

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/44551123>

Bases teóricas de la evaluación educativa / Juan José Monedero Moya

Article

Source: OAI

CITATIONS

20

READS

22,957

1 author:



[Juan Jose Monedero Moya](#)
University of Malaga

16 PUBLICATIONS 86 CITATIONS

[SEE PROFILE](#)

Capítulo II.- La evaluación educativa

Introducción.....	
Capítulo II.1.- Consideraciones epistemológicas de la evaluación educativa: delimitación conceptual e ideas afines	
II.1.1.- Apreciación, tasación, estimación	
II.1.2.- Decisión	
II.1.3.- Indicadores.....	
II.1.4.- Información	
II.1.5.- Investigación educativa.....	
II.1.6.- Juicio, reflexión	
II.1.7.- Medición, medida.....	
II.1.8.- Valoración.....	
II.1.9.- Exámenes, controles, tests o pruebas de rendimiento escolar	
II.1.10.- Calificación, puntuación, nota.....	
II.1.11.- Responsabilidad, rendición de cuentas.....	
II.1.12.- Evaluación	
Capítulo II.2.- Las formas de la evaluación educativa	
II.2.1.- Evaluación inicial	
II.2.2.- Evaluación diagnóstica.....	
II.2.3.- Evaluación continua y evaluación formativa.....	
II.2.4.- Evaluación final y evaluación sumativa	
II.2.5.- Evaluación formal vs. evaluación informal.....	
II.2.6.- Autoevaluación vs. heteroevaluación	
II.2.7.- Evaluación interna vs. evaluación externa	
II.2.8.- Evaluación individual vs. evaluación grupal o colectiva.....	
II.2.9.- Evaluación cuantitativa vs. evaluación cualitativa	
II.2.10.- Evaluación por normas vs. evaluación por criterios.....	
II.2.11.- Indicadores.....	
Capítulo II.3.- Modelos representativos de la evaluación educativa	
3.1.- Modelos clásicos.....	
II.3.1.1.- Modelos basados en la consecución de metas u objetivos.....	
II.3.1.1.1.- Modelo de Tyler.....	
II.3.1.1.2.- Modelo de Adams	
II.3.1.1.3.- Modelo de Taba	
II.3.1.1.4.- Modelo de Carreño.....	
II.3.1.1.5.- Modelo de Hammond.....	
II.3.1.1.6.- Modelo de Metfessel y Michael.....	
II.3.1.1.7.- Evaluación retributiva y comparativa	
II.3.1.2.- Modelos basados en la formulación de juicios.....	
II.3.1.2.1.- Modelo de acreditación	
II.3.1.3.- Modelos facilitadores para la adopción de decisiones.....	
II.3.1.3.1.- Modelo CIPP	
II.3.1.3.1.- Modelo CSE.....	
II.3.2.- Modelos alternativos.....	
II.3.2.1.- Evaluación como crítica artística	

II.3.2.2.- Evaluación libre de metas	
II.3.2.3.- Método del <i>modo operandi</i>	
II.3.2.4.- Evaluación comunicativa	
II.3.2.5.- Evaluación iluminativa.....	
II.3.2.6.- Evaluación respondiente	
Capítulo II.4.- Principales problemas asociados a la evaluación educativa.....	
II.4.1.- Aspectos éticos de la evaluación.....	
II.4.2.- Aspectos político-económicos de la evaluación	
II.4.3.- Aspectos personales de la evaluación.....	
II.4.4.- Aspectos sociales de la evaluación	
II.4.5.- Aspectos técnicos de la evaluación.....	



Introducción

Nos encontramos inmersos en una etapa de continuos y grandes cambios. Se nos va precipitadamente un siglo, repleto de avatares en todos los órdenes -artísticos, bélicos, científicos, ecológicos, económicos, geográficos, ideológicos, políticos, religiosos, sociales, tecnológicos, etc.- y con su extinción nace un nuevo milenio. Son muchos los sectores que en su evolución, o revolución, han cambiado hasta hacerse irreconocibles respecto de sus orígenes. Sin embargo aún se mantienen, parcialmente, estructuras obsoletas que se han «enquistado» en las instituciones sociales. Estos restos «fósiles» cumplen un papel funcional que responde, explícita o tácitamente, a la precisión de variados intereses sociopolíticos y económicos. Obviamente si perduran es tanto por su necesidad como por la incapacidad de implantación de los nuevos modelos, inhábiles para ofrecer una alternativa satisfactoria tal que permita el avance de la humanidad «sin alterar sustancialmente el orden establecido».

Pero como consecuencia de su perdurabilidad no sólo originan disfunciones, sino que igualmente afectan de manera muy profunda a las creencias y a los valores humanos. La diversidad axiológica social, sin parangón en otras épocas históricas mucho más coherentes por simples, crea multitud de situaciones en las que sus miembros, lejos de compartir una filosofía común como forma de explicar la vida y la realidad, se atomizan y segregan. Y cada grupo, a su modo, trata de avanzar frecuentemente oponiéndose y obstaculizando a los demás. Es un desgaste de energías sociales que, por la fuerza de las armas (guerras y dictaduras), en unos casos; o por la del respeto a las mayorías (paz y democracias), en otros; trata de resolverse con la imposición desde el poder de una determinada línea de acción.

Desde esta perspectiva de dialéctica humana hemos enfocado el análisis de la educación en general y, más particularmente, ahora, el de la evaluación educativa. Preguntas clásicas como por qué, para qué, o qué evaluar, o estas otras, cómo es la práctica de la evaluación, qué aspectos intervienen e influyen decisivamente en ella, nos llevan irremisiblemente «a poner el dedo en la llaga» de la mencionada conflictividad social.

Como cabría esperar, el planteamiento de estas dudas nos envuelve en el centro de un inmenso océano en el que para discurrir con cierta seguridad por sus aguas y poder llegar a puerto, hemos debido de plantearnos otras cuestiones más concretas y fáciles de observar y estudiar.

En la primera parte de este capítulo he tratado de delimitar conceptualmente la evaluación educativa en general. He creído oportuno adentrarme previamente en el polisémico universo de los conceptos evaluativos que, aun estando semánticamente tan próximos, pueden, sin embargo, inducir fácilmente a la confusión.

En la segunda sección, trataré de las diferentes formas, que desde mi concepción, puede adoptar la evaluación educativa. Estas formas, que distingo de los modelos, obedecen más a la metodología práctica de la evaluación educativa -como la desarrollan los profesores en la acción de sus aulas- que a los propios modelos en sí, que representan esquemas parciales de las filosofías y paradigmas que hemos visto en el primer capítulo. Obviamente estas formas también comparten con aquéllos sus particulares filosofías, sus peculiares modos de interpretar la realidad educativa.

En una tercera parte, expongo y analizo algunos de los modelos de evaluación usados más comúnmente, porque considero que han ejercido, y aún cultivan, una dilatada influencia en la comunidad académica; aunque sólo haya reparado en los que me llamaron particularmente la atención, o en aquellos otros que se me antojaron más relevantes, dado su extenso tratamiento en la literatura. No trataré sin embargo, deliberadamente, los instrumentos de evaluación utilizados por el profesorado, por considerar que éste aspecto se escapa a las pretensiones y objetivos del presente Proyecto Docente.

Por fin, en la cuarta y última parte, planteo algunas dudas. Especialmente relacionadas con la concepción de la evaluación como el «instrumento» de la selección y de la clasificación de las personas... Porque si quienes evaluamos no somos conscientes y críticos con la servidumbre de nuestras acciones, estaremos defendiendo unos intereses ajenos a nuestros conceptos y valores educativos más preciados: la dignidad y la libertad humanas contra la sumisión acrítica al sistema.

Todo ello, desde el enfoque que representa la evaluación educativa en el ámbito de la Didáctica en general, como desde el que supone la docencia en el perfil de ésta plaza a la que concurso: Diseño, Experimentación y Evaluación de Materiales Curriculares.



Cap. 1.- Consideraciones epistemológicas de la evaluación educativa: delimitación conceptual e ideas afines

Dentro de nuestra cultura educativa, el término evaluación aparece por vez primera de un modo generalizado con la promulgación de la Ley General de Educación de 1970. Desde entonces su concepción se ha ido haciendo más compleja (Taba, 1983, pág. 407) y provocando un mayor grado de confusión, paulatinamente, con su extensión a los diferentes ámbitos de la enseñanza.

La renovación conceptual, en la que se proyectan las aportaciones de numerosas investigaciones educativas de los últimos treinta años, se manifiesta en un considerable enriquecimiento de los ámbitos sobre los que se extiende la evaluación, de los criterios con los que se realizan juicios de valor y de las decisiones que se adoptan en relación con ellos. Queda ya lejos el viejo concepto de evaluación como constatación de un cierto nivel de aprendizaje en el alumno. Es necesario evaluar no sólo conocimientos, sino también habilidades y actitudes. Es necesario conocer las características de los procesos y no sólo de los resultados. Por otra parte, la evaluación ha de abarcar además la tarea del profesor, el desarrollo de los programas, la efectividad de los recursos y la influencia de los contextos organizativos y ambientales. La actividad evaluadora ha de sobrepasar el estudio de lo manifiesto para prolongarse hasta la identificación de las causas. Sólo de esta manera será posible tomar medidas que puedan contribuir al perfeccionamiento de la enseñanza.

Rosales (1990, pág. 13).

Como afirma Forns (1980, pág. 107): “El concepto de evaluación es muy amplio. Se puede evaluar desde el propio sistema educativo hasta cada una de las facetas intervinientes en el proceso de educación.” O como podemos apreciar, de forma más concreta, en la siguiente cita de Ana M.^a Geli:

“La evaluación del proceso docente es un trabajo extenso y complejo puesto que exige la valoración de todos los factores relacionados con el medio educativo: los objetivos planteados y su grado de flexibilidad, los métodos, técnicas y recursos utilizados, los componentes psicológicos y sociológicos que inciden en el aprendizaje de los alumnos, los medios materiales, funcionales y personales del centro, y de manera especial, las aptitudes, características y actitudes de cada profesor.”

Geli (1988, pág. 48).

Nuestra falta de tradición, junto con la ausencia de autonomía de los centros y de los profesores y, consecuentemente, con una administración educativa fuertemente centralizada, se ha dejado sentir en la pobreza conceptual y metodológica de la que ahora empezamos a despegar, especialmente si nos comparamos con otros países de nuestro entorno económico, político y cultural,

principalmente los Estados Unidos de Norteamérica y el Reino Unido, con una mayor historia y atención hacia este importante componente curricular. A diferencia de lo que ocurre en estas naciones, que poseen diferentes vocablos muy específicos según el ámbito en el que se desarrolle, el modo de llevarla a cabo o la finalidad con la que se emplea; nosotros sólo poseemos un único término para describir la evaluación educativa. Si bien, son muchos los conceptos que están relacionados directamente con esta expresión: apreciación, calificación, decisión, estimación, exámenes, indicadores, información, investigación educativa, juicio, medición, responsabilidad, tasación, valoración, etc.

Trataremos de delimitarlos conceptualmente y ofrecer distintas definiciones que en sí mismas encierran visiones muy alejadas de la evaluación.

Precisamente esta ambigüedad es la que nos exige señalar cuáles son las diferentes descripciones que se han dado y qué es lo que debemos de entender por evaluación educativa, idea de la que arrancaremos para comprender el presente trabajo.

1.1.- Apreciación

Uno de los términos que es utilizado frecuentemente como sinónimo de evaluación es el de **apreciación -appraisal-**. Para algunos autores (Popham, 1980, b) es sinónimo de evaluación. De hecho hay un tipo de evaluación **-assessment-** que se conoce así, como **tasación**, y que refiere directamente a lo que se entiende por «evaluación de los alumnos», tests, etc.. Pero para nosotros, como veremos más adelante, la evaluación es un concepto mucho más amplio. Otra idea muy cercana es la de **estimación -appraisement-**, que denota el aprecio y valor que se da y en que se tasa o considera una cosa. No obstante, para Popham habría de hacerse un serio intento por unificar el lenguaje de los evaluadores: “Ya que, si no se aclaran los términos, dos personas podrían utilizar el término *estimación* una queriendo decir medición sin valor y otra evaluación sistemática, lo que originaría confusión.” (Págs.: 19 y 20)

1.2.- Decisión

Consideramos que la evaluación es una parte de la toma de decisiones, que la envuelve en todos sus aspectos y contribuye a su correcto desarrollo -mejorando los procesos y los programas (Cronbach, 1963)⁹; pero son procesos plenamente diferenciados e independientes, como podemos apreciar seguidamente (véase la figura 1). Conviniendo la decisión como una “acción o proceso mediante el cual pretendemos reducir el grado de incertidumbre ante una duda o problema planteados, eligiendo una alternativa de entre las presentes. Toda decisión, en cuanto proceso que es, puede descomponerse en una serie de elementos; tales serían:

- Consciencia de una situación incierta o dudosa (conocimiento del problema).
- Información sobre las causas posibles que han propiciado la aparición del problema.

⁹Cronbach (1963, 673) diferencia tres tipos de decisiones en los que se utiliza la evaluación: a) para mejorar un curso; b) las relativas a las características del alumno; y c) las concernientes a la administración.

- Posibles alternativas para su resolución; salida de la incertidumbre creada, incluso con el asesoramiento.
 - Proceso de selección de aquellas alternativas estimadas como mejores o más idóneas. Estas no tienen por qué referirse sólo a problemas sino a situaciones de elección entre opciones diversas.
 - Ejecución o elección firme de una alternativa.
 - Evaluación de las consecuencias o retroalimentación.”
- (Monedero, 1988, págs. 32 y 33).

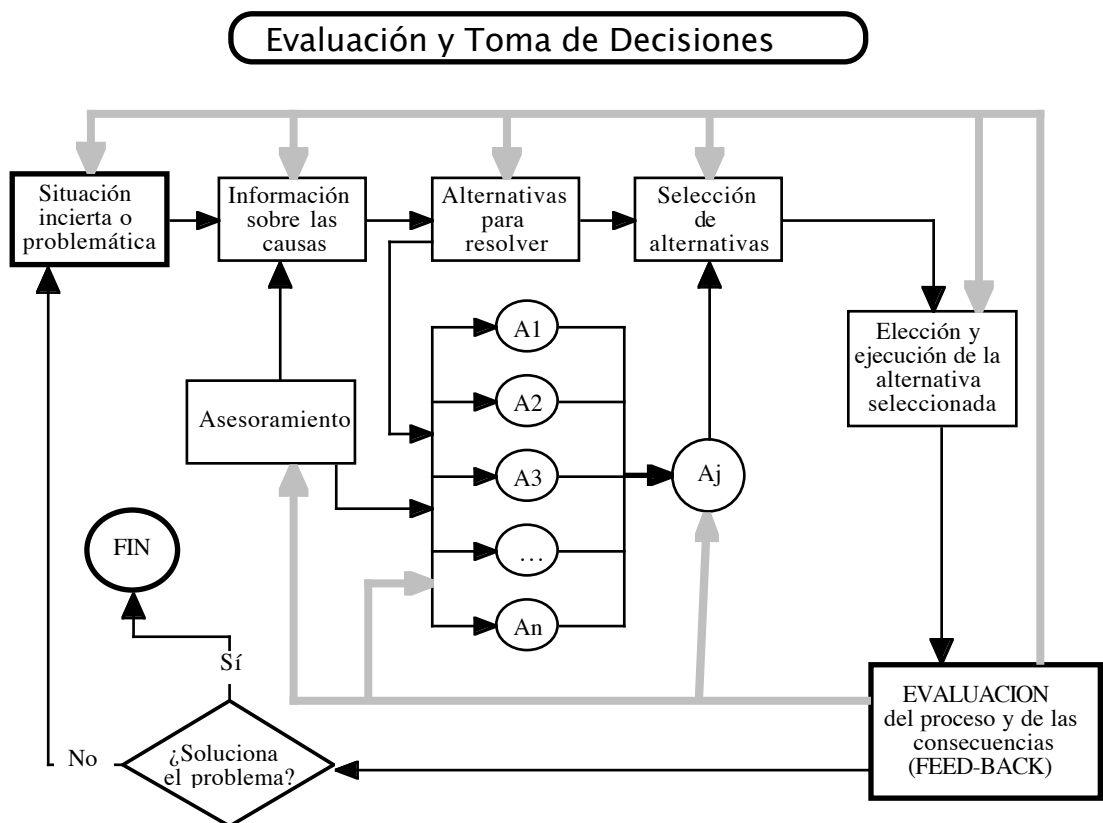


Figura 1.

1.3.- Indicadores

Ultimamente se ha introducido en el campo de la evaluación este nuevo concepto (Norris, 1992; Elliott, 1992; Alvaro Pagé e Izquierdo, 1992; Tiana, 1994; Serrano et alius, 1995), que como veremos en el capítulo siguiente está íntimamente relacionado con ella, pero diferenciándose en que los indicadores sólo son instrumentos al servicio de la propia evaluación educativa, medidores indirectos -de síntomas o factores- que en su conjunto nos dan idea del estado de un objeto.

1.4.- Información

La información es una de las partes esenciales que conforman la evaluación. Sin ella no es posible evaluar. Pero el solo hecho de disponer de los datos no implica ningún proceso de evaluación. No obstante, a partir de la década de los setenta,

entre muchos naturalistas (etnógrafos, cualitativos, hermenéuticos, iluminativos, etc.) existe la tendencia cada vez más acusada de concebir que el receptor de los informes es, en última instancia, el evaluador, la persona que «juzga»; la audiencia a la que estos se dirigen es la que debe de evaluar. Así las cosas, los informes deben ser eminentemente descriptivos y para garantizar su credibilidad deben ir acompañados de las pistas de revisión pertinentes (documentos, observaciones, transcripciones, datos, etc. usados para su elaboración) tales que le permitan al lector juzgar sobre los mismos. Para muchos autores (Alvira, 1982; Goetz y Lecompte, 1984; Cook y Reichardt, 1986; Monedero, 1988) la información debe ser lo más objetiva posible, para lo cual, incluso, se pueden conjugar datos obtenidos a través de métodos cuantitativos con los conseguidos mediante investigaciones cualitativas.

También la tercera acepción del diccionario de la lengua se refiere a la información como “pruebas que se hacen de la calidad y circunstancias necesarias en un sujeto para un empleo u honor.” Semánticamente muy cercano a nuestra concepción de los exámenes y a otra de las acepciones de evaluación identificadas por Hilda Taba (op. cit., págs. 407 y 408) como proceso de recolección de informaciones o evidencias relevantes sobre el logro de los objetivos que permitan la formulación de juicios.

En la actualidad existe una tendencia cada vez más definida a lo que podríamos llamar una caracterización global de la evaluación. Se trataría de encontrar modelos capaces de integrar las diferentes aportaciones, datos cuantitativos y cualitativos, procedentes de la utilización de una diversidad de instrumentos de recogida de información como los planteados por Rodríguez, Gutiérrez y Molledo (1992).

1.5.- Investigación educativa

Desde nuestra perspectiva -siguiendo a Popham (op. cit., págs. 20 y ss.)- hay tres rasgos fundamentales que caracterizan, fundamentan y diferencian tanto a la evaluación como a la investigación educativas. Véase la figura 2.

En primer lugar, en tanto que el objetivo de cualquier investigación radica en conseguir conclusiones veraces sobre el problema o tema de su estudio; la evaluación persigue facilitar la toma de decisiones sobre el objeto evaluado.

En segundo lugar hay una manifiesta diferencia sobre los objetivos de generalización que ambos procesos persiguen: así que mientras la evaluación educativa no pretende extender sus hallazgos a otros contextos, la investigación aspira a una generalización de los mismos. En el caso que nos ocupa los resultados de la evaluación les sirven al profesor expresamente para aplicarlos de un modo directo en su propia aula; contexto del cual los obtuvo, sin importar demasiado si son, o no, aplicables en otras situaciones parecidas. Esta es, claramente, la principal ventaja sobre la validez o funcionalidad directa del conocimiento extraído de la evaluación. Al profesor lo que más le importa son los procesos y acontecimientos que suceden en su propia aula y las informaciones y valoraciones que le permitan tomar decisiones sobre los mismos.

En último lugar, y no menos importante que los ya mencionados, hay también una discrepancia relevante acerca del énfasis valorativo que cada uno de estos procesos presupone. Así mientras que la investigación pretende concluir con afirmaciones sobre las que se prueban la verdad de los hallazgos alcanzados; la

evaluación concluye informando sobre el valor de los mismos. Al profesor lo que le interesa especialmente es el mérito o valor de cada uno de los componentes curriculares implicados en los procesos de enseñanza-aprendizaje de su aula, independientemente de que se pueda demostrar la veracidad del proceso de análisis de datos que corroboren sus hallazgos, a través tanto de la replicabilidad de los mismos como de su transferencia y aplicabilidad a otras situaciones similares.

Evaluador e investigador se distinguen en que para éste no es necesario valorar las cosas. En palabras de Popham (op. cit., pág. 23): “La evaluación es una actividad tremendamente práctica. Los investigadores se contentan, por lo general, con descubrir el mundo; pero los evaluadores quieren hacerlo mejor.” Esta idea de comprometerse con las situaciones para perfeccionarlas, resulta una característica natural de la evaluación resaltada también por otros autores, como Fernández Pérez (1986, pág. 51): “La evaluación es una actividad sistemática integrada en el proceso educativo, cuya finalidad es el mejoramiento, lo más posible, del alumno en todos los aspectos de su personalidad, y de una información ajustada sobre el proceso educativo y sobre los factores personales y ambientales que en éste inciden.” O Rul (1995, pág. 72) “Hay evaluación si a través de la experiencia y de la comunicación humana se produce aprendizaje valorativo, con repercusiones individuales y colectivas, basado en datos e información objetivada.”

Característica: a estudiar	Evaluación educativa	Investigación educativa
Objetivo	Decisiones	Conclusiones
Generalización	Baja	Alta
Enfasis valorativo	Valor	Verdad

Figura 2. Principales diferencias entre evaluación educativa e investigación educativa.
(Extraído de Popham, op. cit., pág. 22)

Igualmente habría que diferenciar evaluación de investigación evaluativa (Rossi y Freeman, 1982) que refiere más a determinadas metodologías sociales de investigación y a los diseños experimentales que comportan que a la evaluación en sí misma. No obstante, determinados autores como Stufflebeam et alius (1971) hacen sinónimos los conceptos de investigación y de evaluación.

1.6.- Juicio y reflexión

Para muchos educadores e investigadores (Taba, op. cit., pág. 407), la evaluación es sinónimo de juicio o la emisión de juicio tras comparar una medida con unas «normas» (Mager, 1984). Sin embargo aunque ambos procesos si están bien fundamentados son el resultado de deducciones lógicas, el primero no exige ninguna acción, en tanto que el segundo se hace además de para conocer el valor

del objeto evaluado, para tomar decisiones sobre dicho objeto, exigiendo la correspondiente intervención bien para modificarlo, o bien para suprimirlo. (Monedero, op. cit., págs. 47 a 50).

Para Juana María Sancho (1990, pág. 8) resulta esencial la calidad de las informaciones que sustentan las premisas de un juicio para que éste sea tal y no se sesgue originando conclusiones equívocas: “la emisión de un juicio, sea sobre el valor de la actuación académica de un alumno, el valor educativo de una institución, la calidad o pertinencia de un programa de formación del profesorado o sobre la práctica docente, que no se sustente en la consideración de un conjunto de datos que permitan una comprensión global del fenómeno bajo consideración, tiene muchas posibilidades de construir un «pre-juicio», es decir, hecho o situación que no es conocido o sin considerar alguno de sus aspectos.”

También se concibe la evaluación como un proceso de reflexión sobre la propia acción con la finalidad de mejorarla. Es el caso de Carlos Rosales (op. cit., pág. 14), quien entiende “la evaluación como una función característica del profesor que consiste básicamente en una actividad de reflexión sobre la enseñanza.” Y aclara más adelante (pág. 15): “Esta reflexión se entiende como evaluación si se sitúa en el origen de decisiones para el perfeccionamiento.”

1.7.- Medición (medida, *measurement*)

Mientras que medir es determinar el grado en el que un objeto contiene una característica o atributo específico, o la asignación de números “a propiedades o fenómenos, a través de la comparación de éstos con una unidad preestablecida” (Livas, 1980, pág. 11); la evaluación determina el grado en que ese objeto posee un valor. Según esta autora habría que resaltar que “la medición siempre es una propiedad y nunca la cosa o persona que la posee”, advirtiendo, acaso, a quienes hacen un isomorfismo entre el alumno y su calificación. Un alumno es considerado «bueno» o «malo» porque sus calificaciones así lo son, independientemente de su «calidad humana». Además esta misma situación se traslada y aplica simplistamente en otros ámbitos extraescolares: con la familia; en el entorno social; etc.; con las repercusiones que para la persona del estudiante pueda conllevar.

Siguiendo con nuestra diferenciación entre la medida y la evaluación, para autores como Coll (1983, pág. 14) “La referencia a unos criterios convierte la simple medición del aprendizaje en una evaluación.” Taba (op. cit., págs. 409 y 410) también reconoce que son muchos los autores que incurren en el error de confundir ambos conceptos y aclara que “la medición es sólo una parte de la evaluación que se refiere al proceso de obtener una representación cuantificada de cierta característica, tal como ciertos tipos de rendimiento o aptitudes escolásticas. El proceso de medición es fundamentalmente descriptivo, en tanto como indica mediante una cantidad el grado en el cual posee una determinada característica.”

Por su parte Stanley, Glock y Wardeberg (1972, pág. 9) opinan que la medición es una parte sustancial de la evaluación, a la que sirve de base. Para estos autores “la medición es fundamentalmente un proceso descriptivo. Aspira esencialmente a describir el grado en que un rasgo es poseído por un alumno mediante cifras y no con palabras.” Crean que los significados de los adjetivos cambian entre las personas, cosa que no le ocurre a las cifras que se les antojan superiores para este fin, si se han obtenido cuidadosamente y mediante los procedimientos adecuados.

Así, ofrecen una explicación de lo que entienden por medida en educación: “la medición educativa tiende a concentrarse en un rasgo o característica limitado, específico y bien definido, y luego procura determinar con precisión el grado en que ese rasgo es poseído por un alumno determinado.”

Sin embargo la propia evaluación se ha visto y, en ocasiones aún se ve, usualmente reducida a esta actividad (Mager, 1984), debido sin duda a la gran influencia cuantificadora de los procesos de medida del rendimiento industrial introducidos por el «taylorismo». Pensemos que durante mucho tiempo, el desarrollo de la evaluación, ha ido en paralelo al de las metodologías de investigación social que la han sustentado. No obstante hay quienes opinan (Thorndike, 1971) que desde comienzos de siglo la medida y la evaluación son procesos plenamente diferenciados e independientes.

Para algunos autores (Fernández Pérez, op. cit., pág. 16) la diferencia central radica que en tanto que el resultado de la medición es un dato, la evaluación implica además la comparación con un entramado “más o menos heterogéneo, de datos ideales interrelacionados (...) dentro de un conjunto de cálculos relacionados” mediante los que se constata su adecuación o inadecuación respecto de un fin. Coincide en su concepción con Adams (1970, pág. 20), para quien “la «evaluación» va más allá de la medición por cuanto supone la existencia de juicios de valor.” Según este autor, la obtención de medidas y el registro de datos se hacen siempre “sobre una muestra de resultados alcanzados bajo condiciones tipo.” Dichas condiciones tipo deberían ser tan sistemáticas y objetivas que pudieran permitir que distintos observadores obtuvieran unos datos “razonablemente comparables.” Pero aun las medidas así obtenidas, no dejarían de serlo, si después no realizamos una adecuada interpretación de estos resultados con respecto a ciertos “patrones” relacionados con la edad, planes de estudios o “cualquier otra base de juicios de valor”. De este modo coincide con Livas (op. cit., 5) que opina que “la verdadera labor del educador no termina con la aplicación de exámenes y la obtención de sus resultados. Para que estos datos puedan conducir y enriquecer el proceso educativo, es necesario interpretarlos.” O Lafourcade (1977, págs. 21 y 22) para quien “la evaluación es una interpretación de una medida (o medidas) en relación a una norma ya establecida.” Pero no sólo este autor considera como fuentes de información para evaluar los datos obtenidos a través de las mediciones cuantitativas, sino también las descripciones cualitativas, siempre que se sometan a interpretaciones que conduzcan a concluir con la emisión de juicios de valor. Además enumera una serie de condiciones exigibles a las conductas que se deseen medir (págs. 41 a 43):

- a) Ser comunes a un grupo o a una clase de sujetos.
- b) Captables por los sentidos.
- c) Definibles con claridad y precisión.
- d) Variables en relación a los fenómenos que las detentan.
- e) Promotoras de reacciones muy similares entre observadores.

Para Popham esta confusión tiene su origen en la propia historia de la evaluación: “Durante muchos años, y especialmente en las primeras décadas de este siglo, los educadores americanos consumieron mucho tiempo midiendo el rendimiento; por ello, muchos consideran que la evaluación es una actividad que atiende a los aspectos cuantitativos.” (Pág. 18).

No obstante algunos autores justifican su analogía, llegando a hacer sinónimos ambos conceptos: “Sin entrar en la polémica de la diferencia entre evaluación y medida vamos a considerar la evaluación como el acto por el cual una persona otorga un valor (una «nota», una «calificación») a una producción.”(Forns, op. cit., pág. 119).

Para Pidgeon y Yates (1979, pág. 18) la medición se hace necesaria para el profesorado no sólo para tomar decisiones sobre los alumnos o para conocer el progreso de éstos, “sino también con el fin de facilitar su salud mental y el desarrollo de sus aptitudes.” Según estos autores “la medida sirve para mejorar la eficacia de la orientación y dirección del proceso de trabajo del alumno, de manera continua, la clase de tratamiento que conviene a sus necesidades y capacidades.” Haciendo sinónimas las numerosas calificaciones con el concepto de evaluación formativa, que tendremos ocasión de ver y analizar más adelante. Sin embargo hemos de tener en consideración que, a pesar del tiempo transcurrido desde que se comenzaron a aplicar las mediciones en educación, las técnicas de recogida de datos, así como las de medida de los mismos, aún no han avanzado lo suficiente como para poder otorgarles una validez ni una confianza de la que hasta el presente, sencillamente, carecen. El profesor José Serrano et alius (op. cit., pág. 2) lo expresan del modo siguiente:

Todavía estamos lejos de tener instrumentación tecnológica para medir en educación. El metro o el litro aún no se ha inventado en educación. Se mide con distintos tests y cuestionarios y se tienen los mismos inconvenientes que cuando se medía a palmos. Eran medidas groseras y no universales. Así, en cada caso, se mide de una forma y con una unidad distinta.

1.8.- Valoración

Desde nuestra perspectiva es el concepto más cercano al de evaluación; pero entraña sólo la parte final de ésta, la referida a la explicitación del valor intrínseco o extrínseco de los objetos. Desde este ángulo la valoración sería más la estimación de la virtud que una ayuda para facilitar la toma de decisiones. Wheeler (1979, pág. 288) relaciona los conceptos de «medición», «valoración» -que curiosamente condiciona al grupo, no al individuo- y «evaluación» del siguiente modo:

La evaluación, por tanto, implica emitir un juicio con respecto a ciertos criterios. «Valoración» será el término que designe el proceso de investigar el nivel de un determinado grupo, normalmente en relación con las conductas esperadas. Se trata, por lo tanto, de un término general que incluye el más restringido de «medición», el cual implica un cierto tipo de escala en la que los individuos se ordenan de acuerdo con lo que saben o pueden hacer.

Para algunos autores, como Rul (op. cit., pág. 70), la disparidad entre la valoración y la evaluación es tan amplia que sitúan a ambos conceptos en dos polos opuestos: “En este sentido, se diferencia de *valorar*, que es el juicio de valor que una persona -subjetivamente- atribuye a una cosa. La valoración y la evaluación son los dos extremos de la acción axiológica de las personas y de los grupos. A menudo se confunden y se utilizan indistintamente; es usual denominar evaluación a la recogida y sistematización de opiniones y valoraciones personales. No obstante, la estandarización de valoraciones subjetivas no deja de ser una

valoración. Por ejemplo, los modelos (e)valuativos de cuestionario.” Este mismo autor profundiza más en sus análisis:

“Valorar es una capacidad humana: la facultad que tenemos las personas para juzgar o atribuir valor a las cosas o a los fenómenos en relación con intereses prácticos. Las decisiones que vamos tomando a lo largo del ciclo vital se fundamentan en valoraciones -expresas o latentes, espontáneas o inducidas, etc.- estrechamente ligadas a factores personales (capacidades, actitudes, experiencias vividas, etc.), socioeconómicos y culturales.

En este sentido podemos decir que valorar es la capacidad humana de emitir juicios de valor sobre las cosas y las situaciones en relación a finalidades prácticas. La valoración se constituye en evaluación individual cuando los valores emitidos se fundamentan en informaciones objetivadas que el individuo ha aprendido: por ejemplo, un médico evalúa la salud de un paciente a partir de los datos e informaciones diagnósticas decantadas por la experiencia profesional. Así pues, podemos afirmar que el fenómeno valorativo tiene dos dimensiones, una subjetiva y otra más objetiva.”

(Rul, op. cit., pág. 71).

Michael Scriven (1983, págs. 75 y ss.) en “The Evaluation Taboo” critica al positivismo lógico haciendo, precisamente, indisociables las valoraciones, de los propios procesos científicos. Afirma que la evaluación es la única actividad esencial de la ciencia. El hombre sabio está obligado a juzgar entre la perfección y la imperfección de las observaciones, entre la validez y la falta de validez de las generalizaciones, y así sucesivamente. Estas distinciones exigen valorar los juicios, no pudiéndolos reducir a meras y simples observaciones. Considera idénticos los conceptos de valoración y de evaluación, y a ésta última como el instrumento que imprime validez al nuevo conocimiento generado en la ciencia.

1.9.- Exámenes, controles, tests o pruebas de rendimiento escolar

Pruebas de lápiz y papel, o de ejecución o desempeño *-performance* o *achievement*, términos ingleses que las describen, traducidos generalmente como «realización»-, que se les proponen a los alumnos para que den respuesta a una serie de cuestiones, generalmente sobre contenidos o destrezas, y a través de las cuales se extrae información sobre su nivel de conocimientos o habilidades con vistas a la evaluación. “... los exámenes sirven, al menos teóricamente y de derecho, a esa clásica función social de selección pero, al mismo tiempo, inevitablemente, «marcan», como función que es realizada por un sistema llamado educativo, al sistema mismo, dotándole de características y rasgos de los que carecería si no realizase aquella función.” (Perrenoud, 1990).

Con anterioridad a la entrada en vigor de la ya mencionada Ley General de Educación de 1970, los exámenes coincidían en su significado con el de evaluación, error de confusión descrito por Taba (op. cit., pág. 410) que aún hoy perdura entre algunos docentes: “Entre todas las innovaciones técnico-pedagógicas aportadas por la reforma educativa española, hoy en marcha, ninguna ha despertado mayor interés -mezcla de zozobra, inquietud y esperanza- entre profesores, estudiantes y familias, que la sustitución del tradicional sistema de exámenes por el nuevo modelo de control denominado «evaluación continua».” (Orden, 1977, pág. 13)

Los exámenes sólo representan una muestra del dominio que pretenden medir. Pueden ser de diferente naturaleza según la forma en que estén elaborados:

pruebas objetivas, de elección múltiple, libres, etc. A pesar de tener grandes inconvenientes técnicos y psicológicos, muy difíciles de solventar, se emplean profusamente, llegando a ser los instrumentos reyes de la evaluación de alumnos. A partir de Piéron (1969), cabría hablar de una «teoría de los exámenes», expresada bajo el nombre de **Docimología**¹⁰. Este autor, junto a Hoyat (1962), Pidgeon y Yates (1979), López, Llopis et alius (1983), Spears (1984), Cardinet (1988), Fernández Pérez, (1986), Fuertes (1990) y Worthen (1993, a), entre otros, hacen una acertada crítica acerca de la falta de fiabilidad y validez de este tipo de instrumentos para la evaluación. Por su parte, Alonso, Gil y Martínez (1992) hacen una crítica de los exámenes de acuerdo al modelo de enseñanza en el que se utilizan, diferenciando entre aquella basada en la transmisión y aquella otra, en la que la investigación del alumno es el núcleo central de las actividades de los procesos de aprendizaje. Por su parte Ennis (1993), admite la dificultad de los exámenes para poder valorar el pensamiento crítico de los alumnos, pero opina que el problema es debido a un diseño deficiente de los mismos y no, por tanto, achacables al instrumento en sí. Propone una serie de criterios para poderlos utilizar de una forma adecuada. Con ese mismo propósito de conseguir que los exámenes sean técnicamente adecuados, eficientes, manejables, totalmente aceptados, comparables y con un bajo costo económico, Madaus y Kellaghan (1993) realizan una revisión de la experiencia británica de los exámenes de selectividad (SAT- Standard Assessment Tasks), y concluyen que examinar consiste, sobre todo, en realizar: hechos, trabajos y productos, que deben ser evaluados teniendo en cuenta su uso y su contexto.

1.10.- Calificación, puntuación o nota

El paso de esto a su investigación sobre la Calidad tuvo lugar a causa de un siniestro aspecto de la calificación que quedó expuesto con la retirada de notas. En realidad, las notas cubren un fracaso en la enseñanza. Un mal profesor puede pasar todo un trimestre sin dejar absolutamente nada memorable en las mentes de sus alumnos, componer las notas mediante una prueba irrelevante, y dejar la impresión de que unos han aprendido y otros no. Pero si las notas son eliminadas, la clase se ve obligada a preguntarse cada día qué es *en realidad* aprender. Las preguntas acerca de qué es ser enseñado, cuál es el objetivo y cómo lo cumplen las clases y las tareas, se vuelven ominosas. La eliminación de las notas expone un vacío inmenso y espantoso.

