

LA EVALUACION EN LA DIDACTICA DE LAS CIENCIAS

F. JAVIER PERALES PALACIOS

RESUMEN

La evaluación constituye una de las actividades educativas cuyo papel está cambiando más rápidamente durante las últimas décadas, al menos desde el punto de vista teórico y desde las instancias educativas oficiales. En este artículo se pretende ofrecer una panorámica sobre su naturaleza, clasificación e investigación, así como de las particularidades que adquiere en el contexto de la educación científica. Finalmente, se aportan algunas recomendaciones para conseguir una potenciación real de la evaluación en la praxis educativa.

ABSTRACT

Along the last years, educational assessment is playing an increasingly important role, at least according to theoretical criterions and official curricula. This paper tries to show a perspective on assessment nature, classification, research and on their specifications in the context of science teaching. Finally, some suggestions to attain a real reinforcement of assessment in educational praxis are offered.

PALABRAS CLAVE

Evaluación, Didáctica de las Ciencias.

KEYWORDS

Assessment, Science Education.

1. INTRODUCCION

La evaluación educativa ha constituido tradicionalmente uno de los elementos del proceso de enseñanza-aprendizaje más contestados, tanto desde la perspectiva del profesor como desde la del alumno. Para el primero evaluar se identifica con tener que preparar exámenes y corregirlos posteriormente -una de las tareas más aburridas de su actividad, a juicio de muchos profesores- y para el segundo tener que estudiar intensivamente, pasar el mal trago del examen y *beber para olvidar*. No obstante, a casi nadie escapa la necesidad de evaluar o, por lo menos, no se concibe una enseñanza sin ningún tipo de control. Lo que sucede es que, en general, se obvia la riqueza subyacente al proceso de evaluación, limitándose a utilizar aspectos muy parciales del mismo y que, en general, implican un menor esfuerzo.

Esta reflexión es en gran medida aplicable a la práctica docente habitual y, en nuestro caso, a la enseñanza de las ciencias. Ciertamente hay que reconocer el esfuerzo investigador realizado esencialmente desde el ámbito pedagógico por crear nuevas fórmulas evaluadoras y dotar de coherencia al proceso, pero sólo muy recientemente los resultados obtenidos están comenzando a extrapolarse tímidamente a la Didáctica de las Ciencias gracias al movimiento renovador que ha supuesto el surgimiento del Modelo Constructivista.

2. CONCEPTO Y TIPOS DE EVALUACION

Si nos interrogamos sobre el significado del término "evaluación" podríamos hallar en el diccionario algo parecido a *hacer un juicio de valor sobre algo*, es decir, *valorar, estimar el valor de una cosa*. Esto inmediatamente nos lleva a constatar la subjetividad inherente a este tipo de actuación que se inscribiría en el viejo dilema filosófico de *lo real y lo conocido* y que aquí se traduciría en *lo real y lo valorado*. Este significado de la evaluación podría ser aplicado, tanto en el campo educativo como en cualquier otra dimensión de la vida cotidiana (p. ej., *la evaluación del impacto ambiental* de una obra pública). Centrándonos en el primero, la evaluación se ha ido extendiendo desde el propio aprendizaje hasta abarcar otros componentes del mismo, tales como el profesor, el currículum, la metodología, el centro, etc. Dada la naturaleza de este trabajo, nos vamos a ocupar prioritariamente de *la evaluación del aprendizaje científico*.

Si nos detenemos en los *objetivos de la evaluación* nos van a interesar especialmente los siguientes, entre los múltiples que se citan para su justificación:

- Para promover y mejorar el aprendizaje.
- Para valorar el grado de cumplimiento de los objetivos educativos.
- Para diagnosticar inicialmente las ideas previas, errores conceptuales, habilidades o actitudes de los alumnos.
- Para analizar el cambio conceptual, de habilidades o de actitudes.
- Para estudiar las posibles causas de un aprendizaje deficiente y modificarlas.

