

Capítulo 8

¿Para qué y cómo evaluar?

La evaluación como instrumento de regulación Y mejora del proceso de enseñanza/aprendizaje

Daniel Gil Pérez y Joaquín Martínez Torregrosa

ALGUNAS CUESTIONES QUE SE ABORDAN EN ESTE CAPÍTULO

- ¿Cuáles pueden ser las ideas, comportamientos o actitudes que los profesores de ciencias podemos haber adquirido por simple impregnación ambiental acerca de la evaluación y que convenga analizar críticamente?
- ¿Qué aspectos de la actividad de los estudiantes deben ser evaluados?
- ¿Qué consecuencias suele conllevar la identificación de la evaluación con la mera calificación?
- ¿Cuáles habrían de ser las funciones de la evaluación?
- ¿Qué características ha de poseer la evaluación para que favorezca el aprendizaje?
- ¿Qué preguntas convendría que nos planteáramos en torno a la evaluación de la enseñanza y el currículo?

EXPRESIONES CLAVE

Influencia de la evaluación en el proceso de enseñanza/aprendizaje; funciones de la evaluación como instrumento de aprendizaje; expectativas de docentes y estudiantes; participación de los estudiantes en la regulación del aprendizaje.

INTRODUCCIÓN

Quizás debamos comenzar planteándonos si merece la pena dedicar mucha atención a esta cuestión de la evaluación. ¿Acaso la evaluación no constituye un aspecto, sin duda necesario, pero rutinario y poco atractivo, en el que más vale no detenerse demasiado? ¿No sería mejor ocupar ese tiempo en otros aspectos más susceptibles de contribuir a la renovación de la enseñanza? En otras palabras:

Propuesta de trabajo

¿Puede considerarse útil dedicar parte del escaso tiempo de que disponemos los profesores a tratar la cuestión de la evaluación?

Conviene detenerse en justificar el interés de esta actividad. En primer lugar, porque no tiene sentido acometer una tarea sin haber reflexionado mínimamente acerca de su relevancia (algo que, desgraciadamente, olvidamos muy a menudo en el trabajo con los estudiantes). Pero, además, porque la atención dada a la evaluación *desde la didáctica de las ciencias* ha sido realmente escasa hasta muy recientemente. Como señala Tamir (1998), en el *International Handbook of Science Education*, si dicho Handbook se hubiera publicado diez años antes, no hubiera contenido una sección dedicada a la evaluación. La situación ha comenzado a cambiar desde finales de los años ochenta y hoy la evaluación constituye una línea de investigación prioritaria en didáctica de las ciencias. ¿A qué se debe este creciente interés? La reflexión de los profesores, cuando se nos plantea una cuestión como ésta, nos lleva a considerar el peso que la evaluación tiene en los estudiantes (“su atención se dirige exclusivamente a lo que puede ser objeto de examen”) y en nosotros mismos (“que vemos condicionada nuestra enseñanza por las pruebas externas”), las dificultades para ajustar la evaluación a las innovaciones que intentamos introducir en la enseñanza, las tensiones y enfrentamientos entre profesores y alumnos que suele generar la evaluación, etc.

Cabe señalar que numerosas investigaciones apoyan estas preocupaciones en torno al papel de la evaluación. Podemos así referirnos a que, en los últimos años, hemos asistido a un importante desarrollo de la innovación en la enseñanza de las ciencias, apoyada en investigaciones sistemáticas, que, sin embargo, encuentra dificultades para ser transferido a la práctica docente. Como ha señalado Briscoe (1991), cada año miles de profesores participamos en seminarios o asistimos a cursos con la intención de perfeccionarnos profesionalmente, y cuando reanudamos nuestras clases creemos estar mejor preparados para utilizar las nuevas técnicas, los nuevos materiales curriculares, las nuevas formas de favorecer la creatividad y el aprendizaje de nuestros alumnos. Sin embargo, muchos de nosotros nos encontramos, antes de que podamos darnos cuenta, enseñando de la misma forma como lo habíamos hecho siempre, adaptando los nuevos materiales o técnicas a los patrones tradicionales. Se genera así una lógica frustración y decepción al percibir que las cosas no han funcionado mejor que los años precedentes a pesar de las nuevas y prometedoras ideas.

Este resultado no es debido, en general, a que las innovaciones contempladas en los cursos de formación carezcan de interés, sino que pone en evidencia que un modelo de enseñanza es algo más que un conjunto de elementos dispersos e intercambiables: posee

una cierta coherencia y cada uno de sus elementos viene apoyado por los restantes (Viennot, 1989; Gil-Pérez, 1991). Se ha empezado así a comprender que los esfuerzos de innovación en la enseñanza de las ciencias realizados estas últimas décadas pierden gran parte de su capacidad transformadora si quedan en aportaciones puntuales, desligadas.

Muy en particular, los investigadores han llamado la atención sobre la necesidad de acompañar las innovaciones curriculares de transformaciones similares en la evaluación para contribuir a consolidar el cambio de modelo didáctico que está teniendo lugar (Linn, 1987). Poco importan, en efecto, las innovaciones introducidas o los objetivos enunciados: si la evaluación sigue consistiendo en ejercicios para constatar el grado de retención de algunos conocimientos conceptuales, éste será para los alumnos el verdadero objetivo del aprendizaje.

Pero hay una segunda razón del fracaso de muchos esfuerzos de renovación curricular, que se relaciona también con la evaluación. Como ha mostrado Cronin-Jones (1991), los diseñadores de currículos no suelen tener en cuenta la fuerte influencia de las concepciones de los profesores en el proceso de implementación curricular. Lo mismo ocurre a menudo, en nuestra opinión, en los programas de formación del profesorado. Dicho de otra manera, para emprender un replanteamiento global de la enseñanza de las ciencias se precisa cuestionar preconcepciones docentes cuya importancia en la actividad del profesorado puede ser tan relevante o más que las preconcepciones de los alumnos en el aprendizaje de las ciencias (Hewson y Hewson, 1987). En efecto, como hemos visto en el capítulo 2, comienza hoy a comprenderse que los profesores tenemos ideas, actitudes y comportamientos sobre la ciencia y su enseñanza debidos a una larga formación “ambiental”, en particular durante el período en que fuimos alumnos, que ejerce una notable influencia por responder a experiencias reiteradas y adquirirse de forma no reflexiva, como algo natural, obvio, “de sentido común”, escapando así a la crítica y convirtiéndose, insistimos, en un verdadero obstáculo. *Y cabe sospechar que la evaluación sea uno de los aspectos de la actividad docente más afectado por estas preconcepciones* (Gil-Pérez et al., 1991; Alonso, Gil-Pérez y Martínez Torregrosa, 1992a).

Nos encontramos, pues, con dos poderosas razones para abordar el estudio de la evaluación y, para comenzar, las posibles preconcepciones docentes que pueden estar obstaculizando su renovación. Es preciso tener en cuenta, a este respecto, otra de las causas que suelen dificultar la apropiación de los nuevos conocimientos didácticos por los profesores: nos referimos a la escasa efectividad de transmitir al profesorado las propuestas de los expertos para su mera aplicación. Como ha indicado Briscoe (1991), es necesario que los profesores *participemos en la construcción* de los nuevos conocimientos didácticos abordando los problemas que la enseñanza nos plantea. Es así como podrá tener lugar la transformación de nuestras concepciones espontáneas. Podemos señalar a este respecto que, si bien estas preconcepciones son muy abundantes y constituyen serios obstáculos (en la medida en que son aceptadas acríticamente), no resulta difícil generar una reflexión descondicionadora que ponga en cuestión estas “evidencias” y contribuya al trabajo de profundización que exige su superación, aproximando las concepciones del profesorado a las adquisiciones de la investigación didáctica. Esto es lo que trataremos de mostrar en este capítulo, que será planteado como una investigación dirigida en torno a las concepciones docentes en este campo.

ANÁLISIS CRÍTICO DE LAS CONCEPCIONES DOCENTES SOBRE LA EVALUACIÓN Y LA PRÁCTICA QUE SE DERIVA

Conviene comenzar planteándonos una reflexión crítica acerca de cuáles son las ideas que los docentes solemos tener en torno a la evaluación, como primer paso de la actividad investigadora que estamos iniciando, para después contrastarlas con las aportadas por la investigación didáctica.

Propuesta de trabajo

Consideremos, a título de hipótesis, posibles preconcepciones que los profesores de ciencias podamos tener acerca de la evaluación y que convenga analizar para no quedar prisioneros de ideas, comportamientos o actitudes incorrectas, asumidas acríticamente.