Pirsig, R. M. (1990, pág. 13).

Para muchos, en especial para los alumnos, son la expresión de los resultados de la evaluación, llegando a hacer sinónimos todos estos conceptos. Popham (op. cit., págs. 18 y 19), tras admitir que durante mucho tiempo esto ha sido así, los diferencia y critica, porque aunque antes de calificar el profesor ha de valorar, esa

¹⁰Término sinónimo de la **docimasiología**, ciencia que estudia las técnicas de investigación o examen, principalmente en medicina. Fernández Pérez (1986, 23) explica su significado del modo siguiente: "Fue propuesto por PIÉRON y deriva del verbo griego: *dokimazo* = examinar. Con esta significación de 'ciencia de los exámenes', 'arte de examinar' o, más ampliamente, 'teoría de la evaluación escolar', el término ha entrado en el vocabulario de las ciencias de la educación."

valoración no se hace con respecto a los procesos educativos en sí mismos, sino sobre el propio alumno como persona, con lo que se pueden juzgar cosas tan indeseadas como su herencia genética [o circunstancias personales del desarrollo psicosocial de cada individuo, añadiríamos nosotros], tras el barniz de los contenidos y de las actitudes.

Para otros son una medida y, consecuentemente, sólo se relacionan con la evaluación en tanto en cuanto proporcionan información. Hay quienes opinan que sirven como sustitutas de los informes y su finalidad es eminentemente social: dirigidas a los padres, a las certificaciones, promociones, etc.

Según Carreño (1985, pág. 27) lo más común entre los profesores es utilizar la evaluación con el propósito exclusivo de constatar los aprendizajes de sus alumnos para emitir calificaciones que son “el *monto* o *volumen* de lo aprendido.” Confundiendo la evaluación con la realización de pruebas, la aplicación de exámenes, la revisión de los resultados y la adjudicación de calificaciones que no son otra cosa que una medición del rendimiento escolar. “La calificación así derivada, aun cuando sea determinada con absoluta justicia, sólo indica *cuánto sabe* el alumno, pero lo deja y nos deja totalmente ignorantes de *qué sabe, qué no sabe, cómo lo sabe* y, lo que es más importante, *gracias a qué sabe lo que sabe*. De ahí que la calificación sirva tan poco educativamente hablando y que sea tan estéril para orientar en el mejoramiento de la enseñanza.”

En el polo opuesto se sitúa Livas (op. cit., pág. 49), que opina que el fin de la evaluación está en poder calificar al alumno, ya que la calificación es un informe “oficial que resume el rendimiento escolar”. Argüye esta autora estadounidense que en la década de los sesenta, “quienes integraban la llamada *corriente progresista de la educación*” se enfrentaron a las calificaciones porque fomentaban el utilitarismo haciendo que los alumnos se vieran motivados por conseguir buenas notas y no por los beneficios inherentes a la educación. Tras desechar esas explicaciones, tachándolas de utópicas, justifica su necesidad, declarándolas imprescindibles, dado que son los parámetros a partir de los que tanto padres, como profesores, o las fuentes del trabajo y los propios alumnos toman las oportunas decisiones. Este mismo parecer es señalado por César Coll (op. cit., pág. 14) cuando expresa:

Poco importa que deje de hablarse de notas o exámenes y que se adopte en su lugar el término más neutro de «evaluaciones»; o que se sustituya la heteroevaluación por la autoevaluación; o, aún, que se abandone el sistema de calificaciones numéricas (de 0 a 10, por ejemplo) en favor de otras más cualitativas (insuficientes, suficientes, bien, ...); en todos los casos, la idea que predomina es que el proceso evaluativo tiene como finalidad última y única precisar el alcance del aprendizaje de los alumnos. Se intenta llevar a cabo esta tarea con el máximo rigor y de precisión posibles, se huye de los errores de apreciación, se cuidan las posibles repercusiones afectivas y motivacionales, pero indefectiblemente la evaluación es considerada, de forma unidimensional, como evaluación del aprendizaje del alumno.

Para Fernández Pérez, (op. cit., págs. 16 y 17) ningún tipo de puntuación o de calificación adquiere el rango “de evaluación hasta qué (...) la mente que toma nota de la medición no introduce a ésta en una relación determinada con algún patrón ideal, estadísticamente elaborado con mayor o menor precisión,

pronunciando un juicio comparativo que, a su vez, suele implicar apreciaciones de valor.”

Pidgeon y Yates (op. cit., pág. 35) consideran que una vez que se han obtenido a través de exámenes u otras pruebas las calificaciones y puntuaciones, éstas aún no son útiles para el profesor hasta que no se analice debidamente su distribución.

Farré y Gol (1982, pág. 12) expresan cómo debe de transmitirse la información a los padres, teniendo en cuenta la totalidad de los trabajos de los alumnos para que éstos puedan verlos, analizarlos y compararlos con los de los demás. No les parece adecuado “dar” notas y con estas diferencias distinguen entre evaluación y calificación. Carreño (op. cit., pág. 28) incide en este mismo aspecto y propone que si la revisión de los resultados de los exámenes se hace conjuntamente entre el profesor y los alumnos, entonces se podrán “enjuiciar y valorar distintos aspectos y momentos del proceso de enseñanza aprendizaje, incluida nuestra propia actuación como educadores.”

Taba (op. cit., pág. 408) acepta el uso sinónimo de la calificación y de la evaluación, pero advierte que “definir la evaluación como calificación, es decir, reducir todo lo que se conoce acerca del progreso de los estudiantes a una simple nota, es el concepto más limitado de evaluación.”

Por su parte, Shavelson et alius (1993) han puesto de relieve la falta de validez concurrente en las calificaciones que emiten los profesores, incluso utilizando diferentes instrumentos de evaluación.

1.11.- Responsabilidad

No hubiésemos reparado en ella a no ser porque hay un término inglés **-accountability**, frecuentemente traducido como «**rendición de cuentas**»- directamente relacionado con la evaluación. Como respuesta a las demandas políticas, sociales y económicas de los contribuyentes, en USA, aparece un tipo de evaluación de los centros o de determinados programas educativos, para «rendir cuentas», sobre la conveniencia de seguir sufragando los gastos originados en su mantenimiento o desarrollo. Sin embargo, no todas las evaluaciones de esta clase son necesariamente «accountability», ya que cualquier profesor (o equipo de profesores) puede hacer este tipo de evaluación sin que le venga impuesta externamente; sino, para conocer el mérito del programa que está desarrollando o sobre el buen funcionamiento de su propio centro educativo.

Respecto a este tipo de evaluación, Sarup (1990, pág. 198) realiza una crítica bastante preocupante desde una posición de justicia social:

“La educación se contempla según el modelo del libre mercado. Un buen sistema educativo es aquel que satisface las demandas de los consumidores. Sin embargo, sabemos que los consumidores que más hacen oír su voz son los integrantes de la clase media. Nos topamos, pues, con una contradicción: por una parte, se sigue proclamando la intención de favorecer la igualdad de oportunidades; por otra, las políticas gubernamentales perjudican gravemente a los más desfavorecidos. Los recortes presupuestarios y la reducción de los servicios y las prestaciones públicas afectan profundamente a los jóvenes y a los desempleados, a las mujeres, a las minorías raciales y a la clase obrera en general.”

1.12.- Evaluación

Así llegamos a lo que se entiende por **evaluación**. Para unos algo muy simple: “evaluar significa establecer una comparación entre lo deseado y lo realizado.” (Alfaro, 1990, pág. 70), o este otro concepto de la misma autora (pág. 72): “Evaluar es comparar lo que nos proponemos en los objetivos con lo que hemos sido capaces de realizar“. Desde el ámbito de la Sociología de la Educación, Perrenoud (1981, pág. 20) define la evaluación como “el criterio social de desviación o conformidad a una norma de buen comportamiento escolar.” Según Stufflebeam (1987, pág. 19) la evaluación coincidiría con “el enjuiciamiento sistemático de la valía o el mérito de algo“ o la “investigación sistemática del valor o mérito de algún objeto.“ (Stufflebeam, 1988, pág. 197). Para otros, algo más complejo: “la palabra evaluación nos induce inmediatamente a pensar en la posibilidad de efectuar cierto número de estimaciones cuantitativas y cualitativas acerca del valor, la importancia o la incidencia de determinados objetos, personas o hechos.” (Forns, op. cit., pág. 108).

Adams (op. cit., pág. 505) considera que evaluar es una actividad cotidiana y muy frecuente entre los seres humanos. Si bien opina que dicha actividad es muy egocéntrica, rápida y sin una reflexión previa y, consecuentemente, se agruparía en el terreno de las opiniones. La auténtica evaluación la formarían “los juicios estimativos que se formulan, o pueden formularse conscientemente siguiendo criterios bien definidos. Por lo general, estos juicios requieren una comprensión adecuada del problema y un cierto análisis que sirva de base al juicio.”

De la Orden (op. cit., pág. 16) concibe la evaluación como “un medio que permite observar y describir con mayor precisión los aspectos cuantitativos y cualitativos de la estructura, el proceso y el producto de la educación. Su finalidad es facilitar una predicción y un control lo más exacto posible del proceso educativo.” Como cabría esperar del contexto histórico en el que la elaboró, expresa una visión puramente conductista de la evaluación. Así admite que el valor de ésta reside precisamente en la comparación entre los objetivos de la educación y los cambios de comportamiento de los alumnos causados por la actividad docente.

Stanley, Glock y Wardeberg (op. cit., pág. 10) definen la evaluación como “un proceso en el que un maestro suele emplear información derivada de muchos orígenes para formular un juicio de valor. La información puede obtenerse mediante el uso de otras técnicas que no dan necesariamente resultados cuantitativos; p. ej., los cuestionarios y las entrevistas.“

Noizet y Caverni (1978), De Ketele (1980) y Coll (1983) opinan que la evaluación se identificaría con la emisión de juicios en función de unos criterios previos para facilitar la toma de decisiones sobre objetos, situaciones o fenómenos. Una concepción representativa de este parecer nos la ofrece Livas (op. cit., pág. 14) que la define como “un proceso que consiste en obtener información sistemática y objetiva acerca de un fenómeno y en interpretar dicha información a fin de seleccionar entre distintas alternativas de elección.”

También Taba (op. cit., pág. 409) emite una definición parecida cuando la describe “como un proceso intrincado y complejo que comienza con la formulación de objetivos, que involucra decisiones sobre los medios para asegurar la evidencia de su cumplimiento, los procesos de interpretación para llegar al significado de esta evidencia y los juicios sobre las seguridades y las deficiencias de los estudiantes y que finaliza con las decisiones acerca de los cambios y las

mejoras que necesitan el currículo y la enseñanza.”No obstante esta misma autora ofrece otra definición de evaluación por la que la identifica con el proceso de determinar en qué consisten los cambios constituidos por los objetivos de la educación y de la estimación de los mismos en función de los valores inherentes a dichos objetivos y de ese modo saber hasta qué punto se están consiguiendo las metas educativas: “la evaluación es parte indivisible de la elaboración del currículo, que comienza con el interés por los objetivos y termina cuando se establece en qué medida éstos han sido alcanzados.” (Pág. 411).

La ya mencionada Juana María Sancho (op. cit., pág. 8) “entiende por evaluación la emisión de un juicio sobre el valor de algo proyectado o realizado por algún individuo o grupo y presupone un proceso de recogida de información sistemática y relevante que garantice la «calidad» del juicio formulado.”

Según Rul (op. cit., pág. 70) “*evaluar* es crear valor en relación a determinados objetos, hechos o fenómenos a partir de datos e informaciones objetivadas en relación a finalidades prácticas. Evaluar es una capacidad humana, en el ámbito del conocimiento axiológico, siempre que el valor atribuido esté fundamentado en datos rigurosos (*medición*) y/o en informaciones objetivadas y plurales.” Para este autor la evaluación origina aprendizajes sociales e individuales a través de la comunicación e incide directamente en la asunción cultural de los valores.

Popham (op. cit., págs. 16 y 17) piensa que “evaluar algo es determinar su valor.” Pero que lo que nos puede interesar como educadores es la evaluación educativa formal o sistemática que “consiste en un juicio formal del valor de los fenómenos educativos.”

Para Carreño (op. cit., págs. 19 y 20) evaluar se corresponde con “la acción de juzgar, de inferir juicios a partir de cierta información desprendida directa o indirectamente de la realidad evaluada, o bien, atribuir o negar calidades o cualidades al objeto evaluado o, finalmente, establecer reales valoraciones en relación con lo enjuiciado.” Y más directamente, en el campo de la enseñanza, evaluar “significa ponderar (colectiva o individualmente, total y parcialmente) los resultados obtenidos de la actividad que conjunta a profesores y alumnos en cuanto al logro de los objetivos de la educación.” O también: “el conjunto de operaciones que tienen por objeto determinar y valorar los logros alcanzados por los alumnos en el proceso de enseñanza-aprendizaje, con respecto a los objetivos planteados en los programas de estudio.”

Dentro de ésta concepción de evaluar la eficacia y el logro de los objetivos previstos se encuentran también autores como Lafourcade (op. cit., pág. 21) que entiende la evaluación “como una etapa del proceso educacional que tiene por fin comprobar de modo sistemático en qué medida se han logrado los resultados previstos en los objetivos que se hubieran especificado con antelación.”

Desde otras ópticas ahora muy en boga en nuestro país (Estrella, 1983; Martín Gordillo, 1994; Rul, 1995), de la mano de las reformas educativas, el proceso de evaluación se entiende de un modo global, referida no sólo a los alumnos o al profesor, sino también al aula, al centro educativo e incluso a la propia Administración. Se complejiza afectando a todos y cada uno de los componentes curriculares y a sus relaciones, a los procesos y no sólo a los resultados: “La evaluación, en este caso, supone la detección de los conocimientos previos del alumno, así como el seguimiento atento del modo en que éste adquiere nociones, emplea estrategias y se enfrenta con dificultades. Pero supone, además, la

evaluación de todas las piezas que componen el engranaje curricular: desde los objetivos y la organización y secuenciación de contenidos, hasta la organización del centro, el uso y calidad de los materiales curriculares y la formación permanente del profesorado.“ (VV. AA., 1990, c, pág. 7).

También Serrano et alius (op. cit., pág. 15) nos ofrecen una visión holística de la evaluación definiéndola como el “único medio de describir y hacer comprensivos fenómenos contextuales, generados, paralelos o marginales en los diversos niveles del sistema educativo implantado.” En este mismo sentido expresa su concepción de evaluación Nieto (1994, pág. 13), definiéndola como “una actividad valorativa e investigadora, que facilita el cambio educativo y el desarrollo profesional de los docentes. La evaluación educativa, en la LOGSE es un elemento básico del currículo. Su finalidad es adecuar o reajustar permanentemente el sistema escolar a las demandas sociales y educativas. Su ámbito de aplicación abarca no sólo a los alumnos -en especial sus adquisiciones instructivas-, sino también a los profesores y los centros educativos, los servicios de apoyo y la misma administración educativa.”

Esta concepción de la evaluación exige un compromiso por parte de los implicados en el desarrollo de los programas educativos, estimulando un debate crítico sobre los mismos y sobre su desarrollo. (Kemmis, 1986, b). Como consecuencia del mismo, la evaluación se transforma en un proceso de investigación en el aula (Estrella, 1983, pág. 48) y de perfeccionamiento curricular y profesional (Simmons y Resnick, 1993; Wooster, 1993).

La evaluación, podemos afirmar empleando un símil, llega a los rincones más oscuros de la casa de la enseñanza. Casi todas las definiciones examinadas hacen especial hincapié en los valores y los juicios unidos a la toma de decisiones. Depende de quién evalúe y para quién se evalúe, para que la evaluación pueda ser vivida como una ayuda o como una auténtica amenaza. Los métodos de evaluación y su concepción misma no se muestran como algo «neutro» o «aséptico», antes al contrario: todo rezuma ideología, creencias, sentimientos, tabús... y al tener repercusiones sociales, la evaluación connota opciones políticas y económicas. Consecuentemente son cada vez más los evaluadores, que comprometidos con su labor, opinan que es hora ya de ir desarrollando una filosofía que conlleve a la conformación de una deontología de la evaluación. Es el caso de House (1986, a, págs. 1 a 26). Este autor se pregunta si realmente es necesaria tal filosofía. En los albores de los procesos evaluativos se hicieron avances sobre la propia evaluación, pero sin necesidad de ella. Posteriormente, en especial durante el desarrollo de algún estudio de campo, fueron apareciendo las principales dudas, escollos y desacuerdos sobre los principios básicos de la evaluación. Nos propone un análisis y reflexión profundas sobre las prácticas evaluativas aplicando los conceptos y técnicas usados en la Filosofía -como disciplina- que nos permitan crear una estructura provechosa para su estudio. Así, elige la línea marcada por recientes trabajos filosóficos arguyendo algunas de sus ideas claves sobre la evaluación fundamentadas en el uso de las metáforas. Sus contenidos en esas metáforas, señalan lo esencial de la estructura de la evaluación.

Cap. 2.- Las formas de la evaluación educativa

¿Cuáles son los modelos de evaluación que se les ofrecen a los profesores? Como iremos viendo a lo largo de esta sección, consideramos que el desarrollo que ha tenido este ámbito del curriculum, especialmente a partir de la década de los setenta, es tan grande e intenso, que fácilmente los docentes pueden sentirse desorientados. A veces, la propia literatura pedagógica de mayor difusión, pero poco específica científicamente, en su afán por ofrecer alternativas y experiencias relativamente viables y asequibles, acaba provocando el efecto contrario al pretendido. Esto es, el maestro se ve inmerso en un auténtico *maremagno* de conceptos y perspectivas que a menudo lo confunden y abruman mucho más de lo que le ayudan.

Nuestro modo de tratar de clarificar esta situación será analizar, en primer lugar, lo que consideramos *formas* de la evaluación. Procuraremos unificar el vocabulario y delimitar su concepción, tan claramente como nos sea posible. En una segunda parte revisaremos, igualmente, los principales *instrumentos* de evaluación. Con la misma intención, en una tercera parte, haremos una exposición de algunos de los *modelos* evaluativos que desde nuestro criterio nos han parecido más relevantes e influyentes para la situación presente y su proyección futura.

2.1.- Evaluación inicial

Para Forns (1980, pág. 111) es un caso particular de la *evaluación diagnóstica*, y se refiere a ella como la que hacen los profesores para conocer cuál es el estado de conocimientos que tienen sus nuevos alumnos. De su resultado el maestro debe de extraer conclusiones, tales, que le permitan hacer cuantas modificaciones considere precisas dentro de la programación de la enseñanza: alterar sus objetivos, tiempos, contenidos, actividades, etc. También adaptarse a las diferencias individuales, eligiendo aquellas estrategias metodológicas y medios y recursos didácticos que mejor sintonicen con los intereses, motivaciones y necesidades reales mostrados, tanto por los alumnos individuales, como

por el conjunto de la clase. Es lo que de acuerdo con la LOGSE se denominan «adaptaciones curriculares».

Por su parte, Farré y Gol (1982, pág. 11) emiten una idea diferente de la anterior. Para ellos la evaluación inicial se hace cada vez que se introduce un nuevo tema o concepto: “Se trata de informarnos sobre el nivel de los alumnos y de saber qué saben sobre ese tema concreto.” Estos autores proponen dos fases diferentes en su aplicación. Una oral, en la que se les plantea problemas relativos a los aprendizajes siguientes para que según las respuestas emitidas, así como las soluciones dadas, el profesor pueda deducir qué conocimiento poseen sobre el tema. La segunda de las fases es de manipulación, con materiales muy variados. La finalidad es la misma que en la fase anterior.

Hilda Taba (1983, pág. 413) justifica la existencia de este tipo de evaluación en la necesidad de determinar qué conocimientos previos sobre materias específicas posee un alumno, su nivel conceptual o qué capacidades podrá manejar. Esto resulta fundamental para eliminar y prevenir muchos de los problemas que se suscitan en la ausencia de conexión y coordinación entre un nivel educativo y el siguiente. En función de todo ello se tomarán decisiones sobre los contenidos, su amplitud y grado de profundización, etc.; adaptándolos a aquello que los alumnos conocían en el curso anterior.

Podríamos concluir admitiendo que el objetivo de la evaluación inicial es describirnos el nivel de conocimientos, destrezas o habilidades que poseen los alumnos -individual o colectivamente- sobre determinados aprendizajes antes de iniciar un programa -por ejemplo, un curso, un tema o un concepto- con el objeto de realizar las adaptaciones curriculares necesarias.

2.2.- Evaluación diagnóstica

Ya hemos visto antes que Forns consideraba a la evaluación inicial como una faceta de la evaluación diagnóstica. Para este autor, (pág. 110) la evaluación diagnóstica “tratará de conocer si el sujeto en cuestión posee un potencial suficiente (intelectual, de aptitudes, de conocimientos, etc.) que le permita llevar a cabo cierto tipo de actividades (o estudios) con un nivel de logro aceptable.” Su finalidad es básicamente el pronóstico o la

predicción del rendimiento académico futuro del alumno. Recomienda que sea realizada por personal especializado, psicólogos u orientadores.

Según Taba (op. cit., pág. 433) “el diagnóstico es el primer paso de la evaluación, ya que establece puntos de referencia acerca de la situación inicial de los estudiantes.” Identifica esta autora la evaluación diagnóstica con la inicial, como ya hemos podido apreciar.

Para Carreño (1985, pág. 39) cuando manejamos la evaluación diagnóstica nos referimos a una evaluación sobre algo que aún no ha sucedido, pero que creemos que sucederá. La define como “aquella forma mediante la cual juzgamos de antemano lo que ocurrirá durante el hecho educativo o después de él.” El momento de su aplicación -tal vez sea ese el origen de la posible confusión de esta modalidad con la evaluación inicial- es siempre al comenzar un plan de estudios, un curso, un programa o, incluso, partes del mismo.

Su finalidad está en realizar una buena planificación, evitando pasos erróneos o innecesarios, identificando “la realidad particular de los alumnos que participarán en el hecho educativo, comparándola con la realidad pretendida en los objetivos y los requisitos o condiciones que su logro demanda.”

Esta «forma» de evaluar también está íntimamente relacionada con la *orientación* y con la *selección profesional de personal*. Así, se tratará de adivinar cuál será la persona más capacitada, a partir de un determinado perfil profesional (exigencias de tareas concretas), para el desempeño de un puesto de trabajo de acuerdo con las características personales de los individuos que aspiren a él, sus intereses, aptitudes, actitudes, personalidad de cada uno de ellos, nivel de formación, conocimientos, etc. Toda una «batería» de pruebas a las que se les pueden unir otras de diversa naturaleza como las entrevistas, dar solución a situaciones problemáticas o cotidianas ficticias, *rol-play*, simulaciones o prácticas reales durante un tiempo más o menos prolongado. Algunos críticos a este tipo de evaluación diagnóstica aplicada a la selección profesional de los individuos, aseguran que los seleccionadores se comportan como auténticos profetas que jamás se equivocan. Así cuando han decidido que un «sujeto» es el idóneo y el empresario lo contrata, se les quita a los demás la oportunidad real de probar su competencia respecto de él. Por su parte el empresario siempre sale satisfecho: si efectivamente el nuevo

empleado es bueno en su oficio, opina que ha sido fruto de la óptima selección llevada a cabo por los seleccionadores; si por el contrario, una vez contratado empieza a apreciársele defectos de desempeño, más o menos graves, el empresario opina que cómo serían los otros candidatos cuando fue éste el seleccionado. Pero los otros aspirantes no estarán ya presentes para mostrar el error. Así que ante una nueva demanda de empleo, el empresario no se fiará de «su olfato» con los candidatos y acudirá nuevamente a los seleccionadores...

Volviendo al tema de la evaluación diagnóstica, como lo veníamos tratando, recientemente se están publicando algunas experiencias en las que se incorporan diversos sistemas informáticos para la realización de este tipo de evaluaciones. En esta línea se encuentran los trabajos publicados por Dassa et alius (1993), en los que exponen este nuevo marco conceptual de la evaluación diagnóstica que, a su vez, sirve de base a la evaluación formativa, ya que sus resultados son útiles tanto para valorar la naturaleza de los errores de los estudiantes, de acuerdo a sus peculiares características; como a la continua mejora de los métodos de enseñanza con ellos empleados. Aumentando, considerablemente, la coherencia y efectividad de la evaluación diagnóstica con el desarrollo curricular.

2.3.- Evaluación continua y evaluación formativa

Por evaluación continua deberíamos de entender una modalidad de evaluación formal, sistemática e integrada plenamente en, y durante, los procesos de enseñanza-aprendizaje. Para Wheeler (1979, pág. 290) debe ser un continuo proceso de retroalimentación que nos permita orientar la enseñanza: "Si al adelantar un paso hacia una meta se hace una valoración, ésta puede servir para modificar o volver a valorar la meta, así como para tomar decisiones sobre el siguiente paso a dar. Este ciclo de pasos hacia delante y hacia atrás puede repetirse hasta alcanzar la meta total, tanto si se trata de la que se estableció originalmente como si ha sido modificada a la luz del *feed-back*." En este sentido, la evaluación continua se identifica con la evaluación formativa, como también argüye Simmons (1994). Para este autor, la evaluación del trabajo de un estudiante es un componente que guía el proceso de aprendizaje. La evaluación continua utiliza las exposiciones, las explicaciones de conceptos por parte de los

estudiantes, composiciones, y también cualquier tipo de trabajo que permita evaluar y reflejar el trabajo realizado por los estudiantes.

Sin embargo, como afirma Miguel Fernández Pérez (1988, pág. 81) desde 1970, año en el que legalmente se incorpora el término a nuestro sistema educativo, tal vez nunca se ha llevado a efecto. La causa, según este autor, habría que buscarla en el peligro real que supone la introducción de reformas sin una preparación previa y una formación adecuada de los docentes. Esto podría entenderse como un «aviso a navegantes» para los responsables de nuestro actual proceso de implantación total de la reforma-LOGSE.

“Desgraciadamente, como ocurre con toda reforma cualitativa que no cuida la preparación del profesorado que ha de aplicar la reforma en sus aulas, al cabo de más de diez años de práctica, teóricamente, de la ‘nueva’ evaluación, los docentes reconocen sinceramente que la diferencia real entre lo que hacían antes y lo que tienen que hacer ahora es meramente cuantitativa, no de calidad: ‘El disparate pedagógico que, por lo visto, hacíamos antes de la reforma, una vez al año, ahora lo hacemos cuatro o cinco veces al año: una cantata de notas que el jefe de estudios va recogiendo’. Se han cambiado las palabras que se utilizan para evaluar el rendimiento de los alumnos, se ha cambiado la frecuencia de las evaluaciones; pero lo que realmente se hace para emitir el juicio evaluativo sigue idéntico. Con un agravante: al haber modernizado el anacronismo pedagógico con la capa de palabras, que no de realidades, más actuales (‘evaluación’, no exámenes; ‘insuficiente’, no suspenso, etc.), lo que se ha hecho es reforzar y prolongar la vigencia nueva de los viejos procedimientos (el típico cambio para no cambiar de todo intento de innovación insuficientemente elaborado).”

Fernández Pérez (op. cit., pág. 81)

Martín Gordillo (1994, pág. 26) se refiere precisamente a la evaluación continua como el aspecto educativo de la renovación de los setenta en el que más se aprecian las contradicciones. En lugar de entenderla como un proceso de retroalimentación siempre presente en los procesos de enseñanza-aprendizaje, se aplica como “la multiplicación de exámenes (del adusto examen de grado se pasa a los más «modernos» controles semanales). También la idea de que siempre (continuamente es posible examinar de toda la materia: lo estudiado memorísticamente en Noviembre podrá ser preguntado de nuevo en Mayo con la coartada de que «lo dice la ley». Cuando estas prácticas se llevan a cabo la memoria episódica de los sufridos alumnos es duramente torturada.” Este estado de cosas, lo único que ha conseguido, ha sido defraudar a nuestra sociedad, que se había dejado seducir por los tan cantados méritos de la

tecnocrática reforma educativa de los setenta. Así, recordamos las palabras de Arturo De la Orden (1977, pág. 13) al iniciar el Prólogo a la edición española de la *Evaluación de los aprendizajes* de Lefebvre, que reflejan ese estado anímico social: “Entre todas las innovaciones técnico-pedagógicas aportadas por la reforma educativa española, hoy en marcha, ninguna ha despertado mayor interés -mezcla de zozobra, inquietud y esperanza- entre profesores, estudiantes y familias, que la sustitución del tradicional sistema de exámenes por el nuevo modelo de control denominado «evaluación continua».” Obviamente no hemos avanzado mucho.

Farré y Gol (op. cit., págs. 11 y 12) aluden a la evaluación continua como «evaluación constante». Para estos autores la observación atenta del trabajo del alumno “es la que nos permite saber cómo trabaja, hasta dónde llega y qué dificultades tiene.” Con esta atención específica, los profesores pueden preparar actividades adecuadas y graduar los distintos objetivos de aprendizaje de una manera eficaz.

Se debe a Michel Scriven, en su famoso artículo “The Methodology of Evaluation” publicado en 1967 la distinción entre lo que llamó evaluación formativa en oposición a la que denominó como evaluación sumativa. Véase en la figura 3.

Ambos conceptos han sido aceptados casi unánimemente por todos los evaluadores educativos. Con estas modalidades distinguía “los papeles que desempeñaban los evaluadores que *formativamente* tratan de mejorar una secuencia pedagógica todavía en desarrollo y la de los evaluadores que *sumativamente* valoran los méritos de las ya terminadas.” (Popham, 1980, b, pág. 36). Para este último autor la evaluación formativa ha sido utilizada por los neo-conductistas (Bloom, Gagné, Glaser, etc.) de manera distinta a la concebida originalmente por Scriven. Así, Bloom, Hastings y Madaus (1971) la definen como el “uso de la evaluación sistemática en el proceso de planificación del curriculum, enseñanza y aprendizaje con el propósito de mejorar cualquiera de estos tres procesos.”

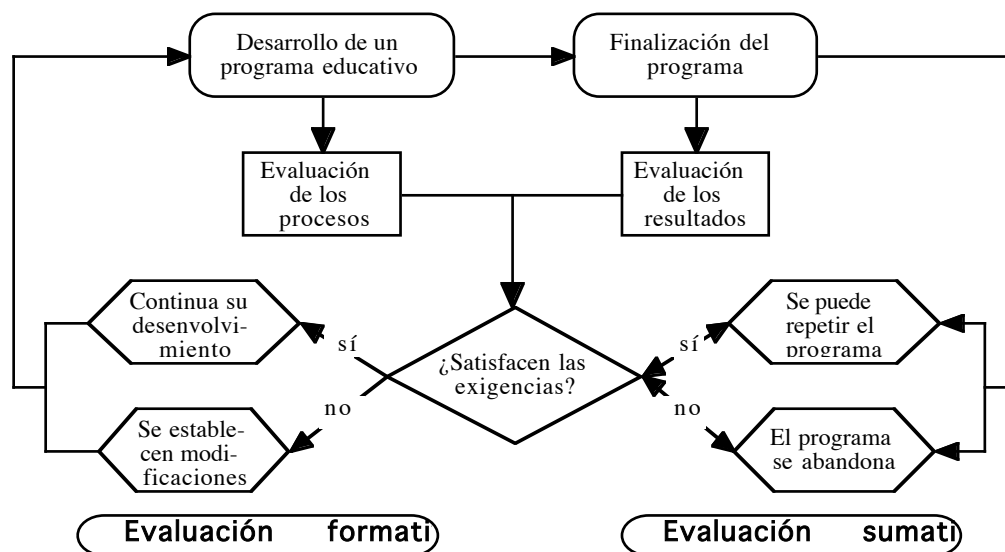
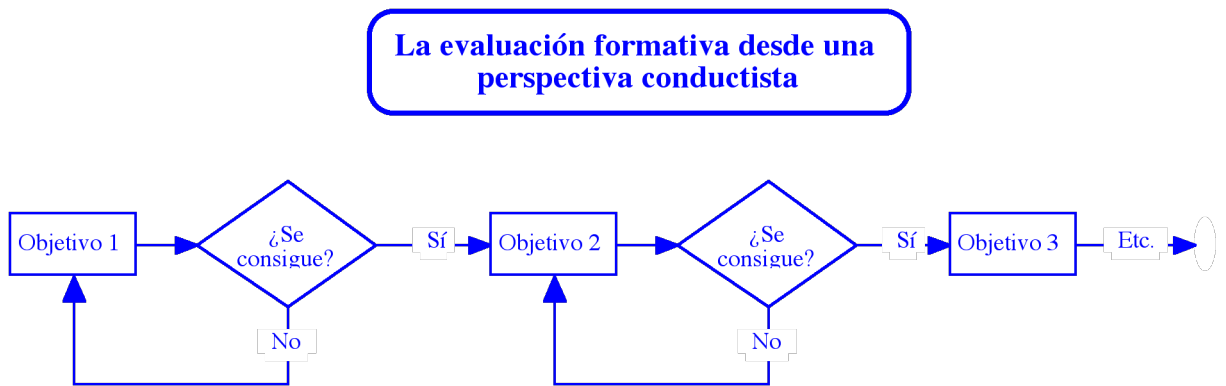


Figura 3: Distinción entre evaluación formativa y sumativa propuesta por Scriven.

Linda Allal (1980, pág. 5) distingue la concepción de evaluación formativa que mantienen los conductistas, de un lado, y los cognitivistas, de otro. Respecto de los primeros, opina que está indicada para realizar “reajustes sucesivos en el desarrollo y la experimentación de un nuevo programa, manual o método de enseñanza. (...) es un componente esencial en la realización de una estrategia de pedagogía de dominio o de cualquier otra aproximación de la individualización de la enseñanza.” Para la perspectiva cognitivista, la recogida de información sobre los resultados del aprendizaje tiene un carácter secundario, enfatizándose primordialmente las informaciones acerca de los procesos de esos aprendizajes. “Los datos de interés prioritario son los que se refieren a las representaciones que se hace el alumno de la tarea o estrategias o procedimientos que utiliza para llegar a un determinado resultado. Los «errores» son objeto de un estudio particular en la medida en que son reveladores de la naturaleza de las representaciones o de las estrategias elaboradas por el alumno.” (Pág. 7).

Un ejemplo de la primera conceptualización podría ser la ofrecida por Carreño (op. cit., pág. 40), quien define la evaluación formativa como el “conjunto de actividades probatorias y apreciaciones mediante las cuales juzgamos y controlamos el avance mismo del proceso educativo, examinando sistemáticamente los resultados de la enseñanza.” En esta

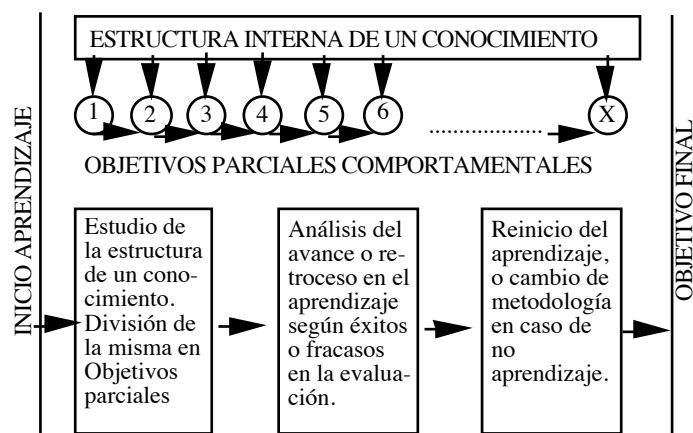
definición se puede apreciar como se confunde la evaluación formativa con la multiplicación de pequeñas evaluaciones sumativas. (Véase la figura 4).



(Figura 4)

También el triple objetivo de la evaluación formativa enumerado por Forns (op. cit., págs. 111 y ss.) se encuadra plenamente dentro de este ámbito conductista: “a) determinar el nivel de logro del alumno respecto al objetivo, b) indicar las principales dificultades de los alumnos en la consecución de los aprendizajes propuestos y c) informar al maestro acerca de la eficacia de su programa en relación al nivel de sus alumnos.” Frente a esta visión, ofrece lo que sería la óptica constructivista, más preocupada por conocer el nivel de organización del individuo y aplicar aquellos procedimientos metodológicos que permitan movilizar su actividad psicológica hasta conseguir niveles de organización superiores (más elaborados). Véase la figura 5.

OPTICA CONDUCTUAL



OPTICA CONSTRUCTIVISTA

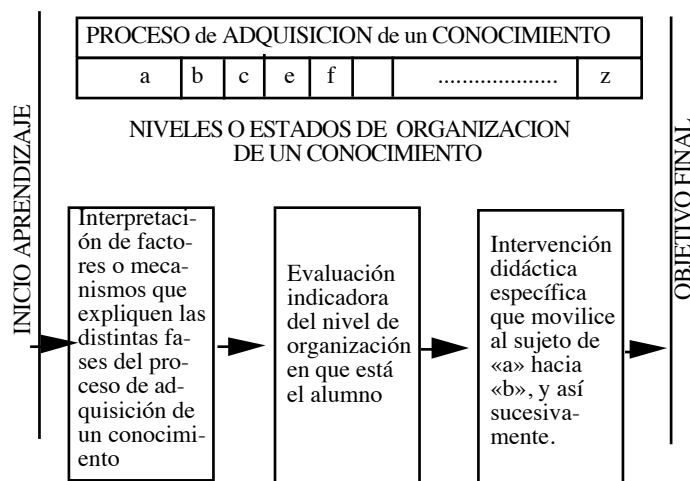


Figura 5: Diagramas para observar las diferencias entre Evaluación Formativa, concebida en términos conductuales y desde una óptica constructivista. (Extraído de Forns, op. cit., pág. 120)

Jean Cardinet (1988, págs. 99 y 100) nos ofrece una perspectiva distinta de las anteriores. Declara que los profesores no tienen por qué sentirse obligados a emitir calificaciones por cada ejercicio que realice el estudiante, “por el contrario, corregir el trabajo y dar una retroinformación detallada al alumno, ayudándole así a progresar, sea cual sea su nivel. Insertar bucles de retroacción en el proceso didáctico es siempre posible.” De este modo la evaluación ofrecería informaciones válidas al alumno y no datos sobre éste. Por su parte, César Coll (1983,

pág. 16) define la evaluación formativa como aquella que sólo tiene “lugar *durante* el desarrollo del proceso educativo y, sobre todo, en la medida que proporcione indicaciones útiles para reconducirlo.” Este mismo autor, junto a Martín (1993), defienden la importancia de la función reguladora de la evaluación desde una perspectiva constructivista.