En cuanto a los *tipos de evaluación*, ya abordados extensamente en la literatura educativa (p. ej., Lafourcade, 1972; Sáenz y col., 1989), vamos a ir deteniéndonos en aquellas pruebas que puedan poseer un interés más concreto para la Didáctica de las Ciencias.

• *Inicial, formativa y sumativa*. Se trataría, respectivamente, de obtener un diagnóstico del nivel de conocimiento de los alumnos antes de abordar el contenido, durante este proceso, o al final del mismo. La evaluación inicial o diagnóstica ha resultado ser la base de múltiples estudios en ciencias a la búsqueda de una descripción de las concepciones que poseen los alumnos de distintos niveles educativos en torno a conceptos científicos básicos, como paso previo a la consecución del cambio conceptual, lo que ha generado un gran volumen de investigación (véase, p. ej., Driver y Watts, 1990).

• *Criterial*. Se elaboran inicialmente unos criterios explícitos de corrección individual de las pruebas, teniendo en cuenta los objetivos instructivos y las distintas posibilidades y "calidad" de las respuestas (véanse, p. ej., George y cols., 1977; Horne, 1984). Esto resulta especialmente recomendable en el caso de pruebas de evaluación típicas de las ciencias, tales como la resolución de problemas o la ejecución de prácticas de laboratorio. En los primeros se podría abordar mediante procesos tales como la expresión de datos, la elaboración de hipótesis, la representación gráfica de los fenómenos, etc. Así, por ejemplo, si de acuerdo con los objetivos de la disciplina se considera esencial que los alumnos sepan utilizar adecuadamente los sistemas de unidades o la formulación y nomenclatura químicas, dichos objetivos constituirán un "criterio" ineludible para abordar la evaluación de un problema. Esta clase de evaluación sólo ha sido tímidamente introducida en la enseñanza de las ciencias durante los últimos años. Un estudio piloto llevado a cabo por Satterly y Swann (1988) entre alumnos y profesores de Bristol para utilizar el nuevo sistema de evaluación condujo a resultados positivos frente a la evaluación normativa, en el

sentido de ser percibido como de mayor utilidad; de cualquier modo, persisten serias incógnitas sobre sus resultados a largo plazo que debieran ser clarificadas por trabajos de investigación ulteriores.

- *Pruebas de lápiz y papel.* Varían desde las que requieren la elaboración de algún tipo de respuesta hasta las que precisan de una selección entre varias opciones. Son las primeras las de mayor utilidad en el campo científico en forma de problemas cualitativos ("cuestiones") o de problemas cuantitativos ("problemas").

Las cuestiones se revelan como un buen test para determinar el grado de dominio en la interpretación de los fenómenos a la luz de la ciencia, objetivo de primer orden en el aprendizaje del alumno. Constan sólo de unos datos cualitativos relativos al fenómeno a analizar y suelen requerir una respuesta del mismo tipo bajo la forma de "interprete", "justifique", "explique", etc. Ejemplo: *¿Cómo podría determinar la velocidad con que caen las gotas de lluvia por las estelas que dejan en las ventanas de un tren en movimiento, si no hace viento en el exterior? Explíquelo.*

En cuanto a los propiamente llamados problemas, constituyen básicamente el eje central de la práctica evaluadora y han generado asimismo una fructífera línea de investigación, tanto en el campo psicopedagógico como en el de la Didáctica de las Ciencias y de las Matemáticas (Perales, 1993). No obstante, esta importancia no suele corresponderse con una atención adecuada por parte de los educadores, cayéndose con demasiada frecuencia en desvirtuar el proceso a través de la resolución de "problemas-tipo", es decir, con unas características similares en su enunciado y resolución, con lo que el acto evaluador se convierte en una reproducción del modelo ya resuelto con anterioridad. (Ejemplo: *Una membrana rectangular tiene como dimensiones 10 cm y 20 cm, su masa es de 5 g y está tensada con 0,1 kg/cm. Calcule las frecuencias para sus cuatro modos de vibración más sencillos y las líneas nodales para estos casos.*)

