Volver a pensar la educación (Vol. II)

Prácticas y discursos educativos (Congreso Internacional de Didáctica)

Morata/Paideia

Madrid, 1995

Este material se utiliza con fines exclusivamente didácticos

ÍNDICE

Vol. 1. Política, educación y sociedad

Introducción
por Jurjo Torres Santomé 11

Política, filosofía y educación 15
Muguerza, Javier: La indisciplina del espíritu crítico (Una perspectiva filosófica) 17
Sotelo, Ignacio: Educación y democracia 34
Lledó, Emilio: Notas históricas sobre un modelo universitario 60
Weinberg, Gregorio: La Universidad, encrucijadas y perspectivas. Una visión latinoamericana 71
Álvarez Uría, Fernando: La escuela y el espíritu del capitalismo 87
Carr, Wilfred: Educación y democracia: Ante el desafío postmodemo 96
House, Ernest R.: La política educativa en una época de productividad 112
Santos Guerra, Miguel Ángel: Democracia escolar o el problema de la nieve frita 128
Young, Michael: La especialización, los intelectuales y la crisis actual de la educación 142

Cultura y educación 151
Apple, Michael W.: La política del saber oficial: ¿Tiene sentido un curriculum nacional? 153
Alfieri, Fiorenzo: Crear cultura dentro y fuera de la escuela: algunos modelos posibles 172
Blanco García, Nieves: El sentido del conocimiento escolar. (Notas para una agenda de trabajo.) 188
Carbonell Sebarroja, Jaume: Escuela y entorno 203
Trilla Bernet, Jaume: La escuela y el medio. Una reconsideración sobre el contorno de la institución escolar 217
Torres Santomé, Julio: La presencia de diferentes culturas en las instituciones escolares. Posibilidades de diálogo y acción 232
Calvo Buezas, Tomás: La educación intercultural en una sociedad pluriétnica 254
Sarup, Madan: Hogar, identidad y educación 268
Fernández Enguita, Mariano: Escuela y etnicidad: el caso de los gitanos 281
Sacristán Lucas, Ana: Apuntes críticos sobre feminismo y educación 294
Arnot, Madeleine: Feminismo y educación democrática 307

Trabajo y educación 327
Morgenstern, Sara: El reparto del trabajo y el reparto de la educación 329
Posner, Charles: El sector informal y la reforma educativa 342
Rodríguez Guerra, Jorge: Exigencias educativas de la producción flexible 371
Segre, Lidia Micaela: Cambios tecnológicos y organizativos y sus impactos sobre la cualificación profesional 386

Notas biobibliográficas de los autores 401

Vol. II. Prácticas y discursos educativos

Racionalidad y utilidad del conocimiento pedagógico 11
Gimeno Sacristán, José: Esquemas de racionalización en una práctica compartida 13
Salinas Fernández, Bernardino: Límites del discurso didáctico actual 45

Varela, Julia: El estatuto del saber pedagógico 61

Vivir y trabajar en las aulas 71
Frabboni, Franco: Un manifiesto pedagógico de la educación ambiental. Por qué y cómo el medio ambiente en la escuela 73
Levin, Henry M.: Aprendiendo en las escuelas aceleradoras (Accelerated Schools) 80
Martínez Rodríguez, Juan Bautista: ¿Tiene el alumnado posibilidad o derecho de realizar innovaciones? 96
San Martín Alonso, Ángel: De la escuela como objeto a los sujetos escolares 119
Jares, Xesús R.: Contexto organizativo y resolución de conflictos en los centros educativos 133
Beltrán Llavador, Francisco: Desregulación escolar, organización y curriculum ... 152

Función de la evaluación 171
Álvarez Méndez, Juan Manuel: Valor social y académico de la evaluación 173
Angulo Rasco, J. Félix: La evaluación del sistema educativo: algunas respuestas críticas al porqué y al cómo 194
Simons, Helen: La autoevaluación escolar como proceso de desarrollo del profesorado: En apoyo a las escuelas democráticas 220

Repensar la función docente 243
Elliott, John: El papel del profesorado en el desarrollo curricular: una cuestión irresuelta en los intentos ingleses de reforma curricular 245
Larson, Magali Sarfatti: "Yo no soy tu madre". Retórica y realidad en una reforma escolar. (Notas para una investigación.) 273
Marrero Acosta, Javier: La cultura de la "colaboración" y el desarrollo profesional del profesorado 296
Martínez Bonafé, Jaume: La estructura del puesto de trabajo del profesorado. Un esquema para la interpretación crítica 312
Pérez Gómez, Ángel I.: Autonomía profesional del docente y control democrático de la práctica educativa 339
Teasley, Cathryn: Hacia una pedagogía para la diversidad en la formación del profesorado de California 354
Cabello Martínez, Mª Josefa: Formación y autoformación de los equipos docentes en educación de personas adultas 366
Zeichner, Kenneth M.: Los profesores como profesionales reflexivos y la democratización de la reforma escolar 385

Notas biobibliográficas de los autores 399

LA EVALUACIÓN DEL SISTEMA EDUCATIVO: ALGUNAS RESPUESTAS CRÍTICAS AL PORQUÉ Y AL CÓMO [*]

Por J. Félix Angulo Rasco
Universidad de Málaga

Introducción

A nadie se le escapa que existen grandes problemas en nuestro sistema educativo, tanto a un lado como al otro del mismo. Aunque son pocos los datos públicos con los que contamos, está claro que, en cuanto se les permite, los docentes tienen serias quejas que presentar de la labor de la Administración, incluido el cuerpo de inspectores, las delegaciones provinciales y la política educativa en general (Pérez Gómez y Gimeno Sacristán, 1992); asimismo es alarmante la escasa participación del profesorado, los padres y los alumnos en el sistema educativo (Feito 1990; CEE 1992), y las fuertes y evidentes desigualdades que lo determinan y que todavía subsisten en él (Muñoz-Repiso y cols. 1992).

Sin embargo, estas apreciaciones y los datos en las que se sustentan no dejan de ser sólo la punta del iceberg. Nuestro país se ha caracterizado siempre, entre otras cosas, por una falta endémica de *información pública* sobre cada uno de los niveles, sectores y responsables del sistema educativo, fomentada, propiciada y alimentada sucesivamente por las distintas administraciones: antes de la democracia y ahora. Podría afirmarse sin temor a equivocarse, que resulta cuando menos arriesgado, si no imposible, alcanzar una apreciación válida, hoy por hoy, del "estado" de nuestro actual sistema educativo. Seguimos desconociendo el entramado de muchos de sus problemas[1], y también sin tener información pública sobre la mera existencia de otros tantos. Con ello no quiero decir que no conozcamos algunos de esos aspectos. Las investigaciones autóctonas (como las citadas antes) y las importadas, nos están ayudando a construir una imagen más fidedigna de estos problemas. Pero ésta, no deja de ser una imagen muy limitada.

A dicha limitación hay que añadir dos problemas paralelos: por un lado el desinterés casi generalizado por demandar más y mejor información y/o la escasa preocupación por las pobremente informativas estadísticas que se nos ofrecen[2]. Por el otro, la existencia de una acrisolada estrategia de oídos sordos alentada en la misma administración educativa; administración, que, al menos en el territorio MEC, todavía no se ha dignado en declarar a la investigación educativa área prioritaria en los sucesivos Planes Nacionales de Investigación, a pesar de encontramos desde 1982 en tiempos de reforma continua.