Dentro de esta concepción podríamos encuadrar la aplicación a los procesos de enseñanza-aprendizaje utilizada por Menoyo et alius (1995, págs. 42 a 54), en su experiencia docente llevada a cabo en la ESO, como un proceso de atención a la diversidad a través del uso de mapas conceptuales. A partir de un primer mapa «diagnóstico», individual, en el que se reflejan los errores del alumno, se le van proponiendo diversas actividades de regulación con la finalidad de corregir las confusiones detectadas. Posteriormente a medida que se profundiza en el conocimiento (regulación por parte del docente) se va elaborando otro mapa conceptual más amplio que se contrasta con el anterior (autorregulación del alumno). El modelo se cierra con la interacción social en el aula, con la construcción por todos de un mapa conceptual definitivo. Con anterioridad los mapas conceptuales habían sido utilizados para la corrección de preguntas por otros autores, entre los que destacamos a Nieda et alius (1985).

Gil, Carrascosa, Furión y Martínez (1991) analizan diferentes concepciones de la evaluación y presentan unas orientaciones básicas que debería de reunir para convertirse en un instrumento de aprendizaje, eminentemente formativo. También Jorba y Sanmartí (1993, 1994 y 1995) exponen de manera muy clara las características en las que se basa esta concepción de la evaluación educativa, desde una perspectiva puramente práctica extraída de la experiencia.

Nosotros hemos considerado que la evaluación continua es, o debe ser, esencialmente formativa. De ahí el título del presente epígrafe. Ya que la «evaluación continua» que se realiza durante los procesos, haciendo que estos puedan ser modificados en el transcurso de su desarrollo, coincide con la concepción de «evaluación formativa» de Scriven. Sin embargo, como hemos visto en la anterior cita de Fernández Pérez, la práctica de la evaluación continua se realiza multiplicando el número de evaluaciones sumativas. Como afirma César Coll (op. cit., pág. 15): “El hecho de multiplicar las evaluaciones tomando procesos de

enseñanza/aprendizaje de corta duración temporal (lecciones, temas, unidades didácticas, etc.) confiere al proceso evaluativo un carácter micro-sumatorio, pero no modifica en absoluto su naturaleza. En este sentido, no es superfluo señalar que las propuestas de evaluación continua equivalen, a menudo, a propuestas de evaluación micro-sumativa.”

2.4.- Evaluación final y evaluación sumativa

Por evaluación final se puede entender la última de las evaluaciones a las que son sometidos los alumnos tras un curso o programa de enseñanza, o aquella evaluación que, al final de un período de aprendizaje, integra y recopila a todas las demás. Este último concepto es el que más se acerca al de evaluación sumativa de Scriven y es utilizado usualmente por algunos autores para definir su concepción de la evaluación (Lafourcade, 1977, pág. 21): “la evaluación es entendida aquí como una etapa del proceso educacional que tiene por fin comprobar de modo sistemático en qué medida se han logrado los resultados previstos en los objetivos que se hubieran especificado con antelación.”

Carreño (op. cit., pág. 41) se refiere a ella como “la forma mediante la cual medimos y juzgamos el aprendizaje con el fin de certificarlo, asignar calificaciones, determinar promociones, etc.” Su finalidad estaría en reflejar el grado en que se han conseguido los objetivos del curso o tema correspondiente a través de una nota o calificación. Coincide con la final por el momento de su aplicación, tras un programa o una unidad didáctica completa.

Farré y Gol (op. cit., pág. 12) hacen referencia a la evaluación final, identificándola con la evaluación sumativa y la consideran como la «menos importante» de las modalidades de evaluación, “porque nos informa de cosas que ya se dan por hechas y acabadas. Ya no habrá remedio para el niño. El maestro, en cambio, podrá extraer de esta valoración final consecuencias para el próximo curso o para el próximo tema.”

Desgraciadamente para los estudiantes la «evaluación final» sí que es muy importante, tal vez la más relevante de todas, y la pueden vivir como un «juicio final bíblico» a pequeña escala. No hay más que echar una ojeada cerca de cualquier centro educativo, en cualesquiera de los niveles

o tipo de estudios, para observar la ansiedad y tensión de los alumnos ante este último y gran obstáculo. Para ellos, la evaluación final, en ese momento, no representa precisamente una ayuda. Cerdá, Moreno y Muñoz (1995, pág. 58) se refieren a ella a través de una metáfora que refleja muy bien estos sentimientos provocados por la evaluación sumativa: “Una autopsia es de gran utilidad para conocer con exactitud las causas de una muerte, pero no consuela mucho al difunto.” Y lo que es peor, en más ocasiones de las que sería deseable, ni siquiera es un apoyo para el profesor, que repetirá un curso tras otro, el mismo rito, sin extraer ningún tipo de consecuencias para realizar modificaciones en una nueva intervención.

Coll (op. cit., pág. 15 y 16) considera esencial la obtención de la máxima información, fiable y válida, sobre el nivel de dominio de los objetivos conseguidos por los estudiantes, y distingue entre cómo se aplica esta modalidad evaluativa según nos movamos dentro del paradigma conductista o en el cognitivista. En el primero, al que denomina tradicional, se trataría de encontrar las técnicas de medición más adecuadas y sensibles a los cambios de comportamiento experimentados con motivo de la aplicación del programa. En el segundo, se establece una distinción entre cambio conductual como resultado de los aprendizajes y los propios aprendizajes concebidos como procesos internos y, consecuentemente, no observables directamente. En este caso las estrategias evaluativas deberán adecuarse a esta concepción y difieren, según este autor, notablemente de las anteriores.

2.5.- Evaluación formal *versus* evaluación informal

Aquella evaluación que se hace con carácter sistemático, fundamentada en una planificación e integrada dentro del curriculum es la que denominamos «formal». Para Cerdá, Moreno y Muñoz (op. cit., pág. 57) es la que se refiere preferentemente al uso de determinados instrumentos y la diferencian de la informal en que ésta se caracteriza por estar “formada básicamente por impresiones e intuiciones”.

Airasian y Jones (1993) han estudiado cómo la valoración informal se utiliza por los profesores para informar de las decisiones adoptadas diariamente en las clases sobre sus alumnos, y cómo estas valoraciones informales influyen en otros juicios y medidas formales, en especial, en la evaluación definitiva. Esta situación plantea un grave peligro ético, pues los docentes, frecuentemente, son «acríticos» con este tipo de valoraciones que, sin embargo, resultan muy potentes. Sobre este aspecto de la evaluación informal, Perrenoud (1990, pág. 257) afirma:

(...) en la escuela, hacer un buen trabajo consiste en realizar un trabajo no retribuido, en gran medida impuesto, fragmentado, repetitivo y supervisado constantemente. Podemos concebir que, en esas condiciones, la energía de los alumnos no se invierta siempre en la búsqueda de la máxima excelencia posible. El papel del maestro consiste en hacer que los alumnos trabajen y mantener su compromiso con la tarea a pesar de la fatiga, del deseo de hacer otra cosa, del aburrimiento o de la falta de sentido de ciertas actividades para el alumno. Pero ese control respecto al compromiso en relación con el trabajo escolar se duplica con el ejercicio sobre su calidad, sobre el grado de excelencia que manifiesta.

Ese es el papel principal de la evaluación informal que el maestro practica a diario.

2.6.- Autoevaluación *versus* heteroevaluación

La evaluación que realiza un alumno, o profesor, sobre su propio trabajo o actividad, sus conocimientos o habilidades, o sobre los cambios acaecidos en sí mismo, es la autoevaluación. Representa una máxima subjetividad por cuanto el sujeto autoevaluado se convierte en «juez y parte» de los juicios y valoraciones emitidos; sin embargo, confiere a éstos una especial sensibilidad -permite la introducción de datos e informaciones difíciles de ponderar y transmitir- que ningún otro instrumento o modalidad de evaluación posee. Además, la autoevaluación, permite corregir determinados sesgos introducidos por los propios instrumentos de valoración, o por los que se deben a las expectativas creadas por el mismo evaluador. En este mismo sentido, Hassmen y Hunt (1994) han realizado una experiencia con personas adultas de ambos sexos, tituladas universitarias, en la que se prueba que las diferencias entre el grupo de hombres, y el de mujeres, era mucho menor cuando ambos grupos respondían a pruebas en las que se podían autocorregir y valorar, que cuando cumplimentaban pruebas objetivas

convencionales de respuesta múltiple que no podían ser corregidas por ellos.

Para Wheeler (op. cit., pág. 297) “puesto que la evaluación es parte integral del proceso de aprendizaje, es preciso contar con el alumno, *en la medida en que sea posible y él sea capaz de hacerlo*, a la hora de evaluar su propio aprendizaje.” Cuando la autoevaluación la realiza el profesor, entra a formar parte de un proceso de reflexión sobre su propia práctica docente con la intención de mejorarla. (Elliott, 1991, cap. 1).

Su finalidad radica en que el sujeto autoevaluado tome conciencia de sus propios aciertos y fracasos de acuerdo a los esfuerzos desarrollados para que le sirvan de autorregulación y acabe siendo el responsable de su propia educación. Psicológicamente favorece la elevación de la autoestima, de la independencia, conciencia del propio desarrollo, etc. Forns (op. cit., págs. 116 y 117) se refiere a ella como la «evaluación compartida».

Por su parte, Jorba y Sanmartí (1995, pág. 77) aunque no la mencionen explícitamente, se refieren a ella cuando afirman:

Es un reto para el profesorado conseguir que los estudiantes adquieran una buena capacidad de autorregularse de manera efectiva, lo que comporta que lleguen a apropiarse de los objetivos y de los criterios de evaluación del profesorado y a tener un buen dominio de las capacidades de anticipación y planificación de la acción e implica incorporar estos aspectos como objetivos prioritarios de aprendizaje. De esta manera de mirar, el proceso de enseñar-aprender-evaluar se convierte en un acto de comunicación social con todas sus exigencias y posibilidades y la evaluación se revela como un elemento primordial en el proceso de auto-socio-construcción del conocimiento.

La autoevaluación también puede ser colectiva o grupal. Esto es, un equipo de trabajo valora su propia actividad, funcionamiento, logros, etc. Cerdá, Moreno y Muñoz (op. cit., pág. 59) lo justifican aduciendo que los diseñadores de una actividad deben responsabilizarse de su valoración, lo que la revestiría de una mayor objetividad (entendida esta “como un crisol de subjetividades”). En este sentido puede considerarse como una evaluación interna. Jesús Ballester (1995) ha establecido experimentalmente unas estrategias de autoevaluación en grupo muy interesantes, dado que conjugan los intereses de los alumnos y los del profesor, tanto de cara a la evaluación como a la calificación.

La heteroevaluación es la evaluación realizada por una o más personas acerca de otra persona, grupo, actividad, resultado, etc. Más objetiva que la anterior por cuanto el evaluador está situado en una posición más distante del objeto de evaluación, lo que le confiere otra perspectiva; si bien, como decíamos antes, pierde información considerada relevante por el sujeto evaluado.

2.7.- Evaluación interna *versus* evaluación externa

Por evaluación interna se entiende aquella que se lleva a cabo por los propios involucrados en el desarrollo de un programa, funcionamiento de una organización, etc.; en contraposición a la evaluación externa, basada en el «modelo de auditorías», que es llevada a efecto por agentes no vinculados, ni implicados, en el desarrollo de tales programas, ni en el funcionamiento de las organizaciones.

Los problemas de este tipo de evaluaciones son, principalmente, éticos y políticos. ¿Cuál es el papel de los evaluados?, ¿A qué audiencia se dirigen los resultados de la evaluación?, ¿Cómo ha de ser la intervención del evaluador?, etc.

2.8.- Evaluación individual *versus* evaluación grupal o colectiva

Generalmente entendemos por la primera aquella que se hace sobre un estudiante, o un profesor, (sus progresos y dificultades personales mostrados en los procesos de enseñanza-aprendizaje). Desde otra óptica -relacionada con los tests de personalidad- Anastasi (1982, pág. 468), se nos ofrecen las «*puntuaciones ipsativas*», en las que el marco de referencia siempre es el individuo respecto de él mismo. Cualquier tipo de modificación en los avances o logros individuales son siempre comparados con los anteriores obtenidos por el propio sujeto.

La evaluación individual según Taba (op. cit., págs. 408 y 409), siguiendo a Shaplin (1961), plantea serias dificultades a la hora de establecer la diferencia entre el progreso individual o la satisfacción de estándares. Propone como más adecuado para la evaluación del progreso

individual la utilización de un diseño de pre- y pos- test. Considera que mientras que esta evaluación apunta hacia la condición del sujeto; la referida a estándares lo hace hacia “un límite superior de aprendizaje o una distribución desde la cual pueda ser determinada una situación relativa. Los usos de los dos tipos de medición son a menudo muy diferentes: el primero es una manera de premiar al estudiante y de decidir sobre las etapas siguientes del aprendizaje; el segundo es un medio para clasificar a los estudiantes en una relación mutua o con un estándar absoluto, a los fines de hacer recomendaciones y de abrir juicios.”

La evaluación grupal, o colectiva, refiere a la acción de evaluar estos mismos aspectos en el ámbito de un equipo de trabajo o de una clase. Frecuentemente se lleva a cabo con referencia a «normas»; esto es, respecto de unos estándares con los que se compara, o bien, con los extraídos del propio grupo.

También se puede entender por evaluación grupal la realizada por el equipo docente sobre un grupo de alumnos, o uno sólo de ellos, tal y como ocurre en las sesiones de evaluación. En estas reuniones evaluativas se manejan simultáneamente ideas, conceptos y actitudes diversas entre el profesorado, junto a exigencias legales, sociales y administrativas. Esta forma de la práctica de la evaluación ha sido estudiada de un modo especial por Fournier y Smith (1993), quienes hacen una valoración de las discusiones que en ella se originan empleando, como herramienta de análisis, las metáforas propuestas por Toulmin (1972) para la comprensión de los diferentes conceptos que más comúnmente se utilizan.

Finalmente, también entraría dentro de esta concepción, cuando tiene lugar, la evaluación que los alumnos conjuntamente hacen de su profesor o del equipo docente que con ellos interviene.

2.9.- Evaluación cuantitativa *versus* evaluación cualitativa

Ambas representan no sólo dos maneras distintas de producir, manejar, procesar y sintetizar datos e informaciones de diferente naturaleza; sino que además son las dos puntas de iceberg tras las que hay una gran divergencia tanto en los valores, como en las metas o en los procedimientos. Entran a formar parte de dos tipos diferentes de paradigmas. La evaluación cuantitativa (Schwartz y Jacobs, 1984, págs. 20

a 36) representa la opción positivista de la ciencia. Sus datos generalmente son el fruto de mediciones numéricas a partir de las cuales se pueden verificar hipótesis. Por el contrario, la evaluación cualitativa no maneja ese tipo de datos ni de análisis. Trata de reconstruir y describir la realidad a través de un lenguaje sencillo, como el utilizado en la prensa periódica. Se realiza a base de juicios de valor sobre los datos y evidencias extraídos de la realidad evaluada.

La evaluación cuantitativa experimentó un gran auge en la década de los sesenta. Fundamentalmente integrada por modelos tecnológicos, basados en datos numéricos y análisis estadísticos. Tyler (1934) sienta las bases paradigmáticas que más adelante serán recogidas por los modelos de Cronbach (1963), Metfessel y Michael (1967), Scriven (1967), Stufflebeam (1971), Popham (1975), etc.

En la década siguiente comienza el desarrollo de la otra perspectiva: la evaluación cualitativa. Parlett y Hamilton (1972) consideran que este nuevo enfoque tiene por finalidad la comprensión holística de la realidad. Opinan que con el nuevo modelo socioantropológico, o iluminativo, basado en los datos recogidos a partir de la observación, las entrevistas, los cuestionarios y el análisis de documentos, se pueden explicar dos planos diferentes de la realidad educativa: el sistema de instrucción y el ambiente de aprendizaje. Otros modelos que se incluyen bajo este mismo epígrafe son los desarrollados en la Universidad de Standford (1971), Stake (1975), MacDonald (1976), Stenhouse (1982) y Elliott (1982), entre otros.

2.10.- Evaluación por normas *versus* evaluación por criterios

La distinción entre la evaluación basada en normas y la basada en criterios se debe a Glaser (1963) que concretaba las diferencias entre ambas en la interpretación que se hace sobre los datos.

Cuando el rendimiento de cada estudiante individualmente es interpretado, comparándolo con el de todos sus compañeros como grupo, estamos ante una evaluación basada en normas. Usualmente es común expresar los resultados en términos ordinales; aunque no necesariamente, ya que si se compara al individuo no con el grupo, sino con el promedio de éste, se obtienen datos numéricos a través de determinados procesos

estadísticos. Precisamente fue el uso de fórmulas estadísticas lo que aumentó su aceptación entre los educadores, al darle una apariencia de exactitud a las calificaciones. (Livas, 1980, pág. 20).

La información extraída de los resultados de este tipo de evaluación es bastante pobre ya que no nos indica, en absoluto, lo que el alumno ha aprendido, sino sólo en qué posición está respecto de los demás «de su grupo». Pensemos que ni siquiera lo podemos comparar con otros grupos diferentes al suyo propio, porque perderíamos toda la garantía que, hasta ese momento, nos ofrecía este tipo de análisis.

Para que los resultados tengan una mayor significatividad, al comparar las puntuaciones, la distribución debe mostrar una gran dispersión. A mayor dispersión, mayor significatividad entre las diferencias de dos puntuaciones cualesquiera. Pero en los procesos educativos, contrariamente, se tiende a homogeneizar a los alumnos en sus rendimientos y habilidades. De tal manera que los profesores de manera «artificial», y en pura contradicción con los procesos de enseñanza-aprendizaje, se ven obligados a introducir en los ejercicios preguntas rocambolescas con la única finalidad de establecer mayores diferencias entre las respuestas del grupo.

Stufflebeam (1988, pág. 197) se refiere a las «normas» como “principios comúnmente aceptados para determinar el valor o la calidad de una evaluación.” Pero en realidad no está proponiendo una evaluación basada en normas, sino, que, lo que en verdad representan esas «normas» son auténticos «criterios» de evaluación.

La evaluación basada en criterios, posee un fuerte carácter tecnológico, como podemos apreciar en la descripción que sobre ella nos ofrece Carreño (op. cit., pág. 46): “consiste en medir y enjuiciar el rendimiento de cada alumno, considerado en términos de *objetivos logrados*, comparándolo con el *volumen total de objetivos especificados para el curso o la porción de curso que se esté examinando*, dejando a un lado, para efectos de contrastación, los resultados del resto del grupo.”

Esos objetivos especificados representan lo que el alumno debe saber hacer. “A este ‘saber hacer’ se le conoce como *dominio*, que implica tanto la clase de tarea que ha de ejecutarse, como el contenido implicado en la ejecución.” (Livas, op. cit., pág. 31).

Con anterioridad a las pruebas, o instrumentos de evaluación utilizados para su medición, se determinarán qué objetivos, o porcentaje de ellos, servirán como criterio para calificar si el alumno los ha alcanzado. En la construcción de esos instrumentos (tests o exámenes, principalmente) deben quedar representados, con una proporción distinta, los diferentes objetivos que se hayan pretendido conseguir, de acuerdo a la importancia relativa de cada uno de ellos respecto de los demás en el curriculum.

Esta modalidad de evaluación se origina a partir de la necesidad de evaluar los modelos tecnológicos de enseñanza (de «caja negra» o conductistas) y supone un avance real respecto a la modalidad basada en normas, ya que ahora el individuo no tiene que ser comparado con el grupo como referencia. Además, no se precisa de un procesamiento estadístico posterior para poder emitir calificaciones. Sin embargo, el proceso de construcción de instrumentos se hace muy complejo, al tener que establecer jerarquías entre el dominio de objetivos a evaluar y, posteriormente, crear distintas cuestiones de acuerdo a la importancia de éstos.

Como puntos débiles hemos de resaltar, además, la propia dificultad de definir y reducir los procesos de enseñanza-aprendizaje a la mera formulación de objetivos con antelación, incluso, a la intervención docente. ¿Por qué esos objetivos y no otros? O, ¿abarca ese dominio, realmente, la conducta deseada? O, ¿cuál es la proporción límite que nos indique si se han conseguido o no, los objetivos evaluados?, etc., representan, a nuestro modo de ver, los flancos más débiles de esta modalidad, sin tener en consideración aspectos relativos a la propia concepción de aprendizaje que implica el modelo tecnológico de enseñanza, o su ignorancia de los procesos (cómo se aprende).

2.11.- Indicadores

Desde finales de la década de los ochenta hasta la actualidad (OCDE, 1987; Wyatt, 1988; Burstein et alius, 1989; Wyatt et alius, 1989; Blank, 1994; Selden, 1990; Fasano, 1994; Nuttall, 1994; Serrano et alius, 1995; etc.), el concepto de «indicador» se ha ido desarrollando y afianzando como una modalidad cada vez más fuerte de evaluación de la enseñanza. Su origen

hay que buscarlo en el mundo económico, de éste se trasladó a las ciencias sociales y, posteriormente, al ámbito de la educación como una forma de medir el rendimiento académico.

Siguiendo a Serrano et alius (1995, pág. 9) los indicadores se utilizan como instrumentos que ofrecen información “sobre el estado de un sistema social” y, de acuerdo a los valores que alcancen, sirven para tomar las decisiones oportunas tales que permitan “modificar la tendencia” por ellos acusada. En el caso concreto de su aplicación a los sistemas educativos, son utilizados con la finalidad de ofrecer información a los poderes públicos para que éstos puedan hacer una política adecuada. No obstante, el uso de los indicadores se está extendiendo incluso para el diseño y confección de pruebas criteriales para evaluar a los alumnos, como es el caso de Buendía y Salmerón (1994).

El concepto de indicador, en sí mismo, es un parámetro: indica la medida indirecta de sólo un aspecto o dimensión de un sistema complejo. Como tal medida aislada no sirve de mucho, ya que sólo informa del grado o nivel adquirido por un factor en concreto; pero cuando puede establecerse un sistema de indicadores diversos, y podemos comparar las mediciones realizadas en éstos, con los estándares de otro, u otros, sistemas diferentes, podemos inferir propiedades para cada uno de dichos sistemas. “Estos estándares pueden ser los objetivos propios del sistema educativo, los resultados anteriores, que permiten estudiar los cambios y tendencias a lo largo del tiempo, y los resultados de otros países.” (Alvaro Pagé, 1992, pág. 13). Si bien, habremos de considerar con César Coll (op. cit., pág. 16) al referirse al condicionamiento ideológico y sociológico que representa la elección de estos estándares que: “la elección de indicadores del aprendizaje y la elaboración de técnicas para su registro son siempre tributarias de una determinada concepción del aprendizaje.”

El indicador, siguiendo a Nuttall (1994) funciona como el piloto de la temperatura de nuestros coches. El hecho de que se encienda la luz roja sólo nos «indica» que la temperatura está alcanzando cotas muy elevadas y que debiéramos parar; pero no nos dice por qué se ha elevado la temperatura de nuestro motor hasta ese extremo. En consecuencia, valga el símil, nunca hemos de esperar explicaciones de causalidad de los indicadores educativos. Para Tiana (1994, pág. 89) los indicadores

deberían permitir valorar qué se esconde detrás de fenómenos muy complejos (como el fracaso escolar) y poder tomar decisiones.

La Organización para la Cooperación y el Desarrollo Económicos (OCDE), junto al Gobierno de EE. UU., han patrocinado un proyecto (Alvaro Pagé, op. cit., págs. 13 a 17) encargado de buscar, definir y señalar grupos de indicadores en diferentes campos de la educación. Pero esta selección ofrece distintas dificultades, muy difíciles de solventar.

Para Nuttall son tres aspectos los que interactúan en la elección de los mismos: políticos, científico-técnicos y prácticos. Blank (1994) afirma que para la correcta selección de los indicadores educativos se debería tener en consideración determinadas características idiosincráticas del sistema educativo de cada país: su política, cultura y sistema educativo. Por su parte Cohen y Spillane (1994) nos advierten de los diferentes problemas inherentes a la confección de los indicadores, que afectarían directamente a los valores que pueden ofrecer: desde la manera concreta de crearlos hasta el modo en que pueden interpretarse por parte de los distintos poderes políticos. O su falta de sensibilidad ante las diferencias existentes en el ámbito de un mismo país o la diferente experiencia y tradición de los «técnicos» encargados de operar con ellos. En esta misma línea Bottani y Tuijnman (1994) señalan que las diferencias entre los países, hace que sea difícil disponer de indicadores válidos para comparar los sistemas educativos.

Mariano Alvaro Pagé y María del Carmen Izquierdo (1994) nos ofrecen una variada bibliografía sobre indicadores, clasificada según se refieran a ámbitos sociales; educativos (teoría, conceptos y usos); políticas educativas; sobre sistemas educativos; sobre el funcionamiento de los centros o para permitir las comparaciones internacionales; que de alguna manera reflejan la gran importancia que está adquiriendo este nuevo concepto en el campo de la evaluación.

Sin embargo, ya han surgido las primeras voces críticas ante este desarrollo de los indicadores de rendimiento. Así, Brick y Hermanson (1994) nos advierten del peligro reduccionista de sacrificar la crítica de los fines y significados de la educación en pro de una mayor eficacia en su medición, ignorando los aspectos humanos, sin valor numérico y sin que puedan ser medidos. Darling-Hammond (1994) previene a los políticos del riesgo de dejarse seducir por el resultado de los indicadores, sin tener

en consideración que deben ser tomados, sólo, como una evaluación igual de relevante que otras, y no como un instrumento que sin más, les dé la «coartada» de tomar decisiones rápidas.

También Norris (1991) y Elliott (1992) proponen la modificación del concepto de «indicadores de rendimiento» por el de «indicadores de calidad». Sustituyendo de este modo, antes de que sea demasiado tarde, este nuevo concepto explotador, reductor y tecnócrata, impregnado de una ideología de «mercado social», por la alternativa de la calidad. No hemos de olvidar la influencia de la política económica en el diseño de este instrumento. Así Rumberger (1994), llega a identificar el Producto Interior Bruto (PIB) de un país con la calidad de la educación del mismo.

Aunque lógicamente los evaluadores educativos quieren informarse sobre la naturaleza de los modelos de la evaluación educativa, no deben dejarse atrapar por el seductor pero debilitante juego conocido como el amasijo comparativo de modelos. Algunas personas se deleitan viendo cómo el Modelo X difiere del Modelo Q, y como es ligeramente parecido al Modelo Z. Este tipo de comparación puede ser una distracción placentera para un evaluador educativo retirado; ya que aunque sea ineficaz no hace daño a nadie. (No hay una base empírica que apoye la hipótesis de que soslayar las comparaciones entre los modelos de evaluación contribuya a una vejez prematura; pero tampoco hay pruebas de lo contrario.) En lugar de enredarse en el juego de "iguales o diferentes" el evaluador educativo debe tratar de familiarizarse lo más posible con los distintos modelos para poder decidir cuál de ellos ha de emplear. Con frecuencia se adoptará el enfoque más ecléctico, extrayendo de los diferentes modelos aquellos procedimientos o constructos que parezcan más útiles.

Popham, W. J. (1980, b, pág. 31).

Cap. 4.- Modelos representativos de la evaluación educativa

Existen una gran variedad de clasificaciones en las que se acoplan, más o menos holgadamente, los diferentes modelos o diseños de evaluación (Castillo y Gento, 1995; Cronbach, 1983; Doctors y Wokutch, 1980; Gardner, 1977; Guba y Lincoln, 1985; Guba, 1978; Guttentag, 1971;; House, 1986,a; Lawton, 1980; Popham, 1975, a; Rafter, 1984; Rossi y Cronbach, 1976; Rossi y Freeman, 1986; Sandefur, Freeman y Rossi, 1986; Stufflebeam y Shinfield, 1987; Stufflebeam y Webster, 1981; Timpane, 1982; Wortman, 1983)¹¹. Sólo hay un rasgo común, previo a la presentación de estos «catálogos», en las manifestaciones de sus autores: que ninguna de las categorías creadas es excluyente respecto de las demás y, en consecuencia, cualquier modelo podría encuadrarse simultáneamente en dos, o más, de ellas. Desde esta perspectiva, no nos ha parecido conveniente entrar en ejercicios de elucubración exagerados y hemos tratado de simplificar la cuestión al máximo. Para ello exponemos aquí dos grandes categorías en las que tan sólo hemos querido señalar una muestra de los existentes. El primero de estos grupos lo denominamos «modelos clásicos». Casi todos los modelos pertenecientes al mismo, en su mayoría, estarían encuadrados en la órbita

¹¹Destacamos entre las más relevantes, sin ningún ánimo de ser exhaustivos, las siguientes: Castillo y Gento (1995): cuantitativos, cualitativos y mixtos; Doctors y Wokutch (1980): auditoría de procesos sociales, diseños experimentales y cuasi-experimentales, libre de objetivos, de sistemas, análisis costo-beneficio y responsabilidad; Gardner (1977): mediciones, juicio profesional, congruencia entre actuación y objetivos, toma de decisiones y libre de objetivos; Guttentag (1971): modelo experimental, jurídico y contextual; House (1986,a): análisis de sistemas, por objetivos conductuales, toma de decisiones, libre de objetivos, crítica artística, acreditación, por adversarios y transaccional; Lawton (1980): clásico, industrial, iluminativo, toma de decisiones, profesor investigador y estudio de casos; Popham (1975,a): consecución de metas, de enjuiciamiento que acentúan criterios intrínsecos, de enjuiciamiento que acentúan criterios extrínsecos y de facilitación de decisiones; Rafter (1984): científicos, interactivos e híbridos; Stufflebeam y Shinkfield (1987): falsas evaluaciones, cuasievaluaciones y verdaderas evaluaciones; Stufflebeam y Webster (1981): evaluaciones orientadas políticamente, orientadas a problemas específicos y orientadas a juzgar; Timpane (1982): experimentales, eclécticas, descriptivas y análisis costo-beneficio; Wortman (1983): experimentación social, metaanálisis, análisis costo-beneficio; etc.

empírico-racionalista y conductista. Manejan datos numéricos y los analizan, usualmente, a través de procesos estadísticos.

La segunda categoría, que hemos denominado «modelos alternativos» está formada por todos aquellos modelos que han aparecido, generalmente, como reacción a los de la categoría anterior. Casi ninguno maneja datos numéricos y/o responden a otro tipo de análisis y enfoque sobre la realidad evaluada.

Para el análisis que hemos efectuado, hemos seguido mayoritariamente las descripciones expuestas por Popham, en su obra publicada en español en 1980, *Problemas y técnicas de la evaluación educativa*, que hemos considerado de un gran interés debido a la imparcialidad y claridad de sus apreciaciones.

1.- MODELOS CLASICOS

A).- Modelos basados en la consecución de metas u objetivos.

B).- Modelos basados en la formulación de juicios.

C).- Modelos facilitadores para la adopción de decisiones.

2.- MODELOS ALTERNATIVOS:

A).- Evaluación como crítica artística.

B).- Evaluación libre de metas.

C).- Método del Modo Operandi.

D).- Evaluación comunicativa.

E).- Evaluación iluminativa.

F).- Evaluación respondiente.

4.1.- Modelos clásicos

4.1.1.- Modelos basados en la consecución de metas u objetivos

Dentro de esta categoría podríamos establecer diferentes grupos, si bien nos ha parecido relevante no hacer nuevas subdivisiones que, sin duda, contribuirían a aumentar su complejidad y no arrojarían una mayor claridad en su comprensión. Señalamos unos cuantos modelos, eminentemente prácticos, que tratan de responder a las demandas de los profesores de acuerdo a sus necesidades.

4.1.1.1.- Modelo de Tyler

Comenzamos, siguiendo a Popham (op. cit., págs. 32 y 33), analizando el modelo de evaluación propuesto por Tyler. A este evaluador se deben muchos de los enfoques de evaluación que durante años se han extendido, e influido, entre la mayoría de los educadores y evaluadores. Tal vez sea él uno de los primeros en establecer, y defender, la necesidad de una definición de los objetivos en términos de conducta -operativos- como el elemento central en el que se debe basar la evaluación. Tan grande ha sido su influencia que de él arranca toda una tradición (llamada «tyleriana»), que con el apoyo y desarrollo del conductismo, se ha instalado fuertemente tanto en los sistemas de enseñanza como en las mentes de muchos profesores. Su propio concepto de educación como proceso complejo e intrincado que va desde la formulación de objetivos, tan claramente como sea posible, hasta la asunción de decisiones sobre los cambios y mejoras necesarios en el curriculum, confiere a la evaluación un carácter dinamizador de los procesos de aprendizaje.

Las actividades básicas de un programa de evaluación serían: la identificación de los objetivos; la determinación de las oportunas actividades y experiencias de aprendizaje en función del conocimiento de los alumnos; y, por último, la evaluación del grado en que han sido alcanzados los objetivos. Para Tyler los objetivos han de formularse a partir del análisis de lo que considera «fuentes de las metas educativas» -el alumno, la sociedad y la naturaleza de los contenidos- y de la filosofía y psicología de la educación que las engloban. Véase la figura 9.

Los objetivos educativos representan los «criterios» en los que se ha de fundamentar la evaluación. La evaluación basada en «normas» es considerada como inadecuada e insuficiente, toda vez que la basada en estos criterios refleja un conocimiento mucho más preciso y objetivo acerca del rendimiento escolar, y posibilita la emisión de informes de

carácter diagnóstico frente a las simples calificaciones, que a lo sumo comparan ordinalmente a los alumnos. Todos estos principios fueron ya reflejados en su famoso «estudio experimental de ocho años» (*Eight Year Study*).

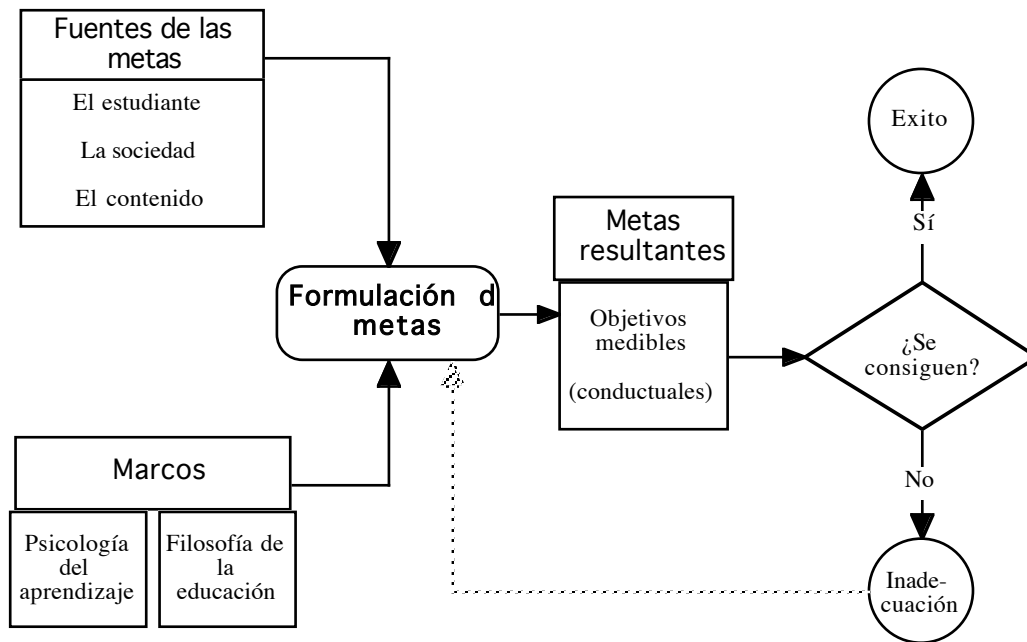


Figura 9. Enfoque de consecución de metas de Tyler.

4.1.1.2.- Modelo de evaluación de Adams.

Todos los modelos propuestos de evaluación, en cuanto procesos que son, se desarrollan siguiendo una secuencia de fases o estadios. Estas etapas poseen algunas diferencias relevantes de unos a otros autores, e incluso en ocasiones cuando se trata del mismo autor. A continuación vamos a exponer algunas de estas fases que debe encerrar cualquier proceso evaluativo. En primer lugar exponemos la secuencia expresada por Adams en su modelo de evaluación (1970, págs. 21 y 22). Véase la figura 10.

Para éste autor, aunque el proceso de evaluación está fundamentado básicamente en las mediciones, como ocurre con

el uso de los tests, las etapas son las mismas cuando el proceso se centra en las observaciones del comportamiento.

En la primera etapa, habrá de tenerse en consideración el tipo de información a obtener, dado que si no fuera significativa afectaría a los juicios que pudiéramos emitir sobre el objeto que se desea valorar.

