

# DOS FORMAS DIFERENCIADAS DE EVALUACIÓN DIDÁCTICA: EVALUACIÓN NORMATIVA PARA SELECCIONAR A LOS ALUMNOS Y EVALUACIÓN CRITERIAL PARA EL DOMINIO DEL CONOCIMIENTO BÁSICO

*Two differentiated forms of didactic evaluation: normative evaluation in order to select students and criterial evaluation to master the basic content*

ALFONSO HEREDIA MANRIQUE  
Universidad de Zaragoza

Este artículo trata de recordar dos formas de evaluación didáctica, la evaluación normativa y la evaluación criterial, para volverlas a utilizar si es preciso, las cuales, con distinta finalidad educativa, se suceden históricamente en el campo de la evaluación. La evaluación normativa responde a la tradición selectiva del alumnado y la evaluación criterial tiene como objetivo que todo alumno domine los contenidos básicos del currículo.

**Palabras clave:** *Evaluación, Evaluación de los alumnos, Método de evaluación, Rendimiento educativo.*

## Planteamientos iniciales para la justificación teórica de la evaluación normativa y criterial

El análisis que vamos a hacer de los dos tipos de evaluación señaladas parte de la tesis que considera que la pedagogía tradicional fagocita los planteamientos didácticos progresistas que van surgiendo permanentemente en el ámbito de la educación, los cuales son transformados con gran éxito según los intereses de dicha pedagogía. Las pedagogías progresistas

han sido contaminadas persistentemente por el pensamiento pedagógico tradicional y el paso del tiempo ha ayudado a ocultar los orígenes y naturaleza de las mismas, creando confusiones educativas. Así, pueden plantearse cuestiones de análisis para desenmascararlas y saber cómo se producen esas confusiones. Una de estas cuestiones confusas en el ámbito pedagógico, y más en concreto de la evaluación didáctica, y técnicamente muy puntual, es la distinción entre «evaluación normativa» y «evaluación criterial».

Alfonso Heredia Manrique

En un sistema de enseñanza basado en objetivos, el cambio de perspectiva que propone la evaluación criterial frente a la evaluación normativa supone incrementar la utilidad de la misma en el ámbito escolar (Lizosoain *et al.*, 2004: 239), aunque es verdad que técnicamente es un cambio a mayor practicidad y de mayor consistencia por la concreción de los aprendizajes que realiza; es también mucho más que un cambio de perspectiva, como apuntan los anteriores autores, interpretación que puede llevar a ese tipo de confusión de la que queremos hablar, porque supone, sobre todo, un cambio de filosofía evaluativa, lo cual queremos resaltar. Éstas son evaluaciones didácticas distintas, cada una pertenece a tradiciones evaluativas diferentes; la primera forma parte de la pedagogía tradicional y la segunda de la pedagogía progresista, por tanto, suponen mentalidades evaluativas distintas. Sin embargo, no todos los autores pensarían igual, ya que consideran a la evaluación criterial como una evaluación conductista que lo es y por ello se situaría dentro de la pedagogía tradicional; esto ocurre porque se hace una identificación total y de continuidad entre la pedagogía tradicional y la pedagogía basada en el conductismo, algo que puede cuestionarse, tal y como veremos. La evaluación criterial pertenece a la denominada pedagogía de dominio o del éxito, la cual técnicamente es un tipo de pedagogía por objetivos, pero también una interpretación progresista del conductismo; diríamos que se hace un uso técnico del conductismo dentro de los planteamientos progresistas. El papel educativo de la pedagogía de dominio es conseguir el éxito total en la instrucción para cada alumno individualmente; de ahí la denominación también como pedagogía del éxito. Así, se busca intencionadamente que la mayoría de los alumnos, si no todos, logren los objetivos educativos establecidos, objetivos que son o pueden ser de tipo conductual, al menos una parte de los mismos, sin que ello provoque ninguna contradicción teórica e ideológica.

