

Evaluación auténtica y "multimedia"

¿De qué manera los estudiantes responden a un modelo de evaluación auténtica? [Notas]

JAN HERRINGTON & ANTHONY HERRINGTON [Edith Cowan University Australia]. En Higher Education research & Development, Vol. 17. Nº 3, 1998, p. 305-322.
Traducción: Pedro D. Lafourcade. - Edición SDI.

- **Introducción**

.....

- **Multimedia y evaluación**

.....

- **Aprendizaje situado y el contexto social de la evaluación**

El contexto social del aprendizaje, ha representado un foco de atención en el pensamiento educativo desde mediados de los 80, y se le otorgó importancia con la publicación de 7 trabajos sobre la teoría del aprendizaje situado o cognición situada [Collins, Brown y Newman, 1989; Brown, Collins y Duguid, 1989; Lave y Wenger, 1991].

La teoría afirma que el conocimiento está "contextualmente situado", e influido por la actividad, el contexto y la cultura en la cual es utilizado [McLellan, 1996].

No sólo el contexto del aprendizaje debe ser auténtico, sino que la evaluación del mismo, también será auténtica y completamente integrada en el ámbito en que se genera. [Young, 1993].

En contraste con la evaluación auténtica, los textos referidos a normas, comúnmente se aplican en el supuesto de que son una medida verdadera de la capacidad de los estudiantes.

Autores como Gardner (1992), han cuestionado la pertinencia de tales pruebas, afirmando que no son lo suficientemente sensibles como para explicar las diferencias culturales. Además, que son empleadas, fundamentalmente, para establecer un orden entre los estudiantes, más bien que para determinar el nivel de comprensión de los mismos.

Este argumento es apoyado por gran parte de la investigación de aprendizajes en el contexto [Lave, Murtagh & de la Rocha, 1984; Rogoff, 1984; Scribner, 1984; Lave y Wenger, 1991]. Estos y otros estudios, han revelado por ejemplo, que quienes exhiben en su vida diaria excelente dominio de habilidades, fracasan a menudo en mediciones formales de cálculo y de razonamiento.

EVALUACIÓN AUTÉNTICA

Si las pruebas referidas a normas deben ser evitadas, cuáles tipos de instrumentos de evaluación serían más adecuados para aprendizajes que hayan sido logrados en ambientes de enfoques de la enseñanza más constructivistas, y en donde los estudiantes trabajan en colaboración, y con una variedad de herramientas y recursos, sobre actividades auténticas.

Mc Lellan (1993), en un artículo sobre la evaluación del aprendizaje situado, sugirió que las mismas pueden adoptar diversas modalidades [que no incluyen las pruebas formales convencionales], tales como: carpetas ["portfolios"]; resúmenes estadísticos de trayectorias de los estudiantes en relación con programas multimedia; diagnósticos; reflexiones y autoevaluación.

Ya existe una amplia literatura sobre evaluaciones *auténticas* y *basadas en el desempeño*, como medios más apropiados para evaluar el aprendizaje. Lajoie [1991], argumenta que en la evaluación auténtica, se requiere información acerca del aprendizaje que los estudiantes deberían realmente poner en evidencia en el mundo real, como opuesta a los tipos de tareas tradicionalmente llevadas a cabo en las clases.

Una definición más común de la evaluación auténtica es la expresada por Torrance en 1995:

"La evaluación auténtica es una expresión genérica que describe una variedad de nuevos enfoques sobre la evaluación. La implicación básica del término, está referida a que las tareas de evaluación diseñadas para los estudiantes, deberían ser más prácticas, realistas, y desafiar lo que uno podría denominar las "tradicionales" pruebas de lápiz y papel". [p.1]

Dicha definición, parecería abarcar el significado general de una variedad de términos empleados en la literatura, para describir formas alternativas de la evaluación, tales como:

- Evaluación auténtica
- Evaluación basada en el desempeño
- Evaluación de materiales contenidos en carpetas ["portfolios"]
- Evaluación del "coursework"

Otros autores e investigadores han intentado ser mucho más específicos acerca de las características de ambos enfoques, y procurando comprobar que las pruebas mismas que definen el segundo enfoque, son esencialmente auténticas.

Así por ejemplo, Wiggins [1990], ha focalizado el proceso de evaluación y ha refinado las características que concurren en el diseño y uso de la evaluación auténtica.

La tabla 1 resume las diferencias entre la evaluación auténtica y la tradicional.

TABLA 1 – Comparación entre evaluación auténtica y tradicional

Evaluación auténtica	Evaluación tradicional
1. Se apunta a una examinación directa del estudiante sobre tareas intelectuales valiosas.	1. Emplea ítemes indirectos.
2. Se requiere de los alumnos que se desempeñen de modo efectivo en relación con los saberes adquiridos.	2. Revela sólo si los estudiantes pueden reconocer, o "enchufar" lo que fue aprendido fuera de contexto.
3. Se presenta a los evaluandos conjuntos completos de tarea.	3. Las pruebas convencionales habitualmente se limitan a lápiz y papel y preguntas de una sola respuesta.
4. Se procura verificar si el alumno puede expresar respuestas justificables y generar productos o desempeños estimados como satisfactorios.	4. Las pruebas convencionales típicamente sólo requieren del alumno que seleccione o escriba la respuesta correcta sin tomar en cuenta las razones.
5. Se intenta garantizar la validez y la confiabilidad en los juicios estableciendo criterios adecuados para calificar los logros observados.	5. Las pruebas tradicionales estandarizan los ítemes y determinan una respuesta correcta por cada uno.
6. Se toma en cuenta que la validez del examen depende en gran medida de, si lo que se requiere que se lleve a cabo, expresa la capacidad para actuar idóneamente en el mundo real.	6. La validez de la prueba se determina apareando los ítemes con el contenido del currículo.
7. Se incluyen desafíos mal estructurados que ayuden a los estudiantes a ensayar cómo operar ante las complejas ambigüedades del mundo profesional. [Notas resumidas de Wiggins, 1990, the case for authentic assesment – Washington, Eric Clearinghouse on test Measurement and Evaluation . Eric doc. Reproduction Service N° DE 328 606].	7. Los textos tradicionales son más parecidos a ejercicios mecánicos que a menudo evalúan elementos estáticos y arbitrarios componentes de tales actividades.

