

Las evaluaciones a gran escala: alcances, limitaciones y desarrollo

Consejo Pedagógico de Evaluación
Educativa (CONPEE)

21 de noviembre de 2014

DOCUMENTO DE TRABAJO

**INSTRUMENTOS DE EVALUACIÓN A
GRAN ESCALA: CARACTERÍSTICAS
GENERALES, USOS, ALCANCES Y
LIMITACIONES.**



Evaluaciones
de aula

Evaluaciones
de gran escala





Aula

- Sabe qué enseñó a sus alumnos: es claro qué va a evaluar
- Conoce a sus alumnos
- Puede realizar evaluaciones en distintos momentos de oportunidad
- Puede desplegar múltiples métodos de evaluación



Gran escala

- Debe definir con cuidado lo que se va a evaluar
- No conoce a los alumnos
- Momentos muy acotados de evaluación
- Métodos de evaluación restringidos por tiempo y costo

Tipos de instrumentos de evaluación

- Respuesta seleccionada (opción múltiple y respuesta corta)
- Respuesta escrita extensa (Ensayo)
- Evaluación de la ejecución (Demostraciones y artefactos)
- Comunicación personal

DOCUMENTO DE TRABAJO

**ELEMENTOS NECESARIOS PARA EL
DESARROLLO DE UN INSTRUMENTO DE
EVALUACIÓN ESTANDARIZADO.**

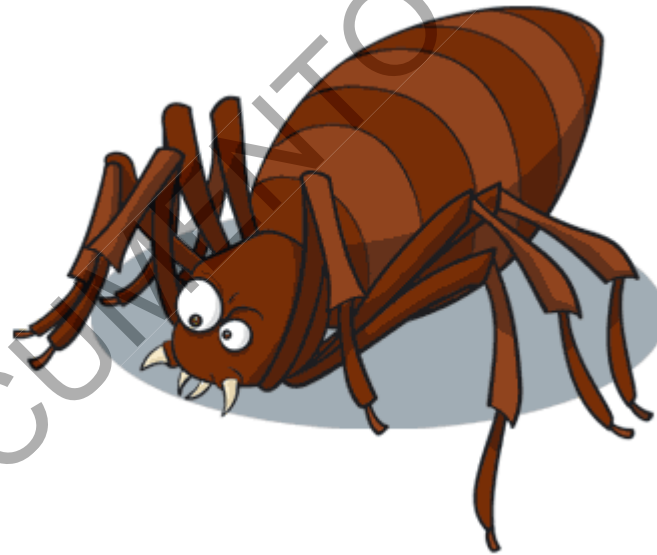
Validez y confiabilidad

Confiabilidad

Precisión de la medida

Validez

Precisión de la interpretación



Precisión

- Tensión entre precisión y oportunidad de la información.
- Las evaluaciones se pueden trivializar si se apresura su desarrollo.
- ¿Información imprecisa es mejor que ninguna información?

Un solo propósito

Una evaluación difícilmente puede servir correctamente a diferentes propósitos. Por ejemplo: la que sirve para promoción tiene alto impacto en las condiciones de vida, lo cual causa distorsiones importantes que no le permiten funcionar también como evaluación para retroalimentación del profesor, de los alumnos o de la escuela.

Ruta de evaluación (toma de decisiones)

- ¿Para qué evaluar?
 - ¿Quién hace la evaluación y por qué?
 - ¿Qué evaluar?
 - ¿Cómo evaluar?
 - ¿Qué significan los resultados?
 - ¿A quienes y cómo reportar los resultados y sus conclusiones?
 - ¿Qué validez, confiabilidad y pertinencia tienen el proceso y sus resultados?
 - ¿Qué mejoras se pueden hacer al proceso?
- Diseño
- Análisis
- Metaevaluación
-
- The diagram illustrates the classification of evaluation questions into three phases. A large, faint watermark 'DOCUMENTO DETRABAJO' is visible across the page. The questions are grouped as follows: 'Diseño' (Design) includes the first four questions; 'Análisis' (Analysis) includes the fifth and sixth questions; and 'Metaevaluación' (Meta-evaluation) includes the seventh and eighth questions. Each group is indicated by a right-facing curly bracket.

Acciones previas a la redacción de reactivos

- Contestar las preguntas de diseño.
- Definir el **universo de medida**.
- Formular una **tabla de contenidos**, que indica la relación entre el **universo de medida** y los instrumentos de medición.
- Elegir el o los tipos de ítems idóneos, hacer especificaciones de reactivos.

El universo de medida

- Casi nunca se pueden enunciar todos sus elementos, pero se pueden establecer criterios que lo describen y limitan
- De este universo se obtendrán los contenidos a evaluar
- Éstos pueden ser extraídos al azar o por conveniencia.

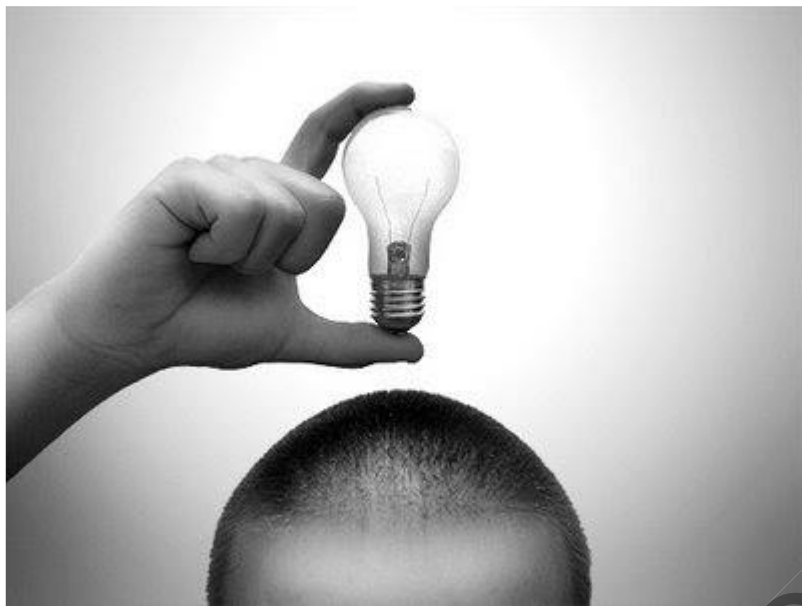
Tabla de contenidos

Debe incluir mínimamente:

- Contenidos que se van a evaluar
- Relaciones que tienen éstos con el universo de medida
- Proporciones en que se van a evaluar.

Especificación de ítems

- Se realiza un análisis del contexto del contenido.
- Se revisa los documentos de justificación.
- Se revisa la tabla de contenidos.
- Se analiza el contenido a evaluar.
- Se elabora la especificación.



DOCUMENTO DE TRABAJO



Normas técnicas para la construcción de los reactivos

Un fragmento, como ejemplo

Del contenido:

- La base del reactivo y los distractores deben redactarse de conformidad estricta con su especificación.
- Los contenidos deben ser vigentes y estar libres de sesgo (cultural, socioeconómico, regional, etc.).
- Evitar:
 - Formular preguntas de manera ambigua.
 - Emplear un vocabulario sofisticado de forma innecesaria.
 - Preguntar sobre conocimientos demasiado específicos.
 - Plantear preguntas que requieran la opinión del examinado.
 - Elaborar reactivos idénticos a los que aparecen en libros, guías y otros documentos.

Elemento fundamental: el trabajo colegiado con personas de distintas instituciones y con diferentes perspectivas

- Ejemplo: “Cecilia camina una cuadra hacia el sur y luego dos hacia el este...”



Algunos elementos de la Validación de reactivos

- ¿El reactivo mide lo que pretende medir?
- ¿La dificultad del reactivo es adecuada para el grado escolar?
- ¿Está libre de sesgos?
- ¿Todas las expresiones usadas son comprensibles?
- ¿Puede expresarse de manera más sencilla sin perder rigor?

¿Para qué sirven los estudios piloto, los micropiloteos y las entrevistas cognitivas?

- Los tres procedimientos provienen de la tradición científica de hacer estudios de campo y experimentales.
- En el contexto de la evaluación, aunque se utilizan para diferentes propósitos específicos, tienen en común que ayudan a mejorar los instrumentos de evaluación una vez que éstos ya pasaron su fase de planeación y se encuentran al final de la de elaboración.

Estudios piloto

- Son experiencias de evaluación muy similares a las que se planea hacer como definitivas o finales, pero que se realizarán en menor escala (menor número de participantes) y que no tienen la intención de arrojar datos válidos, sino de **poner a prueba los procesos e instrumentos** de evaluación.
- Sirven como una especie de “ensayo general”, para ver que los diferentes componentes de la aplicación definitiva sean correctos:
 - Reactivos adecuados
 - Formatos comprensibles
 - Logística adecuada
 - Procedimientos ajustados a la realidad
 - Manejo correcto de imprevistos
- También sirven para poder hacer mejores estimaciones sobre el tamaño que deben tener las muestras representativas de la aplicación definitiva.

Micropiloteos

- Son similares a los estudios piloto, pero en escala aún menor: son aplicaciones que puede hacer una sola persona, para verificar algún aspecto en particular de sus instrumentos o sus procedimientos de evaluación.
- Por ejemplo, puede ser muy útil para mejorar una prueba aplicar 10 reactivos a un grupo escolar de sólo 30 niños, y ver cómo son sus respuestas.

Entrevistas cognitivas

- “una serie de entrevistas individuales semiestructuradas en ambiente controlado con una muestra pequeña de la población meta. Durante las entrevistas, las personas participantes completan el cuestionario en estudio y realizan una serie de pruebas para detectar problemas a la hora de contestarlo” (Willis, 2005, en Smith y Molina, 2011).

Utilidad

Los estudios piloto, micropiloteos y entrevistas cognitivas sólo tienen sentido si, tras llevarlos a cabo, se mejoran los reactivos a la luz de la experiencia que dejaron estos procesos.

La evaluación en contexto

- Conocer información que dé contexto al aprendizaje (¿cómo es el niño? ¿cómo es su familia? ¿qué acontecimientos importantes hay en su vida ahora? ¿viene de otra escuela/entidad/país?, etcétera)
- Tomar en cuenta los contextos específicos al hacer juicios de valor
- Cuidar especialmente la evaluación en contextos vulnerables.

Referencias

- Smith, Vanessa; Molina, Mauricio Cuaderno Metodológico 5. La entrevista cognitiva: guía para su aplicación en la evaluación y mejoramiento de instrumentos de papel y lápiz. San José, CR.: Instituto de Investigaciones Psicológicas, Universidad de Costa Rica. 2011.
- Stiggins, R.; Arter, J.; Chappuis, J. y Chappuis, S. (2004). Classroom assessment for student learning: doing it right – using it well. Assessment Training Institute.
- Suber, E. (2012). Richard Stiggins : Access what? Access how? Clear Targets and Learning Methods. URL:
<http://prezi.com/0npjfipxcgwj/richard-stiggins-access-what-access-how-clear-targets-and-learning-methods/>