Este discurso teórico subyacente y legitimador asume una ideología educativa progresista y la

forma de operativizarlo a través de la evaluación criterial es de carácter conductista. Y esto es lo que puede parecer sorprendente porque el conductismo se ha entendido como continuación y parte de la pedagogía tradicional, pero sin pensar que en su momento podría haber representado un planteamiento progresista. Como ejemplo de lo que decimos es clarificador que Leonard y Utz (1979) en su libro, y a lo largo del mismo, *La enseñanza basada en competencias*, el cual es un desarrollo basado en el conductismo, están constantemente desmarcándose de la pedagogía tradicional y hablando de las diferencias de sus propuestas respecto a la misma, algo que por otra parte, viendo el título del libro, no pensaríamos según la asociación establecida. Igualmente, el libro de Bloom y cols. (1975), *Evaluación del aprendizaje* (volumen 1), surge en un contexto conductista, pero se olvidan u ocultan interesadamente sus planteamientos claramente progresistas. Ya en el primer capítulo son conscientes de que en general siempre se ha pensado que la educación supone distribuir a las personas por categorías y que la evaluación se utiliza en ese sentido para clasificar. Y en su tercer capítulo: «Aprendizaje para el dominio», hacen unas consideraciones interesantes para retomar el sentido original de la evaluación que proponen (su propio título es muy indicativo). Aquí los autores se posicionan en una perspectiva bien contraria a la tradicional y es que la cuestión en educación para ellos no es encontrar una minoría de éxito, sino determinar cómo una mayoría de alumnos pueda aprender lo esencial para el progreso personal en una sociedad compleja; de ahí deriva la necesidad de que el aprendizaje trate de ser exitoso para todos.

Éstos rechazan explícitamente el uso de la calificación para evaluar a los alumnos porque supone la evaluación tradicional que les clasifica; en este sentido, y como muestra de ello, dicen que el uso de la evaluación formativa implica que las pruebas utilizadas deben mostrar el dominio o no del aprendizaje y no deben estar acompañadas por calificaciones. La evaluación formativa sólo debe informar al estudiante (no calificar)

si ha dominado o no la unidad. Y así nos encontramos con dos valoraciones, a saber: «aprobado» cuando se ha aprendido y «reprobado» cuando todavía tiene que aprender (Bloom y cols., 1975: 89, 194 y 155). También consideran que: «Precisar el grado de dominio de una determinada tarea del aprendizaje y señalar con exactitud la parte de la tarea no dominada [...] no es calificar al alumno, sino ayudarlo [...] en la forma particular de aprendizaje necesaria para avanzar hacia el dominio» (Bloom y cols., 1975: 97). Igualmente se critica el uso de las pruebas sumativas para asignación de calificaciones y para comparar resultados de diferentes grupos porque implican una función selectiva y esto no cabe en la concepción que proponen del aprendizaje de dominio.

A todo ello añadiríamos que históricamente la teoría del *mastery learning* (aprendizaje para el dominio) no es más que desarrollar la vieja idea de aprender para dominar lo aprendido y cuyos antecedentes Block (1979: 2-3) los sitúa en la década de los años veinte del siglo pasado, concretamente en el Plan Winnetka de Whashburne (1922) y en Morrison (1926) en la Universidad de Chicago; posteriormente este concepto (*mastery learning*) va desapareciendo hasta que resurge de nuevo hacia finales de los años cincuenta y comienzo de los años sesenta, asociado al conductismo del momento dentro de la instrucción programada. Así, la evaluación criterial surge en el cambio de dicha década por la convergencia de hechos tales como: la difusión de la enseñanza programada, el desarrollo de las teorías del aprendizaje y auge de los objetivos conductuales y el enfoque de la enseñanza basada en objetivos, creándose un clima que puso de manifiesto la necesidad de nuevos instrumentos que aportasen más información que la proporcionada por los instrumentos normativos sobre lo que un sujeto es capaz de hacer (Rodríguez Lajo, 1985: 304).

Concretamente, la enseñanza individualizada desarrollada en los años sesenta en general necesitaba un tipo de pruebas que determinarían

el progreso del alumno, distintas a las utilizadas hasta entonces, como eran las pruebas normativas. Éstas, además, habían generado determinados efectos negativos que debían corregirse, tales como: una competitividad malsana, la desmotivación de los más retrasados, la penalización de alumnos capacitados pero que sufrían las consecuencias de la «astucia» de algunos ítems y la incapacidad de determinar claramente las habilidades dominadas (Mafokozi, 1987: 452). Así, surgen las pruebas criteriales como alternativa de solución a los aspectos comentados. Su objetivo general sería mejorar y optimizar la calidad de la enseñanza, dado que pueden asegurar en mayor medida una evaluación objetiva del conocimiento de cada alumno en distintas fases, etapas o momentos del ciclo formativo (Castejón, Navas y Sampascual, 1998: 266).