Nos referiremos seguidamente a las concepciones señaladas en la literatura, de forma que el lector pueda constatar cuáles de ellas conectan con sus intuiciones y vivencias. Se trata de un conjunto de ideas, comportamientos y actitudes interrelacionados que se apoyan mutuamente entre sí y que podemos resumir en cinco apartados:

- La idea de que resulta fácil evaluar las materias de ciencias con objetividad y precisión, debido precisamente a la naturaleza misma de los conocimientos evaluados.
- La tendencia a limitar la evaluación a aquello que sea más fácilmente medible, evitando todo lo que pueda dar lugar a respuestas imprecisas. En asociación con la creencia en la objetividad de la evaluación, cabe esperar que se limite la práctica evaluadora a lo aparentemente “más objetivo”, es decir, a los conocimientos fácticos, a las leyes a través de ejercicios cerrados o de respuesta unívoca en los que no quepa la ambigüedad.
- La concepción elitista y determinista del aprendizaje de las ciencias, que supone que estas materias no están al alcance de todos y, en consecuencia, que una evaluación bien planteada pondrá de manifiesto el fracaso “inevitable” de un porcentaje importante de alumnos y discriminará a los “buenos” alumnos y a los “mediocres”.
- Muy relacionada con la anterior, la tendencia autoexculpatoria, consistente en achacar los altos porcentajes de fracaso en la evaluación a causas externas a la didáctica empleada. Los resultados negativos obtenidos a menudo por porcentajes elevados de alumnos son imputables a los propios alumnos o a la enseñanza precedente.
- Finalmente, a modo de síntesis de las ideas y comportamientos anteriores, la tendencia a convertir la evaluación en un instrumento de mera constatación, de simple calificación. Ello subyace, insistimos, en las ideas anteriores, pero tiene también sus implicaciones específicas (evaluación terminal, carácter de obstáculo a superar, actuación del profesor como “juez” y “policía”...).

La existencia de estas concepciones y su naturaleza de obstáculo constituye una conjetura (apoyada, sin duda, por la reflexión crítica de lo que solemos o hemos visto hacer, por la coincidencia de distintos autores en denunciarlas, etc.) que debe ser sometida a un cuidadoso análisis, para estimar en qué medida dichas preconcepciones están extendidas

y para poner a prueba su validez. Se trata, pues, de continuar la investigación de las distintas concepciones conjeturadas.

Propuesta de trabajo

Cabe suponer, a título de hipótesis, que los profesores de ciencias podemos pensar que “resulta fácil evaluar las materias científicas con objetividad y precisión, dada la naturaleza misma de los conocimientos evaluados”. ¿En qué nos apoyamos para señalar esa suposición?

Esta actividad pretende facilitar una aproximación cualitativa a prácticas de evaluación que pongan en evidencia la creencia común de que “resulta fácil evaluar las materias científicas con objetividad y precisión, dada la naturaleza misma de los conocimientos evaluados”. Podemos referirnos a hechos como la seguridad con que muchos profesores atribuyen calificaciones numéricas con decimales, tipo 4,75 (¿como si todo lo que sabemos acerca de la imprecisión de los resultados no fuera válido en el caso de la evaluación!) y a la resistencia que solemos ofrecer los profesores a modificar dichas calificaciones (como si realmente nuestra seguridad fuera absoluta). Pero estos indicios han de dejar paso a auténticos diseños experimentales para someter a prueba nuestras hipótesis.

Propuesta de trabajo

Concibamos algunos diseños experimentales para poner a prueba, de una forma más rigurosa, la supuesta objetividad y precisión de la evaluación del aprendizaje de las materias científicas e indiquemos qué resultados cabe esperar al aplicarlos.

Para poner a prueba la supuesta objetividad y precisión de las evaluaciones, una primera idea que surge habitualmente es dar a corregir un mismo examen a diversos profesores. Éste es un diseño clásico que ya fue utilizado por Hoyat (1962) con ejercicios de física del bachillerato francés. Conviene, pues, valorar positivamente esta propuesta y proporcionar algunos resultados para que procedan a su análisis. Puede recordarse así que Hoyat encontró que un mismo ejercicio de física era calificado con notas que iban de 2 a 8 (¡).

Este diseño se ha utilizado numerosas veces con resultados similares; sin embargo, ello no nos permite concluir que queda probada la falta de objetividad y precisión. En efecto, cabe suponer que estas discrepancias en las notas sean simplemente el fruto de distintos criterios (“hay profesores rigurosos, otros con manga ancha...”). Surge así la idea de este otro diseño: hacer corregir de nuevo el mismo examen, al cabo de un cierto tiempo, a los mismos profesores. También este diseño fue ya utilizado por Hoyat, con resultados que mostraban una fuerte dispersión de las notas dadas por los mismos profesores a los mismos exámenes.

Una variante de este diseño hace innecesario esperar varios meses: se pueden incluir dos o tres copias formalmente distintas (letras diferentes, modificación del orden de respuesta de las preguntas...) de un mismo ejercicio en aquellos casos en que los profesores tienen que corregir un número elevado de exámenes. De nuevo aparece una gran dispersión de las notas dadas a un mismo examen.

Se puede concebir un diseño más sofisticado, destinado a ver cómo influyen las expectativas del profesor, planteando a distintos profesores la valoración de un breve ejercicio pidiéndoles una puntuación entre 0 y 10 y, sobre todo, comentarios que puedan ayudar al alumno a comprender mejor la cuestión planteada. El ejercicio que se entrega para corregir es siempre el mismo, con la "única" diferencia de un pequeño texto introductorio, que en la mitad de las copias atribuye el ejercicio a un alumno "brillante" y en la otra mitad a un alumno "que no va demasiado bien". Este pequeño comentario, sin embargo, provoca diferencias en las medias... ¡del orden de dos puntos! (Alonso, Gil-Pérez y Martínez Torregrosa, 1992a).

Una investigación similar, destinada a mostrar el sexismo del profesorado de física, fue realizada por Spear (1984) y consistió en proponer la corrección de un examen a unos 300 profesores de enseñanza secundaria con objeto de que evaluaran toda una serie de aspectos: nivel, precisión científica, capacidad para proseguir estudios científicos, etc. Cada copia del examen fue presentada al 50% de los profesores con la firma de un alumno y al otro con la firma de una alumna. Los resultados mostraron que en cada uno de los 14 aspectos a evaluar el "muchacho" fue calificado por encima de la "muchacha" (¡), con diferencias estadísticamente significativas.

Conviene recordar aquí la célebre experiencia de "Pígalión a la escuela" (Rosenthal y Jacobson, 1968). En una serie de escuelas se hizo creer a los profesores que un test de inteligencia había detectado que determinados alumnos (elegidos en realidad al azar) tenían un cociente intelectual extraordinario, eran una especie de "diamante en bruto". Dos años después se pudo constatar que los alumnos señalados habían experimentado un desarrollo intelectual muy superior al de sus condiscípulos.

Como vemos, todas estas investigaciones cuestionan la supuesta objetividad y precisión de la evaluación en un doble sentido. Por una parte, muestran hasta qué punto las valoraciones habituales están sometidas a amplísimos márgenes de incertidumbre (aunque algunos profesores acostumbremos a escribir notas como 4.75 y creamos en su validez y precisión) y, por otra, dichas investigaciones hacen ver que la evaluación constituye un instrumento que afecta muy decisivamente a aquello que se pretende medir con ella, es decir, al propio proceso evaluado. Con otras palabras, los profesores no sólo nos equivocamos al calificar (dando, por ejemplo, puntuaciones más bajas en materias como la física a ejercicios que creemos hechos por las chicas), sino que contribuimos a que nuestros prejuicios –los prejuicios, en definitiva, de toda la sociedad– se conviertan en realidad: las chicas acaban teniendo logros inferiores y actitudes más negativas hacia el aprendizaje de la física que los chicos, y los alumnos considerados mediocres terminarán siéndolo. La evaluación resulta ser, más que la medida objetiva y precisa de unos logros, la expresión de unas expectativas en gran medida subjetivas pero con una gran influencia sobre los alumnos y los propios profesores.

Pese a los resultados anteriores, es posible seguir apostando por la objetividad y pensar que la solución está en "evitar evaluar lo que puede dar lugar a respuestas imprecisas que dificulten la necesaria objetividad". Ésta podría ser contemplada, siempre a título de hipótesis, como una nueva concepción estrechamente ligada a la anterior y que conviene analizar también cuidadosamente.