Es probable que dichos responsables político-administrativos rechacen estas imputaciones afirmando que se están estableciendo y se establecerán procedimientos, reales decretos e, incluso, instituciones encargadas de ofrecer esa imagen, más válida y comparable internacionalmente, sobre los problemas, la coordinación, la organización, el funcionamiento y, desde luego, los resultados y logros del sistema educativo. Ésta es, se dice, una de las funciones de la LOGSE, una ley más compleja de lo que pudiera parecer. En ella, como es sabido, se dedica todo un título (Título IV. De la Calidad de la Enseñanza; Arts. 61 y 62) a dicho objetivo (MEC 1990).

Aquí quisiera abordar las implicaciones de dicho título, utilizando como referencia lo que en nuestro país tenemos, la información y los precedentes de que disponemos y los ejemplos internacionales que, seguramente, determinarán lo que nos ocurra. Sin duda, tendré que adentrarme en el terreno resbaladizo de la predicción, y por una vez tan siquiera en mi vida, más con ánimo de equivocarme que de acertar. Si es que las negras nubes de lo que ocurre en los así llamados países desarrollados terminan convirtiéndose en el paisaje modelo a partir del cual conformaremos nuestra realidad.

Dado que la moda actualmente predominante, ha reducido el ejercicio intelectual en educación a la formulación de una serie de sencillas preguntas con la equivocada idea de que las respuestas habrán de ser, como mínimo, tan sencillas como las cuestiones, yo quisiera también plantear en forma de epígrafes dos preguntas elementales (¿por qué y para qué evaluar un sistema educativo?; ¿cómo y con qué evaluarlo?), con

[*] Quisiera agradecer a Miguel Sola, Paco Beltrán y José Contreras Domingo, su ayuda, sus consejos y su apoyo durante la realización de este trabajo.

[1] Aunque el informe del CIDE (MEC 1992) aporta información sobre la desigualdad, no poseemos evidencia empírica, ni análisis teóricos, sobre las variables (administrativas, económicas, socio-culturales, escolares, curriculares, pedagógicas e individuales) que inciden en el mantenimiento de la misma.

[2] En parte explicable por una cierta cultura política que ha empobrecido el interés de la sociedad civil por los servicios que paga, mantiene y necesita.

la diferencia, en mi caso, de mostrar que las respuestas son bastante más complejas, enrevesadas y polémicas de lo que en un primer momento podrían parecer.

Pero antes de entrar en materia quisiera introducir una pequeña digresión terminológica que creo que nos podrá ayudar posteriormente.

Las distintas acepciones del término evaluación o por qué no todos los conceptos significan lo mismo

Por razones más políticas que lingüísticas, el concepto de "evaluación" tiene más sinónimos de los que un buen diccionario de la lengua española podría soportar. No puedo, por ello, atreverme a asegurar que las distinciones que yo plantee a continuación sean las únicas pertinentes. En todo caso, no pretendo más que aclarar el enmarañado "mapa conceptual" que el término evaluación padece y soporta.

En castellano es frecuente utilizar el término evaluación para toda una gama de procesos que vienen, sin embargo, diferenciados en la literatura científica anglosajona. Así, por ejemplo, dicho término genérico abarca la evaluación de *currícula*, de centros, alumnos/as y de profesores/as, e incluso, lo que en psicología se denomina "evaluación conductual". Además cuando se hace referencia a los alumnos/as, encontramos otros términos asociados al mismo, utilizados en la misma práctica, como los de "control", "pruebas" y "exámenes" (Álvarez Méndez, 1985; Pérez Gómez y Gimeno Sacristán 1992)[3].

Sin embargo, en los países anglosajones se suele hacer una distinción entre *evaluation* y *assessment* Mientras el primer término puede traducirse sencillamente por "evaluación", la traducción del segundo abriga problemas casi insuperables[4]. Pero podemos aclararnos si situamos a ambos conceptos en dos planos de significación interrelacionados: uno con respecto al sentido y otro con respecto al objeto de ambos conceptos.

Según el primer plano –el sentido–, evaluación hace referencia a todo proceso orientado al conocimiento de la *calidad del servicio educativo prestado*, mientras que *assessment* lo haría con respecto al "impacto" que dicho servicio ha tenido sobre los *receptores* [5]. Según el segundo, la distinción se sitúa en el "objeto" del proceso: la evaluación se ocupa de la influencia en la calidad educativa que puedan tener los distintos y diversos componentes de los procesos de enseñanza/aprendizaje (*curriculum*, metodología, experiencias de aprendizaje, organización de los centros de enseñanza, administración educativa –provincial, autonómica, estatal–, los docentes y los mismos alumnos/as); mientras que *assessment* se orientaría única y exclusivamente a los individuos o grupos de individuos receptores del servicio (alumnos/as) (MacDonald, 1976; Angulo, 1990; Sriven, 1991) (véase Cuadro I).

Podemos acordar desde estas diferencias dos acepciones del término evaluación: una *acepción amplia* que correspondería al proceso por el cual conocemos y valoramos la calidad del servicio y el papel de los distintos componentes en el mismo y otra *acepción estrecha* que señalaría el proceso o procedimiento a través del cual averiguamos su "calidad" únicamente por medio del impacto de dicho servicio en individuos o grupos de individuos receptores del mismo, restringiendo la comprensión de calidad del servicio educativo

CUADRO I

	Sentido	**Objeto**	
Evaluación	Calidad del servicio educativo prestado	Componentes del servicio educativo	Evaluación: Acepción Amplia
Assessment	Impacto en los receptores del servicio (alumnos/as)	Individuos o grupos de individuos	Evaluación: Acepción Estrecha

[3] En lo que sigue se entenderá por exámenes, principalmente, las pruebas estandarizadas. De todas formas, un examen, ya sea el confeccionado por un docente o por un psicómetra es igualmente un tipo de test.

[4] El término *assessment* deriva según Conner (1991) de la expresión latina *assidere*, que significa "sentarse al lado del", aunque en el diccionario latín/español sólo aparece la palabra *assideo* que además de significar "estar sentado junto a", acepta como traducción la de "asediar". En los diccionarios español/inglés dicho concepto viene traducido como evaluación y como tasación. Siendo un tanto críticos, la utilización de dicho concepto en los países anglosajones, teórica y prácticamente, está más cerca de la última, "*tasación*", que de evaluación. Con lo que *Assessment* supondría *sensu stricto* "tasar algo". Pero nadie estaría dispuesto, ni siquiera los anglosajones, a aceptar semejante traducción.

[5] Así es como se ha concebido en Estados Unidos de Norteamérica el *National Assessment of Educational Performance* (Evaluación Nacional de la Actuación /logro/ Educativa), y en el Reino Unido el *Assessment of Performance Unit* (unidad de Evaluación de la Actuación /logro/).

a su incidencia sobre los receptores (alumnos/as)[6]. Hay que recordar, además, que la acepción estrecha de la evaluación subyace a algunos términos usados cotidianamente como los de exámenes y controles, y a otros más técnicos como los de pruebas nacionales, indicadores de calidad educativa, mediciones de rendimiento, sistemas de tests, etc.

Estas primeras aclaraciones concordarían con la idea de "evaluación del sistema educativo" utilizada y manejada en los países que tienen establecidos sistemas y procedimientos institucionalizados al efecto. Pero, resulta necesario profundizar un poco más introduciendo un nuevo concepto: "los tests".

Es bien conocido que un test no es más que un instrumento psicométrico de medida. Pero su importancia en el tema que nos ocupa estriba precisamente en que desde sus orígenes, a finales del siglo XIX, ha sido considerado como el instrumento más objetivo, válido y fiable, es decir científico, de medición de características psicológicas y dominios educativos, como la inteligencia o los conocimientos académicos (Gould, 1984; Madaus, 1985a, 1985b, 1988a).