También en ese momento, según Popham (1980: 146), viendo las posibilidades de la enseñanza programada, un grupo de psicólogos y educadores pensó en las potencialidades de una enseñanza bien concebida; así, se partía de la idea de que con secuencias pedagógicas bien elaboradas todos los alumnos podían dominar cualquier objetivo. Los objetivos educativos de este planteamiento se refieren tanto a conductas desplegadas por los alumnos como a competencias que pueden manifestarse en ellos cuando se necesiten. Así, los tipos de conducta de dichos objetivos son en realidad muy amplios, ya que no es sólo la manifestada después de un periodo de enseñanza, sino también capacidades conductuales, y pone el ejemplo de la adquisición de destrezas en lógica formal como es el razonamiento silogístico, de manera que el alumno pudiera emplear esta capacidad cuando fuese necesario.

Igualmente, es muy significativa la consideración que hace Popham (1980: 59-60), uno de los máximos exponentes del uso de la evaluación criterial, cuando reconoce que se cometió un error estratégico al denominar a los objetivos educativos de los que hablaban como «objetivos conductuales», porque llevó a muchos

Alfonso Heredia Manrique

educadores a confundir la defensa de dichos objetivos como un enfoque mecanicista y deshumanizado de la educación, y dada esa confusión hubiese sido mejor elegir otra expresión que no fuese la de objetivos conductuales, como por ejemplo la de objetivos de ejecución. Asimismo, comenta que ellos no querían hacer referencia ni defender al enfoque psicológico del conductismo cuando hablaban de esos objetivos, sino que lo que querían era resaltar por un lado una manera de discernir si el objetivo se había logrado o no, algo que se ponía de manifiesto con la conducta posterior del alumno, y, sobre todo, por otro lado, señalar que el centro de atención era el alumno, no el profesor ni el contenido. Como vemos, esto último es un indicador básico del planteamiento progresista que tenía esta orientación.

### **Caracterización de la evaluación normativa y de la evaluación criterial**

La distinción entre evaluación normativa y evaluación criterial la hace Glaser en 1963, autor que utiliza por primera vez la expresión «evaluación criterial» (Popham, 1983: 31). Estos dos tipos de estrategias de evaluación hacen referencia explícita al rendimiento del alumno, cada una dando su propia interpretación del mismo, pero en todo caso desde una perspectiva muy técnica y operativa, algo que con frecuencia es olvidado actualmente en el ámbito educativo que se mueve más en la demagogia retórica del lenguaje, y por ello sus diferencias se dirimen a través de aspectos tan concretos como en el tipo y características de las pruebas o test que utilizan. La misma elaboración de test referidos al criterio constituye un ámbito de conocimiento técnicamente complejo, al tenerse en cuenta aspectos como la definición del dominio educativo y la determinación de estándares y puntos de corte. Ahora conviene clarificar la diferencia existente entre evaluación normativa y evaluación criterial, que además de ser una cuestión técnica y estratégica del campo de

la evaluación, también para ver que, sobre todo, corresponden a dos planteamientos evaluativos bien diferenciados teórica e ideológicamente en coherencia con lo dicho en el apartado anterior.