 Propuesta de trabajo

¿Cuáles podrían ser las implicaciones de “evitar evaluar lo que puede dar lugar a respuestas imprecisas”?

Si tenemos en cuenta las características de la actividad científica, ampliamente discutidas en el capítulo 2, comprenderemos el peligro de grave reduccionismo que entraña la búsqueda de máxima objetividad. Como señala Hodson (1992), la preocupación obsesiva por evitar la ambigüedad, y asegurar la fiabilidad de las evaluaciones, distorsiona la naturaleza misma del trabajo científico, esencialmente difuso, incierto, intuitivo. La evaluación debería tener en cuenta dicha ambigüedad, no intentar eliminarla. En el mismo sentido se pronuncia Tamir (1998), criticando las “pruebas objetivas”. De este modo se produce un cambio fundamental de pregunta: en vez de “¿cómo conseguir que la evaluación sea objetiva?”, la cuestión pasa a ser “¿cómo conseguir que la evaluación refleje aquello que se considera importante?”, lo que lleva a plantearnos cuáles son los aspectos del aprendizaje de las ciencias que deberían ser evaluados para evitar visiones empobrecidas que dejen de lado características esenciales de la actividad científica.

 Propuesta de trabajo

Enumeremos aquellos aspectos de la actividad de los estudiantes que deberían ser evaluados y a los que las pruebas de evaluación ordinarias no suelen prestar la debida atención.

La cuidadosa consideración de aquello que debe ser evaluado resulta esencial, insistimos, para evitar reduccionismos empobrecedores y ha de enfrentarse a tradiciones que sistemáticamente han limitado la evaluación a los aspectos más fáciles de medir. La discusión realizada en el capítulo 2 de la naturaleza del trabajo científico, así como el desarrollo del modelo de aprendizaje como investigación propuesto a lo largo del resto de los capítulos, nos permite romper con dichas tradiciones e incluir toda una pluralidad de aspectos que caracterizan al trabajo científico como una actividad abierta y creativa, tal como se muestra en el **cuadro 1**, planteado como una red para el análisis del contenido de las pruebas de evaluación.

Como puede constatarse, el contenido de este cuadro coincide básicamente con el del cuadro 1 del capítulo 2, que recogía los “aspectos a incluir en un currículo de ciencias para favorecer la construcción de conocimientos científicos”. Ello es absolutamente lógico, puesto que todo aquello que merece incluirse en un currículo debe ser evaluado. En caso contrario, los alumnos concluyen rápidamente que los aspectos no evaluados carecen de importancia real. Ahora bien:

Cuadro 1.

Análisis de la presencia de actividades de evaluación coherentes con las estrategias de construcción de conocimientos científicos

1. ¿Se incluyen **situaciones problemáticas abiertas** que permitan a los estudiantes formular preguntas, plantear problemas?
2. ¿Se piden comentarios sobre el posible **interés de las situaciones** propuestas (considerando las posibles implicaciones CTSA, la toma de decisiones, la relación con el resto del programa, etc.)?
3. ¿Se piden **análisis cualitativos**, significativos, que eviten el mero operativismo y ayuden a comprender y acotar las situaciones planteadas?
4. ¿Se pide la **emisión de hipótesis**, fundamentadas en los conocimientos disponibles, susceptibles de orientar el tratamiento de las situaciones y de hacer explícitas, funcionalmente, las preconcepciones?
 ¿Se plantea, al menos, el manejo o el análisis crítico de alguna hipótesis?
 ¿Se incluyen actividades que supongan atención a las preconcepciones (contempladas como hipótesis)?
5. ¿Se plantea la **elaboración de estrategias** (en plural), incluyendo, en su caso, diseños experimentales?
 ¿Se pide, al menos, la evaluación crítica de algún diseño?
 ¿Se presta atención a la actividad práctica en sí misma (montajes, medidas...), dándole a la dimensión tecnológica el papel que le corresponde en este proceso?
 ¿Se presta atención, en particular, al manejo de la tecnología actual (ordenadores, electrónica, automatización...), con objeto de favorecer una visión más correcta de la actividad científico-técnica contemporánea?
6. ¿Se piden **análisis detenidos de los resultados** (su interpretación física, fiabilidad, etc.), a la luz del cuerpo de conocimientos disponible, de las hipótesis manejadas y/o de otros resultados?
 ¿Se pide, en particular, el análisis de los posibles conflictos cognitivos entre algunos resultados y las concepciones iniciales?
 ¿Se favorece la “autorregulación” del trabajo de los alumnos (proporcionando, por ejemplo, indicadores que les permitan comprobar si van o no en una dirección correcta)?
 ¿Se incluyen actividades para que los estudiantes cotejen su evolución conceptual y metodológica con la experimentada históricamente por la comunidad científica?
7. ¿Se piden **esfuerzos de integración** que consideren la contribución de los estudios realizados a la construcción de un cuerpo coherente de conocimientos (al afianzamiento o reforzamiento del mismo, a la superación de supuestas barreras...), las posibles implicaciones en otros campos de conocimientos, etc.?
 ¿Se pide algún trabajo de construcción de síntesis, mapas conceptuales, etc., que ponga en relación conocimientos diversos?
8. ¿Se pide la consideración de posibles **perspectivas** (replanteamiento del estudio a otro nivel de complejidad, problemas derivados...)?
 ¿Se incluyen, en particular, actividades relativas a las **implicaciones CTSA** del estudio realizado (posibles aplicaciones, transformaciones sociales, repercusiones negativas...), a la toma de decisiones, saliendo al paso de la concepción que reduce la tecnología a mera *aplicación* de la ciencia?
 ¿Se incorporan actividades que presten atención a la educación no formal (museos, prensa...)?

¿Se valoran los “productos” elaborados por los estudiantes (prototipos, colecciones de objetos, carteles...), expresión de las estrechas relaciones ciencia-tecnología?

9. ¿Se presta **atención a la comunicación** como aspecto esencial de la actividad científica?

¿Se plantea la elaboración de **memorias científicas** del trabajo realizado?

¿Se pide la lectura y comentario crítico de textos científicos?

¿Se pone en contacto a los estudiantes con la historia de la ciencia?

¿Se presta atención a la verbalización, solicitando comentarios significativos que eviten el “operativismo mudo”?

¿Se hace un seguimiento cuidadoso del dossier de los alumnos como memoria ordenada del trabajo realizado?

10. ¿Se potencia en la evaluación la **dimensión colectiva del trabajo científico** valorando los trabajos realizados en equipo, prestando atención al funcionamiento de los grupos de trabajo, etc.?

¿Se favorece la **interregulación** de los equipos?

¿Se permite (y potencia) el manejo funcional del cuerpo de conocimientos aceptado por la comunidad científica (sin exigir memorizaciones irracionales)? ¿Se presta atención, en ese sentido, a que los prerrequisitos no se conviertan en obstáculo para las tareas propuestas?

Podemos ahora utilizar el contenido del cuadro 1 como red de análisis de las pruebas de evaluación.

Propuesta de trabajo

¿En qué medida las pruebas de evaluación habituales suelen contemplar

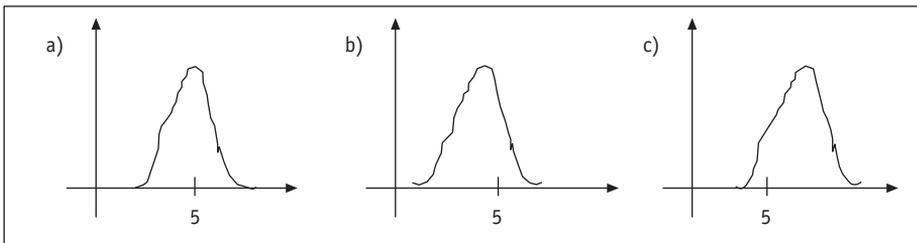
los aspectos fundamentales señalados en el cuadro 1?

Los análisis realizados de los exámenes habituales diseñados por los profesores y de las pruebas externas (como las pruebas de acceso a la universidad) apoyan plenamente la hipótesis de un grave reduccionismo, reforzando la comprensión de que la búsqueda de máxima objetividad y precisión puede derivar (ha derivado) en un peligroso empobrecimiento de la educación científica (Alonso, Gil-Pérez y Martínez Torregrosa, 1992b).