Son múltiples las variables a considerar en la resolución de problemas y afectan a su naturaleza (p. ej., enunciado, complejidad, tarea requerida, etc.), a su contexto de resolución (consulta de fuentes de información, manipulación de objetos, verbalización de la resolución, tiempo, etc.) y al solucionador del problema (conocimiento previo, habilidades cognitivas, creatividad, actitud, etc.). En el primer grupo, es decir la naturaleza del problema, se han acometido, entre otros, estudios relativos al conocimiento del resultado previamente a su resolución (Perales y Cervantes, 1984) o con posterioridad (Escudero, 1979), detectándose influencias positivas sobre su resolución en comparación con su desconocimiento, se han elaborado instrumentos para evaluar el enunciado (Fernández y col., 1984) y se ha investigado asimismo el nivel de información suministrado por éste (Kempa, 1986) y la lectura hecha del mismo (Fauconnet, 1984).

En cuanto al propio solucionador del problema, el volumen de investigación acumulado es muy considerable. Destaquemos algunos resultados (Perales, 1993):

- El conocimiento específico de la materia implicada así como la familiaridad con el tipo de tarea propuesta y las habilidades cognitivas precisas presentan una significativa influencia sobre la capacidad de resolver problemas.

- Especial relevancia han merecido los estudios que relacionan el estilo cognitivo "dependencia-independencia de campo" (Witkin y cols., 1971) con la facultad para resolver problemas, viéndose favorecidos los individuos independientes.

- La creatividad aparece como un buen predictor de esta tarea en el caso de que se trate de problemas de naturaleza abierta, pero no tiene por qué serlo cuando los problemas poseen una solución lógica ya conocida.

- En cuanto al sexo, parece ser que la diferencia entre varones y hembras varía desde la infancia hasta la adolescencia, primando a las segundas en los intervalos de menor edad y a los primeros a partir de la adolescencia.

Para finalizar haremos referencia a las pruebas que requieren el ordenamiento de un contexto y que pueden ser de tipo cronológico, lógico o metodológico. Estas dos últimas han merecido un interés creciente adoptando, respectivamente, la forma de *mapas conceptuales* y *diagramas V* (Novak y Gowin, 1988), cuya importancia ha trascendido al propio papel de instrumento de evaluación al haberse sugerido como una poderosa herramienta de aprendizaje (Perales, 1990).

- *Pruebas prácticas, de ejecución o funcionales.* Tal y como su nombre indica, en el contexto de las ciencias experimentales se trataría de evaluar al alumno mediante técnicas que permitan conocer si domina determinadas habilidades, ya sean de tipo manipulativo o de experimentación, y que son conocidas con otras denominaciones tales como prácticas de laboratorio, trabajos prácticos, etc. Tales actividades de aprendizaje están sometidas asimismo a una profunda revisión de su papel en el proceso de enseñanza-aprendizaje (Perales, 1994) en el sentido de convertirlas en *pequeñas investigaciones* que permitan a los alumnos incorporarse al quehacer metodológico del trabajo científico.

- *Pruebas de observación.* Estas técnicas surgen al amparo del paradigma cualitativo y están aún poco desarrolladas en el ámbito de las ciencias, aunque podrían tener su mayor utilidad como complemento en la evaluación de los trabajos prácticos y de las actitudes de los alumnos. Así, por ejemplo, en el caso de los trabajos prácticos se podría utilizar una escala de calificaciones con el fin de poner de manifiesto la capacidad de un alumno para establecer modificaciones en un dispositivo experimental cuando surge algún imprevisto en su montaje.

- *Otras técnicas directas.* Dentro de éstas habría que resaltar la "entrevista" que, ya sea estructurada, semiestructurada o sin estructurar, ha cobrado un gran protagonismo, especialmente en la investigación didáctica dentro del marco constructivista, ya que se ha revelado como la técnica diagnóstica de las concepciones de los alumnos más útil.