#### **Evaluación normativa**

Esta estrategia de evaluación tiene como objetivo ver la posición relativa de cada alumno respecto a los demás y cumple su función cuando con ella se sabe quién es mejor (o peor) que quién. Así, trata de identificar el talento del alumno, de manera que se puedan establecer comparaciones y luego hacer una selección entre los alumnos (Popham, 1983: 31 y 27). La definición técnica de evaluación referida a la norma lo deja muy claro: «Un test basado en pautas de normalidad está destinado a determinar la posición de un sujeto examinado en relación con el rendimiento de un grupo de otros sujetos que haya hecho ese mismo test» (Popham, 1983: 47). Ése será el grupo normativo respecto al cual se compara el rendimiento de quien haga dicho test. Los test basados en pautas de normalidad son excelentes instrumentos para la finalidad con que fueron creados, a saber: «La discriminación entre los examinados en una situación de números clausus en la que sólo hay posibilidades de admisión para un número limitado de individuos» (Popham, 1983: 47). La evaluación normativa tiene como referencia las realizaciones de los alumnos del mismo grupo, tratando de medir las diferencias individuales de rendimiento, todo ello basado en el modelo de curva normal cuya finalidad es clasificar a los alumnos. La media del grupo se utiliza como criterio de superación o no (punto de corte) y será variable, más alta o más baja, en función de las capacidades demostradas por el grupo de alumnos: siempre habrá un 50% de alumnos máster por encima de la media y un 50% de alumnos no-máster por debajo de la media. Así, el mayor o menor rendimiento individual se disgregará dentro del rendimiento del grupo, el cual, como hemos dicho, puede ser muy variable (Gómez Arbeo, 1990: 12). Ahora

bien, cuando se necesita saber exactamente qué es lo que puede hacer o no hacer el alumno entonces la evaluación normativa es inadecuada, porque no proporciona información clara de los conocimientos y habilidades adquiridas por el alumno, que es precisamente lo que satisface la evaluación criterial dada su naturaleza; así, cada una de estas formas de evaluación tiene su propia finalidad.

Quizá una de las diferencias más destacables es en relación a la forma de tratamiento de los contenidos de conocimiento dentro de las pruebas utilizadas. Primero porque la forma de describir el campo de conocimientos que cubren es distinta en el grado de clarificación de dichos contenidos, muy clara y precisa en la evaluación criterial y poco clara y difusa en la evaluación normativa. Y segundo, como consecuencia de la finalidad que tienen esos contenidos en cada una de ellas, en la evaluación criterial hay que saber contenidos tal cual porque es uno de los fines del aprendizaje, y en la evaluación normativa los contenidos son una excusa para utilizarlos como selección de alumnos; de ahí que Popham (1983) hable de la «trampa psicométrica» que se hace con ellos en las pruebas normativas. Esta trampa que se realiza al construir el test normativo consiste en eliminar los ítems relacionados con los contenidos más importantes de la enseñanza (justo lo contrario de lo que hacen las pruebas criterios) y, por tanto, los que más enseñan los profesores y los que mejor son aprendidos por los alumnos, es decir, aparecen los ítems de contenidos que no tienen la probabilidad de ser respondidos correctamente por la mayoría de los alumnos, porque son los que pueden discriminarlos y diferenciarlos jerárquicamente, que es de lo que se trata. Este tipo de trampa realizada indica claramente que lo importante no es aprender, sino seleccionar con independencia de lo aprendido y de ahí que hablen, más que de aprender contenidos, de con qué tipo de contenidos podemos seleccionar mejor a los alumnos; así los contenidos menos relevantes y más secundarios por ser peor aprendidos son los

que más posibilitan la discriminación entre alumnos. Por tanto, estas pruebas tienden a estimular la comparación de los resultados de los alumnos y la competición consiguiente entre ellos, no proporcionan información precisa sobre el dominio de conocimientos y destrezas importantes y básicas del alumno, y, consiguientemente, no permiten establecer estrategias individuales de corrección y mejora de las deficiencias (Gómez Arbeo, 1990: 12-13). En definitiva, estas pruebas corresponden a una pedagogía tradicional donde la evaluación se entiende y utiliza para seleccionar y clasificar a los alumnos.

### **Evaluación criterial**

La característica fundamental de esta evaluación es apreciar el logro de los objetivos por parte de cada alumno sin compararlo con el de sus compañeros (Gómez Arbeo, 1990: 7). La misma autora proporciona un claro ejemplo donde se muestra la lógica subyacente de la evaluación criterial, a saber, en la obtención del carné de conducir; aquí no importa la comparación del rendimiento de los examinandos, sino la comparación de cada examinando con los criterios objetivos necesarios para ser un buen conductor. La evaluación criterial sería una forma de operativizar la evaluación del aprendizaje individual del alumno respecto a objetivos educativos que actúan como criterio comparativo, y en términos conductuales, oponiéndose a la comparación de aprendizajes entre alumnos. En última instancia podemos comparar a los alumnos en función de los objetivos logrados, pero es lo que hay que evitar hacer porque no es eso lo que se busca en la enseñanza obligatoria, que es la etapa para la que pensamos el uso de la evaluación criterial; otra cuestión sería su posibilidad de aplicación en otros niveles y ámbitos educativos, tal y como ya se ha hecho.

