

La evaluación alternativa como una herramienta para el aprendizaje

Tiburcio Moreno Olivos

Universidad Autónoma del Carmen, México.

RESUMEN: El papel de la evaluación en educación ha sido crucial, probablemente desde las primeras aproximaciones a la educación formal. El cambio más dramático en nuestra visión de la evaluación está representado por la noción de la evaluación como una herramienta para el aprendizaje. Quizá en el pasado, hemos visto a la evaluación sólo como un medio para determinar la medición y después la certificación, ahora hay una comprensión de que los beneficios potenciales de la evaluación son mucho más amplios y afectan a todas las etapas del proceso de aprendizaje. La evaluación de portafolio, la autoevaluación y la revisión de pares son algunas formas de evaluación que motivan a los estudiantes a investigar continuamente y fomentan un enfoque de aprendizaje profundo. Elementos clave de este enfoque son la reflexión, la retroalimentación, y la integración entre el aprendizaje y la evaluación. En este artículo se destacarán algunos de los principales desarrollos en el aprendizaje de la sociedad y sus implicaciones para la práctica educativa y se reflexionará sobre el futuro de la educación dentro de ambientes de aprendizaje potenciados, donde el aprendizaje, la enseñanza y la evaluación están plenamente integrados.

ABSTRACT: The role of assessment in education has been crucial, probably since the earliest approaches to formal education. The most dramatic change in our views of assessment is represented by the notion of assessment as a tool for learning. Whereas in the past, we have seen assessment only as a means to determine measures and thus certification, there is now realisation that the potential benefits of assessing are much wider and impinge on in all stages of the learning process. Portfolio assessment, self-evaluation and peer-review are some forms of assessment which encourage students to engage continuously and foster a deep approach to learning. Key elements of these approaches are reflection, feedback, and integration of learning and assessment. In this contribution, it will outline some of the major developments in the learning society and their implications for educational practice, and it will reflect on the future of education within powerful learning environments, where learning, instruction and assessment are fully integrated.

Palabras clave: Evaluación alternativa. Evaluación como una herramienta para el aprendizaje. Evaluación de la actuación. Reflexión. Retroalimentación. Integración entre aprendizaje, instrucción y evaluación. Ambientes de aprendizaje potenciados. Evaluación de portafolio, auto-evaluación, evaluación de pares y co-evaluación.

Key words: Alternative assessment. Assessment as a tool for learning. Performance assessment. Reflection. Feedback. Integration between learning, instruction and assessment. Powerful learning environments. Portfolio assessment, self-evaluation, peer-review and co-evaluation.

1. LA ERA DE LA INFORMACIÓN

En la Grecia clásica, los científicos eran considerados expertos en todas las ciencias y áreas de conocimiento. Hasta hace una década, algunos científicos podrían declarar una total comprensión o dominio de todo el conocimiento dentro de una cierta disciplina. Sin embargo, esa era hoy ha pasado definitivamente. La información ahora está cambiando muy rápidamente y el conocimiento está creciendo exponencialmente. No hay posibilidad de que un científico individual posea todo el conocimiento dentro de su disciplina. Los científicos actualmente y en el futuro, necesitarán dominar el conocimiento básico de su campo y después las habilidades necesarias para navegar alrededor de su disciplina. Este cambio también es válido para los profesores, quienes tradicionalmente han sido aceptados como la fuente de todo conocimiento y experiencia dentro del proceso educativo. El profesor ahora está mejor representado como una llave para abrir la puerta a dominios de conocimiento y experiencia. La era de la información, que lidera el siglo XXI, está caracterizada por un cambio infinito y dinámico en la masa de información. El funcionamiento exitoso en esta era requerirá habilidades específicas:

1. *Competencias cognitivas:* solución de problemas, pensamiento crítico, formulación de preguntas, búsqueda de información relevante, formulación de juicios informados, uso eficiente de la información, observación de conductas, investigaciones, invento y creación de nuevas cosas, análisis de datos, presentar datos comunicativamente, expresión oral y escrita.
2. *Competencias metacognitivas:* auto-reflexión, o auto-evaluación.
3. *Competencias sociales:* liderazgo de discusiones, persuasión, cooperación, trabajo en grupos, etc.
4. *Disposiciones afectivas:* perseverancia, motivación interna, auto-eficacia, independencia, flexibilidad o arreglo con situaciones frustrantes.

2. MERCADO LABORAL CAMBIANTE

Recientemente, ha habido un rápido cambio en el mercado ocupacional, que demanda una fuerza laboral más flexible con incremento a corto plazo, a tiempo parcial

y con trabajo ocasional. Hay una creciente presión desde la industria hacia la educación superior para formar graduados que sean inmediatamente empleables y efectivos en los negocios y la industria (Moerkerke, 1996). Las organizaciones empresariales han señalado un insuficiente emparejamiento entre los resultados de los programas de estudio y las necesidades del mercado laboral. Si los sistemas de educación académica y vocacional están requiriendo el suministro de graduados que puedan ser utilizados inmediatamente dentro del mercado laboral, esto implica una creciente necesidad de procedimientos de evaluación para evaluar las habilidades relevantes (De Rijke & Dochy, 1995). La educación superior está requiriendo mejorar los resultados a través de procedimientos de evaluación, de modo que los graduados sean equipados para la labor del mercado. Los especialistas en medición educativa (Birenbaum, 1996; Shavelson, Xiaohong, & Baxter, 1996) han reconocido que, en el futuro cercano, sistemas eficaces y eficientes para la evaluación de la actuación tendrán que ser desarrollados.

