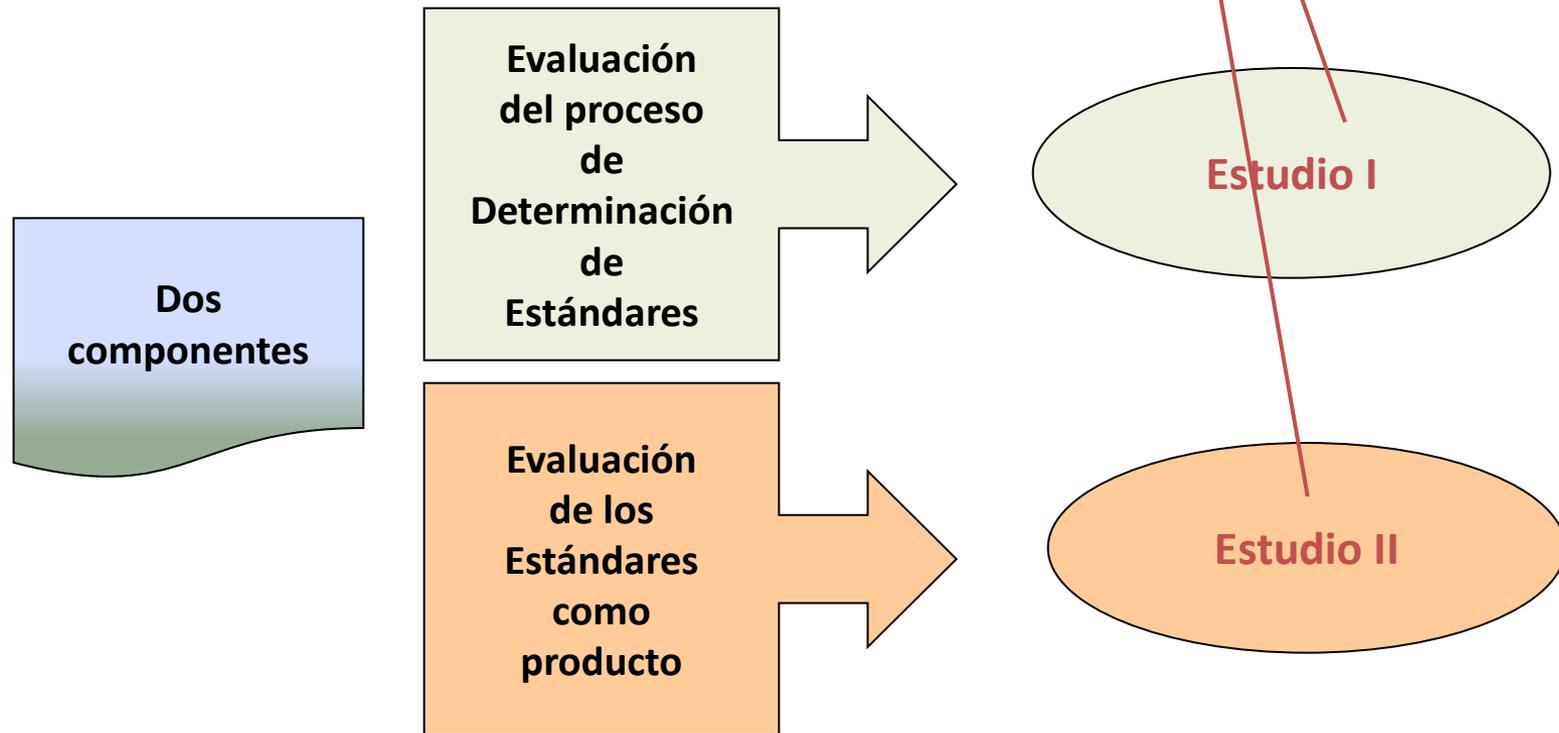
Pretendimos ...:

- Que el sistema de interpretación sea representativo de las opiniones de expertos en educación (**juicio**)
- Que considere la implementación real del sistema educativo a través de una población tan diversa como es la de escolares Mexicanos (**empírico**)
- **Que pudiera ser validado a partir de estudios empíricos**

Racionalidad: se puede evaluar por el grado en el cual las puntuaciones de corte derivadas del proceso de determinación de estándares clasifican a los examinados en grupos de una manera consistente con otra información

Replicabilidad: comparación de la determinación de estándares a partir de la utilización de métodos diferentes y/o comités independientes

Validación de los Estándares de EXCALE.