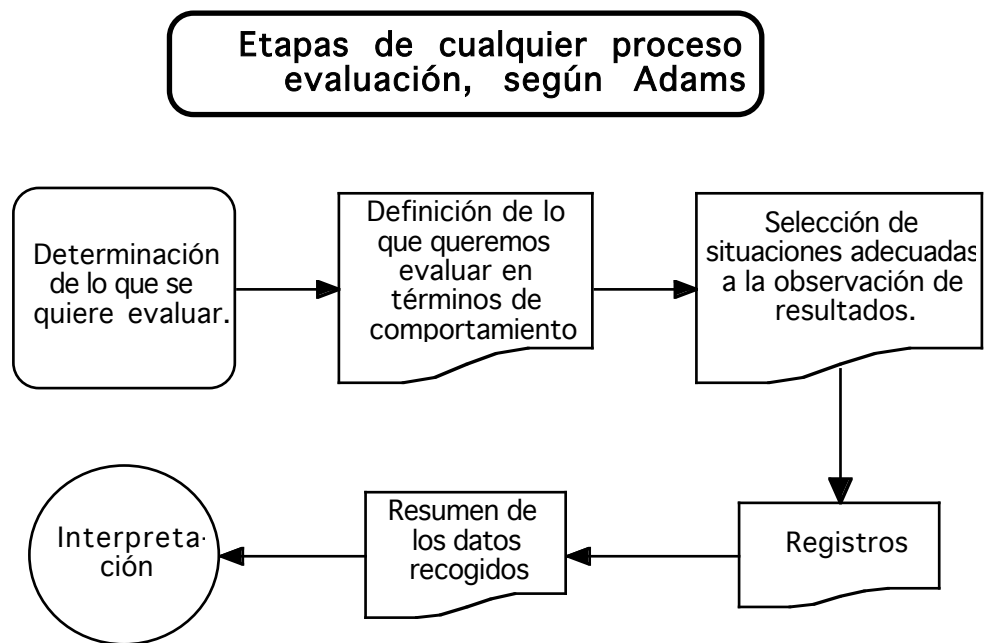


Figura 10.

Para la segunda se propone la operativización del objeto de evaluación, en términos tales que permitan su medición. Normalmente se hará en forma de comportamiento observable.

La tercera etapa se centraría en seleccionar qué situaciones permiten la adecuada observación de los comportamientos anteriormente definidos. Estas situaciones bien pudieran ser las naturales del aula, o recreos, o la emisión de respuestas a determinadas preguntas, lo suficientemente “amplias y representativas como para que las inferencias basadas en los resultados del alumno sobre la muestra suministren un índice medianamente preciso de su nivel habitual”. (Pág. 21)

La cuarta etapa implica la recogida de los datos, para lo que habría de determinarse qué aspectos son susceptibles de puntuación, cuáles las unidades de medida a utilizar, etc. Además se deben utilizar aquellos instrumentos de recogida de información más adecuados a la naturaleza de los propios datos: tests, grabaciones en vídeo, grabaciones en audio, observaciones sistematizadas de acuerdo a determinadas plantillas en unidades de tiempo, como las sugeridas por Flanders (1970), etc.

A esta etapa, continúa la síntesis de todos los datos obtenidos para permitir una adecuada interpretación de los mismos. Con esto finalizaría el proceso.

4.1.1.3.- Modelo de evaluación de Taba.

Nos vamos a fijar a continuación en otra de las secuencias propuestas para los procesos de evaluación. En este caso las sugeridas por Hilda Taba en su modelo, (1983, págs. 410 y ss.). Véanse las figuras 11 y 12.

Para esta autora la educación es concebida como el proceso cuyo principal objetivo consiste en intentar modificar el comportamiento de los alumnos y, consecuentemente, el proceso de evaluación consistirá en la determinación, y estimación, de los cambios producidos en relación a esos mismos objetivos preestablecidos, a través de determinados instrumentos: tests, observaciones, análisis de producciones, etc.

Etapas del proceso de evaluación: propuestas por H. Taba

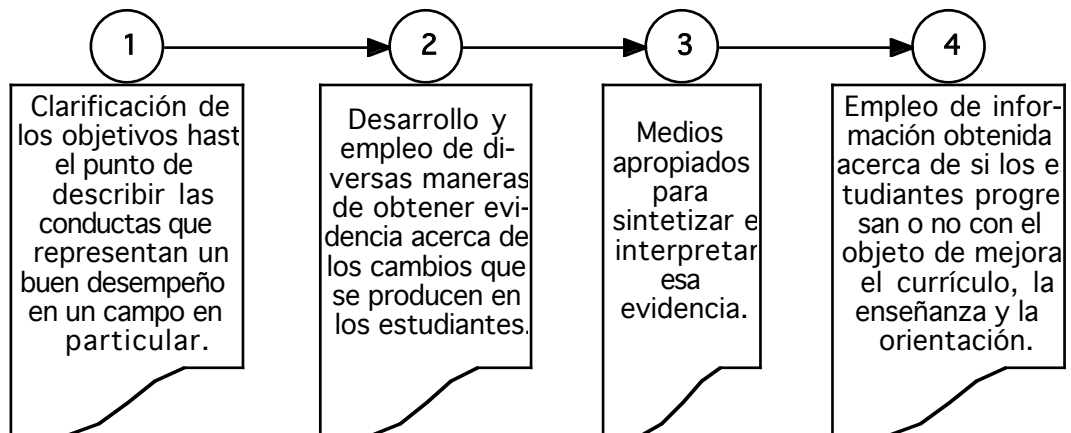


Figura 11.

Por ello, los diferentes estadios que habrán de seguir los programas de evaluación serían: la descripción de aquellos comportamientos que puedan identificarse con los objetivos perseguidos; la creación y el uso de instrumentos de recogida de la información pertinente para saber si se producen, o no, dichos cambios; el establecimiento de las necesarias estrategias que permitan sintetizar e interpretar la información disponible; y la toma de decisiones para mejorar el curriculum de acuerdo al significado de las interpretaciones realizadas.

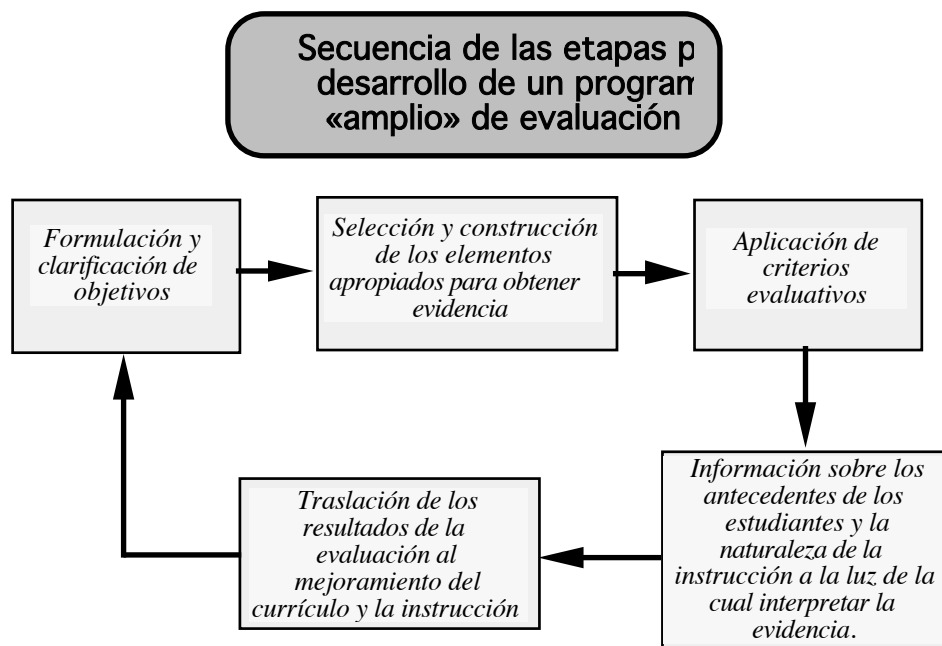


Figura 12.

Más adelante (págs. 424 y ss.), esta misma autora define con mayor precisión el modelo de proceso evaluativo que propone. En primer lugar, habrá que preguntarse acerca de cuáles son los objetivos en que se sustenta el currículum y qué conductas son las que muestran que se han conseguido dichos comportamientos. La segunda fase trataría de responder a cuestiones relacionadas con cuáles serían las condiciones, naturales o controladas, en las que los alumnos podrían expresar mejor sus habilidades. La tercera referiría al establecimiento de criterios relacionados con la cantidad, calidad, madurez, etc. En la cuarta habríamos de preguntarnos sobre cuáles son, y cómo se identificarían, los factores que influyen en la consecución de los objetivos. Por último, el estadio final, lo dedicaríamos a valorar las implicaciones que de los resultados pudiéramos extraer tanto respecto de los objetivos planteados, como de la enseñanza en general. Apréciense cuanto llevamos manifestado en la figura 12.

Pasamos a continuación a enumerar las condiciones que según esta misma autora (págs. 414 y ss.) deben de reunir los programas de evaluación, véase la figura 13. El primer criterio,

«compatibilidad de los objetivos» hace referencia a la congruencia que debe existir entre los planteamientos generales que inspiran el curriculum y aquellos sobre los que se fundamenta el programa de evaluación. Si existe una divergencia entre ambos, la evaluación perderá validez y fiabilidad, toda vez que estaremos valorando aspectos distintos de los pretendidos. Así ocurriría si en un sistema de enseñanza orientado, por ejemplo, hacia los «procesos», llevamos a cabo una evaluación basada en los «logros» (como lo propuesto en nuestro «decreto de mínimos»). Igualmente suele ser muy frecuente este problema de «incompatibilidad», cuando se traslada un modelo de evaluación que nos ha parecido muy interesante, a un sistema para el que no fue ideado. A menudo las divergencias que inspira a uno y a otro, harán fracasar la fiabilidad y validez de nuestras valoraciones. También puede ocurrir cuando los objetivos pretendidos no están lo suficientemente explicitados y claros como para que no originen ambigüedades con los considerados «implícitos» en nuestras prácticas pedagógicas.

La «amplitud» refiere a si los programas de evaluación cubren todo el dominio o espectro formulado en los objetivos educativos. Es fácil observar como para aspectos concretos existen una gran cantidad de instrumentos que permiten su evaluación. Sin embargo no sucede con todos. Así, determinadas capacidades como las actitudes o la creatividad, por sólo poner dos ejemplos, resultan bastante difíciles de valorar, a lo que tendríamos que añadir la dificultad para poder encontrar las herramientas apropiadas. Eso hace, según Taba, que la enseñanza acabe orientándose preferentemente hacia lo que sí puede ser evaluado, en detrimento de aquellos otros aspectos que aun siendo valiosos, se descuidan en la evaluación.

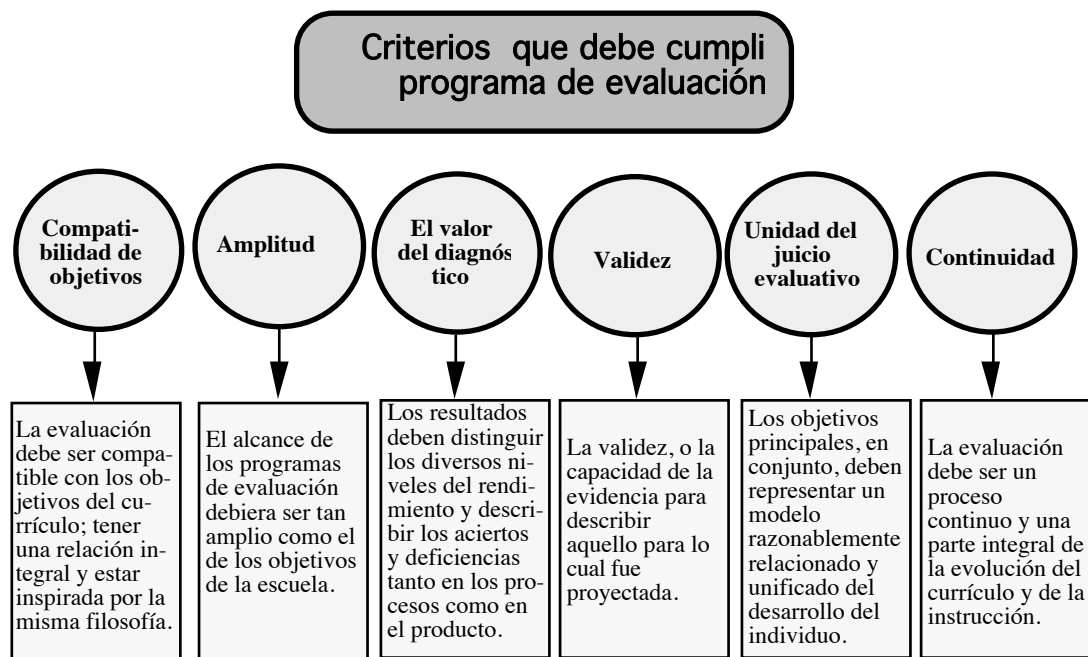


Figura 13.

«El valor del diagnóstico». La evaluación debe facilitar la diferenciación no sólo de los niveles de rendimiento, sino también la distinción entre aquellos procesos que llevan tanto a la provocación de los errores como a la de los aciertos. De este modo la evaluación cobra una dimensión relevante al permitir valorar los niveles de comprensión sobre los que actúa el alumno y la capacidad y profundidad con la que realiza transferencias de sus aprendizajes.

La «validez». Es directamente proporcional a la correspondencia entre los objetivos del currículum y los instrumentos de evaluación. Razón por la que los objetivos deben estar claramente definidos. Su ambigüedad aumentaría la falta de coherencia y por ende, afectaría negativamente a la validez. Cuanto mayor sea la congruencia entre metas y utensilios, más validez tendrá el proceso de evaluación, y viceversa, la validez de la evaluación se verá seriamente afectada a medida que sea mayor la disparidad entre los fines y las herramientas utilizadas para su medición.

Además, si no se tienen presentes las situaciones y los momentos en los que son utilizados los instrumentos, se podría ver muy mermada la validez de todo el proceso evaluativo. De tal modo que se deben crear las oportunidades suficientes para que los alumnos manifiesten la conducta pretendida, excluyendo, y controlando, los obstáculos y variables extrañas que las pudieran afectar.

«Unidad del juicio evaluativo». También afecta a la validez de la evaluación. El principal problema se plantea en la traslación de objetivos generales a específicos, para posibilitar el análisis de las conductas que los componen. Sin embargo como la conducta es algo global, después tendremos que realizar un proceso inverso, de síntesis, tal que permita integrar estas conductas parciales en aquella más amplia que se pretendía evaluar. En estos procesos de análisis y de síntesis cobra una especial relevancia la propia naturaleza del proyecto curricular, y la de los estudiantes, especialmente referida a la condición social de los mismos, ya que puede influir significativamente en las diferencias observadas tanto entre los grupos de alumnos, como entre los propios individuos.

«Continuidad». Refiere a la integración del proceso de evaluación a lo largo de todo el sistema de enseñanza, desde su inicio, con la evaluación diagnóstica, hasta su finalización, con la evaluación sumativa. Al evaluar sólo en determinados momentos, como sucede con los exámenes o controles, se obtienen evidencias parciales que no incluyen la totalidad del dominio pretendido en los objetivos, distorsionando los resultados y disminuyendo la validez y eficacia de la evaluación.

4.1.1.4.- Modelo de evaluación de Carreño

Para este autor (1985, págs. 27 y ss.), los procesos de enseñanza-aprendizaje se dividen en tres fases muy estrechamente relacionadas entre sí, que son las que confieren coherencia y cohesión a su desarrollo: la planificación, la ejecución y la evaluación. Las “operaciones evaluativas”

permiten que las otras fases puedan ser mejoradas, perfeccionándolas, ajustándolas a los cambios según se vayan éstos produciendo, corrigiendo los posibles errores e incidiendo en los aciertos.

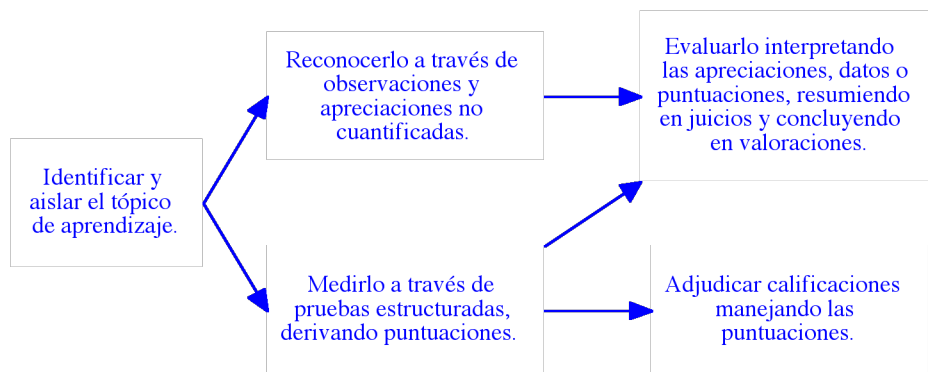


Figura 14: Modelo de evaluación expresado por Carreño (op. cit., 28):

En el proceso de evaluación se deben conjugar tanto los datos obtenidos a través de las mediciones como los recabados de las observaciones, según puede apreciarse en la figura 14. Los primeros otorgarán una mayor objetividad, así como la facilidad que entraña el manejo de datos numéricos para su análisis y para la posterior emisión de calificaciones. Los derivados de las observaciones, eminentemente cualitativos, servirán para la evaluación; aunque representarán mayores dificultades para calificar.

La evaluación debe partir de las interpretaciones explicativas de los resultados y llegar a formular «proposiciones operativas» que permitan la intervención sobre la realidad. El ciclo propuesto, sería el reflejado en la figura 15. (Págs. 36 y 37).

Considera que la funcionalidad de la evaluación está relacionada tanto con los procedimientos de enseñanza como con su selección a través de la medición del rendimiento de los alumnos, pues comparando los resultados se puede preferir objetivamente a unos sobre otros. Desde esta perspectiva,

podría encuadrarse éste dentro de los modelos de evaluación comparativa.

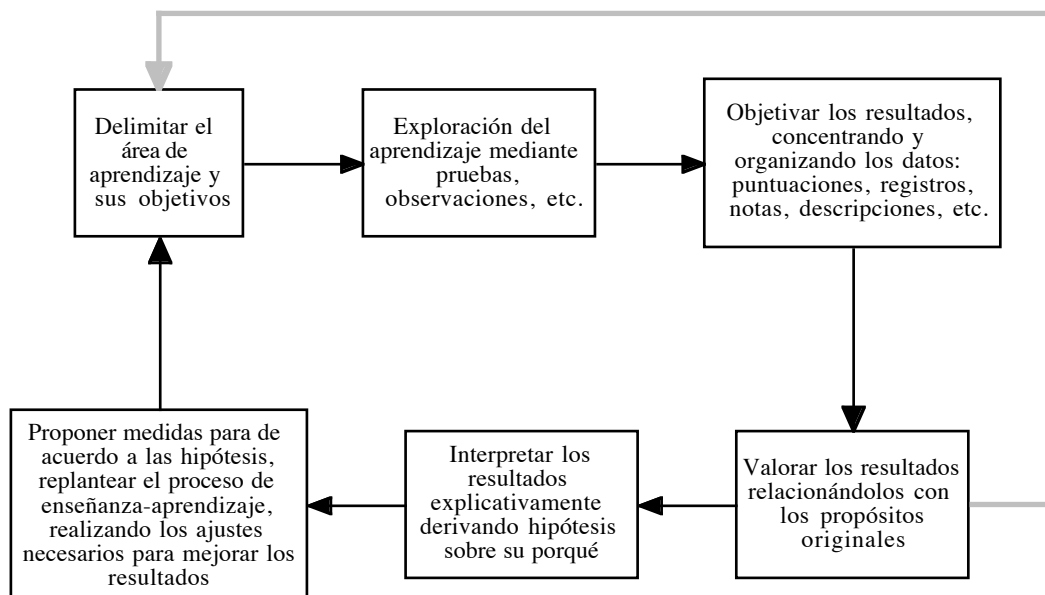


Figura 15: Ciclo propuesto por Carreño para poder realizar proposiciones de carácter operativo.

Un aspecto que Carreño destaca, y que nos ha llamado especialmente la atención, ha sido la importancia ofrecida al tiempo como indicador del grado de adecuación, a la realidad, de un programa. Para él, la evaluación nos informará sobre «el tiempo» “promedio verdaderamente requerido para el logro de cada objetivo. Sumando el tiempo requerido para el logro de los objetivos y cotejando la suma con la del tiempo efectivamente disponible o registrado en los planes, sabremos si los programas son realistas y factibles o ambiciosos e irrealizables.” (pág. 32)

4.1.1.5.- Modelo de evaluación de Hammond.

Para este autor (1973, págs. 157 y ss.) la evaluación debe expresar la eficacia de los programas educativos en función de si se consiguen, o no, los objetivos propuestos. Incide en la importancia que poseen diferentes variables relativas a la

enseñanza y a la institución en donde ésta se desarrolla como condicionadoras de la conducta del alumno. Por ello, resalta la importancia de su descripción y posterior análisis para interpretar de manera correcta las mediciones sobre el rendimiento académico de los estudiantes. Véase la figura 16.

4.1.1.6.- Modelo de evaluación de Metfessel y Michael

Tal vez lo más destacado de estos autores (1967, págs. 931 y ss.) sea la importancia que dan a los estadios primero -tratando de implicar en el desarrollo del programa a todos los participantes- cuarto y quinto de su diseño, con la utilización de múltiples criterios de medida para determinar hasta qué punto se han conseguido los objetivos. Véase la figura 17. Si bien, deberíamos de recapacitar acerca de la realización periódica de medidas en diferentes intervalos de tiempo, que afectaría directamente al criterio de «continuidad» especificado por Taba (véase la figura 13) y, por lo tanto, a la recogida de información de todo el dominio expresado con los objetivos.

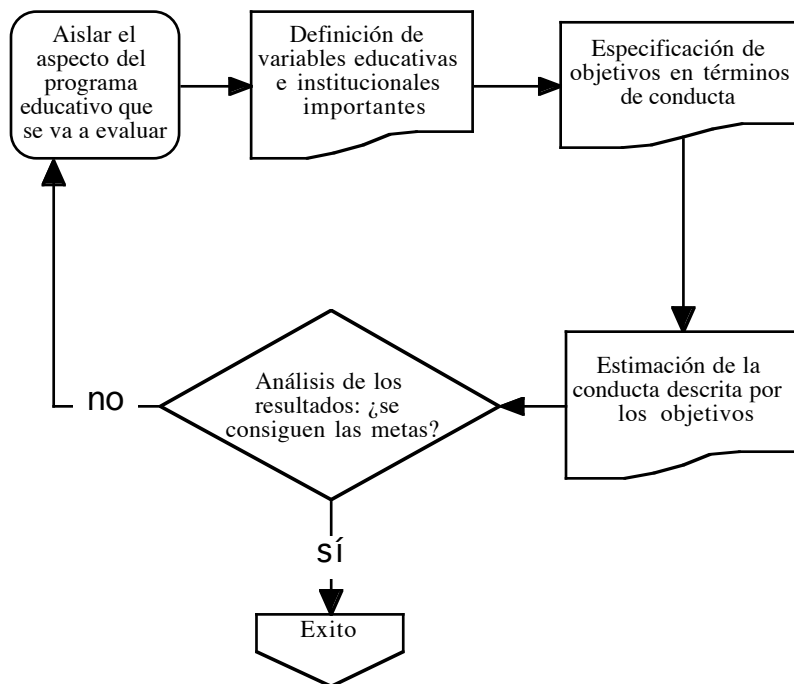


Figura 16: Modelo de consecución de metas de Robert L. Hammond.

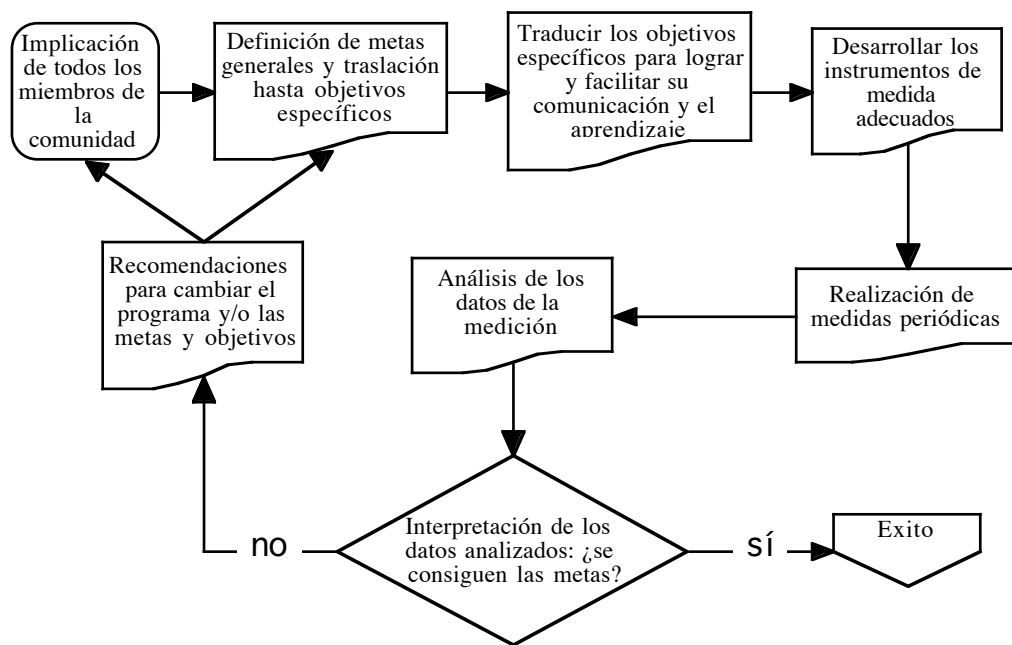


Figura 17. Modelo de consecución de metas de Metfessel y Michael.

4.1.1.7.- Evaluación retributiva y comparativa

Convendría aquí hacer mención de una manera especial a la figura de Michel Scriven, (Popham, op. cit., pág. 35 y ss.). En su artículo de 1967: "The Methodology of Evaluation", incorpora clara y explícitamente, por vez primera, la concepción ética como criterio de evaluación, a través de su idea de «evaluación retributiva» o de evaluación centrada en los efectos producidos por los programas de la enseñanza. Véase la figura 18. Para este autor y filósofo, cuando las metas y objetivos educativos carecen de valor, no tiene sentido conocer hasta qué punto han sido conseguidos. El papel que debe de adoptar el evaluador debe de ser crítico y activo con respecto al valor de dichos objetivos, censurándolos e incluso rechazándolos abiertamente si atentan contra la dignidad de las personas.

Una generalización de este concepto de «evaluación retributiva» podría ser el modelo, igualmente propuesto por Scriven en ése mismo escrito, de «evaluación comparativa», que perfectamente podríamos encuadrar dentro de aquellos otros modelos especialmente diseñados para ayudar a tomar decisiones. La idea clave de Scriven es que si en el ámbito de la educación hay que tomar decisiones sobre diferentes currícula, estas se facilitarán eligiendo desde distintas alternativas. No es necesario comprender las causas por las que se diferencia un programa de otro; sino que para decidir acerca de ellos es suficiente conocer sus desigualdades y el tipo de efectos que cada uno provoca. Véase la figura 19.

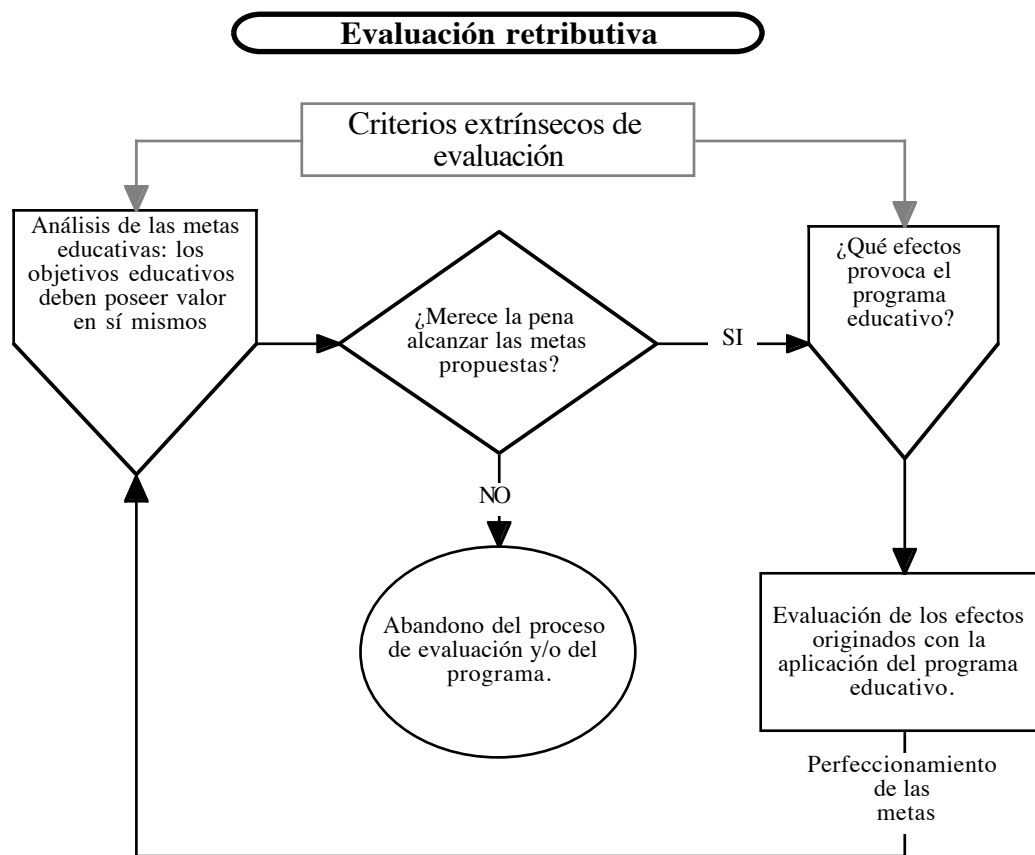


Figura 18: Evaluación retributiva de Scriven. Los criterios extrínsecos hacen referencia tanto al mérito de las metas como al de los efectos originados por los programas curriculares.

Esta opción implicó un enfrentamiento respecto de la postura mantenida desde unos años antes por Cronbach (1963), quien señaló que este tipo de evaluaciones comparativas hacía difícil entender a qué causas se podrían atribuir las diferencias, pues cualesquiera programas pueden coincidir o divergir desde diferentes ópticas, según este último autor.

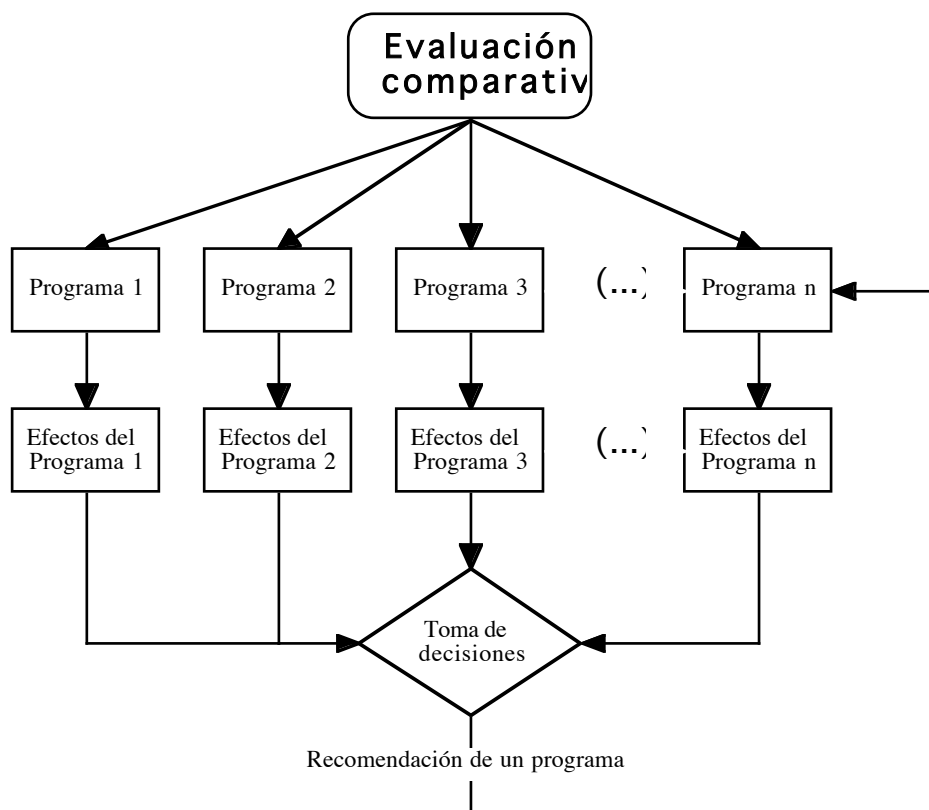


Figura 19: Evaluación comparativa de Scriven.

4.1.2.- Modelos basados en la formulación de juicios

Son diseños de evaluación en los que las conclusiones de los evaluadores son determinantes a la hora de decidir sobre la bondad de una intervención o de un programa educativo. Digamos que la mayor parte de responsabilidad compete a la figura del evaluador que como «experto» emite «juicios» que serán decisivos a la hora de mantenerlo, modificarlo o anularlo. Aunque existan otras «versiones», como la ofrecida por

Stufflebeam y Webster (1981, pág. 6), en las que se suaviza este poder del evaluador trasladando su responsabilidad a aquellos para los que evalúa: "... un estudio de evaluación educativa es aquél que es diseñado y llevado a cabo para ayudar a alguna audiencia a juzgar y mejorar el valor de algún objeto educativo".

Para ello, unos diseños hacen mayor hincapié en unos atributos, que se toman como criterios de la evaluación que pueden ser inherentes al propio programa. Otros, sin embargo, agudizan el peso que ofrecen a los efectos que producen. Los criterios intrínsecos estarían asociados a cualidades del objeto evaluado: color, peso, etc. En nuestro caso atenderían más a las características internas de un programa de enseñanza. Los extrínsecos, por el contrario, se centrarían más en los efectos producidos o en la funcionalidad directa del objeto de evaluación, como ya hemos apreciado en el caso de la evaluación retributiva.

4.1.2.1.- Modelo de acreditación

Como ejemplo de este tipo de evaluación que estuvo bastante generalizado en la década de los sesenta en EE. UU., podríamos citar «el modelo de acreditación», así llamado porque se llevaba a cabo, normalmente, antes de conceder un "crédito". Véase la figura 20.

Con anterioridad a la propia evaluación (Popham, op. cit., pág. 34 y 35), los evaluadores definían los criterios en los que ésta se basaría. De acuerdo con dichos criterios se les pedían a los participantes que redactaran un informe acerca de su propio funcionamiento -autoevaluación-, hecho que generalmente era muy bien recibido entre éstos ya que les permitía realizar una profunda reflexión sobre su labor. Posteriormente recibían la visita del equipo evaluador que obtenía datos, básicamente, de la infraestructura del centro: cantidad y calidad de las instalaciones, de la biblioteca y de otros servicios, del nivel de formación del profesorado, etc.

Con todos estos datos, llamados «de proceso» porque se consideraba que las condiciones en las que se desenvuelven los currícula influirían decisivamente en el desarrollo y calidad de los programas educativos, se emitía un juicio por los expertos otorgándoles, o no, la conveniencia de nuevas inversiones para su aplicación o mantenimiento.

El paulatino abandono del modelo se produjo por la ausencia de evidencias empíricas que permitieran correlacionar los valores relativos a las infraestructuras y los intrínsecos, correspondientes a los programas curriculares, con los resultados finales.

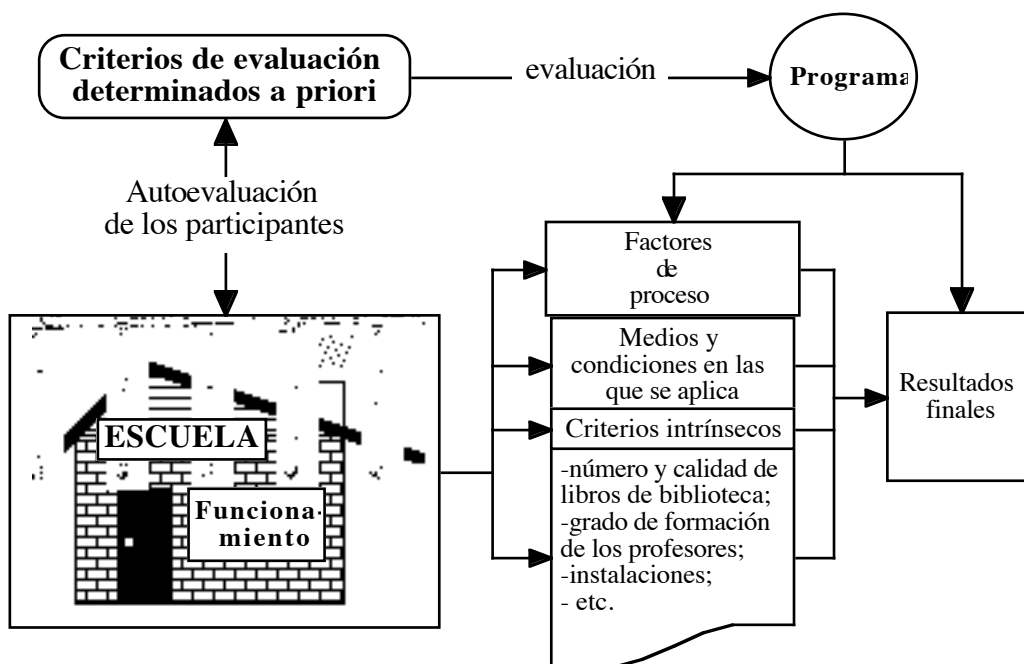


Figura 20. Modelo de acreditación.