Aceptando que los tests representan la esencia de la medición científica en psicología y en educación, los adeptos al concepto estrecho de evaluación creen y defienden que su utilización distingue clara y rigurosamente entre aquellos individuos que han aprovechado las posibilidades que el sistema de enseñanza les brindaba de aquellos que no; discrimina entre los que efectivamente están preparados, y pueden ser seleccionados, para adentrarse en los escalones superiores de dicho sistema y aquellos que tienen que abandonarlo o cursar líneas paralelas y marginales, e, incluso, compara el "valor diferencial" entre componentes del servicio educativo, en razón de su impacto en individuos o grupos de individuos: los "alumnos/as".

Por ello, en muchas ocasiones el empleo masivo de sistemas de tests ha sido sinónimo de evaluación (véase Fig. l). Evaluar un sistema educativo equivale a disponer y emplear tests; los tests que la comunidad de psicómetras ha construido, sin importar, en muchas ocasiones, ninguna otra cuestión. No le falta razón a Madaus (1990) cuando afirma que actualmente se tendría que hablar con propiedad de la tecnología de los tests", como una de las tecnologías más "desarrolladas" y con más importancia social y política de las que podamos conocer hoy día en las ciencias sociales, muy superior si cabe al beneplácito social, académico y político del que gozó en su época (y puede que, lamentablemente, goce todavía) la tecnología de la modificación de conducta.

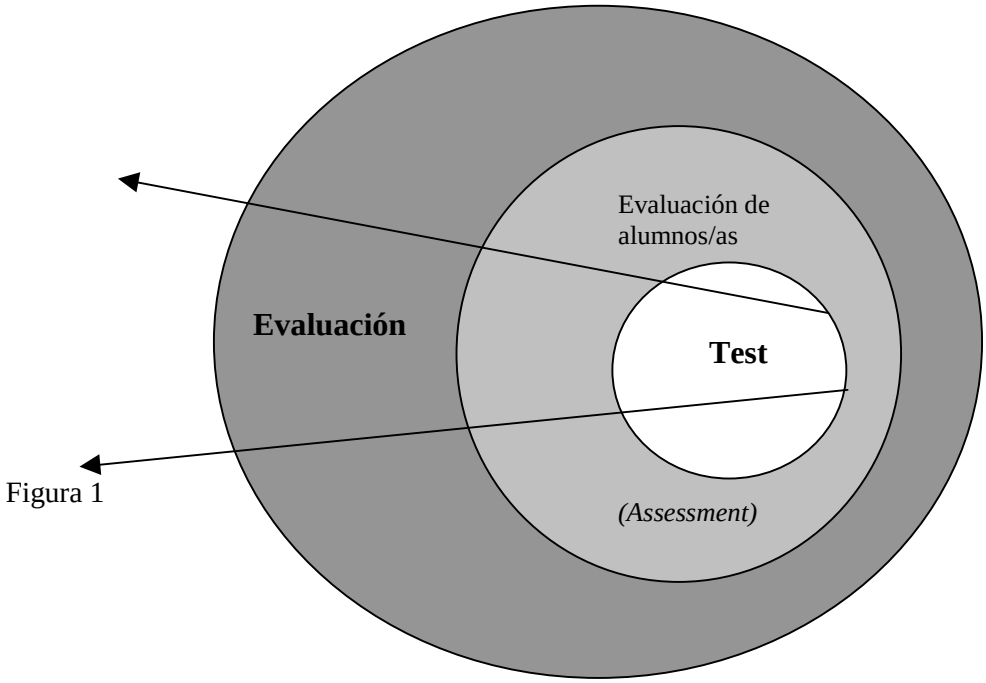

Figura 1

[6] En los países anglosajones se suele hacer uso de una versión extralimitada del concepto de *assessment*, que apoyada en la política conservadora que ha imperado en dichos países en los últimos quince años, ha cobrado un fuerte vigor. Ya no se habla, o se habla apenas, de evaluación del profesorado, sino de "*teachers assessment*", "*assessment*" de profesores/as.

No obstante, en todo ello se suelen olvidar varias cosas. En *primer lugar*, se olvida que, como decía al principio, un test es un *instrumento de medición*, que recoge información muy determinada sobre un individuo, un grupo de individuos o la actuación de una institución (a través de sus individuos). Pero no es, ni muchísimo menos, el único instrumento que tenemos a nuestra disposición. Puede que si sólo nos limitamos a la acepción estrecha de evaluación, los tests sean los instrumentos fundamentales, pero ni siquiera en este caso son los únicos. Si por el contrario nos sentimos cercanos a la acepción amplia, los tests son uno de los muchos instrumentos que podemos emplear (véase Cuadro II) y ni mucho menos, el más importante.

En *segundo lugar*, los test *como instrumentos* sólo son una muestra o selección de "preguntas o situaciones –llamados ítems– tomadas de un dominio de contenido o de un universo de interés" (Madaus, 1988b, pág. 30). El psicómetra tiene que asegurar que dicha selección sea representativa del dominio en general. Lo que técnicamente se denomina validez del test[7]. Lo que los tests no nos señalan son las causas que subyacen a las respuestas (positivas o negativas), es decir, las razones educativas por las que ciertas puntuaciones son altas o bajas, ni tampoco las vías de solución o mejora de los problemas que puedan detectarse.

Al equiparar evaluación y tests limitamos seriamente el significado, la riqueza y la complejidad de la primera instrumentalizándola. Por tanto, decidir sobre el futuro de un individuo o grupos de individuos, sobre la actuación y profesionalidad de un docente, sobre el estado o la calidad de un sistema, una organización, un *curriculum* o unas experiencias de enseñanza/aprendizaje, o el rumbo que han de adoptar en el futuro las políticas públicas en educación, basándose única y exclusivamente en las puntuaciones obtenidas a partir de ciertos tests, por numerosos que ellos sean, es un ejercicio cuestionable e irresponsable educativa y políticamente (Madaus, 1988b, Stake, 1989).

Podemos concluir que la adopción de la acepción estrecha de evaluación o de la tecnología de los tests, como marco teórico para la evaluación del sistema educativo, merma seriamente el conocimiento y la comprensión de la calidad del mismo, reduciéndolo a la exclusiva recogida de información, con unos instrumentos técnica y educativamente limitados, en los puntos terminales del sistema: centros de enseñanza, docentes y alumnos/as. Aunque, desde luego, deja a salvo los iniciales: los primeros escalones administrativos y políticos de la toma de decisiones educativas.

En todo caso, parece sobreentenderse que la calidad misma depende únicamente de los "resultados" del servicio en los alumnos/as. Lo que suele llevar como consecuencia que la responsabilidad de los docentes y de los centros escolares, en la calidad del servicio prestado, venga determinada –casi exclusivamente– por las "tasas de resultados" de sus alumnos/as. Ello ha propiciado, generado y fortalecido concepciones simplistas del cambio educativo y de la toma de decisiones, más cercanas a visiones tecnocráticas y economicistas de su funcionamiento, necesidades y problemas que a un interés educativo por informar inteligentemente a los ciudadanos y actuar democráticamente en el sistema (Angulo, Contreras y Santos, 1991; Angulo, 1992c; Elliott, 1992a, 1992b).

En las páginas siguientes, tomaré como eje de mi discurso y como objeto de crítica la acepción estrecha de evaluación.

¿Por qué y para qué evaluar el sistema educativo?