La evaluación criterial tiene un amplio uso, no sólo con contenidos conceptuales, sino también con contenidos procedimentales y, asimismo, en su extensión a ámbitos como la Educación

Alfonso Heredia Manrique

Infantil y la Educación Especial para lo que son el logro de conductas o habilidades perceptivo-motrices (de carácter procedimental). «Las pruebas criteriales permiten (...) comprobar el uso y aplicación de los conocimientos procedimentales en diferentes situaciones particulares, a la vez que ofrecen unas pautas claras para la observación de la calidad de su ejecución, grado de automatización, generalización a diferentes contextos, grado de conocimiento sobre el procedimiento y grado de acierto en la elección del procedimiento para solucionar una tarea» (Fullan y Pallisera, 1995: 42).

La evaluación criterial proporciona la realización del alumno con independencia de su comparación relativa con los otros y sus pruebas buscan descripciones claras del rendimiento del alumno (Popham, 1983: 131, 134 y 159). Aquí lo más importante es dar una descripción clara de la clase de conducta que puede o no manifestar el alumno; otra cuestión son las dificultades o limitaciones en el rigor descriptivo de las especificaciones de conducta que se hacen. Con las pruebas se determina la posición del alumno respecto a un dominio de conducta perfectamente definido, el cual «consiste en un conjunto de habilidades o de disposiciones que los examinados ponen en juego cuando se les pide que lo hagan en una situación de examen» (Popham, 1983: 134). Por ejemplo, en Matemáticas podría ser la capacidad de un individuo para resolver cierta clase de problemas de ecuaciones.

La evaluación criterial desde el punto de vista didáctico sería una evaluación exclusivamente de aprendizajes de los alumnos, sin salirse de lo que es el ámbito pedagógico. Es una evaluación por objetivos de aprendizaje, los cuales son destrezas académicas conductuales a lograr y por ello comportamientos observables en el alumno. También supone la evaluación individual del alumno en relación a la progresión sobre el dominio de cada objetivo y ajustada a su ritmo de aprendizaje. Una evaluación que puede diagnosticar tanto el nivel de aprendizaje adquirido

por el alumno como las deficiencias del mismo y que en este caso se puede intervenir de manera inmediata dada la información tan precisa que proporciona. Establece un punto de corte de dominio o no dominio del objetivo de aprendizaje, lo que el alumno puede o no puede hacer, y cuya valoración se realiza en términos de pasa/no pasa o apto/no apto, pero sin calificación numérica. Y es una evaluación comparativa con el criterio establecido que decide el dominio de la destreza, pero no una comparación entre alumnos; por esta interpretación se produce la deriva hacia la evaluación normativa.

Como ejemplo de utilidad pedagógica de la evaluación criterial hacemos referencia a la investigación realizada por Rivas y Alcantud (1988) sobre el conocimiento mínimo de EGB en su momento y equivalente hoy a los niveles correspondientes de la Educación Primaria, en cuanto a objetivos, contenidos y criterios de nivel de dominio. Diríamos que los aprendizajes exigidos son básicos, elementales e imprescindibles, no excluyentes de todo lo que además se quiera o se pueda aprender y evaluar con otros instrumentos en dicha etapa. En ella se define un «universo de medida» de entre unos contenidos básicos permanentes, aquellos que están siempre en el currículo consensuados por expertos. Así, tenemos unas áreas curriculares: Matemáticas, Lenguaje, Ciencias Sociales y Ciencias Naturales; dentro de ellas unos sectores curriculares que corresponden a los contenidos exclusivos y exhaustivos en lo fundamental. Ejemplo de Matemáticas: conjuntos y relaciones, numeración, operaciones, solución de problemas, topología y geometría, y medida; también unos objetivos de aprendizaje de carácter conductual para todos los sectores curriculares de las áreas; y los elementos de medida como son las pruebas criteriales de evaluación. Estas pruebas sostienen todo el edificio anterior de manera que determinan el dominio de conocimiento, el cual está traducido a objetivos conductuales, presentándose en forma de rendimiento del alumno con unos criterios de nivel de exigencia. Al final se llega a poder establecer

el nivel de dominio «todo» o «nada» de los aprendizajes, el cual se basa en la idea de que el dominio del aprendizaje se puede establecer con nitidez mediante un punto de corte que separe con claridad a los máster (dominio) de los no-másters (no-dominio) (Rivas y Alcantud, 1988: 78).