3. APRENDIZAJE PARA TODA LA VIDA

En la pasada década, ha habido un centrado aumento en el entrenamiento en la industria y los negocios. La presión económica requirió principalmente de una reestructuración en el mercado laboral, los factores principales fueron dirigidos por el Gobierno y los empleadores, para enfatizar la importancia de la adaptabilidad de la fuerza laboral y la capacidad de los empleados para adquirir nuevas habilidades durante el transcurso de su vida laboral. Es ampliamente aceptada la necesidad de que un aprendizaje para toda la vida se incremente aún más rápidamente en el futuro cercano (Moerkerke, 1996). La investigación en los Países Bajos ha indicado que la sociedad altamente tecnificada de mañana, requerirá personas que sean flexibles y capaces de continuar adquiriendo nuevos conocimientos y nuevas habilidades y aprendizajes (De Rijke & Dochy, 1995).

4. INNOVACIONES EN EVALUACIÓN: CONSECUENCIAS DE LOS DESARROLLOS EN LA SOCIEDAD

Birenbaum (1996) ha planteado una útil distinción entre dos culturas en la medición del logro y las relaciona con los desarrollos del aprendizaje en la sociedad. En la tradicional -así llamada- cultura del test, la enseñanza y la evaluación son consideradas como actividades separadas. La enseñanza es responsabilidad del profesor, mientras que la evaluación es responsabilidad del experto psicómetra, quien puede usar procedimientos elaborados para desarrollar pruebas y modelos psicométricos sofisticados para el análisis de las respuestas de las pruebas. El trabajo de tales expertos dentro de las agencias de pruebas fue estimulado por la demanda de objetividad y justicia en las pruebas y

un alto nivel de estandarización, donde las decisiones importantes estuvieron basadas en las puntuaciones obtenidas en las pruebas. Procediendo de tales demandas, las herramientas comunes para la evaluación en educación eran ítems fáciles de puntuar, tales como los ítems de elección múltiple y los ítems de falso/verdadero. La cultura del test encajó bien con la tradicional concepción de la educación, en la cual la enseñanza es vista como un acto de depositar el contenido que los estudiantes reciben, memorizan y reproducen. La forma más eficiente de evaluar este proceso es mediante pruebas estandarizadas enviadas por agencias de pruebas externas.

El cambio en el aprendizaje de la sociedad ha generado la –así llamada– *cultura de la evaluación* como una alternativa a la cultura del test. La cultura de la evaluación enfatiza fuertemente la integración entre enseñanza y evaluación. Los estudiantes juegan papeles más activos en la evaluación de su logro. La construcción de tareas, el desarrollo de criterios para la evaluación de la actuación y la puntuación de la ejecución podrían ser compartidos o negociados entre profesores y estudiantes. Las tareas de evaluación incluyen todos los tipos de formas, tales como: observaciones, preguntas y pruebas muy enclavadas en el curriculum y el texto, entrevistas, evaluación de la actuación, muestras de trabajos escritos, exposiciones, evaluación de portafolio, y evaluaciones de proyecto y producto. Algunas etiquetas han sido usadas para describir subgrupos de estas alternativas, siendo las más comunes: *evaluación directa*, *evaluación auténtica*, *evaluación de ejecución* y *evaluación alternativa*. Desde un punto de vista psicométrico, pueden ser caracterizadas como evaluaciones mal estandarizadas que dependen en exceso de la instrucción. En estos nuevos contextos, la evaluación será usada más frecuentemente, y tendrá diferentes funciones, diferentes metas y distintas técnicas de administración, además de, nuevas formas de evaluación.

5. NUEVOS ROLES DE LOS PROFESORES EN LOS NUEVOS AMBIENTES DE APRENDIZAJE

En la cultura del test, los profesores eran vistos como cuidadores del conocimiento, el cual tenían que transferir a las cabezas de los estudiantes. De acuerdo con Birenbaum (1996), la cultura de la evaluación está en concordancia con el enfoque constructivista de la educación. En este enfoque, el aprendizaje es visto como un proceso a través del cual el aprendiz crea significado. El profesor no es una persona que transfiere conocimiento, sino alguien que provee oportunidades para que los aprendices usen el conocimiento y las habilidades que ellos ya poseen a fin de lograr la comprensión de nuevos temas. Se espera que el profesor provea tareas interesantes y cambiantes. Cuando los recientes desarrollos, tales como el uso de nuevas herramientas de tecnología educativa, se están tomando en cuenta, parece probable que los cambios aún vayan más

lejos. El profesor tradicional tenía que cumplir un rango de roles diferentes, tales como: diseñador de materiales del curso, tutor del alumno, proveedor de retroalimentación, orientador a través del proceso de aprendizaje, administrador de tests, corrector de tests, etc. En el futuro algunas de estas tareas serán llevadas a cabo por nuevos profesionales altamente especializados tales como el director de un banco de ítems, el diseñador de materiales de curso electrónicos, experto en el desarrollo de multi-media, u orientador/consejero de estudio.