3. LA EVALUACION EN LOS PROYECTOS CURRICULARES

3.1. Proyectos españoles

Si nos detenemos en los últimos currículos imperantes en España, esto es, las *Nuevas Orientaciones Pedagógicas* dependientes de la Ley General de Educación, los *Programas Renovados* y los *Decretos de Enseñanzas Mínimas* de la L.O.G.S.E., observamos a *grosso modo* una grave discontinuidad en la evolución de la importancia concedida al papel de la evaluación. Las *Nuevas Orientaciones Pedagógicas* ya contenían en sus formulaciones un apartado relevante dedicado a la evaluación, donde se hallaban explícitas algunas características retomadas en la actualidad, tales como un sentido global más allá de la propia evaluación del aprendizaje (revisión y mejora del proceso educativo,

orientación del alumnos, ...), evaluación inicial, continua, o la autoevaluación. Asimismo se especificaban las dimensiones del contenido objeto de evaluación (conocimientos básicos o informativos, habilidades para su aplicación, hábitos, destrezas, aptitudes y actitudes, y en sus procesos de comprensión, expresión, valoración y creación).

Por el contrario, tal y como hemos señalado, los Programas Renovados (Ciclo Superior) representan un mínimo en el papel de la evaluación en el proceso educativo, con sólo unas pinceladas referentes a su carácter continuo y a la necesidad de adquirir actividades de apoyo y refuerzo.

Finalmente, el Decreto de Enseñanzas Mínimas (Educación Secundaria Obligatoria) abre definitivamente la evaluación al resto de las componentes educativas (función formativa para el profesorado, función evaluadora del proceso de enseñanza), especialmente en su desarrollo por parte de las Comunidades Autónomas con competencias en educación, contemplando las bases de la evaluación criterial y resaltando su importancia en el proceso interactivo enseñanza-aprendizaje. Por otro lado se contemplan los contenidos agrupados en torno a conceptos-hechos-principios, procesos, y actitudes-valores-normas, lo que repercutirá directamente en la necesidad de desarrollar los instrumentos de evaluación correspondientes. En el caso del área de Ciencias de la Naturaleza (Junta de Andalucía), los criterios de evaluación se extienden a:

- La adquisición de conceptos básicos.
- El planteamiento y la resolución de problemas.
- La expresión y la comprensión.
- La noción de ciencia.
- La participación y trabajo en grupo.

Asimismo y en este último currículo se señala la incorporación de la evaluación como "actividad investigadora" que puede permitir al profesor depurar su proceso de enseñanza y su propia capacitación docente.

3.2. Proyectos extranjeros

- APU (*Assessment of Performance Unit*)

Este proyecto se ha venido desarrollando en el Reino Unido desde el año 1974 y entre sus objetivos se encuentra el *promover el desarrollo de métodos de evaluación y monitorización del rendimiento de los niños (11-15 años) en clase e indagar en la incidencia del bajo rendimiento* (Hein, 1990).

De los distintos contenidos que se reconocen como objeto de aprendizaje (conceptos, procesos y actitudes), el proyecto APU enfatiza los aspectos procesuales de la evaluación científica, es decir, la observación, recogida y análisis de datos, emisión de hipótesis, etc. Proponemos el siguiente ejemplo tomado de la última obra referida:

*Tienes una colección de huesos marcados con letras.
Estos huesos pertenecen a la espalda de distintos animales.
Coloca los huesos en tres grupos de modo que posean algo común en cada uno de ellos.
Escribe las letras de los huesos en las columnas correspondientes a cada grupo.*

La aplicación del APU ha implicado la resolución de tests durante bastantes años y a miles de alumnos, quedando integrado en un organismo estable, *School Examinations and Assessment Council*, a partir del año 1988.