4.1.3.- Modelos facilitadores para la adopción de decisiones

En la década de los años sesenta, en los Estados Unidos, se estaban invirtiendo importantes sumas económicas por parte de la Administración demócrata, entonces en el poder, para el desarrollo de distintos programas educativos que incidieran en una mayor igualdad de oportunidades, o que trataran de

compensar determinados déficits de algunas etnias, o de elevar el status cultural de aquellas clases sociales más desfavorecidas. Dentro de este marco de inversiones públicas en los servicios educativos se hacía patente, por un lado, la responsabilidad política de evaluar la calidad de dichos servicios y, por otro, la necesidad de obtener una adecuada información, tal que permitiera al Gobierno tomar decisiones sobre los mismos en función del binomio costes-beneficios.

Este marco social, político y económico es el caldo de cultivo que desde entonces hace fructificar toda una serie de modelos de evaluación cuya única finalidad radica en facilitar la toma de decisiones sobre nuevas inversiones (o el mantenimiento de las ya existentes) a políticos y a burócratas. Hoy en día, son los «indicadores educativos» los que, con su expansión y desarrollo, están ocupando este decisivo papel. Tal vez podamos considerar como los primeros antecedentes los modelos CSE de Alkin, el denominado PPBS (Sistemas de Planificación, Programación y Presupuesto) descrito por Rivlin (1971, pág. 3) como el proceso sistemático para ir obteniendo información substancial a medida que se toman las decisiones, y el CIPP, de Stufflebeam y Guba.

Nosotros, siguiendo a Popham (op. cit., págs. 43 y ss), sólo vamos a presentar dentro de esta categoría los dos modelos que consideramos más relevantes por la influencia que han podido ejercer en este ámbito: el modelo CIPP y el CSE.

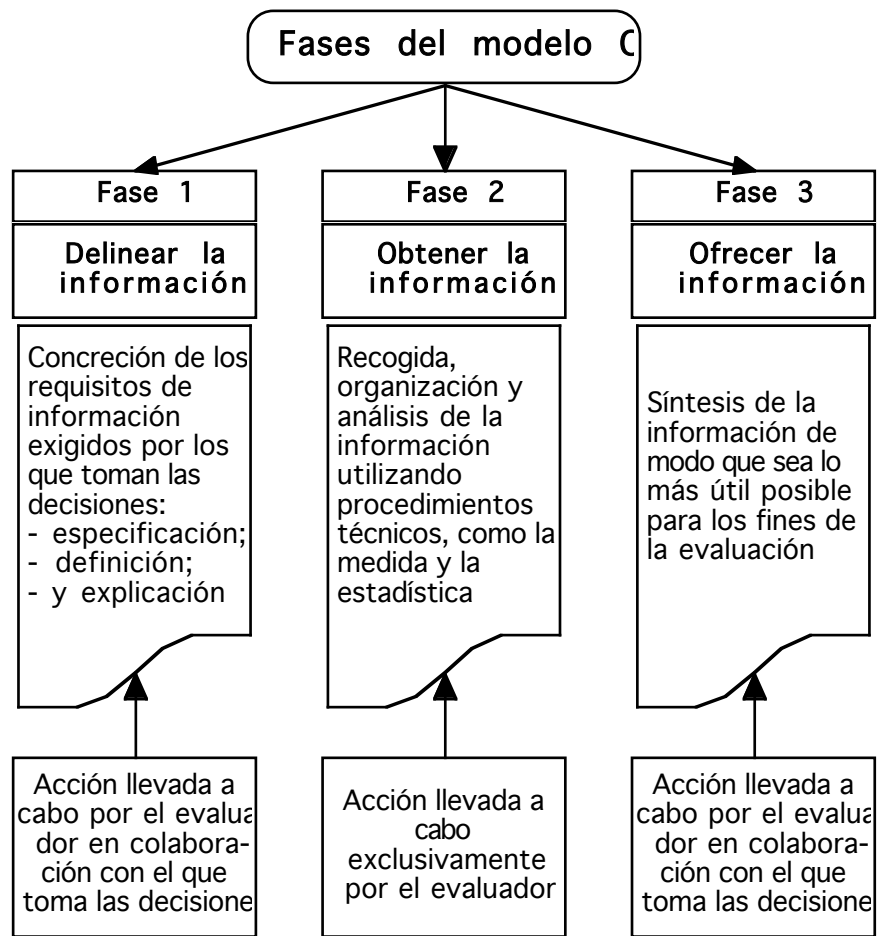


Figura 21.

4.1.3.1.- El modelo CIPP

El modelo CIPP (evaluación del Contexto, de la Información, de los Procesos y de los Productos) fue ideado por Stufflebeam y Guba a comienzos de los años setenta (Stufflebeam et alius, 1971; Stufflebeam, 1974).

La filosofía a la que responde es que la evaluación debe ser una herramienta que debe utilizarse sistemáticamente de un modo continuo y cíclico. Para ello se describen tres estadios para la especificación, obtención y síntesis de la información que se ha de ofrecer a quienes toman las decisiones. Véase la figura 21.

De acuerdo a la variabilidad de los diferentes tipos de decisiones existentes, el programa de evaluación CIPP, distingue otras tantas situaciones en las que habrán de tomarse dichas decisiones. Véase la figura 22. Así, habrá unas exigencias de información distintas según se trate de mantener el equilibrio en el funcionamiento o mantenimiento del sistema (decisiones homeostáticas), de mejorarlo continuamente sin necesidad de grandes inversiones (incrementativas), de exigencia de un fuerte apoyo económico a la innovación (neomovilísticas) o, por último, aquellas otras que implican cambios totales dentro del sistema (metamórficas).

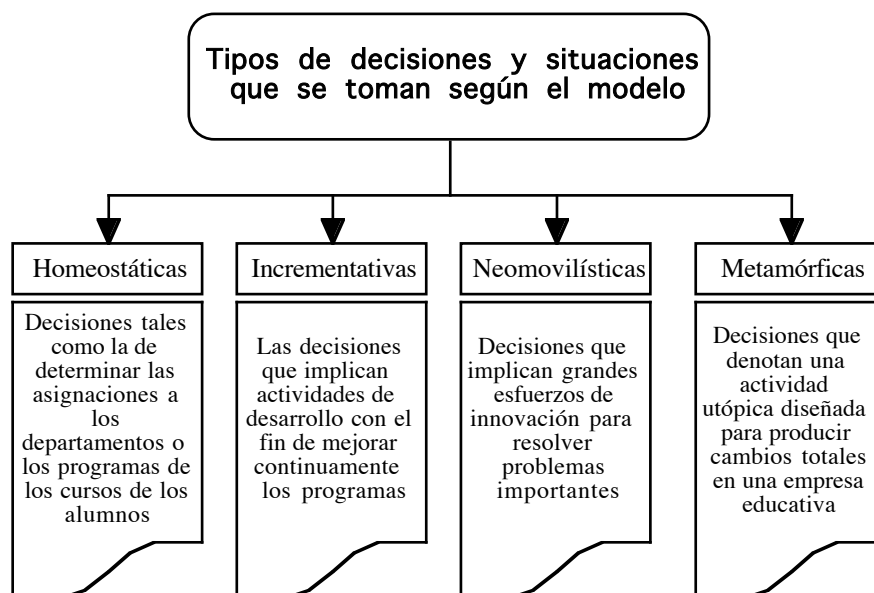


Figura 22.

Además de lo que ya hemos visto sobre el tratamiento de la información y de las situaciones en las que se habrán de tomar decisiones, en general, el modelo CIPP aplicado específicamente al marco educativo posee cuatro estadios que lo caracterizan, constituyendo en sí mismos el núcleo del programa. Es lo que podríamos denominar como las finalidades de los distintos tipos de evaluación. Véase la figura 23. En primer lugar la planificación de las decisiones orientadas

a la determinación de los objetivos. Una segunda fase relativa a la organización de las decisiones de cara al diseño de la instrucción. El tercer paso consiste en la utilización, control y perfeccionamiento de los procedimientos utilizados en el diseño a través de las propias decisiones que interactivamente se vayan tomando. Por último, la reutilización de las decisiones de acuerdo a los resultados obtenidos.

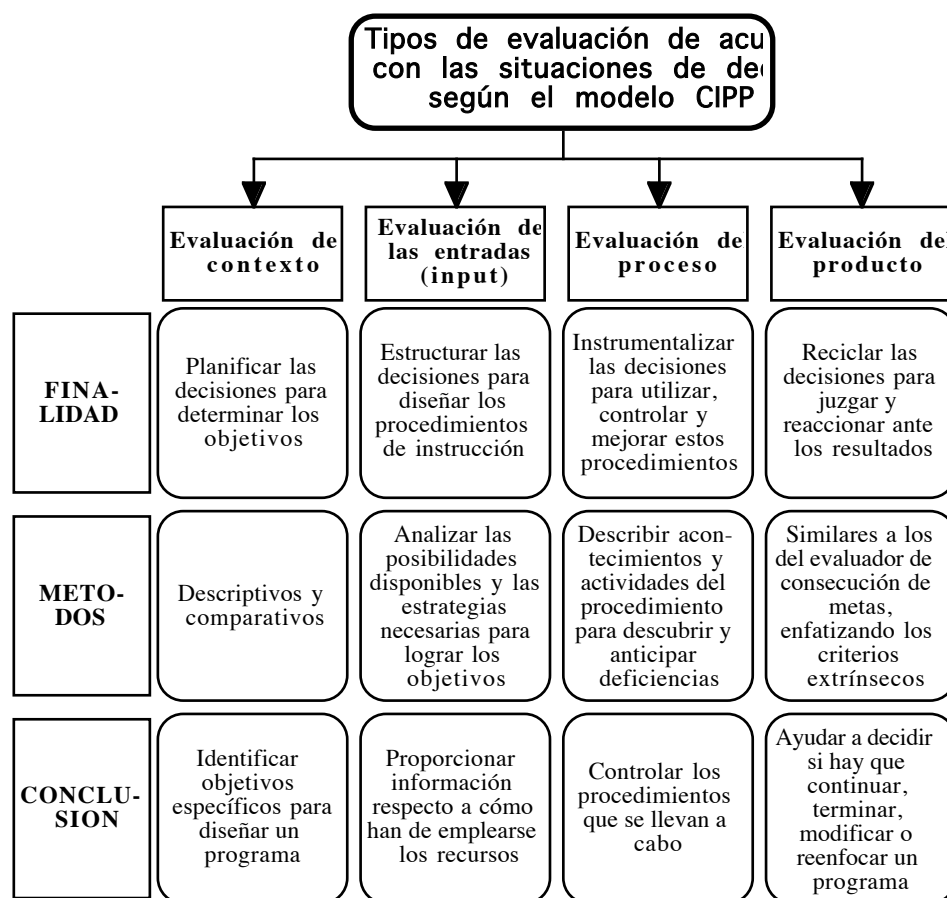


Figura 23.

Stufflebeam relaciona el modelo CIPP con la dicotomía formativa-sumativa propuesta por Scriven. Véase la figura 24. Distingue los que considera dos aspectos esenciales que debe cumplir cualquier programa de evaluación. La evaluación orientada a la facilitación de la información para los que han de tomar decisiones, que adquiere una naturaleza formativa y él denomina «proactiva», y cuando su finalidad es la de rendir

cuentas, en cuyo caso alcanza la característica de sumativa y él llama «retroactiva».

Tipos de evaluación

	Contexto	Input	Proceso	Producto
Toma de decisiones	Evaluación proactiva (formativa)			
Responsabilidad	Evaluación retroactiva (sumativa)			

Figura 24. Extraído de Popham (op. cit., pág. 47).

4.1.3.1.- El modelo CSE

El modelo CSE (Alkin, 1969) debe su nombre al hecho de haberse desarrollado en el Centro para el Estudio de la Evaluación de la Universidad de California, dirigido entonces por Marvin C. Alkin.

Como el modelo CIPP, está igualmente orientado a la facilitación de la toma de decisiones, si bien, la principal diferencia respecto de aquél estriba en su consideración, por igual, tanto de los procesos como de los productos que se van originando durante su desarrollo.

El modelo CSE tiene sentido cuando hay que decidir entre la adopción, mejora o extinción, de un programa educativo que ya está experimentado. Consta de cinco estadios que van desde la determinación de los objetivos necesarios para conseguir el perfeccionamiento del programa, a la evaluación de los resultados finales. Véase la figura 25.

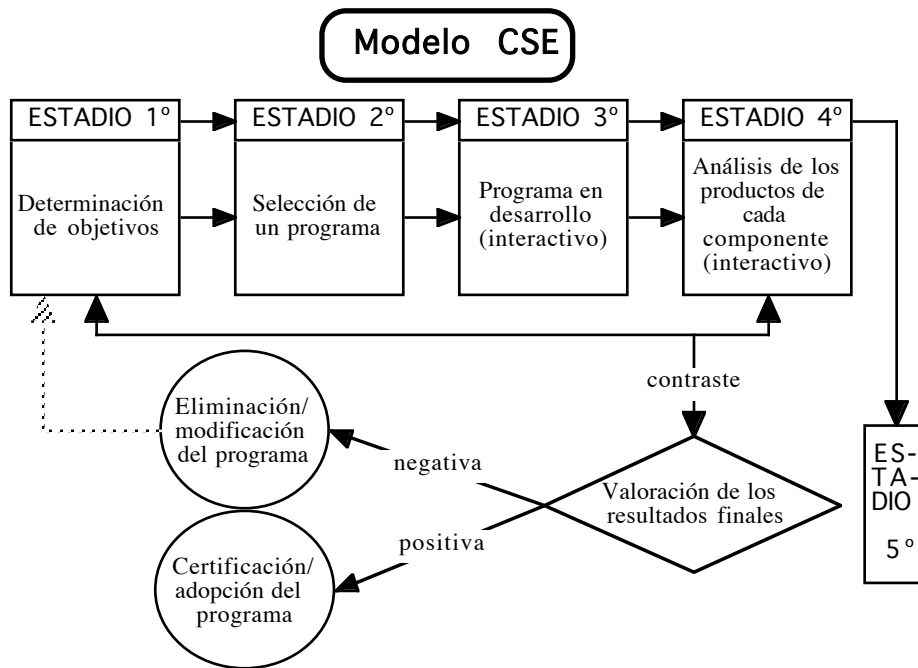


Figura 25.

Para lo cual, en primer lugar, se trata de establecer el nivel real de eficacia del programa y sus posibilidades de mejora. Así, el primer

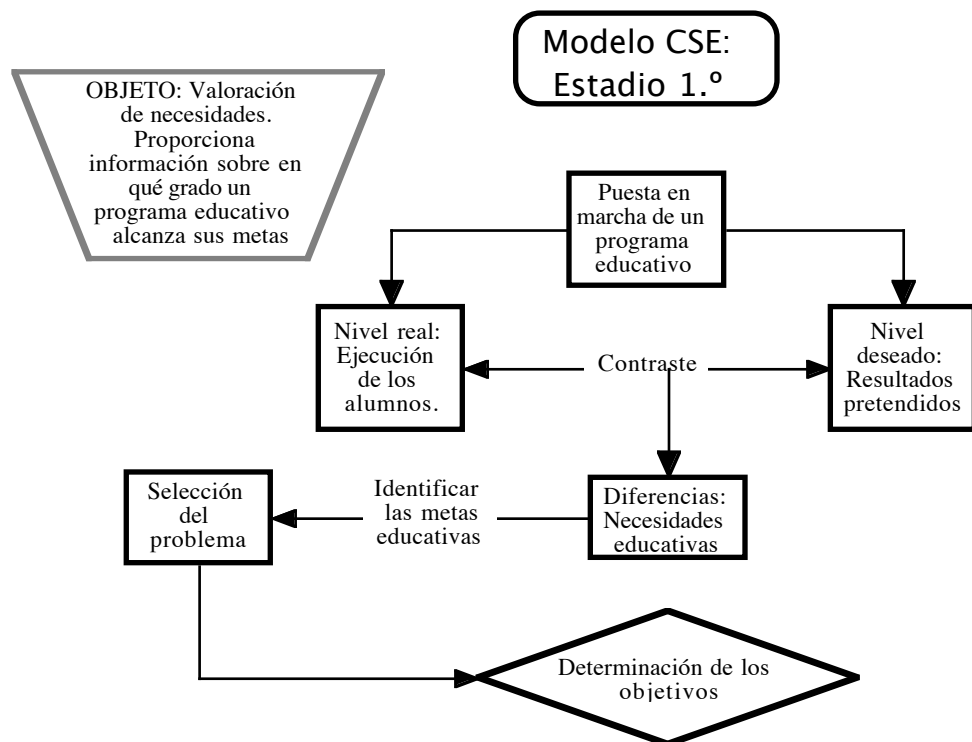


Figura 26.

estadio está enfocado de cara a la determinación de los objetivos que nos propondremos a partir del contraste entre el nivel de ejecución real de los alumnos que siguen el programa educativo y el nivel de ejecución teórico (el deseado o perseguido tras la correcta aplicación de ese mismo programa). Los objetivos obtenidos como resultado de las diferencias entre ambos niveles, real y teórico, representan lo que el modelo llama «necesidades educativas». Véase la figura 26.

En una segunda etapa, se ensayan diferentes «programas de segunda generación», extraídos a partir de la introducción de diversas modificaciones de mejora en el programa original acudiendo, si fuese necesario, a incorporar aspectos concretos de otros programas ajenos al mismo, de acuerdo con las necesidades detectadas en la fase anterior. Se realiza una evaluación comparativa para determinar cuál de todos ellos se adapta mejor a los objetivos planteados inicialmente y, como resultado, se selecciona uno de estos programas. Véase la figura 27.

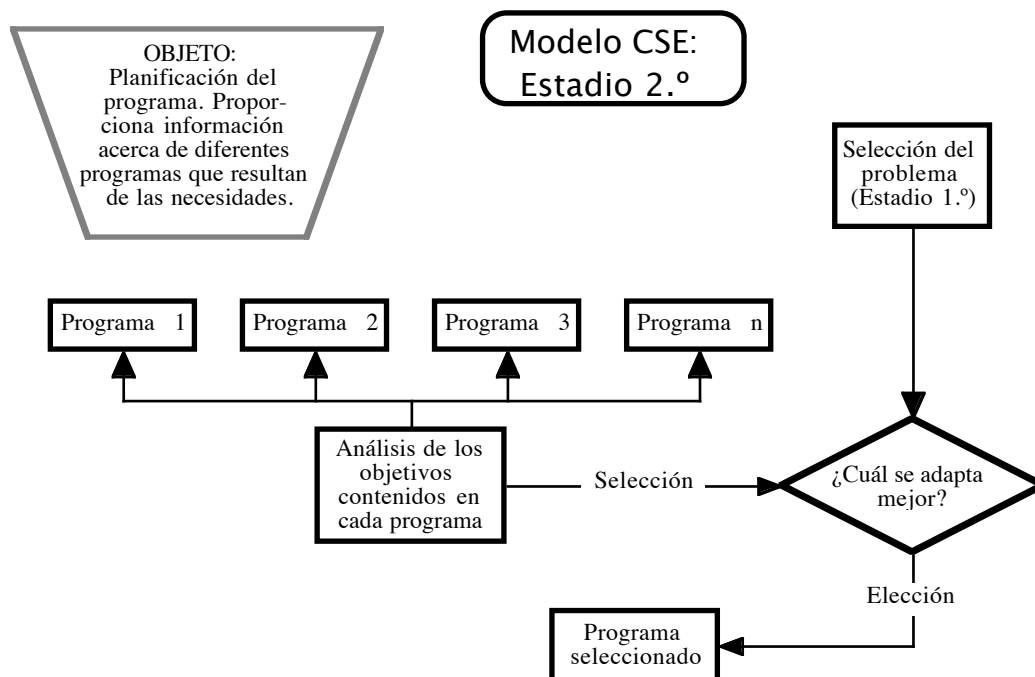


Figura 27.

El tercer estadio trata de detectar, valorar y hacer las modificaciones oportunas para que durante la ejecución del programa seleccionado en la fase anterior se minimicen las posibles diferencias entre lo planificado y su desarrollo real (evaluación de los procesos). Como consecuencia de estos nuevos ajustes, fruto de una continua interacción, se obtendría lo que hemos denominado «plan ajustado». Véase la figura 28.

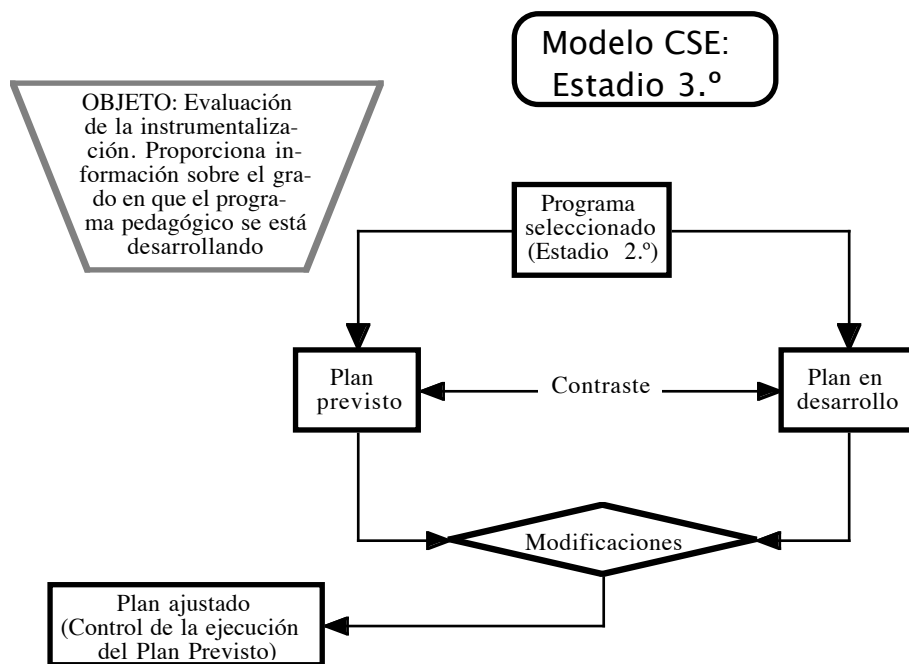


Figura 28.

El estadio 4.º evalúa hasta qué punto se están consiguiendo los objetivos propuestos, no sólo de una manera global, sino también de una manera más específica: en qué grado cada uno de los componentes curriculares está coadyuvando a esa consecución. En el caso de detectarse diferencias en los productos respecto de los esperados, según se vayan produciendo, el evaluador propondrá las modificaciones oportunas del programa para adecuarlo a lo

planificado. De esta manera se obtendría lo que hemos llamado «plan reajustado». Véase la figura 29.

El último estadio trata de evaluar los resultados finales del programa. En el caso de alcanzarse las metas planificadas el programa será adoptado y conservado. Por el contrario, si no se logran los objetivos propuestos y el costo no es excesivo, se pueden establecer

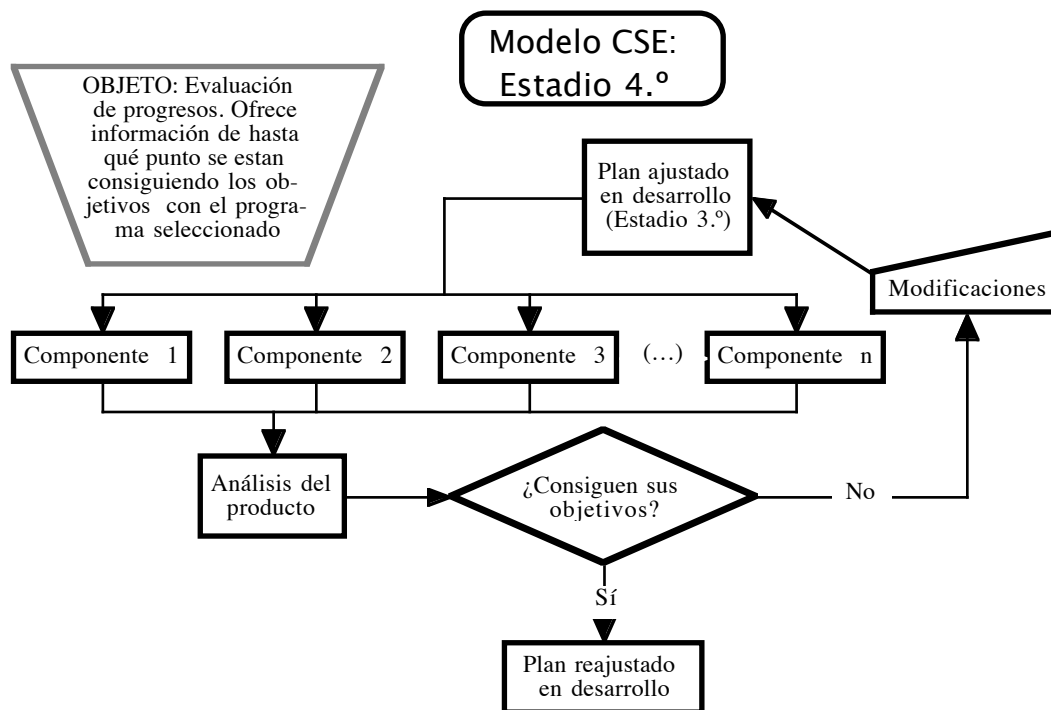


Figura 29.

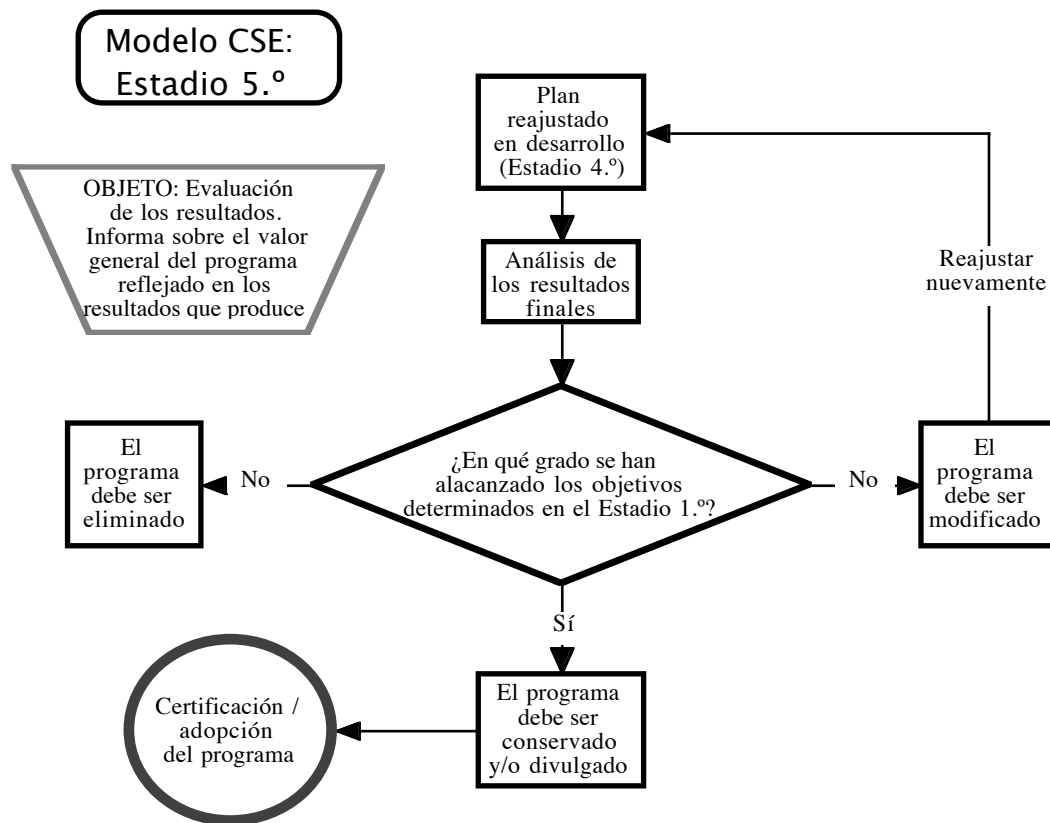


Figura 30.

nuevos reajustes hasta coseguirlos. Cuando esta opción no es viable, el programa se elimina. Véase la figura 30.

4.2.- Modelos alternativos

4.2.1.- Evaluación como crítica artística

Se debe a Eisner (1979, 1985) la visión de la evaluación como un proceso similar al realizado por los expertos cuando enjuician el valor de una obra de arte, especialmente el literario. En nuestro caso sería el propio profesor el experto que juzga los procesos educativos, comprendiéndolos profundamente y como fruto de esta comprensión es capaz de *describirlos, interpretarlos y valorarlos* respetando el desarrollo natural de la enseñanza.

Respecto a la tarea de describir se llevaría a cabo desde dos perspectivas distintas: fáctica, extrayendo datos de la

realidad incluso a través de metodologías positivistas tales que sean capaces de explicar la naturaleza de los hechos, y artística, referida a la forma de expresión de los datos utilizando incluso las metáforas de modo que faciliten su más adecuada comprensión.

La interpretación se realiza a la luz de la teoría y de la práctica, de modo tal que ambas dimensiones se vean mutuamente enriquecidas.

Por último, la valoración, teniendo en consideración la comprensión de las características concretas de la situación en la que se desenvuelven los procesos de enseñanza.

4.2.2.- Evaluación libre de metas

Se debe a la dilatada experiencia como evaluador de Michel Scriven (1973) el concepto de «evaluación libre de metas». Debido a su gran fama como evaluador, formó parte de los principales proyectos de evaluación llevados a cabo en U.S.A. en la década de los sesenta. En estas intervenciones se dio cuenta de cómo los evaluadores se veían condicionados por los objetivos de aquellos programas que pretendían evaluar, sesgando en muchas ocasiones los resultados de sus análisis. Para impedir este hecho y liberarse de las limitaciones impuestas por el conocimiento previo de las metas de los proyectos, propuso que los evaluadores no sólo debían de calibrar el valor de los fines educativos y el de los resultados previstos, sino también aquellos otros productos obtenidos, no pretendidos, pero que pudieran ser incluso más relevantes que los anteriores.

Para el desarrollo de este tipo de evaluación, hemos considerado tres estadios. En el primero de ellos el evaluador tratará de recoger información desconociendo, deliberadamente, los objetivos que se hayan propuesto quienes desarrollan el proyecto, infiriendo los posibles efectos a partir del análisis de los componentes curriculares. En el segundo estadio se trata de contrastar las inferencias realizadas anteriormente a través de la medición de los efectos, bien con

los instrumentos de medida creados por el propio proyecto o bien, creándolos *ex profeso*. Por último, tercer estadio, a través de lo que Scriven llama «artimañas de situación», se tratarán de recoger las evidencias necesarias para corroborar los efectos totales del programa. Véase la figura 31.

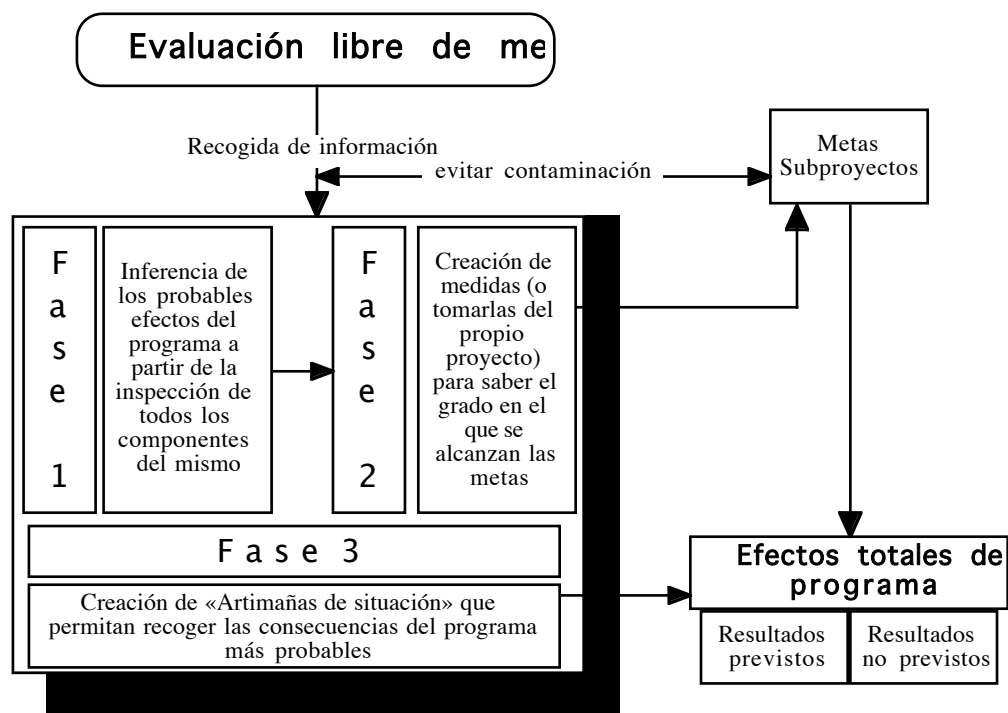


Figura 31.

4.2.3.- Método del *modo operandi*

También se debe a Michel Scriven la creación de este modelo. Para ello se fijó en el modo de investigar la realidad en distintas ramas de la investigación social (antropólogos, historiadores, detectives, etc.). Observó como estos expertos siempre actuaban manteniendo unas pautas comunes. En primer lugar detectaban la existencia de determinados rasgos que hacían concordar, mediante hipótesis de trabajo, con determinados patrones de conducta. Tras recoger las

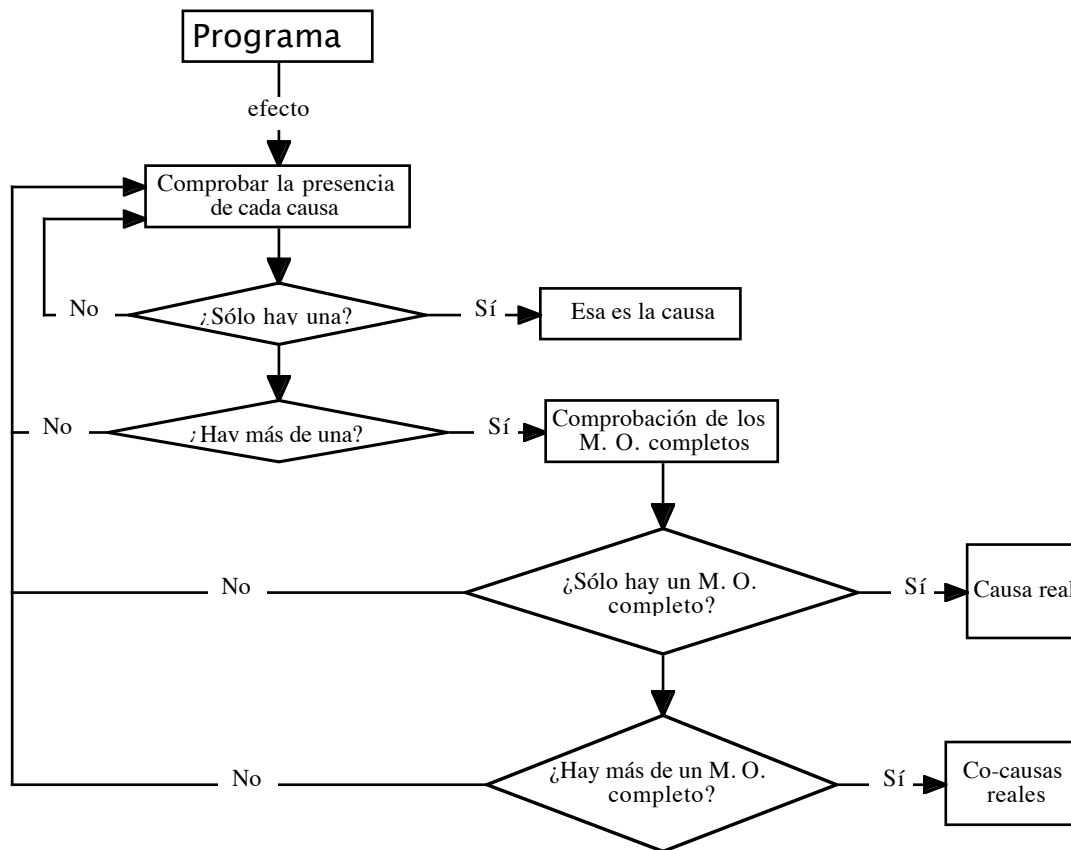


Figura 32.

evidencias necesarias corroboraban, o no, sus hipótesis iniciales. De este modo podían evidenciar determinadas relaciones de causalidad sin tener que acudir a métodos experimentales. En el caso de la evaluación educativa su aplicación serviría para detectar y evaluar las posibles causas que originarían efectos deseables, o perniciosos, en el desarrollo de un programa. Así el evaluador se convertiría en un investigador que trataría de encontrar determinadas cadenas de sucesos características a las que denominó como *modus operandi* (M.O.). La secuencia a seguir sería la indicada en la figura 32.

4.2.4.- Evaluación comunicativa

Nos hemos fijado en este modelo, propuesto por Jesús Rul Gargallo (1994, 1995), dada la frescura que introduce dentro del debate, siempre abierto, acerca de la evaluación. Para este autor, “hay evaluación si a través de la experiencia y de la comunicación se produce aprendizaje valorativo, con repercusiones individuales y colectivas, basado en datos e información objetivada.” (Pág. 70). Considera a la valoración como una dimensión subjetiva sobre un objeto realizada *a priori* y la diferencia de la evaluación en que ésta es una actividad objetiva con una finalidad explicativa realizada *a posteriori*. Pero desde nuestra visión, es la importancia que le ofrece a la comunicación humana, con todo lo que ello conlleva, el aspecto más destacable como diferenciación característica respecto a otros modelos.