Estos cambios están teniendo lugar con el incremento de movimientos de lo que algunos autores llaman *Ambientes de Aprendizaje Potenciados (PLEs)*. (Dochy & McDowell, 1997). Un número de características de aprendizaje que son relevantes para el diseño de PLEs han surgido de la investigación reciente sobre el aprendizaje y la enseñanza (ver De Corte, 1990, 1991). Dos aspectos principales son: la naturaleza constructiva del aprendizaje y la necesidad de ampliar el aprendizaje a contextos y situaciones de la vida real. Un PLEs está caracterizado por un balance adecuado entre aprendizaje por descubrimiento y exploración personal por un lado, e instrucción sistemática y orientación, por el otro. También toma en cuenta las diferencias individuales de los estudiantes en cuanto a capacidades, necesidades y motivación. Las actividades de aprendizaje constructivo del estudiante deberían ser preferentemente insertadas en contextos que son ricos en recursos y materiales de aprendizaje, que ofrecen amplias oportunidades para la interacción social, y que son representativos de los tipos de tareas y problemas en los que los aprendices tendrán que aplicar su conocimiento y habilidades en el futuro.

6. HACIA LA EVALUACIÓN DE HABILIDADES Y COMPETENCIAS

Desde todas partes del mundo ha habido críticas por el empleo de los tests estandarizados como medida del aprendizaje o competencia (Broadfoot, 1984; Darling-Hammond, 1994; Kellaghan, Madaus, & Airasian, 1982; Wolf, 1995). Newmann y Archbald (1990) argumentan que *muchos datos fallan al medir formas significativas de competencia humana y que nuevas formas significativas de evaluación necesitan ser desarrolladas* (p.164). Ha habido muchos intentos que ilustran el desarrollo de aproximaciones de evaluación más *en contexto y auténticas*. Rowe y Hill (1996) identifican los rasgos clave de estos sistemas más recientes, como:

- Una preocupación por mejores métodos de evaluación e información sobre los actuales logros desarrollados y las competencias individuales de los estudiantes;
- el desarrollo de estándares de marco nacional o de sistema amplio para la evaluación, la puntuación y la información del logro del estudiante;
- el uso de un amplio rango de métodos de evaluación, incluyendo las tareas

comunes o estándar emprendidas por los estudiantes, y;

- un rol significativamente destacado del juicio y la observación del profesor en la evaluación de la competencia en relación a estas tareas.

Nisbet (1993) define el término *evaluación auténtica* como: métodos de evaluación que influyen positivamente la enseñanza y el aprendizaje en formas que contribuyen a realizar los objetivos educacionales, requiriendo tareas realistas (o auténticas) que se desarrollan y focalizan sobre habilidades y contenido relevante, esencialmente similar a las tareas implicadas en los procesos regulares de aprendizaje en el aula (p.35).

Hay un acuerdo general de que habilidades tales como la comunicación, la información tecnológica, y habilidades personales o interpersonales tales como el trabajar con otros, mejorar el propio aprendizaje y la actuación, y la solución de problemas, son importantes. Sin embargo, hay diferencias en la terminología y en la conceptualización de tales habilidades. En la educación superior la palabra 'habilidad' es emotiva y tiende a ser empleada como una forma de hacer distinciones entre educación y entrenamiento, y entre metas académicas y vocacionales.

El término capacidad (*capability*) también ha sido usado para enfatizar la noción holística de una capacidad profesional antes que a una persona con un conjunto de habilidades adquiridas.

7. EL EMPLEO DE LA CO-EVALUACIÓN, LA AUTOEVALUACIÓN Y LA EVALUACIÓN DE PARES

Una nueva forma significativa de evaluación es el uso de la autoevaluación, la evaluación de pares y la coevaluación. El tipo de autoevaluación del estudiante referido más frecuentemente en la literatura, es un proceso que incluye criterios establecidos por el profesor y donde los estudiantes por sí mismos llevan a cabo la evaluación y la puntuación. Otra forma de autoevaluación del estudiante es el caso donde un alumno/a se evalúa asimismo/a sobre la base de criterios que él mismo o ella misma ha seleccionado, la evaluación es para la orientación personal del estudiante o para comunicarla al profesor o a otros. Según Hall (1995) hay dos factores críticos para una genuina *autoevaluación*: el estudiante no solamente lleva a cabo la evaluación, sino que también selecciona los criterios en los que la evaluación está basada. Similarmente, la *evaluación de pares* puede indicar que son los estudiantes-compañeros quienes seleccionan los criterios y efectúan la evaluación. Ante cualquier situación en la que el tutor/a y el alumno/a comparten la selección de criterios y/o la realización de la evaluación, será más exacto emplear el término *co-evaluación* (Hall, 1995). Sin embargo, aún es frecuente el caso en que los profesores controlan el proceso de evaluación, algunas veces apoyados por equi-

pos profesionales o expertos en evaluación, mientras los criterios y las evaluaciones hechas por los alumnos son tomados seriamente pero considerados adicionales –antes que reemplazo– a la evaluación emprendida por el maestro (Roger, 1995).