Es ahora pues el momento de plantearnos e intentar responder a la primera pregunta que estructura este trabajo. Si la formulásemos a "expertos" fuera y dentro de la Administración como Walberg y cols. (1990), Alvaro Page e Izquierdo (1992, 1992), sus respuestas serán, desde luego, muy semejantes. La evaluación del sistema trata de ofrecer la información que los gobiernos necesitan para asegurar la eficacia y la eficiencia del sistema, ya que es una exigencia prioritaria conocer la relación "coste-beneficio del producto educativo". Incluso la respuesta de una gestora política Granheim (1990), de un país como Noruega, mucho más descentralizado que el nuestro y con mayor experiencia democrática, es sencillamente que la evaluación de un sistema educativo se ha de realizar para mejorar la calidad de la educación, es decir, para mantener los

[7] Dicho de modo más técnico: la validez de un test estriba en "el grado en el que una inferencia particular y cualquier descripción o decisión resultante sobre un individuo o institución, hecha en razón de los resultados es apropiada o significativa. No existe –añade Madaus– una validez general de los tests, ni un test es válido en abstracto" (1988a, págs. 32-33).

estándares en las escuelas, sostener los valores nacionales y conocer hasta qué punto los objetivos nacionales se están cumpliendo[8].

Lo que tienen en común ambos tipos de respuesta, ya sea la más crudamente economicista (y conservadora) de los primeros, o la más atemperada, pero no menos administrativa y burocrática de la segunda, es su *simplicidad*, su *falsa neutralidad*, y el *juego político de las "necesidades nacionales"* en el que quieren introducimos a poco que lo permitamos. Responder al porqué y al para qué, requiere, por el contrario, adoptar un marco de análisis mucho más amplio y mucho más controvertido.

Para empezar quisiera retomar una vieja idea del hoy olvidado Althusser (1975)[9], según la cual un sistema educativo es fundamentalmente un aparato ideológico del Estado. Me interesa aquí centrarme más en la idea de aparato que en la de ideológico. El sistema educativo está sujeto a los mismos avatares que padece el Estado (Hargreaves, 1989a, 1989b; Dale, 1989); es decir, las crisis de los estados modernos se reflejan en el sistema educativo. Otra manera de enunciarlo es afirmar que los sistemas educativos suelen ser tomados como instrumentos privilegiados para atenuar, paliar o, ingenuamente, solucionar las crisis que los Estados padecen sin solución[10] (Shapiro, 1985; Broadfoot, 1983).

Debe quedar claro que el origen de este tipo de crisis se ha de situar en la propia dinámica de *confluencia/contradicción* en la que se encuentran la esfera estatal de organización, intervención y regulación y la esfera económica capitalista de producción, consumo y apropiación del plus-valor. Incluso tras las políticas supuestamente no intervencionistas de los gobiernos conservadores (en Norteamérica y en el Reino Unido) se esconde la intención, no siempre confesada, de mantener el papel absolutamente indispensable que el Estado tiene para crear las condiciones de acumulación de beneficios del capital y asegurar un grado mínimo de armonía social (Shapiro, 1985). No se trataría, como tanto se ha afirmado, de desmantelar el Estado, sino de encauzar su gestión para reforzar las necesidades de la economía capitalista. Lo que a su manera concuerda con las propias necesidades de supervivencia autónoma que el Estado (como aparato administrativo) requiere.

Pero aunque las crisis de los Estados tengan su origen en la economía, se vehiculan, sin embargo, a través de unos ámbitos particulares que, tal como Habermas (1975) ha señalado, se pueden denominar respectivamente de racionalidad, legitimidad y motivación. La primera está relacionada con la incapacidad estructural de gestión administrativa y organización de la intervención a la que tiene que hacer frente el Estado; la segunda con los déficit de justificación pública, consenso social y aceptación y cohesión cultural necesarios para llevar a cabo las políticas económicas y de gestión del sistema; y, la tercera, lo está con la paulatina falta de confianza en la esfera privada e interclasista de promoción social y económica, en la disolución general de la ideología del rendimiento, del justo-reparto y del esfuerzo individual recompensado (elementos esenciales para la estabilidad de la sociedad capitalista y para la supervivencia misma de la clase media). Como ha recordado McCarthy (1987, págs. 429 y sgs.) estas dos últimas crisis –legitimación y motivación– pueden ser consideradas como dos caras de un mismo ámbito socio-cultural. Pero, no obstante, aquí las trataremos por separado[11].

Las diferentes formas de crisis, por tanto, reverberan en el sistema educativo, al igual que en otras esferas sociales, políticas y administrativas, aunque adoptando en dicho caso características distintivas (Hargreaves, 1989a, 1989b). También puede considerarse de modo propio las *estrategias* que los Estados adoptan para contrarrestar los efectos de la crisis y para satisfacer las necesidades económicas. Según Weiler (1983), las estrategias básicas son tres[12]: "legislación, conocimiento experto y participación". Con la primera se promueven nuevas leyes y reglamentaciones, con la segunda se asegura la legitimidad de las nuevas

---

[8] Compárese con lo que se dice en el *Libro Blanco para la Reforma del Sistema Educativo* (MEC 1990): "Las administraciones educativas requieren un conocimiento preciso de las condiciones del sistema en cada momento para tomar decisiones adecuadas sobre su ordenación y gestión; deben además velar por el cumplimiento de los objetivos de cada nivel y ciclo educativo. La planificación de la educación sólo puede basarse en un conocimiento detallado de la situación educativa en las diferentes zonas y sectores" (pág. 258).

[9] También Habermas (1975) y Foucault (1981), entre otros, han reconocido que los sistemas educativos son parte de la organización de los estados modernos.

[10] Ambas argumentaciones han sido empleadas por los gobiernos conservadores inglés y norteamericano para introducir sus reformas políticas conservadoras en sus particulares sistemas educativos. Un ejemplo se encuentra en el informe del NCEE (1983) "Una Nación en Peligro". Sobre la "crisis" véase O'Connor (1989).

[11] Una excelente crítica de las argumentaciones de Habermas se encuentra en McCarthy (1987, págs. 431 y sgs.). Tengo que agradecer a mi compañero A. I. Pérez Gómez su indicación sobre la importancia de considerar, al igual que hace McCarthy, el sentido socio-cultural de ambas crisis.

[12] Aunque dicho autor las circunscribe al ámbito estricto de la educación, probablemente, estas tres estrategias tengan lugar en otros ámbitos como Sanidad, Justicia y Trabajo.

políticas asegurando su validez a través del uso del conocimiento experto y, con la tercera, se propicia la participación "condicionada» de los sectores sociales, económicos, laborales y ciudadanos implicados[13].

Con estas pocas aclaraciones creo que ya estamos en condiciones de empezar a entender el Cuadro II. En él se intenta reflejar el desarrollo de la política educativa y curricular de los pasados diez años y su orientación en los próximos, combinando los tipos de crisis y las estrategias de solución, en razón de las demandas que el capitalismo avanzado hace al Estado directamente, y al sistema educativo como aparato del mismo.

Estas demandas pueden ser resumidas muy brevemente en las siguientes: a) un funcionamiento del sistema educativo más acorde con criterios económicos de producción, eficacia/eficiencia y control del gasto; b) preparación de mano de obra cualificada o disposición a la cualificación posterior y permanente en el puesto de trabajo[14]; c) diferenciación/selección individual de los sujetos; d) elevación de las tasas de desarrollo de conocimiento científico-técnico; e) cohesión y homogeneización cultural y social de los ciudadanos/as; y f) flexibilización de las estructuras burocráticas de gestión del sistema (Bell, 1976; Levin, 1990).