Una posible justificación del reduccionismo de las pruebas de evaluación sería el intento de evitar la introducción de aspectos “demasiado complicados”. De hecho, una posible concepción a la que nos hemos referido, siempre a título de hipótesis, es que “el fracaso de un número elevado de alumnos es inevitable en materias rigurosas como las ciencias, que no están al alcance de todos”. Para analizar la extensión de dicha concepción se ha concebido, entre otros, el diseño que se describe en la siguiente actividad:

Propuesta de trabajo

Se plantea a una serie de profesores la siguiente cuestión: “Las gráficas adjuntas presentan posibles distribuciones de las notas obtenidas en una prueba por los estudiantes de una clase. Señalad cuál de dichas distribuciones es más correcta y por qué”. ¿Qué resultados cabe esperar? Es decir, ¿cuál de las opciones a), b), c) puede parecer más correcta a la mayoría de los profesores y qué razones pueden aducir para ello?



Los resultados obtenidos al proponer una actividad como la anterior (Alonso, Gil-Pérez y Martínez Torregrosa, 1992a y b) muestran que la mayoría de los profesores optan por las gráficas a o b (aduciendo que “en una clase hay de todo”, alumnos listos y torpes, trabajadores o no, etc.), lo que supone que, en el mejor de los casos, se considera correcto un fracaso del 50% de los estudiantes (;). Si tenemos en cuenta, por otra parte, la influencia de las expectativas del profesor en los resultados de la evaluación y del aprendizaje, se comprende la gravedad de esta predisposición al fracaso generalizado de los estudiantes.

En ocasiones, los resultados negativos obtenidos por porcentajes elevados de estudiantes son imputados también a la enseñanza precedente.

Propuesta de trabajo

En entrevistas realizadas a profesores de facultades de ciencias, la causa de fracaso más mencionada fue “la falta de base de los estudiantes, debida a las insuficiencias de la formación proporcionada en los niveles precedentes”. Para apoyar esta tesis, unos profesores universitarios pasaron, el primer día de clase, a los estudiantes que habían ingresado en la Facultad de Química, un sencillo cuestionario con 20 preguntas consideradas muy elementales por los mismos profesores de secundaria. Los resultados fueron realmente decepcionantes, lo que en opinión de los autores del cuestionario, constituía la prueba de la incorrecta preparación proporcionada a dichos estudiantes. Suponiendo que no aceptemos, por demasiado simplista, la interpretación dada a la experiencia descrita en el párrafo anterior, ¿qué diseños podríamos concebir para mostrar lo incorrecto de dicha interpretación?

Un diseño que puede utilizarse para someter a prueba la idea simplista de que el fracaso de los estudiantes universitarios es debido fundamentalmente a la “falta de base” con que llegan de la enseñanza secundaria, consiste en utilizar el mismo cuestionario que se pasó a los estudiantes de primer curso (en su primer día de clase), a estudiantes que llegan a segundo (con una valoración positiva por parte de sus profesores universitarios), también el primer día de clase. Esta experiencia ha sido realizada con estudiantes de las Facultades de Química y de Física (Calatayud, Gil-Pérez y Gimeno, 1992), y los resultados obtenidos muestran una gran similitud entre los niveles de acierto de ambos colectivos (sin diferencias significativas en la generalidad de los ítems), poniendo así en cuestión la tesis simplista que sólo responsabiliza a la enseñanza precedente.

Las concepciones sobre la evaluación que hemos venido contemplando no constituyen, pues, ideas o comportamientos inconexos, sino que están muy relacionadas y proporcionan una idea de la **evaluación como instrumento de constatación y discriminación de los estudiantes**, es decir, como simple calificación.

 Propuesta de trabajo

¿Qué consecuencias, además de las ya discutidas, suele conllevar la identificación de la evaluación con la mera calificación?

Podemos referirnos al carácter de obstáculo a superar que adquieren las pruebas (con tiempos rígidamente limitados, clima de tensión, etc.), al papel de juez frío y distante y de policía vigilante que ha de asumir el profesor, a la reducción de las evaluaciones a las pruebas terminales, etc. Resulta bastante fácil señalar algunos indicios que apoyan esta visión de la evaluación, pero no consideramos necesario detenernos más en el análisis crítico de las concepciones docentes “espontáneas” (es decir, asumidas acríticamente por impregnación) acerca de la evaluación. Un análisis crítico que puede sin duda ser profundizado, pero que ha mostrado ya, pensamos, las limitaciones de la práctica evaluativa habitual, haciendo ver la necesidad de un replanteamiento global de la evaluación.

REPLANTEAMIENTO DE LA EVALUACIÓN

Con objeto de romper con las limitaciones que suponen las concepciones espontáneas que acabamos de analizar someramente, conviene proceder a un replanteamiento global de la evaluación que no dé nada por sentado y clarifique todos sus aspectos, comenzando por reflexionar acerca de *para qué* evaluar.

 Propuesta de trabajo

Si no se puede aceptar la idea de una evaluación como “juicio absolutamente objetivo” y terminal de la tarea realizada por cada alumno. ¿cuáles habrían de ser las funciones de la evaluación?

El análisis crítico realizado hasta aquí nos ha permitido comprender que carece de sentido una evaluación consistente en el enjuiciamiento “objetivo” y terminal de la labor realizada por cada alumno. Por el contrario, el profesor ha de considerarse corresponsable de los resultados que éstos obtengan: no puede situarse frente a ellos, sino con ellos; su pregunta no puede ser “quién merece una valoración positiva y quién no”, sino “qué ayudas precisa cada cual para seguir avanzando y alcanzar los logros deseados”. Para ello son necesarios un seguimiento atento y una retroalimentación constante que reoriente e impulse la tarea. Eso es lo que ocurre en los equipos de investigación que funcionan correctamente y eso es lo que tiene sentido también, en nuestra opinión, en una situación de aprendizaje orientada a la construcción de conocimientos, a la investigación. Los estudiantes han de poder cotejar sus producciones con las de otros equipos y –a través del profesor/director de investigaciones– con el resto de la comunidad científica; y han de ver valorado su trabajo y recibir la ayuda necesaria para seguir avanzando, o para rectificar si fuera necesario.

La evaluación se convierte así en un instrumento de aprendizaje, es decir, en una *evaluación formativa*, sustituyendo a los juicios terminales sobre los logros y capacidades

de los estudiantes. Pero, aunque ello representa un indudable progreso, éste resulta insuficiente si no se contempla también como un *instrumento de mejora de la enseñanza*. En efecto, las disfunciones en el proceso de enseñanza/aprendizaje no pueden atribuirse exclusivamente a dificultades de los estudiantes y resultará difícil que los alumnos y alumnas no vean en la evaluación un ejercicio de poder externo (y, por tanto, difícilmente aceptable) si sólo se cuestiona su actividad.

Si realmente se pretende hacer de la evaluación un instrumento de seguimiento y mejora del proceso, es preciso no olvidar que se trata de una actividad colectiva, en la que el papel del profesor y el funcionamiento del centro constituyen factores determinantes. La evaluación ha de permitir, pues, incidir en los comportamientos y actitudes del profesorado. Ello supone que los estudiantes tengan ocasión de discutir aspectos como el ritmo que el profesor imprime al trabajo o la manera de dirigirse a ellos. Y es preciso *evaluar también el propio currículo*, con vistas a ajustarlo a lo que puede ser trabajado con interés y provecho por los alumnos y alumnas. De esta forma, los estudiantes aceptarán mucho mejor la necesidad de la evaluación, que aparecerá realmente como un instrumento de mejora de la actividad colectiva.

Las funciones de la evaluación pueden resumirse, pues, en:

- favorecer el aprendizaje,
- contribuir a la mejora de la enseñanza,
- Incidir en el currículo (ajustarlo a lo que puede ser trabajado con interés y provecho por los y las estudiantes).

Nos centraremos, en primer lugar, en el papel de la evaluación como instrumento de aprendizaje, aunque insistimos en la necesidad de romper con los reduccionismos habituales, extendiendo la evaluación a la actividad del profesorado y al mismo currículo (Imbernon, 1990; Rodríguez, Gutiérrez y Molledo, 1992; Porlán, 1993; Santos, 1993).

La evaluación como instrumento de aprendizaje

Propuesta de trabajo

¿Cuáles habrían de ser las características de la evaluación para que se

convierta en un instrumento de aprendizaje?