Por tanto, la evaluación criterial se fundamenta en la precisión de una serie de aspectos: universo de medida, objetivos de aprendizaje, elementos de medida, punto de corte, que tienen como finalidad, esto es lo que se olvida, saber si cada alumno ha adquirido el conocimiento mínimo del área curricular, algo que para nada le perjudica, sobre todo en un momento como el actual, en el cual se detecta de manera masiva la escasez de conocimientos concretos en los alumnos. Con todo esto queremos decir que la evaluación criterial técnicamente es viable, aunque costosa, como modalidad de evaluación, para lo que serían los aprendizajes más básicos y elementales del currículo obligatorio; la escuela primaria «pretende asegurar un mínimo de cultura e instrucción a todos los ciudadanos» (Rivas y Alcantud, 1988: 41). Cuando vemos de qué tipo de aprendizajes estamos hablando, se pueden consultar los extensos anexos de la investigación citada; creo que será muy difícil defender que esos aprendizajes no hay que llevarlos a cabo y a su vez también se disolverán muchos debates retóricos y estériles sobre qué y cuántos contenidos concretos hay que aprender.

### **Deriva de la evaluación criterial hacia la evaluación normativa**

El currículo que plantean la últimas leyes educativas en nuestro país establece una serie de niveles de objetivos engarzados unos con otros (de etapa, de ciclo, de área) que llevan a interpretar que la evaluación criterial es la más adecuada para conseguirlos (Lizasoain *et al.*, 2004; Díaz y García, 2004). Es por ello que no hay que traspasar la fina frontera existente respecto a la evaluación normativa si de verdad queremos

que actúe la evaluación criterial en dicho currículo, pero de hecho se observa esa transgresión. Una ejemplificación concreta del modo como ocurre la deriva entre ambos tipos de evaluación la encontramos en la investigación de Díaz y García (2004). Aquí se hace un planteamiento muy adecuado de la evaluación criterial dentro del área de Matemáticas en los tres ciclos de Educación Primaria, con una completa sistematización de su aplicabilidad, teniendo en cuenta los objetivos, los contenidos y los criterios de evaluación establecidos en el currículo para dicha área. Los autores interpretan a partir del análisis que hacen de la legislación hasta la LOCE que la evaluación criterial es la que subyace en el marco legislativo desde los Programas Renovados de 1981 y esto lo deducen de la red de objetivos de distinto nivel establecidos en el currículo de las distintas leyes educativas promulgadas hasta el momento. Y así hacen una adaptación de la evaluación criterial a dos marcos legislativos distintos, el primero el anterior a la LOCE y el segundo el de la misma LOCE.

Los autores parten de la necesidad de que exista un consenso en el equipo docente sobre los objetivos esenciales mínimos de ciclo, los cuales serían aquellos que permitiesen al alumno continuar el proceso de enseñanza-aprendizaje en el ciclo siguiente. Cuando se sitúan en el marco evaluativo anterior a la LOCE, si el alumno consigue dichos objetivos entonces se le califica como «prograsa adecuadamente» y si no es así, la calificación es «necesita mejorar»; a esta dicotomía evaluativa se le puede hacer corresponder perfectamente con el denominado punto de corte que establece la evaluación criterial respecto al dominio o no dominio del aprendizaje, que en este caso sería respecto a esos objetivos esenciales establecidos. Así, diríamos que aquí hay un encaje técnico perfecto de la filosofía evaluativa que propone la evaluación criterial. Por otro lado, cuando se sitúan en el marco evaluativo de la LOCE con cinco categorías de evaluación, a saber, insuficiente, suficiente, bien, notable y sobresaliente y quieren encajar la evaluación criterial es cuando se produce la