CUADRO II

Ámbitos / Estrategias		Socio/Cultural	
	Racionalidad	Legitimación	Motivación
Legislativas	Leyes de nacionalización del sistema (LOGSE)	Curriculum nacional/estatal centralizado	Sistema de acreditación/selec-ción meritocrático
Conocimiento experto	Estructuras/Racionales: niveles de concreción curricular. DCBs.	Contenidos culturales avalados por expertos. Especialmente: Matemáticas/cien-cias y lenguaje	Pruebas/Exámenes nacionales (Sistema de Tests *Assessment*)
Participación	Leyes de racionalización de la participación (LODE)	Participación constreñida/dependiente	Sociedad Civil como *cliente* del sistema educativo. Competición entre centros de enseñanza

Aunque el cuadro necesitaría mucha más explicación sobre cada una de dichas demandas, quisiera centrarme en algunos de los cruces, teniendo en cuenta una cuestión importante en lo que respecta a nuestro país. Se ha alegado en muchas ocasiones desde 1982 que la crisis nacional provenía de un pasado autoritario, irracional y oligárquico, pero tenemos que tener en cuenta ahora, a principios de la década de los años noventa, que la crisis por la que estamos pasando es, probablemente, producto de las políticas que pretendieron solucionar los problemas estructurales heredados del franquismo.

[13] La primera estrategia, la legislativa, depende en gran medida de la estatalización de la política; en casos como el norteamericano, los tribunales juegan un papel muy importante; en casos como el español, son las reglamentaciones legales, las disposiciones, las órdenes y las circulares las que predominan (Boli, Ramírez y Meyer, 1990; Angulo, 1992a). La segunda estrategia ha resultado prioritaria para el capitalismo en general, pero especialmente para el capitalismo post-industrial. La tercera es una concesión a la democracia formal, tal como puede verse en el informe de la trilateral Crozier, Huntington y Watanuki (1975).

[14] Tengo que agradecer a F. Beltrán que me señalase la importancia de esta idea.

En este sentido, la LOGSE no es una respuesta a la crisis del sistema educativo anterior a la democracia, sino una respuesta legislativa a la actual situación, a las presiones que las necesidades económicas y la crisis mundial del capitalismo avanzado ejercen sobre el Estado y sobre el "aparato educativo" y a las propias necesidades del Estado en tanto estructura administrativo-burocrática.

¿Cuál es el tipo de respuestas que la LOGSE aporta, teniendo en cuenta las demandas de la esfera económica y los distintos ámbitos en los que la crisis se expresa? Si leemos atentamente el cuadro podremos darnos cuenta, en parte, del papel que dicha ley general juega en cuanto a los dos ámbitos de racionalidad y legitimidad:

1. Reforzar la cohesión y socialización de los ciudadanos/as, permitiendo que el Estado continúe ejerciendo una política centralista y *heredada de nuestro inmediato pasado* (Beltrán Llavador, 1991) con respecto a los contenidos curriculares, es decir, *culturales* (determinación por el Estado de los mínimos curriculares).
2. Introducir una estructura curricular racional (los DCB) justificada científicamente con una fuerte capacidad de asimilación administrativa, cuya retórica ha sido, no por impuesta desde arriba menos rápidamente asumida por algunos grupos académicos, editoriales y docentes (Angulo, 1992).
3. Flexibilizar la estructura curricular administrativa permitiendo la participación de docentes en la conformación de los *curricula* de los centros, pero imponiéndoles la dirección y la orientación de la gestión curricular misma. Dirección y orientación que se deduce tanto de los objetivos y contenidos legislados, como de las formas de organización curricular introducidas (Angulo, 1993, Levin, 1990).
4. Reforzar la formación profesional y los conocimientos básicos y especializados en Matemáticas, Lenguaje y Ciencias. (Cuya importancia dependerá en última instancia de la evaluación.)[15]

Si esto es lo que, entre otras cosas, pretende la LOGSE para atenuar las posibles crisis de legitimación y de racionalidad, tenemos que formular ahora la pregunta clave de este apartado: ¿qué papel juega la evaluación del sistema educativo (para qué y por qué evaluar), cuestión que viene inicialmente regulada en dicha ley?

La respuesta a esta pregunta es múltiple y no unívoca, y debe ser hecha teniendo en cuenta los actuales parámetros y exigencias que sobre el sistema educativo realiza el capitalismo avanzado y que antes señalamos. En el Cuadro III he intentado resumir y clasificar las distintas funciones que la evaluación del sistema educativo jugará en los próximos años. Dichas funciones están clasificadas, según su *importancia* real para el Estado y para el capitalismo, en primarias, secundarias y terciarias. En el Cuadro IV estas mismas funciones vienen clasificadas según el grado en que serán asumidas públicamente por la Administración educativa: explícita, intermedia o implícita.

Como puede verse, las funciones terciarias son calificadas en dicho cuadro de explícitas (y retóricas) en razón de que servirán de eslóganes políticos para recoger cierta aceptación general y legitimidad política, dado que su legitimidad técnica dependerá directamente de la validez del conocimiento experto empleado en la construcción de las pruebas o procedimientos de evaluación.

CUADRO III
Funciones de la evaluación del y en el sistema educativo

Funciones primarias	
1. Selección de individuos	Ideología meritocrática. Legitimación del acceso a las posiciones de mayor *status* social y de la desigualdad general según rendimientos.
2. Control administrativo	Rendimiento del sistema. Determinación/estandarización de las prácticas curriculares en los centros. Competitividad entre centros y docentes.
3. Gestión productivista del sistema	Sociedad civil como cliente del sistema educativo. Sobre-

[15] Aunque puedan leerse como cuestiones negativas, la LOGSE tiene puntos positivos, entre los que destaca la extensión de la escolaridad (no plenamente lograda, sin embargo, en los tramos iniciales).

	estimación de los rendimientos del sistema, de su eficacia y de su eficiencia. Disminución de costes/reparto diferencial –según rendimientos– de los recursos.
Funciones secundarias	
4. Reforzamiento de la homogeneidad cultural	Extensión por el sistema de la estandarización cultural común y equiparable nacional e internacionalmente. Depreciación del multiculturalismo. Hegemonía de la cultura de clase media. Sobre estimación de la cultura de elites.
5. Valoración de aprendizajes, contenidos y procesos curriculares	Legitimación y potenciación de ciertos aprendizajes y contenidos concretos, en razón de lo que midan las pruebas de evaluación. Hincapié en contenidos básicos, especialmente matemáticas, ciencias, y lenguaje. Posición marginal de otros contenidos. Especificación externa de las relaciones admisibles entre docentes, alumnos/as y contenidos culturales.
Funciones terciarias	
6. Motivación individual	Certificación. Motivación para el aprendizaje y el logro. Creación de expectativas para la movilidad social. Creación de expectativas para la seguridad y el bienestar económico y profesional.
7. Movilidad social	Posibilidad de ascenso en la estructura social de clases y confirmación/legitimación de la posición social adquirida.
8. Información sobre el sistema educativo	Orientación de la política educativa y curricular. Apreciación de la calidad educativa del sistema. Conocimiento de los problemas, logros y determinantes del mismo.

CUADRO IV

Implícitas (ocultas).	**Intermedias**	**Explícitas (retóricas)**
1. Selección de individuos.	4. Reforzamiento de la homogeneidad cultural.	6. Motivación individual.
2. Control administrativo.	5. Valoración de aprendizajes, contenidos y procesos curriculares.	7. Movilidad social.
3. Gestión productivista del sistema.		8. Información sobre el sistema educativo.