Conseguir que la evaluación constituya un instrumento de aprendizaje, se convierta en una evaluación formativa, supone dotarla de unas características que rompan con las concepciones de sentido común que hemos analizado someramente en el primer apartado. En particular:

- Una primera característica que ha de poseer la evaluación para jugar un papel orientador e impulsor del trabajo de los estudiantes es que **pueda ser percibida como ayuda real, generadora de expectativas positivas**. El profesor ha de lograr transmitir su interés por el progreso de los alumnos y alumnas y su convencimiento de que un trabajo adecuado terminará produciendo los logros deseados, incluso si inicialmente aparecen dificultades. Conviene para ello realizar una planificación muy cuidadosa de

los inicios del curso, comenzando con un ritmo pausado, revisando los prerequisites (para que no se conviertan, como a menudo ocurre, en obstáculos), planteando tareas simples, etc.

- Una segunda característica que ha de poseer la evaluación, para que pueda jugar su función de instrumento de aprendizaje, es su **extensión a todos los aspectos –conceptuales, procedimentales y actitudinales– del aprendizaje de las ciencias, rompiendo con su habitual reducción a aquello que supuestamente permite una medida más fácil y rápida**: la rememoración repetitiva de los “conocimientos teóricos” y su aplicación igualmente repetitiva a ejercicios de lápiz y papel. Se trata de ajustar la evaluación –es decir, el seguimiento y la retroalimentación– a las finalidades y prioridades establecidas para el aprendizaje de las ciencias. La evaluación responde así a unos criterios explícitos de logros a alcanzar por los estudiantes (Satterly y Swann, 1988), al contrario de lo que ocurre con la evaluación atendiendo a la “norma”, basada en la comparación de los ejercicios para establecer los “mejores”, los “peores” y el “término medio”, que es la orientación asumida, lamentablemente, por gran parte del profesorado.

Por otra parte, es preciso no olvidar, a la hora de fijar los criterios, que sólo aquello que es evaluado es percibido por los estudiantes como realmente importante. Es preciso, pues, evaluar todo lo que los estudiantes hacen: desde un póster confeccionado en equipo a los dossiers personales del trabajo realizado. Duschl (1995) ha resaltado, en particular, la importancia de estos dossiers o “portafolios”, en los que cada estudiante ha de recoger y organizar el conocimiento construido y que pueden convertirse –si el profesor se implica en su revisión y mejora– en un *producto* fundamental, capaz de reforzar y sedimentar el aprendizaje, evitando adquisiciones dispersas.

- Si aceptamos que la cuestión esencial no es averiguar quiénes son capaces de hacer las cosas bien y quiénes no, sino lograr que la gran mayoría consiga hacerlas bien, es decir, si aceptamos que el papel fundamental de la evaluación es incidir positivamente en el proceso de aprendizaje, es preciso concluir que **ha de tratarse de una evaluación a lo largo de todo el proceso y no de valoraciones terminales**. Ello no supone –como a menudo interpretamos los profesores y los propios alumnos– parcializar la evaluación realizando pruebas tras períodos más breves de aprendizaje para terminar obteniendo una nota por acumulación, sino insistimos, integrar las actividades evaluadoras a lo largo del proceso con el fin de incidir positivamente en el mismo, dando la retroalimentación adecuada y adoptando las medidas correctoras necesarias en el momento conveniente (Colombo, Pesa y Salinas, 1986). Es cierto que realizar cinco pruebas, aunque tengan un carácter terminal –tras la enseñanza de un determinado dominio–, es mejor que una sola al final del curso; al menos habrán contribuido a impulsar un estudio más regular evitando que se pierdan todavía más alumnos; pero su incidencia en el aprendizaje sigue siendo mínima, o, peor aún, puede producir efectos distorsionantes. En efecto, a menudo la materia evaluada ya no vuelve a ser tratada, por lo que los alumnos que superaron las pruebas pueden llegar al final del curso habiendo olvidado prácticamente todo lo que estudiaron, teniendo conocimientos incluso más escasos que quienes fracasaron inicialmente y se vieron obligados a revisar por su cuenta.

Se acentúa así, además, la impresión de que no se estudian las cosas para adquirir unos conocimientos útiles e interesantes, sino para pasar unas pruebas. Es importante a este respecto ser conscientes de las leyes del olvido (Kempa, 1991) y planificar revisiones/profundizaciones de aquello que se considere realmente importante, para que los

alumnos afiancen dichos conocimientos aunque ello obligue, claro está, a reducir el currículo eliminando aspectos que, de todas formas, serían mal aprendidos y olvidados muy rápidamente.

- Por último, pero no menos importante, hemos de referirnos a la **necesidad de que los estudiantes participen en la regulación de su propio proceso de aprendizaje** (Baird, 1986; Linn, 1987; Jorba y Sanmartí, 1993 y 1995), dándoles oportunidad de reconocer y valorar sus avances, de rectificar sus ideas iniciales, de aceptar el error como inevitable en el proceso de construcción de conocimientos. Pero esto nos remite a las formas de la evaluación, que abordamos seguidamente. Antes plantearemos un problema que suele aparecer cuando se proponen unas características de la evaluación como las que acabamos de discutir. En efecto, a menudo surgen voces que señalan el peligro de que, aunque todas estas nuevas propuestas para la evaluación pueden beneficiar a quienes tienen dificultades, pudieran en cambio perjudicar a los buenos estudiantes, cuyos derechos no deben ser ignorados.

Propuesta de trabajo

¿En qué medida pueden resultar perjudicados los buenos estudiantes por una evaluación con las características propuestas?

Algunos profesores pueden pensar, en efecto, que una retroalimentación constante, un ritmo inicial pausado, etc., han de traducirse en pérdidas de tiempo que perjudicarán a los estudiantes bien preparados, cuyo derecho a aprender no debe ser ignorado. Pero, en realidad, lo que sucede es todo lo contrario: esta aparente pérdida de tiempo inicial permite romper con la rémora que supone a lo largo del curso la existencia de un núcleo importante de alumnos que “no siguen”. Se produce así un progreso global, favorable *también* para los alumnos mejor preparados. Todo esto, por supuesto, debe ser explicitado para evitar inquietudes y tensiones innecesarias y transmitir, en definitiva, expectativas positivas a *todos* los alumnos.

Vistas las características fundamentales que una evaluación habría de poseer para convertirse en un instrumento eficaz de aprendizaje, conviene ahora detenerse en considerar las formas concretas de realizar dicha evaluación.

Propuesta de trabajo

Concibamos formas concretas de realizar la evaluación que permitan incidir positivamente en el aprendizaje de las ciencias. Consideremos, en particular, los pros y contras del examen como forma de evaluación.

Cabe decir, en primer lugar, que una orientación del aprendizaje como la que venimos desarrollando permite que *cada actividad realizada en clase por los alumnos constituya una ocasión para el seguimiento de su trabajo*, la detección de las dificultades que se presentan, los progresos realizados, etc. (Black, 1998). Es ésta una forma de evaluación extraordina-

riamente eficaz para incidir “sobre la marcha” en el proceso de aprendizaje, que se produce además en un contexto de trabajo colectivo, sin la distorsión de la ansiedad que produce una prueba individual. Ello no elimina, sin embargo, la necesidad de actividades de evaluación individuales que permitan constatar el resultado de la acción educativa en cada uno de los estudiantes y obtener información para reorientar convenientemente su aprendizaje. A tal efecto, consideramos muy conveniente la realización de alguna pequeña prueba en la mayoría de las clases sobre algún aspecto clave de lo que se ha venido trabajando. Ello permite:

- impulsar al trabajo diario y comunicar seguridad en el propio esfuerzo;
- dar información al profesor y a los propios alumnos sobre los conocimientos que se poseen, sobre las deficiencias que se hayan producido –haciendo posible la incidencia inmediata sobre las mismas– y sobre los progresos realizados, contribuyendo así a generar expectativas positivas;
- reunir un número elevado de resultados de cada alumno, reduciendo sensiblemente la aleatoriedad de una valoración única.

Conviene discutir inmediatamente las posibles respuestas a la actividad planteada, lo que permitirá conocer si la clase está o no preparada para seguir adelante con posibilidades de éxito. Se favorece así la participación de los alumnos en la valoración de sus propios ejercicios y en su autorregulación (Alonso, Gil-Pérez y Martínez Torregrosa, 1995). Se puede aprovechar también esta discusión –si se realiza al comienzo de una clase– como introducción al trabajo del día, centrando la atención de los estudiantes de una forma particularmente efectiva.