*Alfonso Heredia Manrique*

deriva de la que estamos hablando, saltándose los límites entre los dos tipos de evaluación. Esto se observa ya desde el principio cuando dividen los objetivos en objetivos mínimos esenciales y todos los objetivos por ejemplo de un ciclo; la adquisición de los primeros supondría la calificación de «suficiente» y su no adquisición sería «insuficiente», además estos objetivos suponen el 65% de los objetivos totales. También establecen tres puntos de corte criteriosales más en correspondencia con las otras tres categoría de evaluación restantes, en orden de dificultad creciente, a saber, el primero para el «bien» cuando el alumno consigue el 74% de los objetivos previstos, el segundo punto de corte para el «notable» cuando consigue el 82% de los objetivos y el tercero para el «sobresaliente» a partir del 90% de los objetivos previstos.

Establecer categorías de dominio de aprendizaje asociadas a niveles graduados de calificaciones diferentes, que en este caso corresponden a un porcentaje determinado de objetivos logrados, es precisamente establecer una clasificación de alumnos tal como plantea la evaluación normativa, diferenciando a los alumnos comparativamente por el número de objetivos logrados y por tanto estamos en lo mismo, se compara a los alumnos. No debería ser el caso para una educación primaria y obligatoria y así no se entiende la diferenciación que hacen los autores entre todos los objetivos de ciclo y los objetivos esenciales mínimos de ciclo, dado que partimos de un tipo de educación básica igual para todos y entonces todos los objetivos del currículo son esenciales, no cabe la diferenciación de objetivos. Es más, cuando en su discurso hablan de la necesidad de la precisión de las categorías para asignar adecuadamente a cada alumno a la que le corresponde, están manifestando claramente la deriva entre evaluaciones.

Hay una gran contradicción de fondo, la cual anularía la filosofía de la evaluación criterial, cuando por un lado se considera como muy importante dicha evaluación entendiendo que la propugna la normativa y es más acorde con

la idea actual que se tiene de la evaluación, cuyo objetivo «no es clasificar o calificar a los alumnos, sino articular procesos de enseñanza-aprendizaje rigurosos y sistemáticos para que el alumno aprenda más y mejor, es decir, para conseguir eso que llamamos calidad de la enseñanza» (Díaz Alcaraz, 2003: 228) y, por otro lado, plantear categorías con los objetivos didácticos según el porcentaje de logro para calificar después dentro de una escala (insuficiente, suficiente, bien, notable y sobresaliente), como dicho autor lleva a cabo también en la ESO. Este planteamiento evaluativo podría caber en una enseñanza no obligatoria como pueda ser en la Educación Superior, donde los puntos de corte dicotómicos (pasa, no pasa) o policótomos (no apto, apto, notable...) se refieren «a la puntuación real en la prueba que deben obtener los sujetos para ser asignados a una categoría de conocimiento» (Jornet, Suárez, González y Belloch, 1997: 234).

Aquí el problema surge cuando a alguien se le ocurre que hay que mirar comparativamente los aprendizajes de los alumnos y en esa comparación se observan diferencias entre unos y otros, en realidad siempre las hay. Debido al ritmo individual de aprendizaje, unos logran antes que otros los objetivos, incluso unos logran el dominio del objetivo mejor que otros. Y entonces se considera que hay que establecer diferencias entre mejores y peores alumnos y reflejarlo mediante un tipo de evaluación objetiva adecuada, en este caso utilizando una escala de medida, pero de este modo nos situamos en la evaluación normativa. La cuestión es si esto vale para todos los casos y aprendizajes. La evaluación normativa es anterior históricamente y su objetivo es diferenciar con el objetivo de seleccionar y clasificar en este orden, y esto se ha configurado como una mentalidad evaluativa muy potente en el pensamiento educativo y que responde a intereses sociopolíticos ajenos a lo estrictamente pedagógico. Para Popham (1980: 322) «lo mismo que la masa endurecida, las actitudes endurecidas son difíciles de modificar. Hay algunas actitudes con



respecto a la evaluación educativa bastante endurecidas que nos desafían a que le demos nueva forma. Los profesionales de la educación han sostenido durante tantos años un punto de vista particular sobre la evaluación educativa

bastante endurecido que nos desafía a que le demos nueva forma». Es por eso que la deriva de la evaluación criterial hacia la normativa se vería influida por dicho endurecimiento actitudinal.