Métodos para la determinación de estándares y puntuaciones de corte (10).

Un ejemplo: EXCALE-INEE México

Estudio I: Validación del proceso de determinación de Estándares.



VNIVERSITAT
ED VALÈNCIA

Objetivo principal:

Analizar la calidad del proceso de determinación de estándares y de la identificación de puntuaciones de corte.

Objetivos implicados:

- Documentar y analizar el proceso desarrollado, garantizando la independencia de sesgos en el mismo y valorando su adecuación metodológica.
- Analizar la calidad de los estándares desarrollados y las puntuaciones de corte seleccionadas.
- Analizar la influencia de posibles variables diferenciadoras en el proceso.



Métodos para la determinación de estándares y puntuaciones de corte (11).

Un ejemplo: EXCALE-INEE México

Estudio II: Validación de los Estándares como producto.

Objetivo principal:

Analizar la validez del sistema de interpretación de puntuaciones de los EXCALE, a partir de la comparación de la clasificación de los alumnos realizada aplicando los estándares determinados mediante procesos de juicio y la emanada de un análisis empírico de conglomerados.

Objetivos implicados:

- Analizar la capacidad de discriminación de las pruebas para identificar perfiles diferenciales (al menos cuatro) de respuesta a las pruebas.
- Comprobar si los perfiles diferenciales identificados implican niveles diferenciados de habilidad total (θ).
- Comprobar si la clasificación de sujetos derivada de ambas aplicaciones producen el mismo tipo de resultados.
- Analizar la permanencia de resultados a través de todas las formas de cada prueba.



VNIVERSITAT
ID VALÈNCIA

Hablaremos acerca de...

- Revisión del concepto de estándar.
- Métodos para la determinación de estándares y la identificación de puntuaciones de corte.
- EL CONCEPTO DE VALIDEZ DE LOS ESTÁNDARES.**



El concepto de validez de los estándares (1).

La clave para poder utilizar las pruebas como *indicadores* está en el sistema de criterios (*Estándares*) que se utilicen para interpretar sus puntuaciones.

“Un paso crítico en el desarrollo y uso de algunas pruebas es establecer uno o más puntos de corte dividiendo la escala de puntuaciones para dividir la distribución de puntuaciones en categorías [...] Los puntos de corte plasman las reglas de acuerdo con las que las puntuaciones se usan o interpretan en las pruebas.

*Por lo tanto, **en algunas situaciones la validez de las interpretaciones de prueba podría depender de las puntuaciones de corte**”(APA, AERA, NCME, 1999, p.53).*



VNIVERSITAT
ID VALÈNCIA

El concepto de validez de los estándares (2).

Dos facetas del concepto:

- a) Derivadas del concepto de prueba estandarizada

- b) Derivadas de los usos evaluativos



El concepto de validez de los estándares (3).

*Facetas derivadas del concepto de Prueba Estandarizada
(Jornet y Perales, 2001).*

✓ Validez de Constructo y Validez de Contenido:

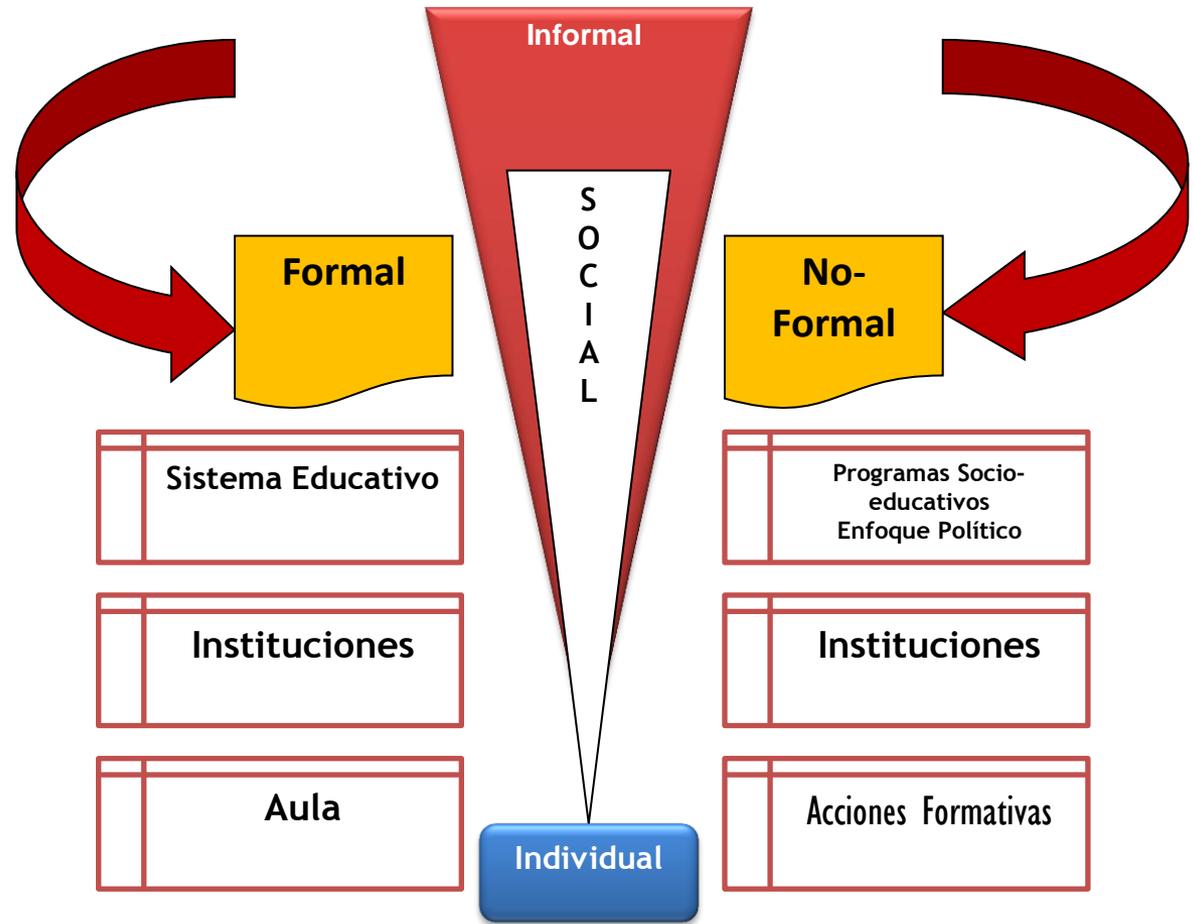
- Dependen de la calidad de la definición del dominio educativo (DE) como universo de medida
- El estándar debe representar de forma adecuada los elementos relevantes del DE, que resultan clave en la descripción de la calidad del aprendizaje

✓ Validez Criterial (Convergente y Predictiva):

- como elementos de contextualización del uso de las puntuaciones
- tienen valor explicativo del funcionamiento del estándar
- comprobación empírica y basada en criterios acerca de la utilidad de la prueba respecto a las decisiones que debe apoyar