Pese al interés y efectividad de estas pequeñas pruebas, consideramos que los exámenes o pruebas más extensas siguen siendo necesarios. Es cierto que el examen es visto a menudo como simple instrumento de calificación de los estudiantes, siendo criticado por lo que supone de aleatoriedad, tensión bloqueadora, etc. (Gould, 1982); sin embargo, un examen, o mejor dicho una *sesión de globalización*, es también ocasión de que el alumno se enfrente con una tarea compleja y ponga en tensión todos sus conocimientos. Por nuestra parte, asumiendo la crítica al examen como instrumento exclusivo de calificación, queremos referirnos al papel de las sesiones de globalización como ocasión privilegiada de aprendizaje si se cumplen algunas condiciones.

En primer lugar, es necesario que la sesión suponga la culminación de una revisión global de la materia considerada, incluyendo actividades coherentes con un aprendizaje por construcción de conocimientos: desde análisis cualitativos de situaciones abiertas hasta el tratamiento de las relaciones ciencia-tecnología-sociedad-ambiente (CTSA); desde la construcción y fundamentación de hipótesis –más allá de las evidencias de sentido común– hasta la interpretación de los resultados de un experimento, etc. Es también necesario que las condiciones de realización de estas actividades globalizadoras sean compatibles con lo que supone una construcción de conocimientos –que conlleva tentativas, rectificaciones, etc.– y, en particular, que los estudiantes no se vean constreñidos por limitaciones de tiempo que sólo son compatibles con la simple repetición de conocimientos memorizados.

En segundo lugar, es muy conveniente que el producto elaborado por cada estudiante en estas sesiones sea devuelto comentado lo antes posible y que se discutan, cuestión por cuestión, las posibles respuestas, las contribuciones positivas y los errores aparecidos, la persistencia de preconcepciones, etc. Los estudiantes, con su producto delante, se

mantienen abiertos y participativos como nunca durante estas revisiones, que constituyen actividades de autorregulación muy eficaces. Y es también conveniente, tras esta discusión, solicitar de los estudiantes que rehagan de nuevo la tarea en su casa con todo cuidado y vuelvan a entregarla. Ello contribuye muy eficazmente a afianzar lo aprendido, como puede constatarse en los días siguientes con la realización de pequeños ejercicios sobre los aspectos que hubieran planteado más dificultades.

No podemos detenernos aquí en una descripción detallada de cada uno de los momentos de evaluación a que nos acabamos de referir, ni entrar siquiera a tocar otros aspectos de importancia como, por ejemplo, el tipo de instrumentos destinados a la recogida de información (Geli, 1995). Insistiremos tan sólo, para terminar, en que *los alumnos y alumnas han de ver debidamente valoradas todas sus realizaciones*, todos sus productos colectivos o individuales –desde la construcción de un instrumento hasta muy en particular, su cuaderno o dossier de clase– y no solamente aquellas planteadas como pruebas. Se incrementa así la información disponible para valorar y orientar adecuadamente el aprendizaje de los estudiantes y se contribuye a que éstos vean reconocidos todos sus esfuerzos, con el consiguiente efecto motivador. Se trata, en definitiva, de *lograr una total confluencia entre las situaciones de aprendizaje y de evaluación* (Pozo et al., 1992), explotando el potencial evaluador de las primeras y diseñando las segundas como verdaderas situaciones de aprendizaje. Es preciso también transformar los enunciados de las actividades de evaluación habituales, para que pierdan su carácter de ejercicio memorístico o simplemente operativo y favorezcan una actividad más significativa. Proponemos, a continuación, un ejemplo centrado en este mismo tema de la evaluación.

Propuesta de trabajo

Procedamos a transformar la siguiente actividad, concibiendo otra más

adecuada: “Enumeren las visiones deformadas de la ciencia que la

enseñanza puede transmitir o contribuir a reforzar (a través, por

ejemplo, de lo que se evalúa o se deja de evaluar)”.

Una actividad que se limita a pedir una “enumeración de visiones deformadas” tiene un carácter puramente memorístico y resulta escasamente significativa. Se puede pedir, a cambio, que se analice críticamente un texto con objeto de indicar las deformaciones en las que puede incidir por acción u omisión o pedir la transformación de un dibujo para salir al paso de las deformaciones en las que incide, etc. Son actividades que hemos utilizado ya en el capítulo 2.

Para favorecer la autorregulación se puede, por ejemplo, presentar directamente al estudiante una respuesta errónea, supuestamente perteneciente a otra persona (pero que es muy probable que coincida con aspectos de su propio pensamiento espontáneo), y pedirle que la corrija. De esta forma se consigue que adopte una actitud de reflexión crítica y no se deje llevar por lo que parece obvio.

Un ejemplo relacionado también con la propia evaluación sería el siguiente: “En un texto pedagógico leemos que la función de la evaluación es constatar en qué medida cada estudiante ha alcanzado los objetivos que se persiguen. Analizar críticamente dicha proposición”.

Por supuesto, otra forma de favorecer la autoregulación de los estudiantes y, en definitiva, de lograr que sean conscientes de lo que se está tratando y por qué, qué se ha avanzado y lo que queda por avanzar, es proponer recapitulaciones periódicas, construcción de mapas conceptuales, etc., y, muy en particular, la cuidadosa elaboración y revisión del cuaderno de trabajo o “portafolio” (Duschl, 1995).

El trabajo realizado hasta aquí nos ha permitido romper con la habitual identificación entre evaluación y calificación de los estudiantes. Cabe plantearse, sin embargo, si la calificación conserva alguna funcionalidad en la nueva propuesta evaluativa.

El papel de la calificación

Hemos fundamentando hasta ahora una propuesta de evaluación como instrumento de mejora del aprendizaje, de la enseñanza y del propio currículo (aunque aquí nos hemos centrado en el primer aspecto). La cuestión a plantearse ahora es si la calificación conserva alguna funcionalidad en la nueva propuesta evaluativa y, en su caso, qué forma de calificación puede resultar coherente con dicha propuesta.

Propuesta de trabajo

Evaluar no es simplemente calificar, pero ¿qué idea de calificación puede derivarse

de las finalidades y características de la evaluación discutidas hasta este momento?

Resumiremos, a continuación, el contenido de las propuestas realizadas en torno a la calificación en el campo de la educación científica (Alonso, Gil-Pérez y Martínez Torregrosa, 1996). Digamos, para empezar, que toda evaluación posee connotaciones valorativas: tanto al indicar la necesidad de profundizar o rectificar aspectos como al aceptar el trabajo realizado sin enmiendas sustanciales, se está expresando implícitamente una valoración, aunque la función esencial no sea la valoración, sino favorecer la mejora del producto. El o la estudiante puede percibir así lo más o menos cerca que se encuentra de haber conseguido un producto satisfactorio. La pregunta a formularse, pues, no es si debe haber o no valoración de la tarea, sino: *¿es conveniente traducir las valoraciones que conlleva toda evaluación a calificaciones explícitas y normalizadas?*

Para contestar a esta pregunta comenzaremos considerando el papel de la calificación en la formación de los investigadores noveles, que es la situación en la que se inspira nuestro modelo de aprendizaje de las ciencias, para después examinar su papel en la formación de los estudiantes de ciencia.

a) La calificación en la formación de los investigadores. Un mínimo análisis de las situaciones de formación de investigadores, como la que supone, por ejemplo, la preparación de una tesis doctoral, permite constatar que la evaluación juega un papel fundamental (en forma de seguimiento constante de la labor del doctorando), pero que la calificación está prácticamente ausente a lo largo de todo el proceso de formación. En efecto, la idea central que subyace en el compromiso del doctorando y del director de la investigación es que ambos son corresponsables de la tarea los dos están comprometidos en lograr un producto satisfactorio. Las críticas y sugerencias del director no constituyen un enjuiciamiento externo, sino una contribución interesada. Y como tal es percibida por el doctorando.

De hecho, durante todo el tiempo que dura la realización de la investigación no hay calificación alguna: sólo cuando el doctorando y el director consideran que el producto es aceptable se somete a la valoración de otros investigadores. Pero incluso esto último ha ido, afortunadamente, evolucionando y en algunas universidades se está generalizando la práctica de someter a los miembros del tribunal el borrador de la tesis con tiempo suficiente para que puedan expresar sus críticas y sugerencias y recogerlas en una nueva versión que cuente con su aceptación. De esta forma, la lectura de la tesis se convierte en un acto protocolario que sanciona un producto que cuenta ya con la aceptación del tribunal (que se ha implicado, en caso necesario, en la mejora del producto).