La implementación de formas de autoevaluación, evaluación de pares y co-evaluación podría disminuir la inversión de tiempo de los profesores, que por otra parte, lo necesitan para hacer evaluaciones más frecuentes. Además de la ventaja de emplear estas formas de evaluación para apoyar el desarrollo de ciertas habilidades de los estudiantes, por ejemplo, habilidades de comunicación, auto-evaluación, observación y auto-crítica.

8. DIFERENTES FUNCIONES DE LA EVALUACIÓN EN EL PROCESO DE APRENDIZAJE

Se ha demostrado que la evaluación tiene un impacto significativo en la enseñanza y el aprendizaje. Un número de autores ha informado de los efectos negativos de la evaluación en la enseñanza y el aprendizaje (Frederiksen, 1984; Ridgway & Schoenfeld, 1994). Las malas prácticas de evaluación, a menudo, pueden ser responsables de una enseñanza y un aprendizaje de baja calidad.

Las pruebas que están vinculadas inadecuadamente a la instrucción conducen a consecuencias indeseables, tales como, información inapropiada acerca del progreso y las dificultades del aprendizaje, reducción de la motivación de los estudiantes para aprender y evaluación incorrecta de la efectividad de la instrucción. Efectos negativos han sido reportados, aún con pruebas que están bien unidas a la instrucción, debido también a la frecuente evaluación formal (Tan, 1992). En los Estados Unidos, donde las pruebas que evalúan la actuación del estudiante estuvieron unidas a inferencias acerca de la competencia del profesorado y de las escuelas, se observó que los centros y el profesorado restringieron sus currícula y sus cursos a la meta de ayudar al alumnado a superar las pruebas de las agencias externas (Baker, 1994).

Para una evaluación efectiva, existe la necesidad de reconceptualizar los instrumentos de evaluación convencionales y la teoría que los fundamenta (Glaser, 1990). Tales instrumentos no siempre son congruentes con las concepciones actuales, basadas en la investigación, acerca de la adquisición de conocimiento y habilidades.

Algunos argumentan que las buenas prácticas de evaluación pueden usarse para mejorar la instrucción. La mayor parte de los especialistas en evaluación, asume la postura de que las evaluaciones educativas usadas apropiadamente constituyen una herramienta potente que realza el proceso instruccional. Una condición importante para el adecuado uso de las pruebas, es que éstas se encuentren unidas o integradas con los procedimientos y materiales instruccionales. Una de las cuestiones principales es la claridad acerca del tipo de decisión instruccional que es apoyado por el procedimiento de eva-

luación. Se identifican cuatro tipos de decisiones instruccionales que son apoyadas por las pruebas:

1. *Decisiones de colocación*: decidir en que secuencia instruccional o en que nivel en la secuencia instruccional un estudiante deberá iniciar a fin de evitar la repetición innecesaria de lo que ya es conocido y permitir un logro más rápido de las nuevas metas. Las pruebas del estado del conocimiento previo se dan antes de que un estudiante inicia una unidad de instrucción. Ellas pueden centrarse en las habilidades y el conocimiento prerrequeridos para el aprendizaje propuesto o en las habilidades y los conocimientos resultantes para determinar si los estudiantes ya han logrado de la instrucción el resultado deseado. Así, los profesores emplean las pruebas del estado del conocimiento previo para ubicar a los estudiantes en modos instruccionales para (a) determinar el grado en el que las habilidades o conductas como prerrequisito de entrada están presentes o ausentes, (b) determinar el dominio de entrada de los objetivos del curso, y también para (c) marcar a los estudiantes modelos de instrucción alternativos, basados en sus características.
2. *Decisiones de monitoreo*: decidir (a) si los estudiantes están aprendiendo apropiadamente, y (b) si la actividad de trabajo asignada está siendo trabajada efectivamente, o debería ser asignada una actividad diferente. La llamada evaluación del progreso o evaluación del crecimiento son los instrumentos para la evaluación del progreso de un estudiante hacia el logro los objetivos de aprendizaje finales. Las evaluaciones del progreso proveen retroalimentación a los estudiantes acerca de como avanza su aprendizaje y pueden dirigir la decisión (formativa) de que hay una necesidad de remediar o re-aprender una cierta unidad de instrucción. Las evaluaciones del progreso son usadas por el profesorado para (a) elegir o modificar subsecuentes actividades de aprendizaje, (b) prescribir acciones remediales para atender las deficiencias de un individuo o de un grupo, y (c) proveer una continua retroalimentación al alumno/a con el propósito de dirigir su avance o su estudio remedial. Las pruebas para evaluar el progreso son aplicadas durante la enseñanza, pero miden los procesos cognitivos y las conductas del estudiante a nivel de los resultados deseados.
3. *Decisiones de logro*: decidir, al final o en un segmento de instrucción particular si un alumno ha conseguido las metas de instrucción deseadas. Las pruebas finales son empleadas para certificar el aprendizaje del alumno/a, y evalúan la efectividad del profesor.
4. *Decisiones de diagnóstico*: decidir que resultados de aprendizaje no ha adquirido un estudiante y la causa (s) probable del fracaso para adquirirlos, a fin de remediar, completar correctamente o corregir los errores del aprendizaje previo.