- NSF (*National Science Foundation*)

Este último proyecto curricular se ha desarrollado recientemente en los Estados Unidos y se generó a raíz de la insatisfacción creciente en el mundo educativo por los sistemas de evaluación habituales, basados en la utilización casi exclusiva de tests compuestos de items de opción múltiple estandarizados (Hein, 1990). Reproducimos seguidamente uno de ellos:

¿Cuál de las siguientes explicaciones sobre por qué las algas marinas están a menudo limitadas a una profundidad máxima de 100 m en el océano es la mejor? :

- *No poseen raíces para sujetarse en el fondo de océano.*
- *Realizan la fotosíntesis y sólo pueden vivir donde exista luz.*
- *La presión es demasiado grande como para sobrevivir por debajo de 100 m.*
- *La temperatura del océano hasta los 100 m de profundidad es ideal para ellas.*

Para ello se tuvieron en cuenta, principalmente, las condiciones iniciales de la instrucción, el currículum implicado y la eficiencia del mismo, medida a través de herramientas afectivas y cognitivas. A partir de los datos obtenidos, el proyecto NSF puso un énfasis especial en actividades tales como la investigación y la resolución de problemas, la identificación de los modos de pensamiento de los alumnos y, con respecto a la evaluación, la elaboración de instrumentos fáciles de aplicar e interpretar por los profesores y la congruencia entre las actividades curriculares y de evaluación.

4. LA INVESTIGACION EN EVALUACION

Ya señalamos en el apartado anterior lo novedoso de considerarse en el Decreto regulador de la Educación Secundaria Obligatoria en España el *papel investigador de la evaluación*. Esto constituye ciertamente la constatación de una evidencia. Baste decir que los estudios de contraste metodológico o los estudios de tipo correlacional dependen básicamente de cómo se plantee y se aplique el proceso evaluador. Pero aún puede irse más allá y así se le va prestando una progresiva atención desde la investigación educativa, rescatándolo de la marginación sufrida durante un largo periodo de tiempo. Esta relativa "juventud" de las investigaciones sobre evaluación impide la extracción de conclusiones relevantes y sólo cabe hablar de "tendencias" que podríamos resumir brevemente en los siguientes párrafos:

- La evaluación se enfrenta con graves problemas de fiabilidad como consecuencia de una elaboración y aplicación excesivamente atomizada, lo que conlleva la dificultad de generalización de sus resultados en la educación. Este problema se hace especialmente constatable en el caso de las pruebas de evaluación corregidas por varios profesores o incluso por el mismo profesor en condiciones diferentes (Rodríguez y col., 1992).

- Si se acepta una clasificación del conocimiento a adquirir como conceptos-hechos-principios, habilidades y actitudes, estas dos últimas componentes adolecen de instrumentos de evaluación suficientemente contrastados.

- Existe una tendencia creciente a asumir el paradigma cualitativo (mediante instrumentos observacionales, estudio de casos, etc.) en detrimento del paradigma cuantitativo (instrumentos "objetivos", inferencias estadísticas, etc.).
- Se intenta dar cabida a nuevos instrumentos de evaluación menos formales pero más sensibles al conocimiento complejo (entrevistas, mapas conceptuales, etc.).

Vamos a centrarnos ahora de un modo más concreto en la *investigación sobre evaluación en el ámbito de la educación científica*, resaltando su especificidad frente a las generalidades apuntadas con anterioridad y dejando constancia, en primer lugar, de la relativa marginación padecida por la evaluación en Ciencia frente a otras áreas, como las Matemáticas o el Lenguaje (Johnson, 1987).

Si damos por supuesta la distinción entre contenidos conceptuales (conceptos, leyes, teorías, principios), procesuales y actitudinales, cada uno de ellos ha despertado distinto grado de interés, aunque no en todos se han alcanzado resultados con un cierto nivel de consenso. La evaluación conceptual ha prevalecido en el uso de modelos didácticos, tanto en el Racional-Tecnológico como en el Constructivista (véase, p. ej., Perales, 1988) haciéndolo la evaluación de habilidades en el de Descubrimiento. Por otro lado, la resolución de problemas como actividad aglutinante del pensamiento conceptual y procesual, ha gozado de interés y tradición en el ámbito de la educación científica (véase, p. ej., Dumas-Carré, 1987; Perales, 1993). Asimismo se ha producido una reciente incorporación de procedimientos propios de otras disciplinas a la evaluación en ciencias (tales como los métodos etnográficos, Rodríguez y col. 1992).