Las funciones secundarias son calificadas de intermedias, queriendo significar que, si no por completo, parte de su temática será defendida y aceptada públicamente. Sin embargo, las implicaciones internas negativas de la homogeneidad cultural (como la estandarización) o de la valoración de ciertos aprendizajes, contenidos y procesos (como pueda ser el reforzamiento de aprendizajes memorísticos, la propensión a centrarse en procesos mentales de bajo nivel cognitivo o el reforzamiento de la recitación) serán ocultados o, en todo caso, expresamente negados.

Por último, las funciones primarias son a mi juicio las más importantes que a principios de los noventa están cumpliendo los sistemas de evaluación en el mundo occidental y especialmente en los países anglosajones; y dado que son funciones *ocultas*, y no sólo negadas, por los gobiernos y administraciones educativas, requieren ser tratadas con mucho más detenimiento que las anteriores, especialmente las de Control administrativo y Gestión productivista del Sistema.

Control administrativo y gestión productivista.
Dos funciones esenciales y ocultas de la evaluación

En realidad la importancia que le concedo a dichas funciones no se encuentra en su carácter oculto, sino en la importancia que han ido adquiriendo en los últimos años, importancia que incluso está haciendo que los gobiernos más conservadoramente radicales (como el inglés) las hagan explícitas en buena medida formando parte de sus políticas educativas.

Las pruebas y los exámenes siempre han existido de una u otra manera en los países occidentales desarrollados y han ido cobrando fuerza e importancia en los países no desarrollados (Broadfoot, Murphy y Torrance, 1990; Moreno Olmedilla, 1992). Sin embargo, en la década de los ochenta un fenómeno vino a resituar políticamente la función de la evaluación dentro de los sistemas educativos.

En general, la entrada del discurso conservador en los dos países anglosajones con más peso internacional supuso la explicitación pública de criterios economicistas de gestión y el reforzamiento del control administrativo de la labor de los distintos sub-sistemas estatales. En algunos casos, esto tuvo como consecuencia la privatización de ciertas competencias del Estado, dejando su funcionamiento en manos de la empresa privada (descentralización por el mercado, para ser más exactos). En otros, como la educación, adoptó perfiles diferentes. La educación no ha dejado de ser un sector (o sub-sistema) de la administración pública (estatal, federal o nacional), pero en él, y en paralelo a la privatización de otros sectores, se han reforzado criterios economicistas de gestión y control administrativo del mismo (Henkel, 1991) (centralización por el mercado).

Ni una ni otra cuestión han sido ciertamente novedades. Al igual que las pruebas y los exámenes, el rendimiento educativo y ciertos tipos de control administrativo han centrado parte de las preocupaciones públicas de los gobiernos occidentales. Lo que ahora resulta significativamente distinto es que mientras el sistema educativo permanece en la órbita de influencia estatal (centralización), la privatización, además de intensificarse, se ha efectuado incrementando y extendiendo a lo largo del sistema y a todos los sectores base del mismo, los criterios economicistas de funcionamiento (por el mercado): centros educativos, docentes y alumnos/as.

Esta nueva orientación, todavía inconclusa pero que perfila nuestro inmediato futuro, se caracteriza por varias cuestiones:

1. Existe una marcada tendencia en las políticas públicas en educación a unir lo que en principio ha podido permanecer relativamente desconexionado: la implantación de un *currículum* Nacional (centralizado) y de un Sistema Nacional de Evaluación. El primer movimiento en este sentido ha sido la defensa pública (de la mano de grupos de expertos y rápidamente asumida por la administración educativa) de la ecuación, según la cual, para reformar el sistema educativo había que reformar –incrementar– el sistema de evaluación (o el sistema de tests). Dicha ecuación se sustenta en la creencia de que un sistema de evaluación (o de tests) nacional representaría los estándares, expectativas y criterios de excelencia que el sistema educativo (escuelas, docentes y alumnos/as) debería alcanzar (Madaus, 1985a, 1985b; Airasian, 1987; Hartnett y Naish, 1990; Linn, 1993; Bryk y Hermanson, 1993). Una vez aceptado este principio, ambas naciones han tomado rumbos convergentes, dependiendo de la tradición e historia política particular.

   Norteamérica y Gran Bretaña han sido naciones que se han distinguido por su descentralización educativa, en donde la expresión "*curriculum* común" ha sido tenida por anatema hasta los años ochenta. En Norteamérica ha existido una amplia aceptación social, política y científica de los sistemas de tests y del mantenimiento de ciertos procesos y ciertas instituciones encargadas de efectuar mediciones sobre el rendimiento de los alumnos. Gran Bretaña, por su lado, junto a su descentralización curricular, se ha distinguido, sin embargo, por un fuerte y selectivo sistema de exámenes al final de la escolaridad, puerta de acceso a la educación superior. Ambos países, aunque mucho más rápida y claramente en el caso británico, han comenzado a plantearse o han legislado, la necesidad de incrementar los sistemas de evaluación y conectarlos a un *curriculum* nacional, común para todos los ciudadanos (Hartnett y Naish, 1990).

   A diferencia de los casos anglosajones, la vía española se encuentra en una posición intermedia aunque orientada al mismo objetivo. Haciendo uso de nuestra histórica tradición, tenemos legalizado y en período de implantación un *Curriculum*, Nacional (centralizado) y estamos en vías (de la mano del Instituto de Calidad y Evaluación) de implantar un fuerte sistema de pruebas nacionales, exámenes o indicadores, que probablemente esté claramente conectado a la parte más

importante del *Currículum* Nacional: los criterios de evaluación de los mínimos curriculares, a lo que se añadirá otro tipo de pruebas y otros mecanismos de medición de rendimientos[16].

2. Además de la conexión *curriculum*-evaluación, los sistemas de evaluación que se institucionalicen o que ya se han institucionalizado tienen dos características: *primero*, se extenderán a lo largo de todo el sistema educativo, enfocándose en lo que se acuerde como "edades" clave de la escolaridad. *En segundo lugar*, se centrarán no sólo en los alumnos/as sino también en los centros y en los docentes; y se emplearán diferentes tipos de mediciones: indicadores de calidad, evaluaciones externas de centros, evaluación del profesorado, cuestionarios, exámenes y pruebas nacionales selectivas, pruebas nacionales (con muestras seleccionadas de alumnos/as) a lo largo de la escolaridad, etc., pero sin salirse de la acepción estrecha de evaluación.