La calificación ha perdido así su función de enjuiciamiento externo y de sanción discriminatoria, y constituyen rarísimas excepciones las tesis que no obtienen la máxima calificación que sanciona una investigación de calidad. Lo esencial, pues, es garantizar que el producto obtenido sea satisfactorio. Ello puede obligar a prolongar el período de realización e incluso puede llevar a algunos doctorandos a abandonar la investigación a la vista de las dificultades encontradas o porque deciden orientar su actividad en otra dirección. Pero, en cualquier caso, ese abandono constituye una opción que, en general, no viene impuesta por unas valoraciones negativas, por un enjuiciamiento “neutral” de quien dirige el trabajo o de un tribunal.

En definitiva, en una situación de formación de investigadores, la calificación explícita no está presente a lo largo del proceso y no conserva otra funcionalidad que el reconocimiento del trabajo realizado cuando éste posee suficiente calidad a los ojos del propio doctorando, del director de la tesis y de los miembros de un “tribunal” que constituye, más bien, una nueva instancia de revisión y ayuda, es decir, de evaluación como instrumento para la mejora del producto. Ésta es, en nuestra opinión, la mejor forma de plantear la evaluación y la calificación. Ahora bien, ¿en qué medida conviene o es posible hacer un planteamiento similar en una clase de ciencias de educación secundaria?

b) La calificación en la formación de los estudiantes de ciencias. En la medida en que nuestro modelo de aprendizaje de las ciencias como investigación se inspira en la metáfora de los estudiantes como “investigadores noveles” y del profesor como “experto”, consideramos adecuado un planteamiento de la evaluación y de la calificación como el que acabamos de describir. Lo esencial, pues, es orientar la evaluación como ayuda para la consecución, en el tiempo que sea necesario, de los logros perseguidos y que la calificación suponga tan sólo el reconocimiento de dichos logros.

Ahora bien, es obvio que, en el caso de una clase de ciencias, el contacto del “director de la investigación” (profesor) con cada “investigador novel” (cada estudiante) no puede ser tan directo, tan personalizado como el que se da en un equipo real de investigadores. Y es obvio también que los estudiantes no pueden centrarse en el trabajo de su clase de ciencias con la misma dedicación de un doctorando en su investigación. Ni puede pedirse a un adolescente la misma responsabilidad y capacidad de autorregulación que a un investigador real. Más aún, no es posible retrasar indefinidamente el momento de la valoración explícita del trabajo de los estudiantes, puesto que un curso tiene una duración determinada y a su término es preciso dar una calificación que indique la posibilidad o no de pasar a un nuevo nivel, etc.

Parece razonable, por todo ello, que la evaluación del trabajo de los estudiantes incluya indicaciones más explícitas y frecuentes del grado de consecución de los logros que se persiguen. Es decir, parece conveniente –y los mismos estudiantes lo reclaman– proporcionar valoraciones de las tareas que ayuden a los estudiantes a conocer si están progresando

adecuadamente o no. Se trataría, en definitiva, de hacer explícitas las valoraciones que en los comentarios de una evaluación formativa aparecen ya implícitamente. En resumen, la calificación puede ser conveniente –además de constituir una exigencia social difícilmente soslayable– como complemento de la evaluación formativa que hemos intentado fundamentar. Pero ello exige también una profunda modificación del uso y sentido de la calificación.

- En primer lugar, la calificación debe ser, como ya hemos señalado, una *estimación* de los *logros de cada estudiante*, una indicación de su grado de consecución de los logros que se persiguen. Más precisamente, la calificación no puede tener, como a menudo ocurre, una función comparativa y discriminatoria, en la que la valoración de un estudiante depende de los resultados de los demás, atendiendo a una “norma” que aproxima las calificaciones a una gaussiana (con, por definición... ¡una mitad de estudiantes fracasados!). Por el contrario, cada estudiante ha de saber que una calificación positiva depende exclusivamente de que alcance los logros que se persiguen. Es más, ha de saber que dichos logros se ajustan a lo que los estudiantes de su edad pueden llegar a realizar y son perfectamente alcanzables.
- En segundo lugar, la calificación ha de constituir una estimación cualitativa que utilice categorías amplias (no tiene sentido una calificación numérica del tipo 6,75), se apoye en una diversidad de elementos y se justifique con comentarios detallados. Cuantos más elementos podamos tomar en consideración (incluyendo, muy en particular, las actividades ordinarias de aprendizaje realizadas en clase) y *cuanto más amplias sean las categorías*, más fiables y fáciles de consensuar resultaran las estimaciones. De hecho, una calificación de estas características permite que no haya discrepancias sensibles entre las valoraciones del profesor y las del propio estudiante (o las que pueden realizar sus compañeros). Esto es algo que hemos constatado reiteradamente a lo largo de más de una década con estudiantes de secundaria y profesores en formación, poniendo de relieve que una evaluación continua (un seguimiento continuo, basado en una pluralidad de elementos como los descritos) proporciona una percepción bastante ajustada del dominio alcanzado por los estudiantes, tanto al profesor como a ellos mismos.
- Toda calificación ha de ser presentada como una indicación provisional y ha de ir acompañada, en caso necesario, de propuestas de actuación para su mejora (y de la comunicación de expectativas positivas en ese sentido). No es lo mismo, por supuesto, dar a un estudiante una valoración de “insuficiente” que explicarle que ha de realizar progresos en tales y cuales aspectos para lograr una valoración global positiva, estimularle a realizar las tareas correspondientes y apoyarle con un seguimiento adecuado.

Esta naturaleza de las calificaciones como indicaciones provisionales, destinadas a favorecer la autorregulación de los estudiantes, puede verse reforzada si se sustituyen las valoraciones negativas, tipo “insuficiente”, por un “pendiente de calificación” *sin connotaciones de rechazo*. Pero no se trata, claro está, de proponer un simple cambio de denominación, sino de plantear con claridad que la evaluación tiene como finalidad favorecer unos determinados logros y que el trabajo ha de continuar hasta conseguirlos *en el tiempo que haga falta*. Incluso si ello implica, en algún caso, continuar los mismos estudios el curso siguiente, es preciso presentar esta prolongación como algo positivo, como una adaptación al ritmo que el estudiante puede llevar en ese momento, con el convencimiento de que así afianzará su preparación para proseguir mejor sus estudios. Se trata, en

definitiva, de introducir aquí la misma flexibilidad que se tiene con el período de formación de un investigador y el mismo convencimiento de que lo esencial es llegar a un producto satisfactorio, transmitiendo expectativas positivas al respecto.

Señalemos, para terminar, que una calificación con las características que acabamos de proponer se integra coherentemente en la propuesta de evaluación como instrumento de aprendizaje y su asunción genera expectativas positivas que se traducen en mejores resultados y en una nueva forma de enfocar las relaciones entre profesores y estudiantes, más de acuerdo con la propuesta de aprendizaje como investigación dirigida.

La evaluación de la enseñanza y el currículo

Después del análisis de la evaluación del aprendizaje de las ciencias y la discusión de su reorientación, cabría ahora proceder a un trabajo similar en torno a la evaluación de la enseñanza y del currículo.

Propuesta de trabajo

¿Qué preguntas convendría que nos planteáramos en torno a la evaluación de la enseñanza y del currículo?

Se trataría de plantearnos cuestiones como, por ejemplo: *¿Qué aspectos de la actividad de los profesores y profesoras de ciencias convendría someter a evaluación? Se trata de contemplar todo aquello que pueda tener una incidencia, directa o indirecta, en la calidad de la enseñanza. ¿Quién y cómo habría de realizar la evaluación de la actividad del profesorado? ¿Qué aspectos del funcionamiento de un centro educativo convendría someter a evaluación?*

No podemos desarrollar aquí esta parte de la evaluación correspondiente a la actividad de los profesores y al currículo, que exigiría una extensión similar a la destinada a la evaluación del aprendizaje de los estudiantes. El tratamiento pormenorizado de ambas cuestiones sobrepasa los objetivos y extensión de este libro. Por lo que se refiere a la evaluación del currículo, nos remitimos, por ejemplo, a los capítulos VIII, IX y X del libro *La enseñanza de las ciencias en la educación secundaria* (Gil-Pérez et al., 1991). En cuanto a la evaluación de la enseñanza, la idea central sigue siendo concebir la evaluación como un instrumento de seguimiento y regulación para la mejora del proceso, evitando reduccionismos empobrecedores, es decir, atendiendo a las distintas dimensiones del acto educativo. En particular, es preciso atender al papel de los docentes como responsables de la creación de un clima de aula y de centro susceptibles de facilitar la implicación de los alumnos en tareas de interés (Fraser, 1994).