- *Evaluación conceptual.* Como acabamos de referir, la evaluación conceptual ha gozado de una situación preferencial en el campo de aplicación de los modelos Racional-Tecnológico y Constructivista, aunque con diferencias básicas. En el primero la evaluación goza de un carácter reproductivo y sumativo: se pretende que el alumno, al final del proceso de enseñanza, sea capaz de "reproducir" fielmente lo impartido por el profesor y sepa, en el caso de las ciencias, resolver "problemas-tipo" también con un carácter similar a los resueltos "transmisivamente" en el aula. Por otro lado, en el Modelo Constructivista, la evaluación ha adoptado nuevas formas: se ha potenciado tremendamente la evaluación inicial o diagnóstica con el ánimo de detectar e interpretar el conocimiento previo de los alumnos, base de la enseñanza posterior, y se ha imprimido un giro a la evaluación sumativa: se considerará que el alumno ha aprendido cuando sea capaz de "aplicar" los nuevos conocimientos en situaciones igualmente nuevas y, a ser posible, cotidianas; asimismo la evaluación criterial debería estar basada en el cumplimiento o no del objetivo primordial de la enseñanza constructivista, es decir, producir un "cambio conceptual" en el alumnado a largo plazo, habida cuenta la persistencia de las concepciones erróneas puesta de manifiesto en numerosos estudios longitudinales. En cualquier caso resultan escasos los trabajos de campo emprendidos para detectar la modificación real de las pruebas de evaluación en esta línea de actuación (p. ej., Alonso y cols. 1992; Lorsbach y col., 1992).

- *Evaluación de habilidades.* Si distinguimos inicialmente entre habilidades cognitivas (p. ej., interpretación de gráficos, elaboración de hipótesis, etc.) (véanse, p. ej., Furió, 1989; Payá, 1990) y destrezas (p. ej., medida con una balanza, soplado de vidrio, etc.) han sido las primeras las que han concitado mayor atención por parte de los investigadores. Aunque tradicionalmente y, a través de las *Techniques for the Assessment of Practical Skills* (TAPS), los instrumentos perseguían evaluar las habilidades de los alumnos de un modo independiente entre sí y respecto del propio contexto teórico, las nuevas

tendencias rechazan esta separación y los presupuestos didácticos subyacentes, apostando por una "enseñanza por investigación" (Gil, 1993), aun a costa de aceptar la ambigüedad implícita a la evaluación de esta dimensión (Hodson, 1992). Tomaremos como ejemplo el siguiente ítem de opción múltiple correspondiente al objetivo: *Dado un problema con una variable dependiente especificada, identificar una hipótesis verificable* (Dillshaw y Okey, 1980).

Algunas gallinas ponen un huevo casi todos los días. Otras gallinas producen pocos huevos. Se plantea un estudio para examinar los factores que afectan al número de huevos producidos por las gallinas. ¿Cuál de las siguientes NO es una hipótesis adecuada para su estudio?:

1. *La mayor cantidad de huevos son producidos por las gallinas que reciben más horas de luz.*
2. *Cuanto mayor cantidad de huevos producen las gallinas más peso pierden.*
3. *Cuanto más grande es la jaula de las gallinas más huevos producen.*
4. *Cuanto más cantidad de proteínas hay en la comida más huevos producen.*

- *Evaluación actitudinal.* Las actitudes hacia la Ciencia y, en general, todas las variables que pueden agruparse bajo el paraguas afectivo se enfrentan con grandes dificultades a la hora de ser evaluadas de un modo válido y fiable (Munby, 1983; Germann, 1988), lo que quizás haya impedido un mayor desarrollo e integración en los modelos de aprendizaje científico. Estas dificultades se tornan particularmente graves cuando el profesor se enfrenta con la necesidad de evaluar las actitudes de los alumnos.

Paralelamente se han emprendido estudios colaterales, como el de Satterly y Swann (1988) y que viene a poner de manifiesto la relación entre el modelo de evaluación de los alumnos y su actitud hacia la Ciencia, favoreciendo en este caso a la evaluación criterial frente a la normativa.