## Referencias bibliográficas

- BLOCK, J. H. (1979). *Cómo aprender para lograr el dominio de lo aprendido (mastery learning)*. Buenos Aires, El Ateneo.
- BLOOM, B.; HASTINGS, J. y MADAUS, G. (1975). *Evaluación del aprendizaje*, tomo I. Buenos Aires, Troquel.
- CASTEJON, J. L.; NAVAS, L. y SAMPASCUAL, G. (1998). Resultados de una evaluación criterial en el área de Ciencias Sociales. Geografía e Historia de la Educación Secundaria Obligatoria, *Revista de Educación*, 315, 251-267.
- DÍAZ ALCARAZ, F. (2003). Un modelo de evaluación criterial para el área de Matemáticas en la enseñanza obligatoria, *Ensayos. Revista de Estudios de la Escuela Universitaria de Magisterio de Albacete*, 18, 225-244.
- DÍAZ ALCARAZ, F. y GARCÍA GARCÍA, J. J. (2004). *Evaluación criterial del área de Matemáticas (un modelo para Educación Primaria)*. Barcelona, Cisspraxis.
- FULLANA, J. y PALLISERA, M. (1995). Aplicación de las pruebas criterioles en la evaluación de los procedimientos, *Bordón*, 47(1), 41-51.
- GÓMEZ ARBEO, B. M. (1990). *Evaluación criterial (una metodología útil para diagnosticar el nivel de aprendizaje de los alumnos)*. Madrid, Narcea.
- JORNET, J. M.; SUÁREZ, J. M.; GONZÁLEZ, J. y BELLOCH, C. (1997). Estrategias de elaboración de pruebas criterioles en educación superior, en MARTÍNEZ MEDIANO, C. (coord.), *Encuentros en la Facultad de Educación sobre evaluación*. Madrid, UNED, 207-240.
- LEONARD, L. D. y UTZ, R. T. (1979). *La enseñanza como desarrollo de competencias*. Madrid, Anaya.
- LIZASOAIN, L. et al. (2004). Establecimiento de estándares y puntos de corte en pruebas de rendimiento. Una aplicación en el área de Matemáticas de 2º y 4º de ESO, *Bordón*, 56(2), 237-252.
- MAFOKOZI, J. (1987). Evaluación criterial, *Bordón*, 268, 451-464.
- POPHAM, W. J. (1980). *Problemas y técnicas de la evaluación educativa*. Madrid, Anaya.
- POPHAM, W. J. (1983). *Evaluación basada en criterios*. Madrid, Magisterio Español.
- RIVAS, F. y ALCANTUD, F. (1988). *La evaluación criterial en la educación primaria*. Madrid, MEC.
- RODRÍGUEZ LAJO, M. (1985). Evaluación del rendimiento criterial versus normativa. Modelo de evaluación FCO, *Revista de Investigación Educativa*, 6(3), 304-321.

## Abstract

### *Two differentiated forms of didactic evaluation: normative evaluation in order to select students and criterial evaluation to master the basic content*

This article focuses on two forms of didactic evaluation, the normative evaluation and the criterial evaluation, in order to use them again if necessary. Both forms, although with a different educative goal, have happened historically in the evaluation field. The normative evaluation responds to the selective tradition of students. The criterial evaluation has a different aim: any student should mastery the basic contents of the curriculum.

**Key words:** *Evaluation, Evaluation of students, Evaluation method, Educational achievement.*

*Alfonso Heredia Manrique*

---

## **Perfil profesional del autor**

---

### **Alfonso Heredia Manrique**

Doctor en Ciencias de la Educación por la Universidad de Valencia. Profesor titular de escuela universitaria del Área de Didáctica y Organización Escolar del Departamento de Ciencias de la Educación de la Universidad de Zaragoza. Actualmente su docencia se desarrolla en las distintas especialidades de Maestro. Sus líneas de investigación se centran sobre evaluación didáctica, recursos curriculares y formación del profesorado.

Correo electrónico de contacto: [aheredia@posta.unizar.es](mailto:aheredia@posta.unizar.es)