Las evaluaciones de diagnóstico son diseñadas para proveer información específica acerca de malas interpretaciones y deficiencias de aprendizaje individual. Cuando las evaluaciones del estado del conocimiento previo, del progreso y finales, son diseñadas e interpretadas apropiadamente, también pueden proporcionar información de naturaleza diagnóstica, que profesores y estudiantes pueden usar para regular los procesos de aprendizaje.

9. EL FUTURO: APRENDIZAJE INTEGRADO, ENSEÑANZA Y EVALUACIÓN A TRAVÉS DE UN APRENDIZAJE TRANSFORMATIVO Y FLEXIBLE EN AMBIENTES DE APRENDIZAJE POTENTES

Antiguamente, y quizá aún hoy en día, los profesores algunas veces eran vistos como lagartos nadando en el agua profunda, esperando atrapar al incauto estudiante que cometiera un error mientras se sometía al examen. Actualmente, hay un movimiento para reducir este aspecto amenazador de los exámenes y para ver a la evaluación como un medio para ayudar a los estudiantes y a los profesores ha alcanzar sus metas. Parece cierto que la enseñanza desarrollará lo que podríamos llamar actividades que apoyan el aprendizaje transformativo del estudiante. En menos de una década, con el uso de más evaluación formativa en las aulas, que actúe como una palanca, podríamos ver mucho más la integración entre enseñanza y evaluación.

9.1 ¿La profecía de la evaluación global?

Por muchos años, Glaser (1981, 1990) ha sido un defensor de lo que se ha dado en llamar la profecía de la evaluación global. Esta profecía sostiene que ya no es posible considerar a la evaluación solamente como un medio para determinar que individuos ya están adaptados a o tienen el potencial para adaptarse a prácticas educativas dominantes. Una meta alternativa es revertir esta secuencia de adaptación; es decir, antes que requerir que los individuos se adapten a los medios de la instrucción, podríamos adaptar las condiciones de la instrucción a los individuos a fin de maximizar su potencial para el éxito.

Este objetivo puede alcanzarse si el aprendizaje puede ser diseñado tomando en cuenta las habilidades y perfil de conocimiento del individuo, pero primero necesitamos saber como lograr este fin. ¿Deberíamos tener como meta una instrucción conducida por la evaluación o una evaluación manejada por la instrucción?. Es cierto el caso de que mucha de la práctica educativa es gobernada por la evaluación; los profesores enseñan para los exámenes. No obstante, esto ha probado ser problemático. Se ha demostrado que la enseñanza dirigida a las pruebas, en muchos casos, sacrifica la profundidad del aprendizaje tanto del conocimiento como de las habilidades.

Nuestra visión es que la valoración de habilidades de orden superior por medio de evaluaciones auténticas dirigiría la enseñanza de tales habilidades y conocimiento de orden superior. Estamos de acuerdo con Knight (1996), que la reforma de la enseñanza hacia la profecía de la evaluación global, puede ser lograda por medio de una cuidadosa reforma de la evaluación. Algunos autores han argumentado (Birenbaum & Dochy, 1996) que el asunto de la enseñanza regida por las pruebas toma un cariz diferente cuando es asociada con las nuevas ideas de evaluación alternativa, de actuación o auténtica. Primero, la asunción hecha de que la evaluación alternativa será la que conduzca la enseñanza usando ítems de pruebas auténticas, relacionadas directamente con la instrucción. Segundo, hay el supuesto de que la evaluación alternativa tendrá un efecto de consecuencias positivas (Biggs, 1996) en la instrucción, haciendo que ésta se encuentre mejor vinculada a experiencias de la vida real y resulte en un aprendizaje más activo.

9.2 Apoyo al aprendizaje del estudiante por medio de la evaluación formativa

Para Knight (1996) las ventajas clave de la evaluación formativa:

- Implica dar retroalimentación a los aprendices, que puede ser usada par mejorar la próxima actuación.
- Está directamente conectada a la noción transformacional de calidad, puesto que ésta es sobre la mejora del aprendizaje del estudiante.
- No es sólo conducida por los tutores, la co-evaluación, la evaluación de pares y la auto-evaluación son posibles, y argumentalmente, necesarias.
- No es una evaluación de alto riesgo.
- Cuando está asociada con un diálogo entre aprendiz y profesor, la fiabilidad inicial de la evaluación no necesita ser alta, lo que pasa es que está dispuesta la base para una conversación. Sin embargo, si habrá una comunicación entre el asesor y el asesorado, aquí sí necesita haber un entendimiento del significado operacional de los criterios relevantes.

La evaluación formativa podría ser más útil cuando provee información de alta calidad sobre las actuaciones de los individuos pero es difícil para los profesores encontrar más tiempo para evaluar uno a uno. ¿Cómo, entonces, podría ser promovida la evaluación formativa?. Tres formas son sugeridas por Knight:

- Haciendo más uso de las ayudas de la enseñanza y la investigación. Sin embargo, ellas necesitan entrenamiento; cuestan dinero, y tienen la presión de sus propios compromisos de carrera y de investigación.
- Hacer más uso de los paquetes de evaluación y aprendizaje basado en ordenador. Los sistemas de evaluación y de aprendizaje basados en los principios de la inteligencia artificial ofrecen gran promesa en algunas áreas, especialmente donde

la actuación y probablemente los errores, pueden ser totalmente definidos con precisión. No obstante, ellos demandan una considerable inversión inicial que solamente podría recuperarse dentro de muchos años.