- El interés sobre el tópico de la evaluación ha continuado creciendo y abarcando otros campos de desarrollo tales como el denominado *Comprensión Pública de la Ciencia*, donde se intenta detectar el nivel de conocimiento que la sociedad posee sobre los conceptos, procesos o hechos científicos. En esta línea se han acometido trabajos puntuales tales como el llevado a cabo por la revista *Newsweek* en 1990 sobre conceptos básicos de la Ciencia o en Cataluña durante 1989 (*El País*, 18-11-1989) sobre nombres de científicos y que revelan un alto grado de ignorancia de la población encuestada.

5. IMPLICACIONES PARA LA ENSEÑANZA DE LAS CIENCIAS

Para finalizar queremos recoger en las siguientes líneas algunas recomendaciones que, a nuestro juicio, podrían ser útiles para implementar el papel de la evaluación en la práctica del proceso de enseñanza de las ciencias:

- Potenciar la evaluación como un elemento esencial y dinamizador del currículum.
- Tomar conciencia de la necesidad de incrementar la preparación del profesorado en este tópico y del esfuerzo que significa asumir el nuevo papel que la evaluación debe jugar en el proceso educativo.

- Contemplar la evaluación como núcleo de investigación-acción para los departamentos de profesores (en temas como la metodología de enseñanza, el profesor, el currículum, los libros de texto, etc.).
- Adecuar la evaluación al modelo didáctico adoptado de un modo consciente.
- Relacionar coherentemente la evaluación con los objetivos planteados en el currículum.
- Considerar, profesores y alumnos, las pruebas de evaluación como actividades de aprendizaje.
- Incrementar, tanto la propia Administración como las editoriales, la difusión de instrumentos de evaluación para las habilidades y actitudes de los alumnos.
- Reflexionar previamente el profesorado sobre la fiabilidad y validez de las pruebas a aplicar.
- Diversificar las técnicas de evaluación con objeto de diluir la posible inadaptación de ciertos alumnos ante algunos instrumentos concretos y diagnosticar un mayor número de variables de conocimiento.
- Elaborar el profesor o el equipo de profesores archivos de pruebas de evaluación contrastadas y mejoradas de un modo continuo.
- Emplear preferentemente evaluación criterial, formativa y de tipo funcional.
- Negociar inicialmente con los alumnos el sistema de evaluación durante el desarrollo de la asignatura.
- Favorecer la autoevaluación a través de la discusión de los exámenes, uso de mapas conceptuales, etc.
- Relajar la tensión propia de la aplicación de las pruebas de evaluación, especialmente en las de tipo sumativo.

AGRADECIMIENTOS

El autor desea expresar su reconocimiento al Prof. Dr. Oscar Sáenz Barrio por sus interesantes sugerencias aportadas en la redacción de este trabajo.

REFERENCIAS BIBLIOGRAFICAS

- AIKENHEAD, G.S. (1988): "An analysis of four ways of assessing student beliefs about STS topics". *Journal of Research in Science Teaching*, 25, 607-629.
- ALONSO, M., GIL, D. y MARTINEZ, J. (1992): "Los exámenes de Física en la enseñanza por transmisión y en la enseñanza por investigación". *Enseñanza de las Ciencias*, 10, 127-138.
- DILLASHAW, F. y OKEY, J. (1980): "Test of the integrated science process skills for secondary science students". *Science Education*, 64, 601-608.
- DRIVER, R. y WATTS, M. (1990): *Research on students' conceptions in science: a bibliography*. School of Education, University of Leeds.