Dentro de su concepción de evaluación, sintetiza toda una serie de factores -véase la figura 33- que, sin duda alguna, influyen y la condicionan, como son: la dependencia respecto a la situación en la que se desarrolla, al contexto, la cultura, los valores (axiología), la subjetividad-objetividad, la sistematización en la recogida y tratamiento de la información, la facilitación para poder tomar decisiones y que éstas repercutan en los fines (teleología) de la evaluación, de tal manera que conlleve al perfeccionamiento. Es por tanto, el suyo, un modelo interpretativo y sistemático que se centra en su dimensión objetiva: “evaluar es la atribución de valores intersubjetivos a un objeto como resultado de operaciones subjetivo-objetivas de tratamiento de datos y de información rigurosa en un contexto específico en relación con finalidades de conocimiento, toma de decisiones, optimización y desarrollo humano, individual y colectivo.” (Pág. 73).

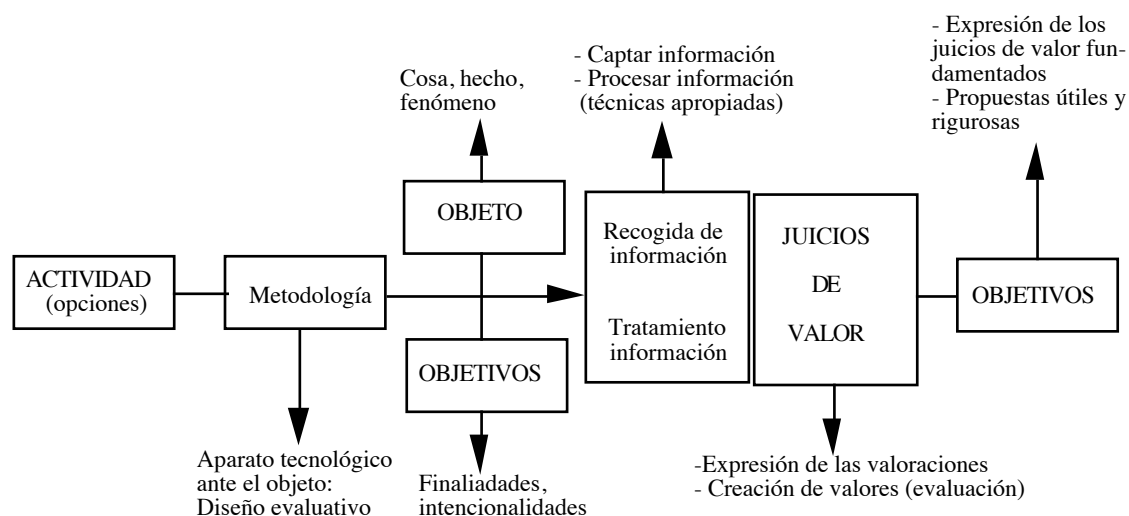


Figura 33: cuadro extraído de Rul Gargallo, 1995, p. 74.

4.2.5.- Evaluación iluminativa

Sus autores, Parlett y Hamilton (1972, 1976), rompen con la tradición empírico-racionalista para trasladar al campo de la evaluación educativa los métodos de trabajo propios de la investigación antropológica. El estudio holístico que proponen debe abarcar desde el análisis pormenorizado del contexto en el que los individuos intervienen, para descubrir la existencia de múltiples realidades que interactúan, asumidas o no por la colectividad, a los resultados, destacando de modo especial los procesos. Precisamente ésta es la causa de su desdén hacia las pruebas psicométricas y de basarse en una metodología cualitativa, observacional, flexible de acuerdo con el momento de la investigación y del contexto -con sus propios condicionantes psicosociales-, negociada con los implicados, centrándose y profundizando progresivamente en aquellos aspectos más significativos de los que en ese «escenario» tienen lugar.

4.2.6.- Evaluación respondiente

Formulada por Stake en 1967¹² (1975, a; 1975, b; 1975, c; 1980) incide en la manera en la que los evaluadores realizan dos tareas básicas: las descriptivas y las relativas a cómo formulan sus juicios. Para ello, propone tres etapas en el desarrollo de la evaluación: los antecedentes (circunstancias anteriores al desarrollo de un programa educativo), las transacciones (las interacciones y modificaciones producidas durante su desenvolvimiento) y los resultados (los productos finales originados tras la aplicación del plan).

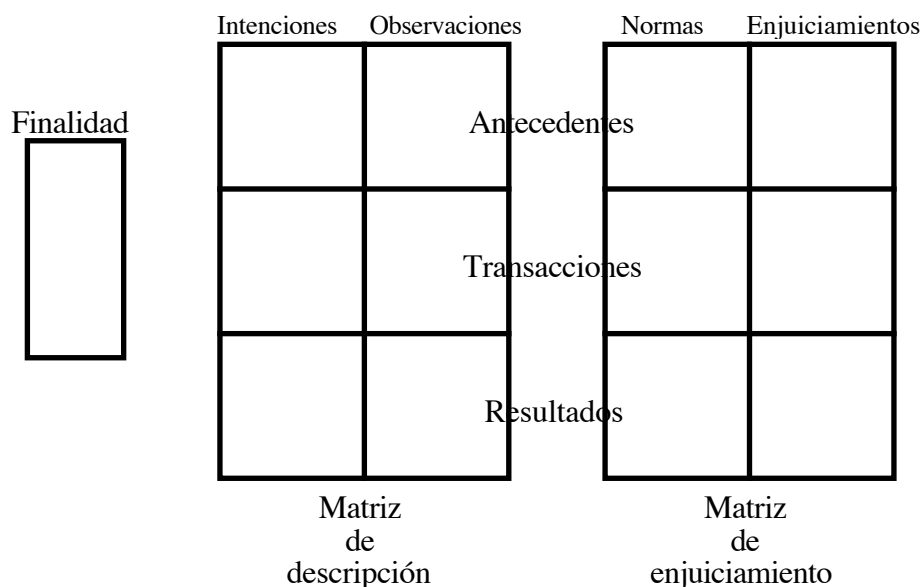


Figura 34: extraída de Popham (op. cit., pág. 41).

Stake combina en un gráfico, que los evaluadores habrán de rellenar, la finalidad educativa perseguida por el proyecto de enseñanza con las tareas descriptivas e interpretativas. Véase la figura 34. Así en la primera columna, matriz de

¹² Fue publicado originalmente en un artículo bajo el título de «El semblante de la evaluación educativa». Por ese motivo también algunos autores lo denominan como «el modelo del semblante». La referencia completa es: "The countenance of educational evaluation". *Teacher's College Record*. N.º: 68. Págs.: 523 a 540. Igualmente se la conoce como el modelo de «réplica» o «respondiente» por cuanto trata de dar respuesta a los intereses y necesidades de todos los implicados en un proceso de evaluación.

descripción, llama la atención la diferenciación que establece entre las intenciones -aquello que se pretende- con respecto a las observaciones -lo que realmente sucede-. Igualmente sucede en la segunda columna, matriz de enjuiciamiento, donde diferencia entre las «normas» en las que los evaluadores se basan para realizar sus juicios de la propia formulación de tales juicios.

Pero tal vez lo más interesante de Stake no sea el cuidadoso aspecto técnico que le confiere a su modelo de evaluación, sino la filosofía que encierra. Stake considera que la evaluación, para que sea tal, ha de responder a los intereses y necesidades de quienes la patrocinan (la administración, generalmente) y de aquellos que están implicados en ella (alumnos, padres, profesores, etc.) de acuerdo con el nivel de conocimientos de cada uno. La manera de «responder» adecuadamente a las distintas, y a veces opuestas necesidades, está en recoger las diferentes visiones de cada sector. Esto le confiere al modelo un avance ético respecto de otros además de, por vez primera, un carácter ciertamente democrático.

Cap. 5.- Principales problemas asociados a la evaluación educativa

A partir de nuestro estudio, nos han surgido multitud de dudas que trataremos de responder con el desarrollo de nuestro trabajo. Para ello, hemos partido de diferentes «cuestiones problemáticas» relativas a la práctica de la evaluación educativa. La principal refiere a la funcionalidad de la evaluación como instrumento reproductor en un subsistema, el escolar, dentro de un sistema más amplio: social, político y económico. De esta cuestión principal derivan las demás.

Prácticamente toda la literatura, como tendremos la oportunidad de ir viendo, reconoce que la escuela es la principal institución encargada de la reproducción social. Desde este punto de vista resulta «natural», como ocurre con los seres vivos, que cualquier elemento trate de reproducirse. Mucho más si se trata de un «sistema» que ha ido evolucionando con el transcurrir del tiempo, y que ha permitido y determinado, a su vez, la evolución humana, tanto individual como colectiva, hasta donde hoy la conocemos.

La vida social no es exclusiva de los seres humanos. Otros animales que nacen y se desarrollan en colectividades estructuradas -sociedades- adquieren, mantienen y defienden unas determinadas pautas de jerarquización y de conducta encaminadas a defender la supremacía y la supervivencia del grupo -como garantía de reproducción de la especie- sobre la de los propios individuos que las conforman. En tanto que la mayoría de esas especies se transmiten la información de una generación a otra, exclusivamente, a través de la vía genética, con sus consecuentes mutaciones y adaptaciones al medio, tales que les permitan subsistir y evolucionar a través de los tiempos; en la especie humana, extremadamente compleja

ontogenética y filogenéticamente, la transmisión de la información se realiza básicamente a través del aprendizaje, quedando la ingente riqueza informativa de los cromosomas relegada a la transmisión de lo puramente biológico.

Estas dos características, estructura social y aprendizaje, nos han permitido ocupar el techo de la evolución de la vida en La Tierra, a gran distancia del conseguido por cualquier otra especie. Pero simultáneamente a su función de desarrollo, también nos condicionan.

La estructura social implica funciones, reparto de las tareas (o del trabajo) y asunción de papeles (del poder, del capital, etc.) que no son repartidos, sino conseguidos y mantenidos. Con su funcionamiento se crea la cultura, la tradición... En determinados momentos de la vida de una sociedad, cuando ésta se hace muy compleja, aparecen unos órganos que facilitan y canalizan la vida del grupo: las instituciones. En principio representan un indudable avance para esa sociedad. Con el discurrir del tiempo también ellas mutan y evolucionan, adaptándose a las nuevas necesidades. Cuando quedan obsoletas, se resisten a desaparecer (Drudis, 1969, págs. 37 y ss.). Crean disfunciones, más o menos graves, y se convierten en rémoras que lejos de su finalidad original, entorpecen el progreso hasta que desaparecen en el tiempo.

Las instituciones, pues, mantienen una vida propia. Crean sus propias exigencias y tratan de hacerse necesarias. En ocasiones pierden, incluso, la referencia a la servidumbre social que las originó y actúan en contra de los intereses de amplios sectores de la población que las mantiene, valiéndose del privilegio que les otorgan el servir al mantenimiento del «orden» social: al poder en cualquiera de sus expresiones.

Desde este punto de vista las organizaciones carecen de ética y del sentido de la justicia. Crean su propia cultura interna y, sistemáticamente, emprenderán un proceso de aculturación -imposición de una cultura sobre otra- en la sociedad en la que están instaladas. En nuestro caso, la institución que se dedica en exclusiva a esta función de aculturizar a los diferentes

grupos de un modo delegado, reconocido, legitimado, sistemático e intencional, es la Escuela. A ella le corresponde el papel -que en otro tiempo tuvieron otras instituciones, como la Iglesia- de «luchar» en contra de otras subculturas para imponer la cultura «oficial» que se expresa en los currícula. Contenidos, actitudes, valores y normas muy específicos, serán la información que ha de transmitir a todos los elementos de la población, para conseguir la pervivencia no sólo de la Sociedad sino de sí misma.

Decíamos antes que junto a la estructura social, como característica que nos había permitido llegar al techo evolutivo que disfrutamos, está el aprendizaje. Los seres humanos hemos creado innumerables instrumentos y herramientas que nos han posibilitado un desarrollo evolutivo único. Acaso el utensilio más eficaz ha sido el lenguaje¹³. Las nuevas generaciones, para incorporarse a la vida social con todas sus implicaciones, han de proveerse del uso extragenético de todos estos instrumentos, así como de su correcto funcionamiento. Eso exige a los nuevos vástagos una sorprendente, y única, capacidad de aprender que los aleja definitivamente de cualquier otra especie animal. Los aprendizajes que deberían hacer, serían puramente «naturales» si fuesen funcionales. Esto es, si los conocimientos que hubieran de adquirir fueran aquellos que el desarrollo de sus vidas en sociedad les va a exigir. Pero nuevamente entramos en contradicción. A los jóvenes no se les pedirá una competencia en nada que no sea la aceptación, y colaboración, con el Sistema social. Es lo que conocemos como *proceso de socialización*.

Aparecen dos problemas asociados con este punto. El primero: no todos los individuos tienen la misma capacidad de aprendizaje «puro»; de, simplemente, aprender. El segundo: los contenidos, actitudes, valores y normas objeto del aprendizaje, frecuentemente, adolecen de falta de funcionalidad

¹³Casi todas estas afirmaciones tienen como punto de partida y fundamento al propio Lev S. Vygotsky (1981, 1990, 1993, 1995); o a alguno de los autores que más profundamente han penetrado y desentrañado su inestimable pensamiento. Así: Wertsch (1981, 1988, 1993); Silvestri y Blanck (1993); Kozulin (1994); etc.

(aplicabilidad de los contenidos hacia la vida) de quien los ha de aprehender y además, usualmente, entran en contradicción con su cultura de origen (actitudes, valores y normas, principalmente).

Es aquí donde se conjugan la estructura social y el aprendizaje. La Escuela y la aculturación; la Escuela y la eficacia (papeles). En definitiva, la orientación, selección y clasificación social. Y todo este proceso, sinérgicamente con otras instituciones, lo consigue la escuela a través, muy especialmente, de la evaluación junto con el resto de los demás componentes curriculares. Es pues, el momento de comenzar a hablar del objeto de nuestro problema antes planteado: la funcionalidad de la evaluación como instrumento reproductor en un subsistema, el escolar, dentro de un sistema más amplio: social, político y económico. Para ello diferenciamos cinco ámbitos o aspectos de la evaluación educativa:

- a) éticos;
- b) político-económicos;
- c) personales;
- d) sociales;
- e) técnicos.

5.1.- Aspectos éticos de la evaluación

La evaluación cumple una doble finalidad. Por un lado sirve como comprobación de estar en posesión de unos determinados criterios. Por otro, como acreditación social, a través de los títulos académicos. Esta acreditación social lleva aparejada el reconocimiento socialmente legitimado para poder vivir por encima de las condiciones en las que lo hacen la mayoría de los demás miembros de la colectividad.

Desde el ángulo de la desigualdad social ante la escuela, la diferente asimilación de un mismo programa presenta pues generalmente un doble aspecto:

-es, por una parte, una desigualdad actual de competencias reales o reconocidas (por el alumno, por sus familiares, por la escuela), desigualdad que puede como tal acarrear ciertas consecuencias independientemente de la carrera escolar ulterior. Así, saber o no saber leer correcta y rápidamente al término de la escolaridad primaria implica toda una serie de consecuencias

inmediatas, por ejemplo en cuanto a la lectura en los ratos de ocio. En la medida en que la escolaridad ulterior no corrija las diferencias, las consecuencias se manifiestan también a más largo plazo, por ejemplo, en la vida de adulto.

-es, por otra parte, una desigualdad *virtual* de competencias uno, cinco o diez años más tarde en la medida en que la desigualdad actual provoque distintos destinos escolares: el que no domine suficientemente la lectura, será orientado hacia especialidades en que la literatura o la lingüística sean componentes marginales de la enseñanza de la lengua materna o de las lenguas extranjeras, donde se aprenda a redactar o descifrar textos simples con fines esencialmente prácticos y durante pocos años. No saber leer bien a los doce años es pues un factor que excluye de un conjunto de enseñanzas que seguirán, por el contrario, los que lean bien. De este modo, la desigualdad del dominio de la lectura contiene en germen muchas otras desigualdades, algunas de las cuales se refieren a conocimientos totalmente ajenos a la lengua, por ejemplo las matemáticas.

(Perrenoud, 1981, pág. 24).

En este mismo sentido habrían de considerarse las obras, publicaciones e investigaciones realizadas por Lerena, 1980; Apple, 1982; Popkewitz, 1987; Monedero, 1987; Fernández Enguita, 1988; Sarup, 1990; Baudelot y Establet, 1990; Giroux, 1990, etc.. De tal modo que se cambia la herencia por consanguinidad (o de padres a hijos), por la herencia, «más justa», expresada por la capacidad y dada por los méritos académicos (que deja de ser estimada como tal «herencia», para convertirse en «conquista» personal del individuo). Se pasa así a la consagración de las diferencias. “Explicar la desigualdad ante la escuela consiste *en primer lugar* en mostrar cómo una definición *particular* de la cultura y de las normas escolares y un funcionamiento *particular* del sistema de enseñanza transforman las diferencias y desigualdades extraescolares de todo tipo en desigualdades reales de aprendizaje o de capital escolar. *El análisis de la fabricación de los juicios de excelencia sólo da cuenta de la última fase de la génesis de las desigualdades, fase en cuyo transcurso se hacen visibles gracias a la evaluación escolar.*” (Perrenoud, 1990, pág. 263).

Estos análisis no evidencian la existencia simplista de una ética universal y única, basada en la igualdad. Antes al contrario, otros autores como Berne y Stiefel (1984) hacen que el tema se complique aún más distinguiendo dentro del concepto-valor de igualdad de oportunidades, entre lo que

ellos consideran como igualdad horizontal (o igual tratamiento entre iguales) y su concepción de igualdad vertical (o apropiado y desigual tratamiento de desiguales).

Respecto de los criterios, cabría preguntarse “si el problema de fondo reside en la propia acreditación o en la naturaleza de los criterios de exigencia con relación a los cuales se acredita o no a los alumnos.” (Coll, 1983, 15). Hemos de tener en consideración, de acuerdo con Perrenoud¹⁴ (1990, 260) que los referidos criterios son creados “*por organizaciones que detentan el poder de evaluar las conductas y de construir una representación legítima que pasa por ser la realidad.*” Otro aspecto radica en si los criterios deben aplicarse a todos los alumnos sin diferenciación, o como defiende Eisner (1993) se deberían adaptar a las peculiaridades de cada sujeto.

A este respecto, algunos autores, en lugar de cuestionarse la validez ética de tales criterios, están más preocupados por encontrar la manera de administrar los mismos con carácter universal¹⁵, mejorando los aspectos puramente técnicos¹⁶: “Para poner a punto la práctica de la evaluación educativa hace falta más y mejor preparación, más presupuesto, más especialistas,

¹⁴Más adelante afirma Perrenoud: “La fabricación de los juicios de excelencia y de las correspondientes jerarquías no conduce nunca, por tanto, a un puro «reflejo», a una simple «fotografía» de las desigualdades reales. Proporciona una representación selectiva, situada, fechada, dramatizada y, así, en parte *arbitraria*. Arbitraria, pero no inexplicable; en cada sistema de enseñanza, la fabricación de las jerarquías escolares tiene una razón de ser. Cumple determinadas funciones en el interior de la organización escolar, a veces, sin que sus agentes se den cuenta.” (Pág. 262)

¹⁵Lafourcade (1977, pág. 32) llama la atención del profesorado, relativizando el valor de la evaluación, para lo cual invita a los docentes a tener en consideración:

- que la evaluación es un medio y no un fin.
- que la evaluación carece de técnicas y procedimientos infalibles.
- que para asegurar categóricamente la efectividad de un proceso de evaluación hay que demostrarlo estadísticamente.

¹⁶En este sentido, hay autores que se refieren a la evaluación como un proceso en el que hay que ser especialmente cuidadosos para que los alumnos no engañen al profesor tratando de aparentar que saben más de lo que realmente conocen. Tal es el caso de Coll, Pozo, Sarabia y Valls (1994, pág. 77) cuando afirman: “Terminaremos este capítulo con un resumen de las cautelas que es necesario guardar durante el proceso de evaluación para evitar que los alumnos «nos den gato por liebre» y nos hagan creer que comprenden lo que apenas tienen memorizado con alfileres.” Con esta filosofía, no resulta extraño encontrar estudios como el de Booth (1993) en el que se expresa la separación existente entre curriculum, alumnos y evaluación como cosas distintas, cuando deberían estar todas ellas integradas.

mejores diseños e instrumentos para la recogida de datos, y un perfeccionamiento de los sistemas institucionales de evaluación. Aún más básica es la necesidad de clarificar las normas a usar por todas las partes implicadas por guiarse y comprobar su propio trabajo evaluador.” (Stufflebeam, 1988, b, pág. 195). Sin tener en consideración que “la evaluación es una actividad esencialmente valorativo-estimativa que requiere, tanto en su enfoque como en su aplicación, rigurosidad ética y crítica social, dada la gran repercusión que una evaluación positiva tiene en la acción formativa y en la imagen de la institución.” (Medina, 1995, pág. 147) Es decir, la evaluación educativa resulta determinante para las vidas de las personas. Razón por la que no sólo concierne a los políticos o evaluadores, en general, la crítica de tales criterios, sino a la totalidad de la sociedad, y de un modo más concreto, a los propios profesores que, a fin de cuentas, son los ejecutores de la misma. Los maestros habrán de asumir su responsabilidad indagando mejores formas de llevarla a cabo. Aunque sobre este aspecto del profesor como investigador, Stenhouse (1981, pág. 168) expresó cierto escepticismo, al no saber si los «evaluadores profesionales» permitirían finalmente a los docentes realizar tal cambio: “Lo que más me interesa aquí es expresar mi temor de que una excesiva atención a las dimensiones políticas de la evaluación y a la presentación de resultados en formas bien patentes y aceptables, pero con frecuencia oblicuas, puede llegar a hacer que la discusión de tales criterios constituya meramente parte de la conversación privada entre los evaluadores. Las cuestiones planteadas son demasiado importantes como para ser consideradas como puramente domésticas dentro de la economía evaluativa. Este reconocimiento de la evaluación como una economía doméstica, con sus límites, me lleva a mi planteamiento de la probabilidad de que la postura asumida por los evaluadores bloquee muy probablemente el progreso de una innovación basada en la investigación.”

Resulta, sin embargo, un tanto «chocante» que todo el mundo esté de acuerdo en la necesidad de evaluar al alumno¹⁷; sin embargo, cuando a quien se puede evaluar es al profesor, sí se levantan voces exigiendo criterios éticos, clarificación de fines, por qué, para qué y para quién se evalúa, etc., como una salvaguarda del derecho de aquellos. Así, Bolívar Botia (1992, pág. 80) afirma a este respecto, refiriéndose a la evaluación de los propios colegios: “Pero como sabemos, una evaluación no es neutral, todo conocimiento sobre el funcionamiento de un centro implica una relación de poder (Bates, 1988; Bates y otros, 1989; Godd, 1988), por lo que una «ética de la evaluación» (House, 1990, b) exige un plano contractual (acuerdo negociado) que respete la autonomía, imparcialidad e igualdad, lejos de instrumentalizar a las personas o instituciones que se evalúan. Además, más allá de la negociación, los profesores tienen derecho a saber para qué va a servir la evaluación y por qué se realiza, pudiendo oponerse o rechazarla; al tiempo que tienen derecho a participar activamente en ella.” Da la sensación de que los únicos que no tienen «ningún» derecho en la evaluación son los alumnos. Incluso como afirma Wheeler (1979, pág. 297), ni en el caso siquiera de concederles la oportunidad de autoevaluarse, dejan de ser supervisados.¹⁸

López (1975)¹⁹, Rosenthal y Jacobson (1980)²⁰, Sarup (1990)²¹, etc., han explicado de diferente manera cómo se

¹⁷Según Fernández Pérez (1986, pág. 57) “el profesor «sabe» que dispone de un instrumento (quizá sería mejor «arma») para hacer valer su autoridad en última instancia, a través del papel delegado de juez que la sociedad, a través del sistema educativo vigente, ha adscrito al profesorado en general. Y esto el profesor lo «sabe» no sólo de forma consciente, sino que puede ser peculiarmente aguda su fuerza en casos en que se escapa a los estratos conscientes del análisis.” Sin embargo esto, no parece ser un buen tema para su discusión.

¹⁸“Es cierto que a través del proceso de concesiones mutuas de los métodos tradicionales del aula se permite a veces que sean los niños los que califiquen su propio trabajo o el de los otros. Pero, incluso entonces, el profesorado comprueba el trabajo. Esta situación se produce a causa de los roles de profesor y alumno. En un aspecto de su rol, al profesor se le considera como un juez; es asunto suyo señalar hasta qué punto ha habido un cambio de conducta, especialmente en el campo cognoscitivo. No se reconoce, sin embargo, que en el rol de alumno pueda ser éste quien valore su propio cambio educativo o el de sus compañeros, aunque ambos tipos de valoración son frecuentes fuera de la escuela.”

¹⁹“Las opiniones de todos los niños reflejan la existencia de una grave contradicción entre su mundo y la escuela, es decir, entre lo que son sus verdaderos

produce la oposición de muchos alumnos a la cultura académica. En general, podemos afirmar que el papel del alumno resulta cada vez más difícil, lejos de lo que se pueda creer en la sociedad.

Desde nuestra perspectiva, la democratización de la evaluación, entendida como participación significativa de los evaluados en la interpretación y uso de los informes, es una necesidad, no sólo ética, sino de rentabilidad política si se quiere asegurar la viabilidad futura de un programa de innovación educativa. De lo contrario, cabe esperar la oposición de los evaluados a poder serlo y a seguir adelante los planes maquilladores de nuevas innovaciones curriculares impuestas por los gobiernos.

5.2.- Aspectos político-económicos de la evaluación

Cuando hablamos de los aspectos político-económicos, estamos haciendo referencia a las medidas legales de la evaluación. Es decir, que hubiésemos podido simplemente haber dicho aspectos legales, entendiendo que la obligada

intereses y los que le propone la institución. [...] aparecen claramente definidas dos actitudes muy distintas.

Las respuestas del grupo no fracasado reflejan una postura de aceptación ante todos los requerimientos de la escuela” Y continúa afirmando que las respuestas del grupo fracasado están divorciadas de los oficiales. Este grupo posee otros valores más auténticos y espontáneos respecto a sus intereses. (Pág. 23).

²⁰“... el adolescente solicitado por otras fuerzas que las de la familia y las de la escuela [...] puede restringir sus ambiciones intelectuales y dejar los logros escolares con la finalidad de contestar al mundo del adulto o de oponerle una contracultura.” (Pág. 48).

²¹“En esencia, Hargreaves afirma que el actual sistema de educación secundaria, principalmente a través del *curriculum* oculto, destruye la dignidad de muchos alumnos hasta el punto de que sólo una pequeña proporción logra recuperarse. Los alumnos reaccionan generando medios alternativos para alcanzar dignidad y *status*. Crean una contracultura. En general, los docentes culpan de las prácticas de oposición en el aula al niño, al hogar o al ambiente social, explicaciones que distraen la atención del papel que la *escuela* puede desempeñar en el desarrollo de dichas prácticas. Eludiendo las críticas sobre sí mismos, los docentes favorecen sus intereses y evitan tener que analizar su propia actuación.

A menudo, la contracultura del alumnado no es sino un intento (vulnerable y de resultados inciertos) por recuperar la propia dignidad. Puede caracterizarse como una oposición creativa a la escuela. Como pone de manifiesto Willis en su trabajo *Aprendiendo a trabajar*, los alumnos varones suelen apoyarse en la cultura de la masculinidad para llevar a cabo una inversión de la distinción intelectual-manual desde la que actúan los profesores. Los *lads* asocian con la feminidad y lo afeminado el trabajo intelectual y todo lo que suene a teórico, es decir, justamente aquello que es necesario para triunfar en el medio escolar.” (Págs. 202 y 203).

adaptación a esas normas jurídicas se puede tomar en dos sentidos diferentes: como garantía de que con su cumplimiento se está velando por la igualdad de todos a través de su generalización (igualdad horizontal); o como salvaguarda de privilegios de los grupos sociales con capacidad para legislar (igualdad vertical).

La pretendida igualdad horizontal que la imposición de un currículum común pretende resulta una falacia, una auténtica quimera. Se trata de una trampa política tras la cual se esconden otros intereses ocultos, pero que desde luego no resuelven el problema, antes al contrario, lo agudizan:

Decir que los alumnos siguen la misma enseñanza o el mismo programa equivale en principio a designar una realidad familiar para todos los que viven en una sociedad fuertemente escolarizada. Basta con imaginar un inicio de curso: en un pasillo, 25 alumnos -ó 18, ó 43- esperan al maestro que, durante todo el año escolar, les tomará a su cargo para hacerles recorrer el programa de un curso, de una disciplina particular o del conjunto de las disciplinas. En el mismo edificio o en otras escuelas de la ciudad o del país, otras clases esperan a un maestro que les enseñará durante un año el mismo programa.

Si se reflexiona un poco sobre ello, uno se da cuenta de que el lenguaje utilizado es terriblemente ambiguo. ¿La misma enseñanza? No totalmente, sin duda, puesto que los grupos-clase se componen de alumnos diferentes confiados a maestros diferentes. Los alumnos difieren por su origen, sus adquisiciones anteriores, sus proyectos, su número, las relaciones que se establecen entre ellos. Los maestros difieren por su formación, su experiencia, su ideología, su concepción pedagógica, su personalidad. Las condiciones materiales y el entorno difieren de una clase a otra. ¿Cómo pretender entonces que todos los alumnos del 3^{er}. grado primario, por ejemplo, reciben la misma enseñanza?

(Perrenoud, 1981, pág. 26).

Respecto a la igualdad vertical, se da incluso con las reformas curriculares de los llamados sistemas comprensivos. Como ejemplo, veamos las indagaciones realizadas por Sarup (pág. 195) para el Reino Unido:

En resumen, considero que el Estado está desarrollando nuevas formas de control. El sistema educativo experimenta actualmente un proceso de reestructuración que provoca que la diferenciación entre grupos de alumnos tenga lugar a una edad temprana. La selección de una élite se está convirtiendo en la función primordial de la escuela. Existen objetivos diferentes según la *clase*. Ante unos pocos se abre la vía hacia las ocupaciones mejor retribuidas y de más prestigio a través de los niveles A; la mayoría queda relegada a la rama técnico-profesional, de la que se está haciendo cargo progresivamente la *Manpower Services Commission*. Este poderoso organismo está financiando el aumento de la oferta de cursos de formación técnica y profesional por parte de

escuelas y autoridades locales (a través, por ejemplo, de la Iniciativa para la Educación Técnica y Profesional). Señalaré, por último, que, en la actualidad, ambas vías, la académica y la técnico-profesional, no se desarrollan en instituciones diferenciadas, pero siguen presentes *en el interior* de las escuelas comprensivas.

Todas estas injerencias gubernamentales en el ámbito de la evaluación educativa la realizan los estados -sea cual sea la ideología en la que se sustenten- con una doble finalidad. De un lado controlar el sistema educativo. De esta manera se vigila y mantiene el orden social establecido. De otro, hacerlo más eficaz. Esto es, minimizar los gastos a la vez que se aumenta su rendimiento, aunque esto se haga a base de establecer la competitividad²² como norma dentro del sistema y no tan sólo para los alumnos, como sucede ahora.

Desde la perspectiva de control del sistema se tratará de evaluar al sistema en sí mismo: programas, centros, profesores, alumnos... Los currícula los confecciona la propia Administración (Sarup, pág. 195). La principal tarea aquí será velar por su estricto cumplimiento (Torrance, 1993). Los centros deberán ser evaluados porque representan fuertes inversiones económicas de las que se habrá de obtener una máxima rentabilidad²³ (eficacia). Los alumnos serán valorados por los profesores con los criterios establecidos en los currícula.²⁴ Los

²²Según Tiana Ferrer (1994, pág. 90) sólo por decisión política se elegirán muestras de centros para ser evaluados y no fomentar la competencia entre los centros y entre los profesores; sin embargo resulta razonable pensar que con la evaluación, por un lado, y, por otro, "con el descenso de la natalidad, la oferta de centros educativos para los alumnos será mayor. De tal modo que dependiendo de los resultados de la evaluación de éstos, los padres podrán preferir a un centro sobre otro. De este modo, gracias a la demografía y a la evaluación, se vislumbra un futuro en el que se pasará de la competencia entre los alumnos a la competencia entre los centros." (Martín Gordillo, 1994, pág. 30)

²³Según Gilly (1978, pág. 20), "para explicar el bajo rendimiento de la escuela primaria, podemos hacer referencia a tres órdenes de razones: razones que dependen de la escuela y de las condiciones pedagógicas, razones que dependen del niño y razones que dependen de su medio familiar."

²⁴El profesorado todavía posee cotas de autonomía tanto en su intervención pedagógica como en la evaluación. Afirma Sancho Gil (1990, pág. 8): "Los profesores, además de realizar valoraciones de «sentido común» tienen un amplio campo, aunque específico, de evaluación educativa, que va desde la elección de libros de texto y los materiales y medios de enseñanza, hasta la calificación de alumnos, que es la tarea de evaluación por excelencia en nuestro Sistema Educativo, que representa una relativa cuota de poder para los profesores y una fuente de controversia permanente." Por su

únicos elementos que podrían «escapar», serían los propios docentes; pero a éstos también les ha llegado su hora para ser «controlados»²⁵ (Bolívar Botia, pág. 80).

Respecto a la eficacia a la que aludíamos, se trata de introducir patrones ajenos a la educación -incluso contrarios-, trasladados de las ciencias y teorías económicas neomercantilistas e industriales (Fernández Pérez, 1986, pág. 51; Castillo Arredondo y Gento Palacios, 1995, pág. 25), al ámbito de la evaluación, como es el caso de los indicadores educativos. En ocasiones se trata de legitimar de algún modo este proceder. Así, no resultan extraños comentarios como el siguiente, en la boca, nada más y nada menos, del exdirector del nuevo Instituto Nacional de Calidad y Evaluación (INCE) refiriéndose a las líneas de actuación que se proponía en este organismo: “En una situación de globalización de la economía y de la educación, cada día está adquiriendo mayor relevancia este tipo de estudios. Todos los sistemas educativos quieren saber qué pasa en su interior, cuál es su situación real, qué pueden ofrecer en comparación con otros.” (Tiana Ferrer, 1994, pág. 90).

Sin embargo, como afirman Serrano et alius (pág. 23) refiriéndose a la evaluación en un tramo educativo intocable hasta hace bien poco: “A nivel universitario se plantea en estos momentos el tema de la evaluación institucional. Sin embargo, este interés reciente por la evaluación está más motivado por la

parte Cerdá, Moreno y Muñoz (1995, pág. 56) refiriéndose precisamente al poder de los profesores en las aulas, afirman: “Como es lógico, en la escuela este proceso está tradicionalmente desequilibrado, pero también existe un cierto contrapoder en el alumnado. Los docentes no ignoran esto y pretenden influir, actualizar su poder sobre los discentes, preferentemente, o sobre otros docentes, pero conocen igualmente que su poder no es absoluto y se encuentra ligado a complejos procesos de negociación y de evitación de conflictos. De hecho la principal preocupación de casi cualquier docente consiste en mantener *el orden* en su aula. [...] De hecho, los excesos de poder pueden ser severamente castigados y la evaluación representa el espacio históricamente reglado en el que se plasman más perentoriamente las relaciones de poder. No es otro el motivo de la corriente confusión entre el todo, la evaluación, y la parte, la calificación.”

²⁵De acuerdo con las manifestaciones de Tiana Ferrer (op. cit., pág. 91), en las que declara que la evaluación que se va a realizar no amenaza ni cuestiona las posiciones individuales, porque se dará publicidad de los resultados de tal centro o profesor. Como si el mero hecho de hablar de la publicidad de los datos, en lugar de su intimidad, no fuera en sí mismo una amenaza.

necesidad de renovar el sistema de financiación, y disminuir una parte del déficit público, que por razones pedagógicas que guíen una mejora de la enseñanza superior.” En consecuencia, podríamos afirmar con Sarup (págs. 194 y 195) que el desarrollo de la evaluación responde a exigencias de rendimientos ajenas a los procesos educativos. Para ello hemos de tener en consideración que “la educación estatal está siendo reestructurada en beneficio del capital. Ello no significa afirmar que Estado y clase dominante sean realidades homogéneas.”

5.3.- Aspectos personales de la evaluación

Las contradicciones son frecuentes en el campo educativo, pero se amplifican en el de la evaluación educativa. Así, hoy en día, bien podríamos hablar de cierta hipocresía burguesa que trata de buscar «chivos expiatorios», según los intereses del momento, sólo en algunos agentes sociales, no en todos. Juana María Sancho (1990), como muchos otros autores, responsabiliza a los docentes de ser los que hasta ahora han mantenido el sistema de la reproducción social²⁶ -¿por qué no a los que detentan el poder económico, a los políticos, a los jueces, a las fuerzas de seguridad, a los medios de comunicación, etc., en fin, al propio Sistema?- y les propone «subirse» al carro de las innovaciones como manera de paliar semejante aberración, como si con ello, efectivamente se eliminara el problema. Sin embargo se encubre el otro motivo, el de la eficacia, disfrazándolo de «justicia» y de «calidad» educativas: “... si los profesores no sólo quieren cumplir con su función reproductora del Sistema, quieren aprender de su propia práctica y, además, quieren aportar elementos de reflexión para la comprensión y mejora de los sistemas de

²⁶La verdad es que no le faltan razones. Alonso, Gil y Martínez (1995) tienen realizadas algunas investigaciones sobre las ideas que sobre evaluación educativa poseen los profesores de Física y que ellos, ahora, tratan de generalizar al ámbito de las Ciencias. Con una ojeada a sus afirmaciones, nos daríamos cuenta de que muchos docentes aún están en la primera mitad del presente siglo en cuanto a concepciones acerca de la evaluación educativa.

enseñanza, la evaluación del curriculum en el centro (con o sin ayuda de un asesor) puede constituir una fuente de reflexión para el profesorado que posibilite el aumento de su criterio profesional y la mejora de su práctica como individuos y como miembros de un equipo docente, a la vez de ayudar a situar las diferentes problemáticas implicadas en el ámbito de la educación institucional allí donde se generan.” (Pág. 11). Sin embargo, esta misma autora reconoce que es el propio Sistema quien origina esta selección-clasificación²⁷ de los individuos: “En la propia esencia del Sistema Educativo, en su faceta más clara de reproducción social, se encuentran la valoración y clasificación de las personas tras exámenes más o menos rigurosos y de muy distintos tipos, lo que les permite acceder o no a determinados ciclos del propio Sistema y a un tipo u otro de trabajo.” (pág. 10).