Según Nevo (1997), la función formativa es la más constructiva de la evaluación, pero para ello se deben cumplir ciertas condiciones: (1) tiene que ser muy detallada y brindar información específica para la mejora, en vez de dar puntuaciones aisladas, (2) ha de ser lo menos agresiva posible para el alumno y para el profesor, en caso de que la lleve a cabo otra persona, para no fomentar la autodefensa y el rechazo, (3) debe ser oportuna si los datos se quieren utilizar para mejorar.

La opción de la auto-evaluación, la evaluación de pares y la co-evaluación parece atrayente. No solamente asegura que los aprendices comprendan el significado de los criterios y los apliquen seriamente, también promete importantes habilidades que vienen con el juicio y la auto-valoración (Boud, 1992). Donde los procesos de aprendizaje son evaluados, la auto-evaluación, la evaluación de pares y la co-evaluación podrían ser las únicas formas válidas de hacerlo, y el valor e importancia de evaluar los procesos de aprendizaje fue mostrado en los hallazgos de la investigación, en cuanto a las habilidades que la industria valora en los nuevos graduados (Harvey & Knight, 1996).

Un problema de la auto-evaluación, la evaluación de pares y la co-evaluación es que podrían cambiar la carga del tiempo del profesor al estudiante. Hasta que los estudiantes no sean asesores prácticos les tomará más tiempo que a un profesor con experiencia hacer un juicio formativo. Esto implica, que son necesarias muy buenas herramientas o marcos de evaluación para auxiliar a los estudiantes como asesores, o quizá, que tendrá que ser cubierto menos material, a fin de dejar tiempo disponible para que los estudiantes empleen estas nuevas estrategias de evaluación. Sin embargo, Knight (1996) argumenta que mientras algunos autores aceptan el destacado aprendizaje en profundidad que la auto-evaluación, la evaluación de pares y la co-evaluación ofrecen, otros valoran más la amplitud del aprendizaje y llegan a preocuparse si se restringe el contenido del programa que hay que cubrir. Aquí parece haber una necesidad por:

- La consistencia entre las metas del programa, la enseñanza, el aprendizaje y los procedimientos de evaluación;
- el uso de un amplio rango de métodos y modos de evaluación;
- la provisión de oportunidades para el desarrollo de significados compartidos de los criterios de evaluación;
- una plena evaluación formativa, tanto a nivel de aula como a nivel individual;
- el uso de la auto-evaluación, evaluación por pares y co-evaluación;
- una preocupación por la profundidad de comprensión en preferencia a la amplitud o extensión del conocimiento, y
- provisión de oportunidades a todas las facultades para aprender de formas de

enseñanza que son consistentes con un énfasis en el aprendizaje del estudiante y en el potencial transformativo de la educación superior.

9.3 Evaluación y conducta del estudiante

La investigación en la enseñanza y el aprendizaje muestra que en el aprendizaje 'la cola mueve al perro', esto quiere decir, que lo que los estudiantes hacen está muy guiado por las formas en las que son evaluados (Gibbs, 1992). Las pruebas y exámenes tradicionales han tendido a fomentar la memorización y el aprendizaje mecánico aún cuando el alumnado se haya propuesto algo más bien diferente. Los métodos de evaluación alternativa tales como tareas escritas variadas, periódicos científicos, Pruebas Globales, portafolios, proyectos de grupo, presentaciones orales, uso de estudios de caso y simulaciones, ofrecen oportunidades para hacer de la evaluación una experiencia de aprendizaje valiosa, además de permitir niveles para que sea asignada (Birenbaum & Dochy, 1996, Brown & Knight, 1994). McDowell (1995) sugiere que los estudiantes encontrarán formas más nuevas de evaluar, intrínsecamente interesantes y motivantes. Ellos se mantendrán conscientes de la necesidad de lograr altos niveles, pero no tenderán a centrarse exclusivamente en sus logros. Hay evidencias de que ellos aprenden y se comportan diferentemente en lo que hacen en cursos donde hay un examen o prueba tradicional. Así, la evaluación alternativa podría contribuir a cambiar la cultura entre los estudiantes desde una cultura del test a una cultura de la evaluación (Birenbaum & Dochy, 1996). Un mayor empleo de la evaluación formativa es una forma de convencer a los estudiantes de que la evaluación tiene los propósitos de indicar sus fuerzas, sus debilidades y sus progresos, y de guiarlos hacia sus metas de aprendizaje, además de verificar sus niveles de logro.