- DUMAS-CARRE, A. (1987): *La resolution de problemes en Physique au Lycée*. Tesis, Universidad de París 7.
- ESCUADERO, T. (1979): *Enseñanza de la Física en la Universidad: La evaluación periódica como estímulo didáctico*. Madrid, M.E.C.
- FAUCONNET, S. (1984): "Etude de resolution de problemes analogues". En: *Atelier International d'été: Recherche en didactique de la Physique* (La Londe les Maures, 1983). París, C.N.R.S., 261-269.
- FERNANDEZ, E., PERALES, F.J. y PLAZA, R. (1984): "Una experiencia sobre didáctica de los problemas". *Cuadernos de Física y Química*, V, 47-50.
- GEORGE, K.D., DIETZ, M.A., ABRAHAM, E.C. y NELSON, M.A. (1977): *Las Ciencias Naturales en la Educación Básica. Fundamento y métodos*. Madrid, Santillana.
- GERMANN, P.J. (1988): "Development of the attitude toward science in school assessment and its use to investigate the relationship between science achievement and attitude toward science in school". *Journal of Research in Science Teaching*, 25, 689-703.
- FURIO, C. (1989): "Selección bibliográfica sobre evaluación de prácticas de laboratorio". *Enseñanza de las Ciencias*, 7, 187.
- GIL, D. (1993): "Contribución de la Historia y de la Filosofía de las Ciencias al desarrollo de un modelo de enseñanza/aprendizaje como investigación". *Enseñanza de las Ciencias*, 11, 197-212.
- HEIN, G.E. (ed.) (1990): *The assessment on hands-on elementary science programs*. University of North-Dakota.
- HODSON, D. (1992): "Assessment of practical work: some considerations in philosophy of science". *Science and Education*, 1, 115-144.
- HORNE, S. (1984): "Criterion-referenced testing: pedagogical implications". *British Educational Research Journal*, 10, 155-174.
- JOHNSON, S. (1987): "Assessment in science and technology". *Studies in Science Education*, 14, 83-108.
- KEMPA, R.F. (1986): "Resolución de problemas de Química y estructura cognoscitiva". *Enseñanza de las Ciencias*, 4, 99-110.
- LAFOURCADE, P.B. (1972): *La evaluación de los aprendizajes*. Madrid, Cincel.
- LORSBACH, A.W., TOBIN, K., BRISCOE, C. y LA MASTER, S.V. (1992): "An interpretation of assessment methods in middle school science". *International Journal of Science Education*, 14, 305-317.
- MUNBY, H. (1983): *An investigation into the measurement of attitudes in science education*. Columbus, SMEAC Information Reference Center, Ohio State University.
- NOVAK, J.D. y GOWIN, B. (1988): *Aprendiendo a aprender*. Barcelona, Martínez Roca.
- PAYA, J. (1990): "Los trabajos prácticos de Física y Química: una revisión bibliográfica". *Enseñanza de las Ciencias*, 8, 181-185.
- PERALES, F.J. y CERVANTES, A. (1984): "Influencia del conocimiento del resultado numérico en la resolución de problemas". *Enseñanza de las Ciencias*, 3, 211-219.
- PERALES, F.J. (1988): "La instrucción científica en el marco de un modelo constructivista para la enseñanza". *Revista de Educación de la Universidad de Granada*, 2, 115-133.
- PERALES, F.J. (1990): "Las representaciones simbólicas del conocimiento y su aplicación en la enseñanza de las ciencias". *Enseñanza de las Ciencias*, 8, 79-82.
- PERALES, F.J. (1993): "La resolución de problemas en ciencias y matemáticas: una revisión estructurada". *Enseñanza de las Ciencias*, 11, 170-179.
- PERALES, F.J. (1994): "Los trabajos prácticos y la didáctica de las ciencias". *Enseñanza de las Ciencias*, 12, 121-125.
- RODRIGUEZ, L.M., GUTIERREZ, F.A. y MOLLEDO, J. (1992): "Una propuesta integral de evaluación en ciencias". *Enseñanza de las Ciencias*, 10, 254-267.
- SAENZ, O. (ed.) (1989): *Didáctica General*. Madrid, Anaya.
- SATTERLY, D. y SWANN, N. (1988): "Los exámenes referidos al criterio y al concepto en ciencias: un nuevo sistema de evaluación". *Enseñanza de las Ciencias*, 6, 278-284.
- WITKIN, H.A., OLTMAN, P.K., RASKIN, E. y KARP, S.A. (1971): *A manual for the embedded figures tests*. Palo Alto, Consulting Psychologists Press.