**Situación
educativa:
Escenarios**

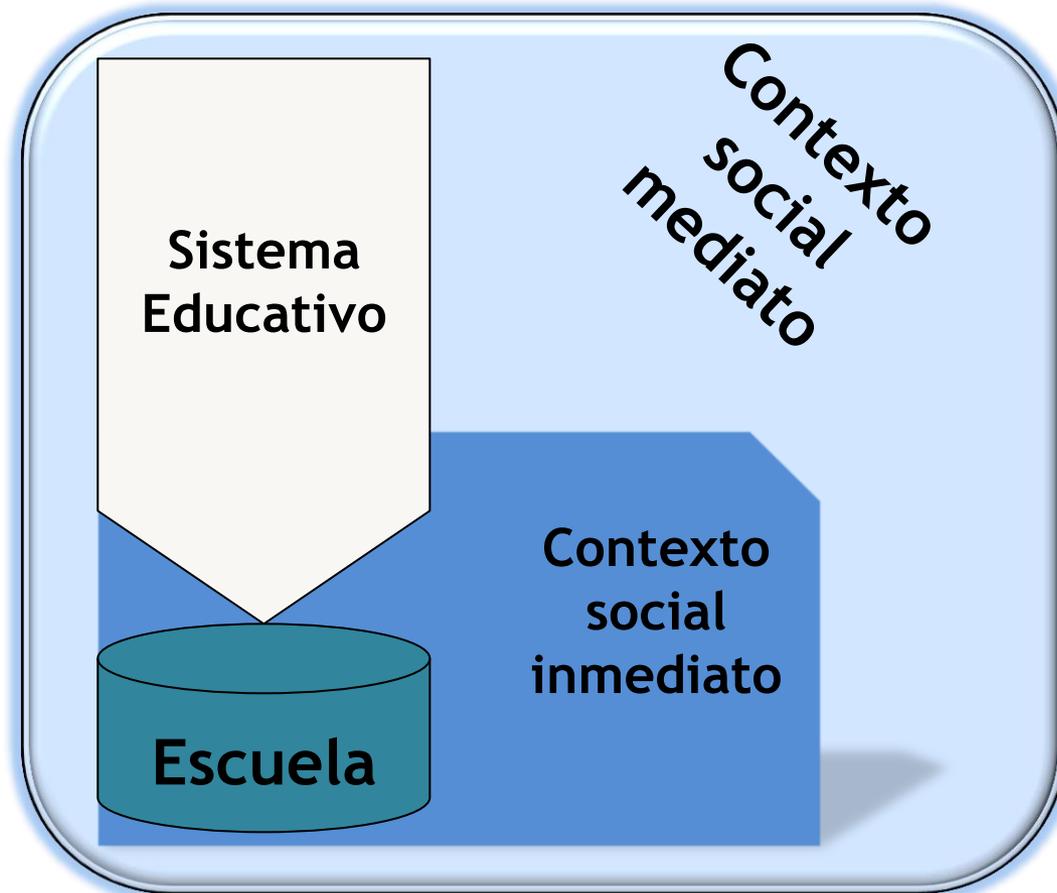


Condicionantes de los enfoques de evaluación de Escuelas.

Del concepto de Escuela.



VNIVERSITAT
DE VALÈNCIA





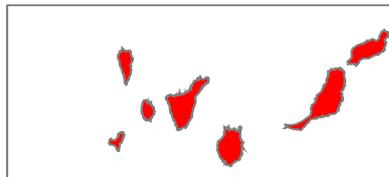
VNIVERSITAT
DE VALÈNCIA

España PISA y DESEMPLEO

>PISA <PARO

¿?

<PISA >PARO





VNIVERSITAT
DE VALÈNCIA

El concepto de validez de los estándares (4).

Facetas derivadas de los usos evaluativos (1).

(Jornet y Perales, 2001).





VNIVERSITAT
ID VALÈNCIA

El concepto de validez de los estándares (5).

Facetas derivadas de los usos evaluativos (2).

(Jornet y Perales, 2001).

✓ **Credibilidad.** Se basa en la adecuación del procedimiento de desarrollo de la prueba en general, y del Estándar en particular.

Factores:

- Composición del comité de expertos
- Selección de los expertos
- Adecuación de las técnicas de emisión de juicio
- Congruencia de resultados
- Transparencia del proceso
- Comité meta-evaluador
- Estudios y evidencias de la adecuación de los estándares



VNIVERSITAT
ID VALÈNCIA

El concepto de validez de los estándares (6).