Esta generosa situación de reparto gratuito de responsabilidades nos parece, cuando menos, falaz; en la que se les pide a los profesores -desde luego de manera muy barata- que mantengan una actitud positiva para la autoformación y continuo reciclaje, en lugar de preguntarse siquiera por qué necesitan tal formación y no responsabilizar a quienes, en su día, los acreditaron socialmente como tales, o a quienes tuvieron parte en la elaboración y desarrollo de los currícula conducentes a la obtención del título que los legitima en su ejercicio profesional. O, por último, a quienes pueden tomar, ahora, las decisiones y financiar una adecuada «formación permanente».

Esta posición, repito, crea cierta desorientación en el profesorado que no encuentra referentes respecto a los que intervenir. Para algunos profesores la evaluación es desagradable; para otros la base de su autoridad. Sus divergencias muestran formas profundamente distintas de

²⁷Para Pérez Juste (1995, pág. 74): “De los resultados de tal tipo de evaluación, las *calificaciones*, se han seguido consecuencias de naturaleza *selectiva* para los estudiantes; con menos frecuencia se han utilizado con carácter *orientador* para su futuro o como elemento *facilitador* de su mejora o progreso presente.”

entender la enseñanza. Sin embargo, estas polémicas se silencian en sus relaciones con el resto de la comunidad educativa. El profesorado trata de aparentar una seguridad sobre los procesos de evaluación que realmente no tiene. (Martín Gordillo, 1994, pág. 23). Pero si agobiante puede ser para los profesores; imaginemos como la pueden vivir la mayoría de los alumnos. Saben que van a ser observados y evaluados²⁸ por sus acciones y omisiones, por sus yerros y aciertos, por sus destrezas y torpezas, etc. y, si pudieran, harían suyas las palabras de Rosenthal y Jacobson (págs. 36 y 37) para describir donde transcurre una parte importante de sus vidas: "... en el campo escolar un mundo cerrado donde la situación vivida es angustiada." No obstante se sigue evaluando entendiendo que viven esta situación de un modo «natural». A fin de cuentas, ya es historia aquel eslogan pedagógico de «la letra con sangre entra». Se ha cambiado el castigo directo, por el «ya verás...» Son muchos los estudios realizados recientemente a este respecto.²⁹

²⁸Rosenthal y Jacobson (1980, pág. 69) describen como se va formando el conocido «efecto Pigmalión», también llamado «efecto Rosenthal» o «efecto halo» (las expectativas del profesor sobre el alumno son como profecías que se cumplen por sí mismas): "... en el transcurso de las interacciones maestro y alumnos se produce furtivamente un juego de confirmación del juicio inicial y de reiteración de las expectativas. De aquí viene una actitud diferencial del enseñante con relación a los alumnos y comportamientos diferenciales según la percepción que tiene de cada uno de ellos." Fernández Pérez (op. cit., pág. 23) citando a Piéron, afirma: "uno de los perturbadores más claros de la objetividad en los exámenes tradicionales: el factor subjetivo del llamado «efecto de halo» (especie de «prejuicio» global, es decir, pre-examen, pre-calificación ya asignada al individuo por parte del examinador en virtud de una actitud personal o juicio previo respecto del examinado)." Cardinet (1988, pág. 94) se refiere a este efecto en los siguientes términos: "De aquí resultan unas variaciones muy grandes de exigencia entre las clases, de las que los profesores no son conscientes. Como distribuyen sus notas de la misma manera en todas partes, independientemente del nivel efectivo de su clase, se ven involuntariamente llevados a sobreestimar a sus mejores alumnos y a menosvalorar a los más flojos." Marzano (1994) en sus hallazgos, ratifica la existencia del «efecto Pigmalión», según el cual los profesores se ven muy influidos por sus expectativas sobre los alumnos en la evaluación. A la misma conclusión llegan Altschuld y colaboradores (1993).

²⁹Destacamos la investigación realizada por Davis y Felkner (1994) sobre los problemas técnicos y personales que puede originar la evaluación: bajo rendimiento, ansiedad, desmotivación, etc. Para Martín Gordillo (op. cit., pág. 23): "Los alumnos perciben la evaluación con preocupación. De hecho tienden a verla como «veredicto» sobre sus capacidades y esfuerzo, con lo que manifiestan ante ella en unos casos miedo (cuando sospechan que tal «veredicto» les va a «condenar») o esperanza (cuando lo reciben como «premio» a sus cualidades). En todo caso suelen percibir como externo el origen de ese veredicto (lo dan los profesores) así como sus consecuencias (la reacción de

En tal observatorio humano³⁰ resulta relativamente fácil establecer diferencias, especialmente cuando nacen de la comparación de unos frente a otros. No importa que no sean relevantes, como afirma Perrenoud (1990, 262): “La evaluación escolar, cuando procede de manera sobre todo comparativa, corre siempre el riesgo de olvidar *todo aquello por lo que los alumnos se asemejan* para marcar diferencias acusadas que, desde un punto de vista antropológico, pueden parecer menores ante el trabajo de uniformidad cultural que lleva a cabo el sistema educativo.” Lo más sutil de todo este asunto -y por qué no, escalofriante- ha sido la introducción explícita de determinados elementos relevantes a tener en consideración en la evaluación educativa de los alumnos, sin que ni tan siquiera puedan ser observables: “Los nuevos currículos educativos han introducido las actitudes y los valores como un contenido de enseñanza junto a los conceptos y los procedimientos. Como tal contenido de enseñanza forman parte de todas las materias, no constituyen una asignatura separada, y van a ser evaluados de forma conjunta con los otros tipos de contenidos. Esa evaluación global se refleja finalmente en una calificación, por lo general un número o un término que lleva asociado un número.” (Sarabia, 1994, pág. 189). Según este mismo autor³¹:

los padres). De ahí que sus dudas respecto de la evaluación se planteen comúnmente sobre la objetividad y justicia con que son calificados.” López (1975, pág. 24) describe la situación en los siguientes términos: “para ser aceptado en la escuela, la primera regla de oro consiste en que el niño entregue a la entrada la capacidad de pensar, a unas personas muy sabias que allí hay y que lo van a hacer por él. De este modo, debe someter sus intereses a lo que los adultos, ajenos a su mundo, organizaron un día y que sus maestros le repiten para que pueda convertirse en «buen alumno». Todo esto se cumple gracias a que existe una completa escala de valoración: puntos, notas, premios, etc. Junto al aprendizaje queda también reglamentada su vida. Las relaciones sociales se anulan en una estructura competitiva. Lo que debe o no debe hacerse lo deciden otros.”

³⁰Pozo (1994, pág. 67), refiriéndose a la evaluación del aprendizaje de hechos y de conceptos, afirma: “El concepto de «evaluación implícita» -por oposición a la evaluación explícita tradicional- hace referencia precisamente a la conveniencia de usar con frecuencia sistemas de evaluación que se hallen integrados en las actividades cotidianas del aula y que, por tanto, no sean percibidas por el alumno como actividades de evaluación.” Con lo cual, si pasa desapercibido el acto de evaluar, el alumno no tendrá «defensas» para salir airoso. La evaluación se puede convertir en una trampa continua.

³¹El mismo manifiesta la falta de validez de la evaluación de las actitudes cuando afirma: “En la escuela las actitudes pueden evaluarse principalmente a partir del comportamiento y atendiendo a las *respuestas de los individuos ante el mensaje persuasivo*; es decir, se trata de evaluar si los alumnos han aprendido las actitudes observando si

“... las actitudes son constructos hipotéticos creados por los psicólogos sociales, por lo que no son directamente observables.” ¿Está preparado el profesorado, en general, para evaluar este aspecto, o se dejará llevar por el «efecto halo»? La verdad, creemos que acaso se evalúen otras cosas que, en su parte positiva, correlacionarán con la pertenencia a determinada clase social y, en su parte negativa, con otras.

5.4.- Aspectos sociales de la evaluación

Como venimos viendo hasta ahora, el Sistema se encarga de «crear» determinados valores que legitiman a unos individuos sobre otros. En palabras de César Coll (pág. 14): “... la naturaleza y el grado de exigencia que fijan los criterios es el fruto de una decisión previa. Quizás sea éste el punto en el que se manifiesta con mayor claridad la naturaleza intrínseca social de la educación y, consecuentemente, de las prácticas evaluativas; la elección de los criterios con relación a los cuales se compara el aprendizaje de los alumnos traduce, inevitablemente, un sistema de valores y, a través de él, unas opciones ideológicas y culturales más o menos coherentes y conscientes. Es una razón de peso para desconfiar de los planteamientos *exclusivamente* técnicos o psicológicos del proceso evaluativo, que descontextualizan la educación, situándola al margen del complejo entramado de factores sociales en el que tiene lugar.”

Aunque en realidad se diga que se valoran determinados aspectos en el ámbito educativo, las evaluaciones escolares resultan relativamente simplistas³², al parecer las únicas medidas reales que se hacen no van más allá de lo que es

manifiestan de hecho los comportamientos que se les pretendía enseñar. [...] Pero puede ocurrir que un comportamiento dado no se corresponda con la actitud que se podría inferir a partir de él.” (pág. 191)

³²Shoemaker y Lewin (1993) han creado unas estrategias tendentes a facilitar a los profesores la evaluación de los contenidos del curriculum. Contenidos y evaluación resultan inseparables. Ambos aspectos los describen con la metáfora de las dos caras de una misma moneda. Del mismo modo, Stenhouse (op. cit, pág. 171) mostraba su rechazo a la separación entre el desarrollo curricular y la evaluación: “La evaluación debería, (...), guiar al desarrollo curricular e integrarse con él. Así quedaría destruida la distinción conceptual entre desarrollo y evaluación, manifestándose ambas como investigación.”

relativamente fácil de cuantificar: "... en la práctica, lo que realmente parece importar a unos y a otros no es sino ese número o palabra mágicos colocados frente a la casilla de la matemática, el inglés o la historia. ¿Qué alumno es obligado a «repetir curso» por insociable, por incapaz de un pensamiento independiente o por reiterado quebrantador de la sinceridad? A la hora de la verdad, lo que nuestro sistema educativo mide, por desgracia, no va más allá de una memorización, más o menos verbal y habilidosamente expresada, de viejas informaciones." (Fernández Pérez, pág. 61).

También Apple y Jungck (1990, pág. 155) expresan esta misma opinión de la que responsabilizan a la sociedad, en general: "En la medida en que el control sobre los contenidos, la enseñanza y la evaluación sale de la órbita del aula, se hace cada vez más hincapié sólo en aquellos elementos de los estudios sociales, la lectura, la ciencia, etc. que pueden medirse fácilmente mediante pruebas estándar. Los focos primarios de interés son el saber «qué» y, ocasionalmente, el saber «cómo» de bajo nivel. El resto es considerado cada vez más como intranscendente. Es malo, desde luego, lo que sucede, pero aún más lo es el hecho de que hasta el saber «qué» que se enseña se vuelve «más seguro», menos controvertido, menos crítico."

El «plato fuerte» de la evaluación lo conforman otros aspectos que se valoran de manera acrítica, oculta e informal³³, pero que sirven a la finalidad selectiva del Sistema. Martín Gordillo (pág. 26) refiriéndose a este mismo tema, afirma: "Así el darwinismo social (también pretendidamente «científico» desde premisas positivistas) tiene en el sistema educativo, especialmente a través de la evaluación, su instrumento principal de actuación. La escuela adopta la función de filtro de los individuos para la ulterior incorporación al proceso productivo."

³³Airasian y Jones (1993) en su artículo titulado "The teacher as Applied Measurer: Realities of Classroom Measurement and Assessment", señalan cómo la valoración informal es utilizada por los profesores para informar de las decisiones diarias de la clase y cómo estas valoraciones informales influyen en otros juicios y medidas formales.

Para Rul Gargallo (1995, pág. 70) en la actualidad, la evaluación se realiza de una manera bastante pobre: “No obstante, acercarse a la prácticas evaluativas en el mundo de la educación es descubrir una realidad escindida; por una parte encontramos mediciones (denominadas evaluaciones) subsidiarias de una concepción evaluativa centrada en los «resultados» con las limitaciones del enfoque e interpretación positivistas, y por otra parte, prácticas valorativas centradas en los procesos que no van más allá de expresar visiones puramente subjetivas de lo que acontece.”

Sancho Gil, en su obra ya citada, también se refiere a este aspecto del Sistema: “... el tema de la evaluación es extremadamente complejo, no sólo porque existen diferentes formas de entenderla y abordarla, sino y sobre todo, porque una de las funciones del Sistema Educativo ha sido decidir qué individuos eran susceptibles de seguir una determinada vía de formación y esto ha propiciado una serie de concepciones sobre el proceso de enseñanza y ha generado unas determinadas formas de conceptualizar sus resultados.” (pág. 11)

Para algunos autores como Perrenoud (1981, pág. 48), es la evaluación la que sin ningún género de dudas no sólo detecta y mantiene las desigualdades sociales, sino que las crea:

¿Es la evaluación, en tanto que tal, responsable de la desigualdad ante la misma enseñanza? Para resumir diré:

1.- Que la evaluación no es simplemente reveladora de la desigualdad, sino que participa en su génesis.

2.- Que lo hace directamente cuando subestima las competencias reconocidas de los alumnos de las clases desfavorecidas o sobreestima las competencias de los alumnos de las clases privilegiadas (desigualdad ante la evaluación en sentido estricto; ante la evaluación como instrumento de medida, en principio equitativo de las adquisiciones reales de los alumnos.

3.- Que a título de componente importante, omnipresente en el sistema de enseñanza tomado en su conjunto y en la organización pedagógica en el marco de una enseñanza colectiva, la evaluación participa en la génesis de las desigualdades de aprendizaje y éxito.

Los alumnos que no tienen éxito³⁴, obviamente son los fracasados. El fracaso escolar se ha estudiado desde muy

³⁴Para Stenhouse (op. cit, pág. 164) “El mejor modo de asegurar el «éxito» es mantener bajo el nivel de aspiración. Cualquier currículo ambicioso producirá éxitos y

diversas perspectivas. Así las investigaciones y estudios realizados por Alvarez y colaboradores, 1979; Carabaña, 1983; Hernández Ruíz y Gómez Dacal, 1982; López, 1975; Lurçat, 1979; Martínez Muñiz, 1982; Pérez Serrano, 1981; Rodríguez Espinar, 1982; Samper, Samper y Soler, 1982; Monedero, 1996; etc.

Son muchos los autores que hacen un paralelismo entre el fracaso escolar y el social. Para Julio Carabaña³⁵ (pág. 29) las causas que lo producen no son el origen social³⁶, la socialización familiar, las diversas cualidades del sujeto, o determinadas prácticas pedagógicas, "... esto es apariencia, pues se sabe que las causas se encuentran propiamente más allá de lo que los individuos hacen, en una realidad que los

fracasos en toda institución que se lleve a cabo." Para Gilly (op. cit., pág. 19) "El éxito escolar del niño depende en mucho de cómo la escuela modele la personalidad del alumno." Es decir, de en qué grado esté dispuesto a socializarse; esto es, a expresar abierta y claramente la asunción del modelo humano preconizado por la escuela.

³⁵El fracaso escolar está implícito en la normativa legal: "Pues tal y como está planteada la EGB desde la Ley, la introducción de la competitividad y la selección en ella es una tendencia inevitable. Al final de la EGB se premia a los adelantados con el paso al BUP, y se castiga a los retrasados limitándolos a la FP; pero; durante la EGB, se pretende que todos actúen como si no lo supieran. Aquí se trata sólo de señalar a los alumnos el camino más adecuado a sus capacidades, un camino, por así decirlo, que le es connatural. ¿Iba nadie a poner interés en forzar su naturaleza, estropeando la armonía entre el sistema escolar y sus capacidades? Pues resulta que sí; todos actúan como si, efectivamente, hubiera una salida mejor y otra peor, y los objetivos de la pedagogía se orientan a conseguir que el alumno domine los contenidos necesarios para poder pasar al BUP. Lo contrario es, precisamente, el «fracaso escolar»." Para este autor, "Hablar de fracaso escolar puede llevar a la aceptación *acrítica*, implícita, casi inconsciente, de la legitimidad de la competitividad y la selectividad en la EGB." Pág. 28.

³⁶La introducción del contexto en la evaluación es una exigencia, no sólo ética, sino también técnica. Wiggins (1993) critica el sistema de evaluación centrado en exámenes. Para este autor, las dos claves asumidas del diseño tradicional de los exámenes, la separación en categorías de los conocimientos y la descontextualización de lo que se sabe, son falsas. Una interpretación competente de los datos requiere tanto del contexto como del juicio. No tiene sentido examinar para comprobar los conocimientos como si fueran una respuesta invariable a un estímulo ambiguo. Zazzo (1978, pág. 13) afirma que "no es necesario ser un psicólogo eminente para admitir que los factores que intervienen en primer lugar en el éxito escolar son: la inteligencia, las condiciones pedagógicas y el medio sociocultural en el que vive el niño." Más adelante, continua: "La indigencia sociocultural del medio familiar frena el desarrollo intelectual del niño y su progreso escolar. Es el factor que, ante todo, debe retener nuestra atención, porque es el más gravoso de todos y porque prefigura y determina la desigualdad social (en la enseñanza secundaria, los niños nacidos en medios populares tienen una media de más de un año de retraso sobre sus compañeros de otros medios) y también porque podemos, debemos actuar en él. No es una fatalidad." Worthen (1993, b) propone formas críticas y alternativas de evaluar, que introduzcan la valoración del contexto.

controla, en una facticidad social que los determina, en último término, en la «sociedad» o en el «sistema».”

Para López (pág. 18), la psicología tiene también su parte de culpa con el pretexto de la ayuda que le brinda a los niños fracasados, «clasificándolos» según qué deficiencia. Para este autor: “El fracaso escolar es la primera situación en que el individuo se ve marginado de su vida social. Esta marginación es consecuencia de la discriminación ejercida sobre aquellos niños que no se ajustan al modelo que la sociedad impone, cuando lo considera apropiado. La constitución de grupos según el rendimiento escolar se ha justificado alegando la existencia de dificultades personales en aquellos individuos que no siguen los caminos, únicos y uniformes para todos, establecidos por la escuela.”

En definitiva, podemos afirmar que los resultados de la evaluación educativa son trasladados y aceptados por la sociedad, que los utiliza para legitimar el *status quo* de cada individuo.³⁷

5.5.- Aspectos técnicos de la evaluación

Después de cuanto llevamos dicho no resulta inconveniente afirmar que la evaluación, sirviendo a intereses a menudo ajenos al mundo educativo, se ha convertido en uno de los componentes del currículum más relevantes en la actualidad. Como afirma Martín Gordillo (pág. 28) con referencia a la LOGSE: “vuelve a quedar patente la preeminencia de la evaluación como elemento central en el currículo en el nuevo sistema educativo.” Los modelos técnicos de evaluación educativa se han ido desarrollando y perfeccionando tanto, especialmente en los Estados Unidos,

³⁷Farré y Gol (1982, pág. 12) se refieren a esta aceptación social de los resultados de la evaluación educativa, en los siguientes términos: “La nota, socialmente, adquiere un valor absoluto para medir el valor del niño como persona. El niño que obtiene «buenas notas» es alabado; el que las tiene «malas» incluso es menospreciado. A menudo, en torno a la «nota», padres y maestros se asocian frente al niño que no responde a las exigencias escolares.”

que hoy podemos hablar de una teoría de la evaluación con naturaleza propia, al margen de los demás ámbitos de la educación. Pero cabría hacer un alto en este punto para reflexionar y preguntarnos en nombre de qué ha avanzado tanto esta «tecnología evaluativa». Evidentemente la respuesta suele ser la misma: un servicio social. Pero tal vez lo que deberíamos de comenzar a hacer antes que seguir perfeccionando y desarrollando nuevos modelos, sea “humanizar la técnica para poder actuar más «con» las comunidades que «por» las comunidades.” (Alfaro, 1990, pág. 66). Creemos, como afirma esta autora, que con la evaluación hemos conseguido, incluso, desfavorecer a la mayoría de los grupos humanos a los que declaramos querer ayudar: “Si alguna vez caemos en la equivocación de querer programar lo que no es programable y ponemos las personas al servicio de la técnica concreta de programar, es decir, acomodar las personas a nuestros esquemas teórico-técnicos y no al revés, la técnica al servicio de las personas, no estaremos más que reproduciendo lo que ocurre en tantos otros ámbitos de lo humano: la técnica en contra de los intereses de las comunidades.”³⁸ Pensamos, con Perrenoud (1990, pág. 265), que todos los implicados en la educación somos corresponsables de esta lamentable situación: “La escuela está muy comprometida en la génesis de las conductas que somete a las normas de excelencia, porque, antes de evaluar, enseña.”

Habrá quien argumente en contra, aduciendo que sin la mencionada calidad técnica de la evaluación, se perderá toda referencia. No se tendrán orientaciones precisas acerca del

³⁸Como afirma Perrenoud (1981, pág. 48): “No olvidemos que las diferencias culturales se transforman en factores de desigualdad social únicamente en función de determinados objetivos pedagógicos.” Sin embargo, este autor (1990, pág. 36) reconoce que los profesores en la evaluación no tienen presentes los objetivos: “Cuando evalúan, los maestros no se refieren a objetivos muy detallados; convierten de forma bastante intuitiva lo que han enseñado en pruebas, cuestiones para controles, pruebas escritas o preguntas orales. A menudo se trata de una rutina, tanto más económica cuanto que no obliga a reflexionar en lo que se está haciendo. El docente experimentado dispone, en su cabeza o en su archivo, de un repertorio de cuestiones o de ejercicios para escoger con gran celeridad.” Por su parte, Yen (1993) tiene realizado un trabajo de investigación sobre la independencia de los objetivos y sus repercusiones en la evaluación.

rendimiento y los «niveles» educativos bajarán inexorablemente. Resulta curiosa la aceptación unánime de la palabra «nivel», como explican Baudelot y Establet (1990) en su obra *El nivel educativo sube. Refutación de una antigua idea sobre la pretendida decadencia de nuestras escuelas*. En estos dos últimos siglos, ha sido utilizada socialmente siempre con fines, cuando menos, bastante conservadores.³⁹ Siguiendo a estos autores, “la metáfora de la caída y la connotación desdeñosa gravitan desde hace dos siglos sobre la palabra. Para mantener es preciso medir y regularizar; la preocupación por el nivel toma cuerpo en la organización de un sistema burocrático de examen con el que acaba por confundirse.” (Pág. 141). Sin embargo, no se ha experimentado un aumento de la validez⁴⁰ de estas pruebas cotidianas de evaluación en cuanto a la no discriminación de los individuos. A veces no tanto porque no exista una técnica, como por la falta de su correcta aplicación o por puro desconocimiento debido a una formación insuficiente. (Pérez

³⁹Uno de los múltiples ejemplos que ponen estos autores (pág. 134) sobre el empleo del uso del «nivel», es el análisis de las estadísticas del paro laboral que afectan principalmente a los jóvenes. “Afirmar que el nivel baja es encontrar un fundamento de derecho al orden de cosas; la justificación por el mérito llega muy a punto para fortalecer el poder de los viejos.”

⁴⁰Curiosamente es la sociedad, como ya hemos visto, la que confiere una «validez» a las evaluaciones de las que carecen. En este sentido Fernández Pérez (op. cit., pág. 46) asegura: “Para la sociedad se ha inaugurado la garantía de que el sistema educativo está evaluando lo que dice evaluar, es decir, que cuando la cartilla escolar o la libreta de calificaciones dice estar midiendo el civismo, la adaptación social, el interés o aplicación, la conducta o los conocimientos (idioma moderno, matemáticas, etc.), son realmente esos aspectos y no otros los que se miden.” Además este autor, refiriéndose a la falta de validez de las calificaciones asegura (págs. 28 a 30) que tampoco dos o más correctores se ponen de acuerdo en la calificación dada a un mismo examen. Por su parte, recientemente, Ruíz Primo (1993) ha estudiado el tema de la falta de «estabilidad» de las calificaciones. Este mismo tema es tratado igualmente por Serrano y colaboradores (1995, pág. 3): “La evaluación de alumnos es algo que todos los profesores tenemos que hacer de una forma o de otra. En los casos extremos de notas, o calificaciones, muy buenas o muy malas, posiblemente coincida más de un evaluador pero, por lo general, si un alumno es evaluado de la misma materia por profesores distintos seguramente obtendría calificaciones cuantitativamente distintas. Incluso dos evaluaciones distintas por el mismo profesor podrían dar resultados diferentes. ¿El alumno sabe más cuando lo evalúa un profesor u otro? La respuesta es obvia, las diferencias las produce el instrumento de evaluación o el propio evaluador. Es difícil determinar una solución al problema, pero hoy por hoy es lo que tenemos.” Sin embargo, Lafourcade (op. cit., pág. 41) afirma: “Que las mediciones que se obtengan de ciertas formas de conducta no tengan una fiabilidad perfecta, no significa que tales valores registrados carezcan de significación y que sobre ellos no se puedan realizar predicciones bastante posibles dentro de ciertos márgenes.”

Juste, 1995).⁴¹ Pero también, en la mayoría de las ocasiones, por la propia naturaleza de la población escolar a la que se evalúa.⁴²

Creo que habríamos de comenzar un proceso valiente, sin duda, mediante el que se trataran de transmitir nuevos valores a la educación, tales que nos permitieran desenmascarar todos los intereses no declarados, pero decisivamente influyentes, en la evaluación. Tal vez deberíamos ser un poco iconoclastas con la figura de los «evaluadores profesionales», y tratar de cambiar la formación⁴³ y el pensamiento de los profesores. Como propuso Stenhouse (pág. 170) cuando afirmaba que: “la existencia del evaluador implica la existencia del que desarrolla el currículo, otro papel sobre el cual me siento escéptico. Deseo argumentar contra la separación entre ambos papeles y en favor de una investigación integrada en el currículo.” El encargado de esa investigación es el propio profesor. Será a él, en cuanto profesional directamente involucrado en las tareas

⁴¹Según este autor, pág. 73, “los profesores consumen una parte significativa de su dedicación profesional en la evaluación, si bien, con frecuencia, no aprovechan el conocimiento técnico elaborado y los resultados de la investigación a la hora de diseñar su sistema evaluativo.”

⁴²Perrenoud (1981, pág. 28) afirma: “Si estos grupos no están compuestos por igual proporción de alumnos de las diferentes clases sociales, sino que, por el contrario, unos acogen en su mayoría a alumnos de clases populares y los otros a alumnos de clases burguesas, el sistema de evaluación habrá incrementado la desigualdad social ante el mismo curriculum formal.”

⁴³Schafer (1993) en su artículo “Assessment Literacy for Teachers”, considera la necesidad de dotar de un mayor *estatus* cultural a los profesores. El artículo examina si los profesores necesitan o no asesoramiento, si hay cursillos de formación, qué cursos se deben realizar, qué temas deben cubrir y, sobre todo, qué está pasando con las calificaciones en los colegios y cuáles son sus principales necesidades formativas. Otra experiencia realizada con este mismo motivo de perfeccionamiento del profesorado es la relatada por Jametz (1994) sobre la creación en 1990 del California Assessment Collaborative. Por su parte, Impara y colaboradores (1993) han realizado una investigación en la que señalan la falta de una formación adecuada de los profesores para poner «calificaciones» a sus alumnos, además de dar a conocer las opiniones que los profesores tienen acerca de la funcionalidad que los estudiantes otorgan a esas calificaciones, con la finalidad de mejorar tanto los procesos de enseñanza-aprendizaje, como los de la evaluación. Baker (1994) también ha estudiado los elementos básicos de lo que debería ser un programa de formación del profesorado en evaluación. Para Martín Gordillo (op. cit., pág. 32), con la aplicación de la LOGSE, los docentes deberán estar formados suficientemente, pues “los instrumentos de evaluación que los centros, las áreas y los profesores han de diseñar serán ahora necesariamente plurales. Registros de observación en el aula, revisión de trabajos y cuadernos, pruebas, protocolos para la autoevaluación... serán, entre otros muchos, los instrumentos que, en cada situación, pueden servir para la evaluación de los aprendizajes de los alumnos y, con ellos, de los diseños curriculares.”

evaluativas, a quien corresponda «poner las cosas en su sitio» no dejándose utilizar por pretendidas teorías solucionadoras, a través de su propia reflexión y formación. Como afirma Ballester (1995, pág. 48): “Nosotros mismos no podemos evitar que cualquier intento por explicar para predecir, y predecir para explicar, en un terreno como éste, nos suene cuando menos a sospechoso y nos recorra un escalofrío al pensar en el éxito de una tal tarea, aunque se haga inicialmente con la mejor de las pretensiones. De ahí el recelo con el que acogemos, dentro ya de nuestra profesión, las soluciones caídas del cielo en forma de teorías psicopedagógicas que llegan con el marchamo de ser lo último y definitivo, lo mejor, lo que ahora está de moda, lo que se vende y lo que se lleva.”

En síntesis, concluimos el presente capítulo afirmando, de acuerdo con nuestros análisis anteriores, que la evaluación se introduce en los sistemas educativos como respuesta a exigencias sociopolíticas y económicas; pero no pedagógicas.

Esta afirmación nos lleva a plantearnos la siguiente cuestión: ¿son los profesores conscientes de su papel social cuando evalúan a los alumnos?

Si la respuesta es positiva, consideramos que los docentes tratarán de «innovar para la liberación»; la evaluación al servicio del evaluado. En caso de ser negativa, lo harán sólo para la «reproducción».

A partir de aquí, se nos plantea otra duda: ¿coincide la Reforma de la enseñanza -LOGSE- con esta idea de «escuela innovadora para la liberación»?

Podremos responder afirmativamente sólo cuando -desde presupuestos éticos, legales, personales, sociales y técnicos- podamos afirmar que la evaluación educativa no se utiliza con fines seleccionadores y clasificatorios. Consecuentemente habrá diferencias significativas respecto a la escuela tradicional ordinaria.

Al contrario, la respuesta será negativa si no existen diferencias relevantes entre la finalidad y la forma en que es utilizada la evaluación -desde los mismos presupuestos

anteriores- en la escuela ordinaria (anterior a la LOGSE) y la de la Reforma.

Referencias bibliográficas

- ADAMS, G. S. (1970): *Medición y evaluación en educación, psicología y "guidance"*. Herder, Barcelona.
- AIRASIAN, P. W. y JONES, A. M. (1993): "The teacher as Applied Measurer: Realities of Classroom Measurement and Assessment." *Applied Measurement in Education*. Vol.: 6. N.º: 3. Págs.: 241 a 254.
- ALFARO, M. E. (1990): "Aspectos prácticos del proceso de programación y evaluación." *Documentación Social*. N.º: 81. Págs.: 65 a 80.
- ALKIN, M. C. (1969): "Evaluation Theory Development." *Evaluation Comment*. Vol.: 2. N.º: 1. Págs.: 2 a 7.
- ALKIN, M. C. (1980): "Evaluación del currículo y la instrucción". *Temas de Educación*. Facultad de Humanidades y Educación. U.C.V. Caracas.
- ALKIN, M. C. (1990): *Debates on Evaluation*. SAGE Publications. Beverly Hills, California. U.S.A.
- ALONSO, M.; GIL, D. y MARTINEZ, J. (1992): "Los exámenes en la enseñanza por transmisión y en la enseñanza por investigación." *Enseñanza de las Ciencias*. Vol.: 10. N.º: 2. Págs.: 127 a 138.
- ALONSO, M.; GIL, D. y MARTINEZ, J. (1995): "Concepciones docentes sobre la evaluación en la enseñanza de las ciencias". Monografía: La evaluación de los aprendizajes. *Alambique, Didáctica de las Ciencias Experimentales*, Abril. N.º: 4, Págs. 6 a 15.
- ALTSCHULD, J. W. et alius (1993): "The Utilization of Needs Assessments Results." *Evaluation and Program Plannig*. Vol.: 16. N.º: 4. Págs.: 279 a 285.
- ALVAREZ, V. B. et alius (Equipo VII) (1979): "Las raíces del fracaso escolar". *Cuadernos de Pedagogía*. N.º 57. Págs.: 42 a 47.
- ALVARO PAGE, M. (1992): "El proyecto de la OCDE. Indicadores educativos". *Cuadernos de Pedagogía*. N.º: 199. Págs.: 13 a 17.
- ALVARO PAGE, M. e IZQUIERDO, M. C. (1992): "Para saber más". *Cuadernos de Pedagogía*. N.º: 199. Págs.: 25 a 26.
- ALVIRA, F. (1982): "Perspectiva cualitativa-perspectiva cuantitativa en la metodología sociológica". *Revista Española de Investigaciones Sociológicas*. N.º 22. Págs.: 53 a 57.

- ALLAL, L. (1980): "Estrategias de Evaluación Formativa. Concepciones Psicopedagógicas y Modalidades de Aplicación". *Infancia y Aprendizaje*. N.º: 11. Págs.: 4 a 22.
- ANASTASI, A. (1982): *Tests psicológicos*. Aguilar, Madrid.
- APPLE, M. W. (1982): *Educación y Poder*. Paidós, Barcelona. 1987.
- APPLE, M. W. y JUNGCK, S. (1990): "No hay que ser maestro para enseñar esta unidad: la enseñanza, la tecnología y el control del aula". *Revista de Educación*. N.º: 291. Págs.: 149 a 172.
- BAKER, E. L. (1994): "Making Performance Assessment Work: The Road Ahead." *Educational Leadership*. Vol.: 51. N.º: 6. Págs.: 58 a 62.
- BALLESTER MARTINEZ, J. (1995): "Evaluación del trabajo en grupo." *Cuadernos de Pedagogía*. N.º: 234. Págs.: 48 a 53.
- BATES, R. (1988): *Evaluating schools: a critical approach*. Deakin University Press.
- BATES, R.; SMYTH, W. J. et alius (1989): *Teoría crítica de la administración educativa*. Serveis de Publicacions de la Universitat, Valencia.
- BAUDELLOT, Ch. y ESTABLET, R. (1990): *El nivel educativo sube. Refutación de una antigua idea sobre la pretendida decadencia de nuestras escuelas*. Morata, Madrid.
- BERNE, R. y STIEFEL, L. (1984): *The measurement of equity in school finance*. Johns Hopkins University Press, Baltimore.
- BLANK, R. K. (1994): Developing and implementing education indicators. En *Making Education Count: Developing and using international indicators*. OCDE, París. Págs.: 123 a 136.
- BLOOM, B. S.; HASTINGS, J. T. y MADDAUS, G. F. (1971): *Handbook on Formative and Summative Evaluation of Student Learning*. McGraw-Hill, New York. Existe una traducción al español, titulada: *Evaluación del aprendizaje*. Troquel, Buenos Aires. 1975. Vol. 1.
- BOOTH, M. (1993): "Students' Historical Thinking and the National History Curriculum in England." *Theory and Research in Social Education*. Vol.: 21. N.º: 2. Págs.: 105 a 127.
- BOTTANI, N. y TUIJNMAN, A. (1994): International education indicators: framework, development and interpretation. En OCDE (1994): *Making Education Count. Developing and Using International Indicators*. OCDE, París. Págs.: 21 a 36.
- BRICK, A. y HERMANSON, K. (1994): "Observations on the structure, interpretation and use of education indicator systems." En *Making education Count: Developing and using international indicators*. OCDE, París. Págs.: 37 a 53.