La investigación muestra que los métodos de evaluación alternativa son menos amenazantes para la mayor parte de los estudiantes que los exámenes tradicionales. Ellos son percibidos como pruebas más justas (Sambell, McDowell & Brown, 1997). Además, los estudiantes parecen encontrar significado en tareas tales como los proyectos, los ejercicios en grupo y los portafolios, quizá porque son más vistos como actividades de la vida real, más apropiadas a Ambientes de Aprendizaje Potentes, que los exámenes. Tal evaluación enfatiza la motivación intrínseca de los estudiantes, aunque los niveles de motivación extrínseca permanezcan presentes. Cambiar los métodos de evaluación es una forma efectiva de motivar a los estudiantes a cambiar sus métodos de aprendizaje. Cuando la evaluación sigue siendo la misma, a menudo, los estudiantes no aceptan la necesidad de cambiar sus enfoques de aprendizaje; por ejemplo, los estudiantes frecuentemente se preparan para los exámenes mediante un aprendizaje maquinal aún cuando éste no sea apropiado. La evaluación es una de las herramientas más efectivas para innovar tanto la enseñanza como el aprendizaje.

10. Conclusiones

Creemos que la cultura de la evaluación puede ser usada para cambiar la instrucción desde un sistema que transfiere conocimiento a las cabezas de los estudiantes a uno que trate de desarrollar a los estudiantes, quienes son capaces de *aprender como aprender*. La sociedad actual y el contexto tecnológico requieren que la educación haga tal cambio. Aunque han sido reportados resultados prometedores y entusiastas, los primeros intentos para introducir nuevos procedimientos de evaluación no han sido totalmente positivos. Madaus y Kellaghan (1993) informaron de problemas con la organización, el tiempo y los costes de los programas de evaluación.

Numerosas lecciones pueden aprenderse de las primeras aplicaciones de los programas de evaluación alternativa. En primer lugar, uno no debería tirar al bebé fuera junto con el agua de la bañera. Las pruebas objetivas son muy útiles para ciertos propósitos, tales como la evaluación sumativa de alto-riesgo, del logro de un individuo, aunque éstas no deberían dominar un programa de evaluación. Cada vez más, los especialistas en medición recomiendan los denominados programas de evaluación pluralísticos o balanceados, donde sean empleados formatos de evaluación múltiples (Birenbaum, 1996; Clarke & Stephens, 1996; Ridgway & Schoenfeld, 1994). Hay varios motivos para estos programas de evaluación pluralísticos: un sólo formato de evaluación no puede servir a varias tomas de decisiones y propósitos diferentes, y cada formato de evaluación tiene su propio método de discrepancia, que interactúa con las personas.

Hay una necesidad de establecer un sistema de valoración de la calidad de la evaluación alternativa y del control de la calidad implementada (Birenbaum, 1996). Varios autores recientemente han propuesto formas para ampliar los criterios, técnicas y métodos usados en la psicometría tradicional (Cronbach, 1988; Kane, 1992; Linn, Baker & Dunbar, 1991). Otros como Messick (1984; 1995), oponen la idea de que los criterios deberían ser especificados y declaran que el concepto de validez de constructo aplicado a todas las mediciones psicológicas y educativas, incluya a la evaluación de actuación/ejecución.

Hay una necesidad de heurísticos bien diseñados y bien evaluados que ayuden a los profesores a implementar y diseñar procedimientos de evaluación alternativa de alta calidad. Los profesores necesitan criterios claros para el desarrollo y uso de procedimientos de evaluación, y técnicas fáciles de usar para el control y mejora de la calidad.

No hay ningún formato de evaluación ideal para todos los casos. Todos los formatos de evaluación pueden tener efectos negativos en la enseñanza y el aprendizaje. Desde esta perspectiva, la crítica de Frederiksen (1984) puede ser interpretada de forma más amplia que simplemente como una exposición de los efectos negativos de los ítems de opción-múltiple; también muestra que algunos aspectos de cualquier evaluación pueden

tener consecuencias educativas tanto positivas como negativas. Sin embargo, la evaluación alternativa con su potencial para desarrollar y evaluar los logros de aprendizaje que son congruentes con la visión contemporánea de la naturaleza del conocimiento y las necesidades de la sociedad de personas educadas, es una avenida que debemos continuar explorando.

Bibliografía:

- Baker, E. (1994). Learning based assessments of history understanding. *Educational Psychology* 29 (2), 97-106.
- Biggs, J. (1996). Assessing learning quality: Reconciling institutional, staff and educational demands. *Assessment & Evaluation in Higher Education*, 21 (1), 5-15.
- Birenbaum, M., & Dochy, F. (Eds.) (1996). *Alternatives in assessment of achievements, learning process and prior knowledge*. Boston: Kluwer.
- Birenbaum, M., & Dochy, F.J.R.C. (1996). Introduction. In M. Birenbaum & F.J.R.C. Dochy (Eds.), *Alternatives in assessment of achievements, learning process and prior knowledge*. (pp.xiii-xv). Boston: Kluwer.
- Birenbaum, M. (1996). Assessment 2000: Toward a pluralistic approach to assessment. In M. Birenbaum & F.J.R.C. Dochy (Eds.), *Alternatives in assessment of achievements, learning process and prior knowledge*. (pp.3- 30). Boston: Kluwer.
- Boud, D. (1992). The use of self-assessment schedules in negotiated learning. *Studies in Higher Education*, 18 (5), 529-549.
- Broadfoot, P.M. (1984). From public examination to profile assessment: The French experience. In P.M. Broadfoot (Ed.), *Selection certification and control*. Sussex: Falmer.
- Brown, S., & Knight, P. (1994). *Assessing learners in higher education*. London: Kogan Page.
- Clarke, D., & Stephens, M. (1996). The ripple effect: The instructional impact of the systematic introduction of performance assessments in mathematics. In M. Birenbaum & F.J.R.C. Dochy (Eds.), *Alternatives in assessment of achievements, learning process and prior knowledge*. (pp.63- 93). Boston: Kluwer.
- Cronbach, L.J. (1988). Five perspectives on validity argument. In H. Wainer & H.I. Braun (Eds.), *Test validity* (pp.3-17). Hillsdale, NJ: Erlbaum.
- Darling-Hammond, L. (1994). Performance based assessment and educational equity. *Harvard Educational Review*, 64, 5-29.
- De Corte, E. (1990). Toward powerful learning environments for the acquisition of problem solving skills. *European Journal of Psychology of Educational*, 5, 519-541.