Facetas derivadas de los usos evaluativos (3).

(Jornet y Perales, 2001).

✓ **Utilidad.** Hace referencia al grado en que los resultados de la evaluación se usan para la toma de decisiones acerca del sistema educativo.

ES CONSECUENCIA DE LA CREDIBILIDAD Y LA RETROALIMENTA

Niveles:

- Orientación política
- Diseño curricular
- Mejora de la actuación docente

Comprobación: impacto de la evaluación



- Concepto de calidad a evaluar: debe estar muy asociado a la orientación de la Educación
- Los estándares representarán ese concepto: Necesidad de definir para qué se quiere disponer de los estándares
- Valorar si respetan la diversidad cultural y social: contextualización
- Los estándares deben estar desarrollados de manera que se asegure la Validez (en especial, cultural y consecuenencial)
- Necesidad de informar acerca del modo en que se han determinado los estándares



- **ANÁLISIS DEL DE (JUICIO)**
 - Identificar Dimensiones
 - Identificar elementos relevantes de cada dimensión
- **DISEÑAR LOS ESTÁNDARES (JUICIO)**
 - Ajustados a la definición de calidad
 - Considerar elementos de contextualización
 - Definir el uso/rol de los estándares (validez consecucional)
- **ANÁLISIS DE ESTÁNDARES (JUICIO)**
 - Alineación de los estándares al Concepto de Calidad
 - Alineación de los estándares respecto al DE
 - Implicaciones de los estándares respecto al enfoque de la enseñanza (prever sesgos vinculados al enfoque metodológico-didáctico)
 - Validez Cultural



VNIVERSITAT
ID VALÈNCIA

A modo de conclusió

Acerca del Proceso de Desarrollo de Estándares:

Lecciones Aprendidas

- **DISEÑAR LOS REACTIVOS (JUICIO)**
- **REVISIÓN LÓGICA DE REACTIVOS (JUICIO)**
- **ANÁLISIS DE ALINEACIÓN DE REACTIVOS (JUICIO)**
 - Respecto al DE/Estándares
 - Enfoques de Enseñanza
 - Modos de evaluación en el aula
- **ENSAYOS PILOTO (EMPÍRICOS)**
 - Adecuación de los Estándares (validez criterial convergente)
 - Determinación/validación de puntuaciones de corte
 - Alineación empírica de Estándares a enfoques de enseñanza
 - Alineación empírica de Estándares a modos de evaluación en el aula



VNIVERSITAT
ID VALÈNCIA

Validez de las pruebas depende de muchos factores

Validez y Estándares son dos facetas de un mismo problema



VNIVERSITAT
ID VALÈNCIA

Para concluir...

Cuando abordemos el diseño de procesos de evaluación, recordad...

Julio Cortázar escribía:

'La coma, esa puerta giratoria del pensamiento'



Para concluir...

Lean y analicen la siguiente frase:

“Si el hombre supiera realmente el valor que tiene la mujer andaría a cuatro patas en su búsqueda.”



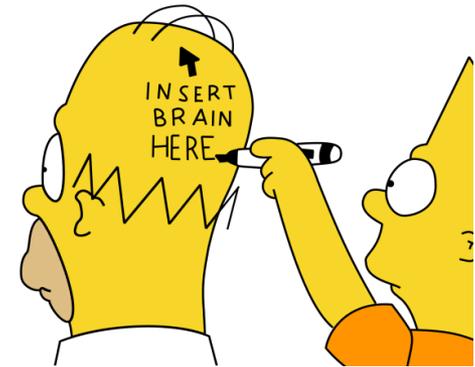
Para concluir...

Las trampas de la evaluación...

¿Quién escribió cada frase?...

*“Si el hombre supiera realmente el valor que **tiene**, la mujer andaría a cuatro patas en su búsqueda.”*

*“Si el hombre supiera realmente el valor que tiene la **mujer**, andaría a cuatro patas en su búsqueda.”*





VNIVERSITAT
DE VALÈNCIA



Muchas gracias por su atención

jornet@uv.es