- BUENDIA EISMAN, L. y SALMERON, H. (1994): "Construcción de pruebas criterios de aula." *Revista de Investigación Educativa*. N.º: 23. Págs.: 405 a 410.
- BURSTEIN, L. et alius (1989): *Stages in the Construction of Indicators Systems for Assessing Educational Quality*. Presentado en el Simposio sobre: "Developing State Indicator Systems: From Theory To Practice". American Educational Research Association Annual Meeting, San Francisco.
- CARABAÑA, J. (1983): "Le llamaban fracaso escolar". *Cuadernos de pedagogía*. N.º: 103. Págs.: 26 a 30.
- CARDINET, J. (1988): La objetividad de la evaluación. En *Temas actuales sobre Psicopedagogía y Didáctica*. Narcea, Madrid. Págs.: 92 a 102.
- CARREÑO H., F. (1985): *Enfoques y principios teóricos de la evaluación*. Trillas, México, D. F.
- CASTILLO ARREDONDO, S. y GENTO PALACIOS, S. (1995): Modelos de evaluación de Programas Educativos. En MEDINA RIVILLA, A. y VILLAR ANGULO, L. M. (Coord.) (1995): *Evaluación de programas educativos, centros y profesores*. Ed. Universitas, Madrid. Págs.: 25 a 72.
- CERDA, R.; MORENO, J. y MUÑOZ, A. (1995): "Experiencia del mes: De la evaluación como finesse." *Cuadernos de Pedagogía*. N.º: 232. Págs.: 55 a 61.
- COHEN, D. y SPILLANE, J. (1994): National education indicators and traditions of accountability. En *Making Education Count: Developing and using international indicators*. OCDE, París. Págs.: 323 a 337.
- COLL, C. (1983): "La evaluación en el proceso de enseñanza-aprendizaje." *Cuadernos de pedagogía*. N.º: 103-104. Págs.: 13 a 17.
- COLL, C.; POZO, J. I.; SARABIA, B. y VALLS, E. (1994): *Los contenidos en la Reforma. Enseñanza y aprendizaje de conceptos, procedimientos y actitudes*. Aula XXI, Santillana, Madrid.
- COOK, T. D. y REICHARDT, Ch. S. (1986) (Eds.): *Métodos cualitativos y cuantitativos en investigación evaluativa*. Morata, Madrid, 1986.
- CRONBACH, L. J. (1963): "Course improvement through evaluation". *Teachers College Record*. N.º: 64. Págs.: 672 a 683.
- CRONBACH, L. J. (1983): *Designing Evaluations of Educational and Social Programs*. Jossey-Bass, San Francisco.
- DARLING-HAMMOND, L. (1994): Policy uses and indicators. En *Making Education Count: Developing and using international indicators*. OCDE, París. Págs.: 357 a 375.
- DASSA, C. et alius (1993): "Formative Assessment in a Classroom Setting: From Practice to Computer Innovations." *Alberta Journal of Educational Research*. Vol.: 39. N.º: 1. Págs.: 111 a 125.

- DAVIS, A. y FELKNOR, C. (1994): "The Demise of Performance-Based Graduation in Littleton." *Educational Leadership*. Vol.: 51. N.º: 6. Págs.: 64 a 65.
- DOCTORS, S. I. y WOKUTCH, R. E. (1980): "Social Program Evaluation: Six Models", en *New Direction for Program Evaluation*. N.º: 7. Jossey Bass Inc. U.S.A.
- DRUDIS BALDRICH, R. (1969): *Apuntes de Sociología de la Educación*. Ediciones Eurolibro, Madrid.
- EBEAM, D. L. (1974): *Alternative Approaches to Educational Evaluation*. En *Evaluation in Education: Current Applications*. McCutchan Publishing, Berkeley, California.
- EISNER, E. W. (1979): *The Educational Imagination*. MacMillan Publish. New York.
- EISNER, E. W. (1981): "On the Differences between Scientific and Artistic Approaches to Qualitative Research". *Educational Re-searcher*. Vol.: 10. N.º: 4. Págs.: 5 a 10.
- EISNER, E. W. (1982): *An Artistic Approach to Supervision*. En SERGIOVANNI, T. J. (Ed.) (1982): *Supervision of Teaching*. Alexandria, Virginia. Association for Supervision and Curriculum Development. Págs.: 53 a 66.
- EISNER, E. W. (1982): *Cognition and curriculum*. Longman, New York.
- EISNER, E. W. (1983): "Los objetivos educacionales veinte años después". *Temas de Educación*. Vol. 11, N.º 2. Facultad de Humanidades y Educación. U.C.V. Caracas.
- EISNER, E. W. (1985): *The Art of Educational Evaluation. A Personal View*. The Falmer Press, London.
- EISNER, E. W. (1993): "Why Standards May Not Improve Schools." *Educational Leadership*. Vol.: 50. N.º: 5. Págs.: 22 a 23.
- ELLIOTT, J. (1982): *Self Evaluation and The Teacher...* Schools Council, London.
- ELLIOTT, J. (1991): *Action Research for Educational Change*. Open University Press, Milton Keynes, Filadelfia.
- ELLIOTT, J. (1992, a): "¿Son los «indicadores de rendimiento» indicadores de la calidad educativa?" *Cuadernos de Pedagogía*. N.º 206. Págs.: 56 a 60.
- ELLIOTT, J. (1992, b): "¿Son los «indicadores de rendimiento» indicadores de la calidad educativa?" *Cuadernos de Pedagogía*. N.º 207. Págs.: 44 a 47.
- ENNIS, R. H. (1993): "Critical Thinking Assessment." *Theory into Practice*. Vol.: 32. N.º: 3. Págs.: 179 a 186.
- ESCUDERO, T. (1995): "La evaluación de las actitudes científicas." Monografía: La evaluación de los aprendizajes. *Alambique, Didáctica de las Ciencias Experimentales*, Abril. N.º: 4, Págs. 33 a 41.
- ESTRELLA, B. (1983): "La investigación en la escuela: un arma contra el fracaso". *Cuadernos de Pedagogía*. N.º: 103-104. Págs.: 44 a 48.

- FARRE, J. y GOL, T. (1982): "La evaluación". *Cuadernos de Pedagogía*. N.º: 94. Págs.: 11 a 12.
- FASANO, C. (1994): *Knowledge, ignorance and epistemic utility: Issues in constructing and organising indicator systems*. En OCDE (1994): *Making Education Count. Developing and Using International Indicators*. OCDE, París. Págs.: 55 a 78.
- FERNÁNDEZ ENGUITA, M. (1988): "Entre la Esperanza del Cambio y el Estigma de la Reproducción." *Revista de Educación*. N.º: 286. Págs.: 151 a 165.
- FERNANDEZ PEREZ, M. (1988): Evaluación y fracaso escolar. En *Cuestiones de Didáctica*. Soc. Esp. de Pedagogía, Barcelona. Pags.: 77 a 92.
- FEUER, M. J. y FULTON, K. (1993): "The Many Faces of Performance Assessment." *Phi Delta Kappan*. Vol.: 74. N.º: 6. Pág.: 478.
- FLANDERS, N. A. (1977): *Análisis de la interacción didáctica*. Ediciones Anaya, Madrid.
- FORNS, M. (1980): "La evaluación del aprendizaje." En COLL, C. y FORNS, M. (Ed.): *Áreas de intervención de la Psicología*. Horsori, Barcelona. Págs.: 107 a 141.
- FOURNIER, D. y SMITH, N. L. (1993): "Clarifying the Merits of Argument in Evaluation Practice." *Evaluation and Program Plannig*. Vol.: 16. N.º: 4. Págs.: 315 a 323.
- FRISBIE, D. A. et alius (1993): "An Evaluation of Elementary Textbook Tests as Classroom Assessment Tools." *Applied Measurement in Education*. Vol.: 6. N.º: 1. Págs.: 21 a 36.
- FUERTES, J. F. (1990): "Sobre exámenes." *Enseñanza de las Ciencias*, 8 (2). Pág.: 199.
- GARDNER, D. E. (1977): "Five Evaluation Frameworks. Implications for Decision Making in Higher Education" *Journal of Higher Education*. Vol.: 48, N.º: 5. Págs.: 571 a 593.
- GELI, A. M. (1988): "Paradigmas contemporáneos de evaluación y su relación con la enseñanza de las ciencias". *Investigación en la Escuela*. N.º: 5. Págs.: 47 a 51.
- GELI, A. M. (1995): "La evaluación de los trabajos prácticos." Monografía: La evaluación de los aprendizajes. *Alambique, Didáctica de las Ciencias Experimentales*, Abril. N.º: 4, Págs. 25 a 32.
- GIL, D.; CARRASCOSA, J.; FURIO, C. y MARTINEZ, J. (1991): *La enseñanza de las ciencias en la educación secundaria*. Horsori, Barcelona. Cap. VII.
- GILLY, M. (1978): *El Problema del rendimiento escolar*. Oikos-Tau, Barcelona.
- GIMENO SACRISTAN, J. y PEREZ GOMEZ, A. I.(Eds.) (1983): *La enseñanza: su teoría y su práctica*. Akal, Madrid.
- GIROUX, H. A. (1990, a): "Curriculum Theory, Textual Authority, and the Role of Teachers as Public Intellectuals" *Journal of Curriculum and Supervision*. Vol.: 5, N.º: 4. Págs.: 361 a 383.

- GIROUX, H. A. (1990, b): *Los Profesores como Intelectuales. Hacia una Pedagogía Crítica del Aprendizaje*. Paidós/M.E.C. Barcelona.
- GLASER, R. (1963): "Instructional Technology and the Measurement to Learning Outcomes: Some Questions". *American Psychologist*. N.º 18. Págs.: 519 a 521.
- GODD, J. A. (1988): *Knowledge and control in the evaluation of educational organizations*. Deakin University Press.
- GOETZ, J. P. y LeCOMPTE, M. D. (1984): *Etnografía y Diseño Cualitativo en Investigación Educativa*. Morata. Madrid.
- GUBA, E. G. (1978): *Toward a Methodology of Naturalistic Inquiry in Educational Evaluation*. Center for the Study of Evaluation, University of California. Los Angeles.
- GUBA, E. G. y LINCOLN, Y. S. (1985): *Effective Evaluation. Improving the Usefulness of Evaluation Results through Responsive and Naturalistic Approaches*. Jossey-Bass Publishers, San Francisco.
- GUTIERREZ, R.; MARCO, B.; OLIVARES, E. y SERRANO, T. (1990): *La enseñanza de las Ciencias en la educación intermedia*. Narcea, Madrid. Cuarta Parte.
- GUTTENTAG, M. (1971): "Models and Methods in Evaluation Research" *Journal of Theory in Social Behaviour* Vol.1, N.º: 1. Págs.: 75 a 95.
- HABGOOD, M. K. (1993): "Alternative Assessment Practices as New Rules in the Whole-Language Game." *International Journal of Social Education*. Vol.: 8. N.º: 2. Págs.: 81 a 87.
- HAMMOND, R. L. (1973): Evaluation at the Local Level. En WORTHEN, B. R. y SANDERS, J. R. (1973): *Educational Evaluation: Theory and Practice*. Charles A. Jones Publishing. Worthington, Ohio. Págs.: 157 a 169.
- HASSMEN, P. y HUNT, D. P. (1994): "Human Self-Assessment in Multiple-Choice Testing." *Journal of Educational Measurement*. Vol.: 31. N.º: 2. Págs.: 149 a 160.
- HERNANDEZ RUIZ, S. y GOMEZ DACAL, G. (1982): *Fracasos escolares*. Escuela Española. Madrid.
- HOUSE, E. R. (1986, a) (Ed.): *New directions in educational evaluation*. The Falmer Press, London.
- HOUSE, E. R. (1990, a): "Trends in Evaluation". *Educational Researcher*. Vol.: 19. N.º: 3. April. Págs.: 24 a 28. También en HOUSE, E. R. (1992): "Tendencias en Evaluación." *Revista de Educación*. N.º: 299. Págs.: 43 a 55.
- HOUSE, E. R. (1990, b): Voz «Ética de la evaluación», en *Enciclopedia Internacional de Educación*. Vicens Vives/MEC, Barcelona. Vol. IV. Págs.: 2477 a 2481.
- HOYAT, F. (1962): *Les examens*. Institut de l'Unesco pour l'Education. Bourrelier, París.

- JAMETZ, K. (1994): "Making Sure That Assessment Improves Performance." *Educational Leadership*. Vol.: 51. N.º: 6. Págs.: 55 a 57.
- JORBA, J. y SANMARTI, N. (1993): "La función pedagógica de la evaluación." En *Aula de Innovación Educativa*. N.º: 20. Págs.: 20 a 30.
- JORBA, J. y SANMARTI, N. (1994): *Enseñar, aprender y evaluar: un proceso de regulación continua*. MEC, Madrid.
- JORBA, J. y SANMARTI, N. (1995): "Autorregulación de los procesos de aprendizaje y construcción de conocimientos." Monografía: La evaluación de los aprendizajes. *Alambique, Didáctica de las Ciencias Experimentales*, Abril. N.º: 4, Págs. 59 a 78.
- KEMMIS, S. (1986, b): Seven Principles for Programme Evaluation in Curriculum En HOUSE, E. R. (Ed.) (1986): *New directions in educational evaluation*. The Falmer Press, London. Págs.: 117 a 140.
- KEMPA, R. (1989): *Assesment in science*. Cambridge University Press, Cambridge.
- KETELE, J. M. de (1980): *Observer pour éduquer*. Peter Lang, Berne. Existe traducción al castellano: KETELE, J. M. de (1984): *Observar para educar*. Visor, Madrid.
- KOZULIN, A. (1994): *La psicología de Vygotski*. Alianza, Madrid.
- LAFOURCADE, P. D. (1977): *Evaluación de los aprendizajes*. Cincel, Madrid.
- LAWTON, D. (1980): The Politics of Curriculum Evaluation. En McCORMICK, R. (Ed.) *Calling Education to Account*. Heineman Educational Books. London. Págs.: 169 a 184.
- LEARY, N. (1993): "An Introduction to Whole Language from a Secondary Social Studies Perspective." *International Journal of Social Education*. Vol.: 8. N.º: 2. Págs.: 12 a 32.
- LERENA, C. (1980): *Escuela, Ideología y Clases Sociales en España*. Ariel, Barcelona.
- LIVAS GONZALEZ, I. (1982): *Análisis e interpretación de los resultados de la evaluación educativa*. Trillas, México, D. F.
- LOPEZ, A. (1975): "Patología del niño. Patología de la escuela". En *Cuadernos de Pedagogía*. N.º 12. Págs.: 18 a 24.
- LOPEZ, N.; LLOPIS, R.; SALINAS, B. y SOLER, I. (1983): "Análisis de dos modelos evaluados referidos a la Química de COU y selectividad." *Enseñanza de las Ciencias I (I)*, Págs.: 21 a 25.
- LURÇAT, L. (1979): *El fracaso y el desinterés escolar en la Escuela Primaria*. Gedisa, Barcelona.
- MacDONALD, B. (1976, a): "Evaluation and the control of Education". En TAWNEY, D.: *Curriculum evaluation today: Trends and implications*. McMillan, Londres. Trad. al castellano en MacDONALD, B. (1980): La evaluación y el control de la educación.

- En GIMENO SACRISTAN, J. y PEREZ GOMEZ, A. I. (1983): *La Enseñanza: su Teoría y su práctica*. Ed. Akal. Madrid. Págs.: 467 a 478.
- MacDONALD, B. (1976, b): *Who's Afraid of Evaluation*. Centre for Applied Research in Education. University of East Anglia. Norwich. U. K. (Multicopiado).
- MADAUS, G. F. y KELLAGHAN, T. (1993): "The British Experience with "Authentic" Testing." *Phi Delta Kappan*. Vol.: 74. N.º: 6. Págs.: 458 a 469.
- MAGER, R. (1984): *Measuring instructional results*. Pitman Learning. Belmont, California. U.S.A.
- MARTIN GORDILLO, M. (1994): "Evaluar el aprendizaje, evaluar la enseñanza." En *Signos*. N.º: 13. Págs.: 22 a 37.
- MARTINEZ MUÑIZ, B. (1982): *Causas del fracaso escolar y técnicas para afrontarlo*. Narcea, Madrid.
- MARZANO, R. J. (1994): "Lessons from the Field about Outcome-Based Performance Assessments." *Educational Leadership*. Vol.: 51. N.º: 6. Págs.: 44 a 50.
- MENOYO, M. P. et alius (1995): "La evaluación formativa como instrumento de atención a la diversidad: una experiencia en Secundaria." Monografía: La evaluación de los aprendizajes. *Alambique, Didáctica de las Ciencias Experimentales*, Abril. N.º: 4, Págs. 42 a 54.
- METFESSEL, N. S. y MICHAEL, W. B. (1967): "A Paradigm Involving Multiple Criterion Measures for the Evaluation of School Programs". En *Educational and Psychological Measurement*. N.º: 27. Págs.: 931 a 943.
- MONEDERO MOYA, J. J. (1987): "¿Qué es la selectividad?" *Puertanueva*. N.º: 3. Págs.: 6 a 8.
- MONEDERO MOYA, J. J. (1988): *La Toma de Decisiones en la gestión del Centro Escolar de EGB: estudio de casos*. Tesis de Grado. Departamento de Didáctica y Organización Escolar. Universidad de Málaga. Málaga. Inédita.
- MONEDERO MOYA, J. J. (1996): *Métodos y Técnicas de la Evaluación Educativa en la EGB: Reforma versus No Reforma*. Tesis Doctoral. Departamento de Comunicación Audiovisual y Publicidad, y Métodos de Investigación. Universidad de Málaga. Málaga. Inédita.
- NIEDA, J. et alius (1985): "El uso de mapas conceptuales en la corrección de preguntas abiertas de Biología." *Enseñanza de las Ciencias*. Vol.: 3. N.º: 2. Págs.: 91 a 95.
- NIETO GIL, J. M. (1994): *La autoevaluación del profesor. Cómo puede el profesor evaluar su propia práctica docente*. Ed. Escuela Española, Madrid.
- NOIZET, G. y CAVERNI, J. P. (1978): *Psychologie de l'évaluation scolaire*. Presses Universitaires de France, París.

- NORRIS, N. (1992): *Evaluation, Economics and Performance Indicators*. Paper presented at a symposium on *Judging Quality in Education*. 15th Annual conference of the British Educational Research Association. Nottingham Polytechnic.
- NUTTALL, D. (1994): "Choosing Indicators." En OCDE (1994): *Making Education Count. Developing and Using International Indicators*. OCDE, París. Págs.: 79 a 96.
- OCDE (1987): *International Educational Indicators: A conceptual working plan*. OCDE, París.
- OLIVARES, E. (1995): "Tipos de contenidos e instrumentos de evaluación." Monografía: La evaluación de los aprendizajes. *Alambique, Didáctica de las Ciencias Experimentales*, Abril. N.º: 4, Págs. 16 a 24.
- ORDEN, A. de la (1977): Prólogo a la edición española. En LAFOURCADE, P. D. (1977): *Evaluación de los aprendizajes*. Cincel, Madrid. Págs.: 13 a 17.
- PARLETT, M. R. y HAMILTON, D. (1972): *Evaluation as Illumination: a new approach to the study of innovatory programmes*. Centre for Research in the Educational Sciences. Universidad de Edimburgo.
- PARLETT, M. R. y HAMILTON, D. (1976): La evaluación como iluminación. En GIMENO, J. y PEREZ, A. I. (1983): *La Enseñanza: su Teoría y su práctica*. Ed. Akal. Madrid. 450 a 466. También: PARLETT, M. y HAMILTON, D. (1972): *Evaluation as Illumination*. TAWNEY, D. (1976): *Curriculum Evaluation today Trends and Implications*. McMillan London. O también en PARLETT, M. y HAMILTON, D. (1976): "Evaluation as Illumination: A New Approach to the Study of Innovative Programs". En *Evaluation Studies*. Vol.: 1. Págs.: 140 a 157.
- PEREZ JUSTE, R. (1995): Evaluación de programas educativos. En MEDINA RIVILLA, A. y VILLAR ANGULO, L. M. (Coord.) (1995): *Evaluación de programas educativos, centros y profesores*. Ed. Universitas, Madrid. Págs.: 73 a 108.
- PEREZ JUSTE, R. (1995): Metodología para la evaluación de programas educativos. En MEDINA RIVILLA, A. y VILLAR ANGULO, L. M. (Coord.) (1995): *Evaluación de programas educativos, centros y profesores*. Ed. Universitas, Madrid. Págs.: 109 a 146.
- PERRENOUD, Ph. (1981): "De las diferencias culturales a las desigualdades culturales: La evaluación y la norma en una enseñanza indiferenciada". *Infancia y Aprendizaje*. N.º: 14. Págs.: 19 a 50.
- PERRENOUD, Ph. (1984): *La fabrication de l'excellence scolaire dans l'enseignement primaire: du curriculum aux pratiques d'évaluation*. Droz, Ginebra.
- PERRENOUD, Ph. (1990): *La construcción del éxito y del fracaso escolar. Hacia un análisis del éxito, del fracaso y de las desigualdades como realidades construidas por el sistema escolar*. Morata, Madrid.

- PIDGEON, D. y YATES, A. (1979): *Evaluación y medida del rendimiento escolar*. Anaya/2, Madrid.
- PIERON, H. (1969): *Examens et docimologie*. Presses Universitaires de France, París.
- PIRSIG, R. M. (1990): Zen y el arte del mantenimiento de la motocicleta. En AA. VV. (1990): *Equipos Directivos. Curso de Formación. Dossier sobre evaluación*. Delegación Provincial. Consejería de Educación y Ciencia, 1990-1991, Junta de Andalucía.
- POPHAM, W. J. (1975, a): *Educational Evaluation*. Englewood Cliffs, Prentice-Hall, New Jersey.
- POPHAM, W. J. (1975, b): *Evaluation*. Prentice-Hall, New Jersey.
- POPHAM, W. J. (1980, a): "Educational Measurement for the Improvement of Instruction" *Phi Delta Kappan*. Vol.: 61, April. Págs.: 531 a 534.
- POPHAM, W. J. (1980, b): *Problemas y Técnicas de la Evaluación Educativa*. Anaya/2, Madrid.
- POPKEWITZ, T. S. (1987): *Educational Reform: Rhetoric, Ritual and Social Interest*. Dept. of Curriculum and Instruction. University of Wisconsin-Madison. Madison.
- POZO, J. I. (1994): La evaluación del aprendizaje de hechos y de conceptos. En COLL, C.; POZO, J. I.; SARABIA, B. y VALLS, E. (1994): *Los contenidos en la Reforma. Enseñanza y aprendizaje de conceptos, procedimientos y actitudes*. Aula XXI, Santillana, Madrid. Págs.: 67 a 78.
- RAFTER, D. O. (1984): "Three Approaches to Evaluation Research". *Knowledge: Creation, Diffusion, Utilization*. Vol. 6, n.º 2. Págs. 165 a 185.
- RIVLIN, A. M. (1971): *Systematic Thinking for Social Action*. The Brookings Institution. Washington, D. C.
- RODRIGUEZ ESPINAR, J. (1982): *Factores del rendimiento escolar*. Oikos-Tau, Barcelona.
- RODRIGUEZ, L. M.; GUTIERREZ, F. A. y MOLLEDO, J. (1992): "Una propuesta integral de evaluación en Ciencias." *Enseñanza de las Ciencias*. Vol.: 10. N.º: 3. Págs.: 254 a 267.
- ROSALES, C. (1990): *Evaluar es reflexionar sobre la enseñanza*. Narcea, Madrid.
- ROSSI, P. H. y FREEMAN, H. E. (1986): *Evaluation. A Systematic Approach*. Sage Publications. Beverly Hills, California. U.S.A.
- RUIZ PRIMO, M.ª A. et alius (1993): "On the Stability of Performance Assessments." *Journal of Educational Measurement*. Vol.: 30. N.º: 1. Págs.: 41 a 53.
- RUL GARGALLO, J. (1994): "L'avaluació comunicativa, factor de creixement professional, organitzativa i curricular". *Perspectiva Escolar*, N.º: 183. Págs.: 23 y ss.
- RUL GARGALLO, J. (1995): "La evaluación comunicativa, factor de desarrollo humano, organizativo y curricular. Valorar, medir y evaluar". *Aula*, N.º: 39. Págs.: 70 a 77.

- RUMBERGER, R. (1994): Labour market outcomes as indicators of educational performance. En *Making Education Count: Developing and using international indicators*. OCDE, París. Págs.: 265 a 286.
- SAMPER, L.; SAMPER, I. y SOLER, I. (1982): *Perspectivas psicológicas y sociológicas sobre el fracaso escolar*. ICE. Univ. Autónoma de Barcelona. Barcelona.
- SANCHO GIL, J. M.^a (1990): "De la evaluación a las evaluaciones". *Cuadernos de Pedagogía*. N.º: 185. Págs.: 8 a 11.
- SANDEFUR, G. D.; FREEMAN, H. E. y ROSSI, P. H. (1986): *Workbook for evaluation: a systematic approach*. SAGE Publications. Beverly Hills, California. U.S.A.
- SARABIA, B. (1994): Evaluación del aprendizaje de actitudes. En COLL, C.; POZO, J. I.; SARABIA, B. y VALLS, E. (1994): *Los contenidos en la Reforma. Enseñanza y aprendizaje de conceptos, procedimientos y actitudes*. Aula XXI, Santillana, Madrid. Págs.: 189 a 197.
- SARUP, M. (1990): "El curriculum y la reforma educativa: hacia una nueva política de la educación". *Revista de Educación*. N.º: 291. Págs.: 193 a 221, Madrid.
- SCRIVEN, M. (1967): "The Methodology of Evaluation." En STAKE, R. W. y otros (Ed.) (1967): *Perspectives of Curriculum Evaluation*. AERA Monograph Series on Curriculum Evaluation. N.º: 1. Chicago. Rand McNally. Págs.: 39 a 83.
- SCRIVEN, M. (1973): Goal Free Evaluation. En HOUSE, E. R. (1973): *School Evaluation*. McCutchan, Berkeley.
- SCRIVEN, M. (1974): Evaluation Perspectives and Procedures. En *Evaluation in Education: Current Applications*. McCutchan Publishing. Berkeley, California.
- SCRIVEN, M. (1976): "Evaluation Bias and Its Control". En *Evaluation Studies*. Vol.: 1. Págs.: 119 a 139.
- SCRIVEN, M. (1976): *Bias Control Systems in Evaluation*. AERA, San Francisco.
- SCRIVEN, M. (1981): *Evaluation Thesaurus*. Edgepress. Inverness. California.
- SCRIVEN, M. (1983): "Evaluation Ideologies." En MADDAUS, G. F, SCRIVEN, M. S, y STUFFLEBEAM, D. L. (Eds.) (1983): *Evaluation Models Viewpoints on Educational and Human Services*. Kluwer-Nijhoff Publish, Boston, Mass
- SCRIVEN, M. (1983): "The Evaluation Taboo". En *Philosophy of Evaluation*. E. R. HOUSE, San Francisco. *New Direction for Program Evaluation*. N.º: 19. Jossey Bass Inc. U.S.A. September. P. 75 a 82.
- SCRIVEN, M. (1986, a): "New Frontiers Evaluation". En *Evaluation Practice*. Vol.: 7. N.º: 1.
- SCRIVEN, M. (1986, b): "Evaluation as a Paradigm for Educational Research". En HOUSE, E. R. (1986, a) (Ed.): *New directions in educational evaluation*. The Falmer Press, London.

- SCRIVEN, M. y HOUSE, E. R. (1973): *School Evaluation*. McCutchan, Berkeley.
- SCHAFER, W. D. (1993): "Assessment Literacy for Teachers." *Theory into Practice*. Vol.: 32. N.º: 2. Págs.: 118 a 126.
- SCHWARTZ, H. y JACOBS, J. (1984): *Sociología cualitativa. Método para la reconstrucción de la realidad*. Trillas, México.
- SELDEN, R. W. (1990): "Developing Educational Indicators: A State-National Perspective". *International Journal of Education Research*. Vol. 66. Págs.: 383 a 393.
- SERRANO, J. et alius (1995): "Aportaciones recientes en medición y evaluación educativa." Ponencia presentada al Séptimo Seminario de Modelos de Investigación Educativa: *Avances de la Investigación Educativa en la Intervención Psicopedagógica. Aportaciones teórico-prácticas*. Valencia, Septiembre, 1995.
- SHAPLIN, J. T. (1961): "Practice in Teaching." *Harvard Educational Rev.*, N.º: 31.
- SHAVELSON, R. J. et alius (1993): "Samplig Variability of Performance Assessments." *Journal of Educational Measurement*. Vol.: 30. N.º: 3. Págs.: 215 a 232.
- SHOEMAKER, B. J. E. y LEWIN, L. (1993): "Curriculum and Assessment: Two Sides of the Same Coin." *Educational Leadership*. Vol.: 50. N.º: 8. Págs.: 55 a 57.
- SILVESTRI, A. Y BLANCK, G. (1993): *Bajtín y Vygotski: la organización semiótica de la conciencia*. Anthropos, Barcelona, pp. 17 a 24.
- SIMMONS, R. (1994): "The Horse before the Cart: Assessing for Understanding." *Educational Leadership*. Vol.: 51. N.º: 5. Págs.: 22 a 23.
- SIMMONS, W. y RESNICK, L. (1993): "Assessment as the Catalyst of School Reform." *Educational Leadership*. Vol.: 50. N.º: 5. Págs.: 11 a 15.
- SINGER, A. (1994): "The Impact of Industrialization on American Society: Alternative Assessments. Classroom Teacher's Idea Notebook." *Social Education*. Vol.: 58. N.º: 3. Págs.: 171 a 172.
- SPEARS, M. G. (1984): "Sex bias in science teachers' ratings of work and pupils characteristics." *European Journal of Science Education*. Vol. 6. Págs.: 369 a 377.
- STAKE, R. E. (1967): "The countenance of educational evaluation". *Teacher's College Record*. N.º: 68. Págs.: 523 a 540.
- STAKE, R. E. (1969): "Needed Concepts and Techniques for Utilizing More Fully the Potencial of Evaluation". En TYLER, R. W. (1969) (Ed.): *Educational Evaluation: New Roles, New Means* (The sixty-eight Yearbook on the National Society for the Study of Education). Part II. University of Chicago Press, Chicago, Illinois. Págs.: 370 a 390.
- STAKE, R. E. (1973): Measuring what Learners Learn. En HOUSE, E. (Ed.) (1973): *School Evaluation: The Politics and Process*. McCutchan. Berkeley, CA, 193 a 223.

- STAKE, R. E. (1975, a): *The Responsive Evaluation of Educational Programs*. University of Illinois at Urbana-Champaign. Center for Instructional Research and Curriculum Evaluation. College of Education. (Multicopiado).
- STAKE, R. E. (1975, c): To Evaluate an Arts Program. En STAKE, R. E. (Ed.) (1975): *Evaluating the Arts in Education: a Responsive Approach*. Charles E. Merrill. Columbus, Ohio. Págs.: 91 a 108. (También publicado en (1976): *Journal of Aesthetic Education*. Vol.: 10, N.º: 3-4. Págs.: 115 a 133.
- STAKE, R. E. (1976): "A Theoretical Statement of Responsive Evaluation" *Studies in Educational Evaluation*. Vol.: 2, N.º: 1. Págs.: 19 a 22.
- STAKE, R. E. (1978): "The Case Study Method in Social Inquiry" *Educational Researcher*. Vol.: 7, N.º: 2. Págs.: 5 a 7.
- STAKE, R. E. (1980): La evaluación de programas: En especial la evaluación de réplica. En STAKE, R. E.; DOCKRELL, W. B. y HAMILTON, D. (Eds.) (1980): *Nuevas Reflexiones sobre la Investigación Educativa*. Ed. Narcea, Madrid. Págs.: 40 a 52.
- STAKE, R. E. (1981): "Setting standards for educational evaluators". *Evaluation New*. Vol. 2. N.º 2. Págs.: 148 a 152.
- STAKE, R. E. (1986): *Quieting Reform: Social Science and Social Action in an Urban Youth Program*. University of Illinois Press, Illinois.
- STAKE, R. E. (1987): An Evolutionary View of Programming Staff Development. En WIDEEN, M. y ANDREWS, I. (Ed.) (1987): *Staff Development for School Improvement. A Focus on the Teacher*. Lewes, East Sussex. The Falmer Press. Págs.: 55 a 69.
- STAKE, R. E. (1988): *Case Study Methods in Educational Research*. Seeking Sweet.
- STAKE, R. E. (Ed.) (1975, b): *Evaluating the Arts in Education: a Responsive Approach*. Charles E. Merrill. Columbus, Ohio.
- STAKE, R. E. y PEARSOL, J. A. (1981): Evaluating Responsively. En BRANDT, R. S. (Ed.) (1981): *Applied Strategies for Curriculum Evaluation*. Association for Supervision and Curriculum Development. Virginia. Págs.: 25 a 33.
- STAKE, R. E.; DOCKRELL, W. B. y HAMILTON, D. (Eds.) (1980): *Nuevas Reflexiones sobre la Investigación Educativa*. Ed. Narcea, Madrid.
- STANLEY AHMANN, J.; GLOCK, M. D. y WARDEBERG, H. L. (1972): *Evaluación de los Alumnos en la Escuela Primaria*. Aguilar, Madrid.
- STENHOUSE, L. (1982): *Investigación y Desarrollo del Curriculum*. Morata, Madrid, 1984.
- MEDINA RIVILLA, A. (1995): Elaboración de un modelo para valorar el Proyecto Formativo de centro. En MEDINA RIVILLA, A. y VILLAR ANGULO, L. M. (Coord.) (1995): *Evaluación de programas educativos, centros y profesores*. Ed. Universitas, Madrid. Págs.: 147 a 176.

- STUFFLEBEAM, D. L. (1987): *Evaluación sistemática*. Paidós-MEC.
- STUFFLEBEAM, D. L. (1988, a): *Normas de evaluación para programas, proyectos y material educativo*. Trillas, México.
- STUFFLEBEAM, D. L. (1988, b): Normas para evaluadores. En *La gestión educativa ante la innovación y el cambio*. Narcea, Madrid.
- STUFFLEBEAM, D. L. (1989): *The personnel evaluation standards: How to asses for evaluating educators*. SAGE Publications. Beverly Hills, California. U.S.A.
- STUFFLEBEAM, D. L. et alius (1971): *Educational Evaluation and Decision Making*. F. E. Peacock Public. Inc. Itasca. Illinois.
- STUFFLEBEAM, D. L. y SHINKFIELD, A. J. (1987): *Evaluación sistemática. Guía teórica y práctica*. Paidós/MEC. Barcelona. También STUFFLEBEAM, D. L. (1985): *Systematic evaluation: a self-instructional guide to theory and practice*. KUWER-Nijhoff cop. Boston.
- STUFFLEBEAM, D. L. y WEBSTER, W. (1981): An analysis of Alternative Approaches to Evaluation, in H. FREEMAN y M. SALOMON: *Evaluation Studies Review Annual*. SAGE Publications. Beverly Hills, California. U.S.A. También en MADAUS, G. F, SCRIVEN, M. S, y STUFFLEBEAM, D. L. (Eds.) (1983): *Evaluation Models Viewpoints on Educational and Human Services*. Kluwer-Nijhoff Publish, Boston, Mass. Págs.: 23 a 43. Este mismo artículo apareció publicado en 1980, en *Educational Evaluation and Policy Analysis*. Vol.: 2. Págs.: 5 a 20.
- STUFFLEBEAM, D. L.: The relevance of the CIPAGS. evaluation model for educational accountability.
- TABA, H. (1983): La evaluación de los resultados del currículo. En TABA, H.: *Elaboración del currículo*. Troquel, Buenos Aires. (1974)
- THORNDIKE, R. L. (1971): *Educational measurement*. American Council of Education, Washinton.
- TIANA FERRER, A. (1994): "Los nuevos retos de la evaluación". *Cuadernos de Pedagogía*. N.º: 231. Págs.: 87 a 91.
- TORRANCE, H. (1993): "Combining Measurement-Driven Instruction with Authentic Assessment: Some Initial Observations of National Assessment in England and Wales." *Educational Evaluation and Policy Analysis*. Vol.: 15. Págs.: 81 a 90.
- TOULMIN, S. (1972): *La Comprensión Humana I. El Uso Colectivo y la Evolución de los Conceptos*. Alianza, Madrid, 1977.
- TYLER, R. W. (1934): "Techniques for Evaluating Behavior" *Educational Research Bulletin*. Vol.: 13, N.º: 1, Págs.: 1 a 11.

- TYLER, R. W. (1969) (Ed.): *Educational Evaluation: New Roles, New Means* (The sixty-eight Yearbook on the National Society for the Study of Education). Part II. University of Chicago Press, Chicago, Illinois.
- VV. AA. (1990, c): Monográfico. "Las evaluaciones". *Cuadernos de Pedagogía* . N.º: 185. Págs.: 5 a 36.
- VYGOTSKI, L. S. (1981): *Pensamiento y Lenguaje*. La Pléyade, Buenos Aires.
- VYGOTSKI, L. S. (1990): *Obras escogidas* (Vol. I). Visor, Madrid.
- VYGOTSKI, L. S. (1993): *Obras escogidas* (Vol. II). Visor, Madrid.
- VYGOTSKI, L. S. (1995): *Obras escogidas* (Vol. III). Visor, Madrid.
- WERTSCH, J. V. (1981): *The Concept of Activity in Soviet Psychology*. Sharpe, New York.
- WERTSCH, J. V. (1988): *Vygotski y la formación social de la mente*. Paidós, Barcelona.
- WERTSCH, J. V. (1993): *Voces de la mente. Un enfoque sociocultural para el estudio de la Acción Mediada..* Visor, Madrid.
- WHEELER, D. K. (1979): *El desarrollo del Curriculum escolar*. Santillana, Madrid.
- WIGGINS, G. (1993): "Assessment: Authenticity, Context, and Validity." *Phi Delta Kappan*. Vol.: 75. N.º: 3. Págs.: 200 a 208.
- WOOSTER, J. S. (1993): "Authentic Assessment: A Strategy for Preparing Teachers to Respond to Curricular Mandates in Global Education." *Theory into Practice*. Vol.: 32. N.º: 1. Págs.: 47 a 51.
- WORTHEN, B. R. (1993, a): "Critical Issues That Will Determine the Future of Alternative Assessment." *Phi Delta Kappan*. Vol.: 74. N.º: 6. Págs.: 444 a 454.
- WORTHEN, B. R. (1993, b): "Is Your School Ready for Alternative Assessment?" *Phi Delta Kappan*. Vol.: 74. N.º: 6. Págs.: 455 a 456.
- WYATT, T. (1988) (ed): "Education Indicators: An Overview", en *Reporting on Educational Progress Monograph Series*, Sidney: Conference of Director's General of Education.
- WYATT, T. et alius (1989): *Reporting on Educational Progress: Performance Indicators in Education*. Informe sobre la Conferencia de Directores Generales de Educación. Sidney.