- De Rijke R., & Dochy, F. (1995). *Assessment centres in onderwijs, opleiding en HRM* (Assessment centres in Education, training, and Human Resources Management). Utrecht: Lemma.
- Dochy, F.J.R.C., & McDowell, L. (1997). Assessment as a tool for learning. *Studies in Educational Evaluation*, 23 (4), 279-298.
- Frederiksen, N. (1984). The real test bias, influences of testing on teaching and learning. *American Psychologist*, 39 (3), 193-202.
- Gibbs, G. (1992). Improving the quality of student learning through course design. In R. Barnett (Ed.), *Learning to effect*. Milton Keynes, UK: SRHE & Open University Press.
- Glaser, R. (1981). The future of testing. A research agenda for cognitive psychology and psychometrics. *American Psychologist*, 36 (9), 923-936.
- Glaser, R. (1990). Toward new models for assessment. *International Journal of Educational Research*, 14, 375-483.
- Hall, K. (1995). *Co-assessment: Participation of students with staff in the assessment process*. Invited address at the 2nd EECAE Conference (European Electronic Conference on Assessment & Evaluation), EARLI-AE list, March 10-14.
- Harvey, L., & Knight, P. (1996). *Transforming higher education*. Buckingham. Open University Press.
- Kane, M.T. (1992). An argument-based approach to validity. *Psychological Bulletin*, 112 (3), 527-535.
- Kellaghan, T., Madaus, G.F., & Airasian, P.W. (1982). *The effects of standardized testing*. Boston: Kluwer-Nijhoff.
- Knight, P. (1996). *Quality in higher education and the assessment of student learning*. Invited paper at the Third European Electronic Conference on Assessment and Evaluation, March 4-8, EARLI-AE list European Academic & Research Network (EARN).
- Linn, R.L., Baker, E.L., & Dunbar, S.B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher*, 20 (8), 15-21.
- Madaus, G.F., & Kellaghan, K. (1993). British experience with authentic testing. *Phi Delta Kappan*, 74, 458-469.
- McDowell, L. (1995). The impact of innovative assessment on student learning. *Innovations in Education and Training International*, 32 (4), 302-313.
- Messick, S. (1984). The psychology of educational measurement, *Journal of Educational Measurement*, 21, 215-238.
- Messick, S. (1995). Validity of psychological assessment. *American Psychologist*, 50 (9), 741-749.
- Moerkerke, G. (1996). *Assessment for flexible learning*. Utrecht: Lemma.

- Nevo, D. (1997). *Evaluación basada en el centro. Un diálogo para la mejora educativa*. Ediciones Mensajero. España.
- Newmann, F.M., Archbald, D.A. (1990). Organizational performance of schools. In P. Reyes (Ed), *Teachers and their workplace: Commitment, performance and productivity*. Newbury Park, CA: Sage.
- Nisbet, J. (1993). Introduction. In *OECD- curriculum reform: Assessment in question* (pp.25-38). Paris: Organisation for Economic Cooperation and Development.
- Ridgway, J., & Schoenfeld, A. (1994). *Balanced assessment: Designing assessment schemes to promote desirable change in mathematics education*. Invited paper at the First European Electronic Conference on Assessment and Evaluation, February 21-23, EARLI-AE list European Academic & Research Network (EARN).
- Roger, P. (1995). *Validity of assessments*. Contribution to the 2nd EECAE conference (European Electronic Conference on Assessment & Evaluation), March 10-14, EARLI-AE list European Academic & Research Network (EARN).
- Rowe, K.J., & Hill, P.W. (1996). Assessing, recording and reporting students' educational progress: the case for subject profiles. *Assessment in Education*, 3 (3), 309-352.
- Sambell, k., McDowell, L., & Brown, S. (1997). 'But is it fair?': An exploratory study of student perceptions of the consequential validity of assessment. *Studies in Educational Evaluation*, this issue.
- Shavelson, R.J., Xiaohong, G., & Baxter, G. (1996). On the content validity of performance assessments: Centrality of domain-specifications. In M. Birenbaum & F.J.R.C. Dochy (Eds.), *Alternatives in assessment of achievements, learning process and prior knowledge*. (pp.131-142). Boston: Kluwer.
- Tan, C.M. (1992). An evaluation of the use of continuous assessment in the teaching of physiology. *Higher Education*, 23, 255-272.
- Wolf, A. (1995). *Competence-based assessment*. Buckingham: Open University Press.