CRÍTICAS A LA EVALUACIÓN AUTÉNTICA

Una de las principales críticas a la evaluación auténtica y a la evaluación basada en el desempeño, es que la validez se logra a expensas de la confiabilidad. [Wolf, 1995].

Linn, Baker y Dunbar [1991], identificaron 8 dimensiones frente a las cuales puede ser útilmente juzgada la validez de la evaluación basada en el desempeño:

- ✓ consecuencias
- ✓ imparcialidad
- ✓ transferencia
- ✓ cobertura del contenido
- ✓ complejidad cognitiva
- ✓ significatividad
- ✓ juicio
- ✓ costo y eficiencia

No obstante, es mucho más difícil, asegurar la confiabilidad de la evaluación. Linn y otros, expresan que: *“es inevitable una mayor dependencia de los juicios críticos sobre el desempeño de las tareas”*.

Esto puede ser un problema mayor, en una evaluación en gran escala, efectuada para determinar la responsabilidad por los resultados, por ejemplo, en la escuela secundaria, que para los propósitos de contextos más pequeños y específicos, comúnmente existentes en las instituciones terciarias [Gipps, 1995] y cuando el estudiante sea evaluado en relación con las actividades de clase, en donde el aprendizaje y la evaluación, son esencialmente inseparables. [Reeves & Okey, 1996].

Gipps [1995] ha señalado, sin embargo, que las pruebas estandarizadas aún en evaluaciones en gran escala, no garantizan mayormente una confiabilidad más elevada.

Ella cita una prueba nacional estandarizada de lectura y de comprensión, aplicada en Gran Bretaña, en donde una re-calificación de la misma, fue similar a la asignada inicialmente en sólo el 55% de los casos. En otra de ortografía, la recalificación fue del 72%. Por el contrario, la confiabilidad entre los calificadores en una prueba de matemáticas basada en el desempeño, fue superior a la obtenida en una prueba escrita. En realidad, en ambos casos la evidencia no es definitiva.

Dos críticas frecuentemente citadas en relación con la evaluación auténtica [Reeves y Okey, 1996], son, por un lado, que la misma no permite efectuar sencillas comparaciones entre los estudiantes y no proporciona información que podría constituir base de principios generales a ser aplicados a otros contextos. Los autores citados, conceden que la primer crítica es válida y que debe ser resuelta mediante un consenso más general acerca del propósito de la evaluación. Este consenso podría refocalizar útilmente la noción de comparación con un estándar o un resultado, como opuesto al de otro estudiante.

En cuanto a la segunda inquietud, se refiere a principios generales, y Reeves y Okey afirman que, los proponentes de la evaluación auténtica, los deberían desterrar sobre la base que ellos, buscan **situar el aprendizaje** dentro del contexto del mundo real: *“un mundo en el cual la jactada generabilidad de las pruebas estandarizadas, puede tener escasa relevancia”*.

Este punto es apoyado por Young [1995] quien cree que las necesidades de la evaluación, pueden ser percibidas de un modo más funcional y validado –no sólo por su estabilidad como un instrumento psicométrico- sino más críticamente, por su real utilidad en el mundo real.

LOS ELEMENTOS ESENCIALES DE LA EVALUACIÓN AUTÉNTICA

Teniendo en cuenta las principales lecturas de la bibliografía arriba citada, y lo expresado por otros teóricos en el campo, una lista de las características principales de la evaluación auténtica puede ser agrupada en 4 categorías:

- ❖ el contexto
- ❖ el estudiante
- ❖ la autenticidad de la actividad
- ❖ y los indicadores

Empleando estas orientaciones, se afirmará que la evaluación, con más probabilidad será auténtica, si satisface los siguientes criterios:

CONTEXTO

- Requiere fidelidad del contexto, de modo que refleje las condiciones bajo las cuales operará el desempeño. Ello más bien que condiciones artificiales, o descontextualizadas. [Meyer, 1992; Wiggins, 1993; Reeves & Okey, 1996].

ROL DEL ALUMNO

- Se requiere de los alumnos, que sean actuantes efectivos con conocimiento adquirido y produzcan refinados desempeños o productos [Wiggin,s 1989, 1990, 1993].
- Se requiere significativo tiempo y esfuerzo del estudiante en colaboración con otros [Linn y otros, 1991; Kroll, Masingila y Mau, 1992].

ACTIVIDAD AUTÉNTICA

- Implica desafíos complejos, mal estructurados, que requieren del juicio y de un completo conjunto de tareas. [Wiggins (ya citado), Linn (ya citado) y Torrance, 1995].
- Requiere que la evaluación esté inextricablemente integrada con la actividad. [Young, 1995; Reeves & Okey, 1996].

INDICADORES

- Supone la existencia de múltiples indicadores del aprendizaje. [Lajoie, 1991; Linn y otros, 1991].
- Se procura validez y confiabilidad con criterios adecuados para calificar la variedad de productos. [Wiggins, 1990; Lajoie, 1991; Resnick & Resnick, 1992].

Nota: Este artículo incluye además, un estudio de evaluación auténtica más referido a multimedia interactivos, programados para formar docentes primarios y secundarios.

Bibliografía