   Es previsible, además, dada su extensión y periodicidad en el sistema, que los procesos de evaluación se sedimenten no sólo como artefactos externos impuestos a las instituciones educativas y a los sujetos dentro del sistema mismo, sino que lleguen a formar parte de los hábitos cotidianos, es decir, artefactos subjetivos que delimiten la concepción de enseñanza, educación y escolaridad que aceptemos de modo natural (Foucault, 1981).
3. La conjunción de la primera y segunda característica conlleva la entronización de una cultura política en la sociedad civil íntimamente relacionada con la tecnocracia (Angulo, 1992c). Esta cultura supone concebir que los sujetos, organizaciones y relaciones que dan cuerpo a la sociedad civil, antes que ciudadanos/as (y organizaciones ciudadanas) son *clientes del sistema*. A ellos y a ellas se les vende un producto y un rendimiento a un coste relativamente bajo: los resultados y logros concretos del sistema. Los clientes podrán elegir y criticar a las escuelas el trabajo docente y aceptar los resultados y la selección de sus hijos e hijas en razón de un sistema múltiple de evaluación, claramente conectado con los estándares curriculares y de rendimiento. A su vez, los distintos Estados podrán comparar internacionalmente su situación y, en caso de ser negativa, presionar de modo legítimo y en pro del mercado educativo, a la máquina escolar. La participación en el sistema educativo dejará de ser, por ello, participación ciudadana, convirtiéndose paulatinamente en mercantil (Angulo, 1993).
4. La conversión de ciudadanos/as en clientes está relacionada con un efecto importante que la evaluación introducirá en el funcionamiento del sistema educativo. El ejercicio de control administrativo y la gestión productivista son mucho más sutiles que lo que la unión entre sistema de evaluación nacional y *curriculum* pudiera dar a entender. Tanto una como otro se regirán por lo que he denominado anteriormente centralización por el mercado, que en paralelo a las nuevas formas de organización industrial (Angulo, 1993), significa centralizar la gestión y los objetivos de rendimiento descentralizando las responsabilidades. De tal manera que cada institución escolar, cada grupo de trabajo, ha de asumir sus propias decisiones curriculares y su forma interna de organizarse para alcanzar los objetivos de rendimiento previstos. Las escuelas pueden competir en el mercado de la educación porque son autónomas en la selección de las vías, aunque no lo son en la de los resultados que se deben lograr. Los sistemas de evaluación serán los encargados de hacer visibles y "controlar administrativamente" el bien "público" que los centros de enseñanza y los docentes estén produciendo en cada momento y a lo largo de la escolaridad. Los malos resultados serán debidos, no a una mala gestión administrativa, sino a una mala y defectuosa organización en las unidades de producción básicas: las escuelas y los docentes.
5. Por ello es muy probable que, tal como ha señalado Broadfoot (1986), la creciente atención a los sistemas de evaluación como instrumentos de la política pública y el predominio de las funciones primarias en particular, consigan realizar una vieja aspiración fracasada en los años setenta: introducir un control más directo y centralizado del *curriculum* y la cultura, el trabajo docente, la organización de los centros de enseñanza y las aspiraciones de los futuros ciudadanos/as. Es así como debemos entender la complementariedad evaluación/*curriculum*, la multiplicidad de las pruebas y los receptores de las mismas, y el sentido mercantil y productivista que cualifica a todas estas tendencias en conjunto.

[16] Entre las funciones señaladas para el Instituto Nacional de Calidad y Evaluación, el art. 3.1. señala: "Evaluar el grado de adquisición de las enseñanzas mínimas establecidas para los diversos niveles, etapas, ciclos y grados del sistema educativo, en el marco del currículo establecido por las administraciones educativas y de su desarrollo en los centros educativos".

Puede que todo lo que he expuesto sobre las funciones de la evaluación, lleve al pesimismo, pero en absoluto debe ser tomado como justificaciones de la inactividad y el derrotismo. A pesar de todo, evaluar un sistema educativo es un acto importante y necesario del que nadie puede excluirse y desinteresarse. Y aunque sólo sea porque en un futuro se tratará de un acto impuesto, a todos los que nos importa nuestro sistema educativo (es decir, ciudadanos/as, docentes, alumnos/as, asociaciones de padres y madres) deberíamos, cada uno desde nuestras respectivas responsabilidades, intentar movilizar con la evaluación del sistema otros fines distintos a los que seguramente éste adquirirá en manos exclusivas y monopolizadoras de expertos y de la Administración.

Aunque luego se expondrán otras estrategias alternativas, con respecto a las funciones mismas que acabo de analizar, propongo desde estas páginas las siguientes:

Primero, criticar sistemáticamente la ejecución de las funciones primarias y secundarias expuestas utilizando, entre otras medidas, la investigación teórica y empírica sobre los procesos que la Administración y los expertos decidan introducir: el impacto en el *curriculum*, los procesos de enseñanza/aprendizaje, en los centros escolares, en los docentes y en los alumnos/as de los procedimientos de evaluación que se implanten.

Segundo, radicalizar y potenciar la función informadora y motivadora de la evaluación, situándola en el terreno de la democracia participativa, es decir, en el terreno del debate ciudadano en el que amplios sectores sociales se sientan informados, y puedan participar. Como ha afirmado House (1990, pág. 23) "la evaluación pública debería ser una institución para democratizar las decisiones públicas, y para que dichas decisiones, los programas y las políticas estuvieran mucho más abiertas a la crítica y la deliberación pública".

Tercero, exigir y promover el cumplimiento de ciertas normas éticas, de justicia e igualdad social en todas aquellas pruebas que vayan a determinar el futuro de los individuos, con la intención de que no perpetúen y agraven las ya de por sí marcadas desigualdades sociales que todavía persisten en nuestro sistema educativo y en nuestra sociedad.

¿Cómo y con qué evaluar el sistema educativo?

Como ya expuse al principio, es seguro que la acepción de evaluación que predominará será la estrecha. Ya tenemos algunos ejemplos nacionales que confirman esta idea haciéndola más verosímil que hipotética. Dos de los más claros son la evaluación de la reforma de enseñanzas medias (Álvaro Page, 1988) y la participación española en el proyecto internacional de indicadores de calidad (Álvaro Page, 1992)[17], que ya han sido criticados en otro lugar (Angulo, 1992a, 1992b)[18].

Ambos ejemplos nacionales y la acepción de evaluación a la que remiten estarían englobados en lo que Berlak (1992) ha denominado "paradigma (o enfoque) psicométrico", de evaluación (MacDonald, 1988). Un enfoque que actualmente representa la tendencia hegemónica en evaluación (Wood, 1987), aunque su historia, lamentablemente no ha sido un dechado de virtudes éticas ni de justicia social. Desde sus inicios, la craneometría, de la mano de sus representantes más cualificados como Galton, Catell, Terman, Thornidike, Brigham, Yerkes, Spearman, o Cyril Burt, ha estado plagada de intencionales falsedades estadísticas, así como implicada en la promoción y defensa política del racismo, el elitismo y el hereditarismo (Gould), 1984; Karier, 1974)[19].

---

[17] Esto en realidad también concuerda con las contradicciones de los textos –legales o de principio– que han instituido la reforma de la LOGSE. Si en el artículo 62 de dicha ley se afirma que la evaluación se aplicará sobre los alumnos/as, profesorado, centros, procesos educativos y la propia Administración, al igual que en el Libro Blanco (MEC, 1989, pág. 259), en otro lugar se especifica que la "evaluación de la calidad del sistema educativo requiere la realización de estudios sobre muestras representativas de alumnos, profesores y centros", lo que viene a decir que, en la calidad del sistema, la Administración educativa no tiene ninguna responsabilidad.

[18] Otro ejemplo es el de las famosas "Cajas Negras" de riguroso secreto desvelado. Las "Cajas Negras" son un conjunto de instrumentos que se supone están evaluando en territorio MEC la implantación de la LOGSE.

[19] Un personaje como Spearman justificó las leyes norteamericanas de restricción de la inmigración de 1924 (Gould, 1984, pág. 285). El mismo Terman (quien adaptó la escala de inteligencia de Binet dando lugar a lo que hoy se conoce como los "tests de inteligencia") también llegó a afirmar que el mejor servicio que podrían hacer "los inferiores y los viciosos" (*i.e.* los pobres) sería evitar que sus genes se esparcieran por la sociedad (cit. por Karier, 1974, pág. 297). Otro ejemplo aún más inquietante nos lo ofrece el estado de Virginia en donde se practicaron, desde 1924 hasta 1972, 7.500 esterilizaciones de personas que habían sido consideradas idiotas según la escala Stanford-Binet, apoyándose en una sentencia de la Corte suprema que aprobaba como científicas y necesarias dichas esterilizaciones. Con estas citas no se trata de hacer acusaciones fáciles, ni de desprestigiar a una disciplina científica, se trata de no olvidar nuestro pasado para no repetirlo (Gould, 1984, págs. 354 y sgs.).