Procederemos ahora, para finalizar, a recapitular el trabajo realizado acerca de la evaluación.

POSIBLES TAREAS DE GLOBALIZACIÓN DEL TRABAJO REALIZADO EN TORNO A LA EVALUACIÓN

La naturaleza de los aspectos tratados en este capítulo, asociados a comportamientos y actitudes que hemos asumido, muy a menudo, de forma acrítica, como algo “natural”, “de sentido común”, recomienda formas de globalización que permitan recapitular los debates que se han producido en la comunidad científica. Por ejemplo:

Propuesta de trabajo

Elaborar un cuadro comparativo de la evaluación habitual y de la propuesta alternativa que hemos construido.

O bien:

Propuesta de trabajo

Construir un mapa conceptual que refleje la evaluación habitual de las ciencias y otro para las propuestas renovadoras.

Invitamos a los lectores y lectoras a realizar su propia recapitulación de este aspecto central del modelo de aprendizaje como investigación que hemos desarrollado en esta segunda parte del libro.

Ahora dedicaremos la tercera parte a presentar ejemplos concretos de desarrollos de temas con las orientaciones que se han venido fundamentando.

NOTA:

Este capítulo ha sido preparado a partir de los siguientes trabajos:

GIL-PÉREZ, D., CARRASCOSA, J., FURIÓ, C. y MARTÍNEZ TORREGROSA, J. (1991). *La enseñanza de las ciencias en la educación secundaria*. Barcelona: Horsori. (Capítulo 7: "La evaluación en la enseñanza/aprendizaje de las ciencias").

ALONSO, M., GIL-PÉREZ, D. y MARTÍNEZ TORREGROSA, J. (1992). Concepciones espontáneas de los profesores de ciencias sobre la evaluación. Obstáculos a superar y propuestas de replanteamiento. *Revista de Enseñanza de la Física*, 5(2), 18-38.

ALONSO, M., GIL-PÉREZ, D. y MARTÍNEZ TORREGROSA, J. (1996). Evaluar no es calificar. La evaluación y la calificación en una enseñanza constructivista de las ciencias. *Investigación en la Escuela*, 30, 15-26.

Referencias bibliográficas en este capítulo

ALONSO, M., GIL-PÉREZ, D. y MARTÍNEZ TORREGROSA, J. (1992a). Concepciones espontáneas de los profesores de ciencias sobre la evaluación. Obstáculos a superar y propuestas de replanteamiento, *Revista de Enseñanza de la Física*, 5(2), 18-38.

ALONSO, M., GIL-PÉREZ, D. y MARTÍNEZ TORREGROSA, J. (1992b). Los exámenes en la enseñanza por transmisión y en la enseñanza por investigación. *Enseñanza de las Ciencias*, 10(2), 127-138.

ALONSO, M., GIL-PÉREZ, D. y MARTÍNEZ TORREGROSA, J. (1995). Actividades de evaluación coherentes con una propuesta de enseñanza de la Física y la Química como investigación: actividades de autorregulación e interregulación. *Revista de Enseñanza de la Física*, 8(2), 5-20.

ALONSO, M., GIL-PÉREZ, D. y MARTÍNEZ TORREGROSA, J. (1996). Evaluar no es calificar. La evaluación y la calificación en una enseñanza constructivista de las ciencias. *Investigación en la Escuela*, 30, 15-26.

BAIRD, J.R. (1986). Improving learning through enhanced metacognition: A classroom study. *European Journal of Science Education*, 8(3), 263-282.

BLACK, P. (1998). Assessment by teachers and the improvement of students' learning. En Fraser, B. J. y Tobin, K. G. (Eds.), *International Handbook of Science Education*. London: Kluwer Academic Publishers.

BRISCOE, C. (1991). The dynamic interactions among beliefs, role metaphors and teaching practices. A case study of teacher change, *Science Education*, 75(2), 185-199.

CALATAYUD, M. L., GIL-PÉREZ, D. y GIMENO, J. V. (1992). Cuestionando el pensamiento docente espontáneo del profesorado universitario: ¿las deficiencias de la enseñanza secundaria como origen de las dificultades de los estudiantes?. *Revista Interuniversitaria de Formación del Profesorado*, 14, 71-81.

COLOMBO DE CUDMANI, L., PESA DE DANON, M. y SALINAS DE SANDOVAL, J. (1986). La realimentación en la evaluación de un curso de laboratorio de Física. *Enseñanza de las Ciencias*, 4 (2), 122-128.

CRONIN-JONES, L. L. (1991). Science teaching beliefs and their influence on curriculum implementation: two case studies. *Journal of Research in Science Teaching*, 38(3), 235-250.

DUSCHL, R. (1995). Más allá del conocimiento: los desafíos epistemológicos y sociales de la enseñanza mediante cambio conceptual. *Enseñanza de las Ciencias*, 13(1), 3-14.

FRASER, B. J. (1994). Research on classroom and school climate. En Gabel, D. L. (Ed.), *Handbook of Research on Science Teaching and Learning*. New York: McMillan Pub Co.

- GELI, A. (1995). La evaluación de los trabajos prácticos. *Alambique*, 4, 25-32.
- GIL- PÉREZ, D. (1991). ¿Qué han de saber y saber hacer los profesores de ciencias? *Enseñanza de las Ciencias*, 9(1), 69-77
- GILPÉREZ, D., CARRASCOSA, J., FURIÓ, C. y MARTÍNEZ TORREGROSA, J. (1991). *La enseñanza de las ciencias en la educación secundaria*. Barcelona: Horsori.
- GOULD, S. J. (1982). *La falsa medida del hombre*. Barcelona: Bosch.
- HEWSON, P. W. y HEWSON, M. G. (1987). Science teachers' conceptions of teaching: implications for teachers education. *International Journal of Science Education*, 9(4), 425-440.
- HODSON, D. (1992). In search of a meaningful relationship: an exploration of some issues relating to integration in science and science education. *International Journal of Science Education*, 14(5), 541-566.
- HOYAT, F. (1962). *Les Examens*. Paris: Institut de l'UNESCO pour l'Éducation. Ed Bourrelier.
- IMBERNON, F. (1990). La formación del profesorado. *Cuadernos de Pedagogía*, 178, 88-97.
- JORBA, J. y SANMARTÍ, N. (1993). La función pedagógica de la evaluación. *Aula de Innovación Educativa*, 20, 20-23.
- JORBA, J. y SANMARTÍ, N. (1995). Autorregulación de los procesos de aprendizaje y construcción de conocimientos. *Alambique*, 4, 59-77.
- KEMPA, R. F. (1991). Students' learning difficulties in science. Causes and possible remedies. *Enseñanza de las Ciencias*, 9(2), 119-128.
- LINN, M. C. (1987). Establishing a research base for science education: challenges, trends and recommendations. *Journal of Research in Science Teaching*, 24(3), 191-216.
- PORLÁN, R. (1993). *Constructivismo y Escuela. Hacia un modelo de enseñanza-aprendizaje basado en la investigación*. Sevilla: Diada.
- POZO, I., GÓMEZ, M. A., LIMÓN, M. y SANZ, A. (1992). *Procesos cognitivos de la ciencia: Las ideas de los adolescentes sobre la Química*. CIDE, MEC, Colección Investigadora.
- RODRÍGUEZ, L. M., GUTIÉRREZ, F. A. y MOLLEDO, J. (1992). Una propuesta integral de evaluación en Ciencias. *Enseñanza de las Ciencias*, 10(3), 254-267.
- ROSENTHAL, R. y JACOBSON, L. (1968). *Pygmalion in the classroom*. New Jersey: Rineheart and Winston.
- SANTOS, M. A. (1993). La evaluación: un proceso de diálogo, comprensión y mejora. *Investigación en la Escuela*, 20, 23-35.
- SATTERLY, D. y SWAM, N. (1988). Los exámenes referidos al criterio y al concepto de ciencias: un nuevo sistema de evaluación. *Enseñanza de las Ciencias*, 6(3), 278-284.
- SPEAR, M.G. (1984). Sex bias in science teachers' ratings of work and pupils characteristics. *European Journal of Science Education*, 6(4), 369-377.
- TAMIR, P. (1998). Assessment and evaluation in science education: opportunities to learn and outcomes. En Fraser, B. J. y Tobin, K. G. (Eds.), *International Handbook of Science Education*. London: Kluwer Academic Publishers.
- VIENNOT, L. (1989). L'enseignement des sciences physiques objet de recherche. *Bulletin de l'Union des Physiciens*, 716, 899-910.