Pero como decía, la importancia social y política adquirida por la tecnología de los tests en evaluación (Madaus y McDonagh, 1979; Wood, 1987; Madaus y Kallagham, 1992) y su impacto en una sociedad y en sus individuos es actualmente mucho más importante de lo que pudiera parecer y de lo que su pasado podría justificar. Conviene, pues, que expongamos brevemente algunos de los supuestos más importantes que aún hoy subyacen a su desarrollo y a sus propuestas (Berlak, 1992)[20].

a) Certidumbre técnica

Quizá una de las razones más importantes para entender el influjo que han tenido la tecnología de los tests y las evaluaciones que hacen uso de dicha tecnología se encuentra justamente en su halo de certidumbre y rigurosidad (lo que puede denominarse cientificidad, y no carácter científico) con el que se ha solido presentar. En este sentido, el poder cuantificar los estados mentales, el funcionamiento de la psique, la inteligencia, el aprendizaje o ciertas características de la realidad individual, a través de ciertos instrumentos o procedimientos técnicos, supuso en su momento, y supone aún hoy día, la posibilidad de seleccionar, decidir y comparar rigurosa y científicamente (Gould, 1984).

Esta aceptación de la certidumbre técnica y la mística de la cuantificación que lleva aparejada, y cuya persistencia procede de los primeros logros psicométricos (como el "g" de Spearman)[21] se encuentran relacionadas directamente con ciertas creencias actuales, algunas de las cuales, como señala Wood (1987) paradójicamente bordean la creencia mágica:

- Se afirma y se defiende, que los sistemas de test mantendrán y aumentarán los estándares educativos, en razón de que lo que dichas pruebas miden es, sin discusión, aquello que la sociedad ha de aceptar racionalmente por valioso. Sin comprender que el concepto de estándar, de logro, y aun de rendimiento adquiere significaciones cambiantes histórica, social y políticamente.

CUADRO V
Características del enfoque psicométrico de evaluación

Certidumbre técnica
Unidad (estandarización) del significado
Fragmentación y descontextualización del aprendizaje y del conocimiento
Neutralidad valorativa de la técnica
Necesidad de control centralizado y dominio experto

- Se afirma que los rendimientos de los alumnos/as (el nivel) han descendido o bajado según los resultados que nos muestran los tests aplicados, sin reconocer que incluso en los mejores casos, cuando pueda demostrarse que un test tiene una alta validez (algo bastante más complejo de lo que parece), lo más que podemos asegurar es que el alumno/a ha respondido correcta o incorrectamente a la pregunta, pero no podemos asegurar si el alumno/a ha aprendido o domina el ámbito de contenido en general sobre el que se ha construido el test (Stake, 1989).

[20] En lo que sigue utilizo el trabajo de Berlak (1992), como referencia básica.
[21] Una excelente crítica del "g" de Spearman se encuentra en Gould (1984). Véase también sobre la política de la Medición de la Inteligencia el trabajo de Zigler y Seitz (1982).

– Se afirma que los sistemas educativos pueden compararse entre sí y año tras año según indicadores económicos y educativos determinados, sin tener en cuenta la historia social, política y educativa de cada uno, sin apreciar las diferencias culturales, su desarrollo económico, etc.[22]

El enfoque psicométrico ha justificado la aceptación de una especie de *ingeniería social* según la cual los instrumentos y ciertos algoritmos matemáticos nos permiten cuantificar el mundo social y psíquico, de tal manera que podemos prescindir incluso de nuestro juicio.

b) Unidad del significado

Este principio señala que existe o puede ser establecido un significado único que trascienda los contextos sociales e históricos, para las características, atributos o estados objeto de medición, llamados en el discurso psicométrico "constructos" (Cronbach y Meehl, 1955).

Constructos frecuentemente invocados son los de habilidad lectora y matemática, rendimiento, competencia, logro académico (en un área particular) o inteligencia. Está claro que si no se pudiera postular la univocidad del significado de tales constructos, es decir, si no pudiéramos cuantificarlos, las dificultades para construir instrumentos adecuados serían insuperables. En este tipo de premisas, por ejemplo, se sustenta la llamada evaluación criterial.

Pero además, la existencia de constructos está también relacionada con la validez de los propios instrumentos. Si la validez de un test pregunta si mide lo que pretende medir, está claro otra vez que si no disponemos de un significado estable y único de lo que mide, es decir, del constructo, su validez puede resentirse considerablemente o ser inviable.

Tomemos un ejemplo. Para construir y validar un test de logro académico según el enfoque psicométrico se necesitaría asegurar, como afirma Berlak, que "el constructo 'logro académico' posea un significado universal y estable, o que cierta unanimidad sobre significado es posible y deseable (Berlak, 1992, pág. 14). Pero, sin embargo, un constructo tan aséptico como el citado, exige un proceso bastante serio de clarificación inicial. Tendríamos que formular preguntas como las siguientes: ¿nos referimos a habilidades y conocimiento de hechos?, o ¿nos referimos a complejos procesos cognitivos (como comprensión, pensamiento crítico, o solución de problemas)? Incluso si hablamos de un constructo tan engañosamente sencillo como el de habilidad lectora[23], ¿nos estamos refiriendo a la suma de habilidades individuales, al reconocimiento de letras y palabras, a la asociación de letras y sonidos, a la pronunciación de palabras en una secuencia, al significado de una frase, a la utilización de estrategias de autocorrección en el proceso de lectura, etc.? (Cole, 1990)[24]

Los constructos, en general, son formas de concebir los "objetos de medición", y en principio como tales formas (o si se prefiere teorías) están sometidos a los mismos influjos que cualquier otra concepción sobre la realidad social y psíquica. Sin embargo, cuando "los objetos de medición" son términos no sólo psicológicos sino educativos o con repercusiones educativas, la pluralidad de significados, de creencias y suposiciones que les subyacen y los determinan son todavía mucho más cambiantes, conflictivas y contradictorias (Cole, 1990; Berlak, 1992).

Por tanto, resulta cuando menos preocupante la tendencia de las administraciones educativas a suplantar y suprimir la pluralidad de significados que albergan los "constructos" educativos, imponiendo sus particulares censuras (legales) de significado, como por ejemplo a través de los objetivos (contenidos mínimos), los criterios de evaluación o los instrumentos legalmente autorizados. Es probable que, actuando de esa manera la implantación de un sistema (psicométrico) de evaluación sea más fácil y factible; haciendo que la censura legal se convierta en racional. No obstante, el precio a pagar, no es sólo la simplificación de nuestras concepciones sobre la educación, la escolaridad, el aprendizaje, la enseñanza y la docencia, sino la

[22] Muchos docentes, que aceptan los supuestos del enfoque psicométrico, admiten que los "expertos psicómetras" utilicen sus herramientas científicas para poder segregar *racionalmente* a los alumnos difíciles, simplemente esperando que sus puntuaciones medidas por cierto test clasifiquen a dicho niño o niña como "deficiente".

[23] Otro constructo importante es el de inteligencia. Cuando se nos afirma que se está midiendo la inteligencia deberíamos preguntar, al menos, cuál es la que se está teniendo en cuenta y por qué: lógica-matemática, lingüística, musical, espacial, cinética, interpersonal e intrapersonal (Gardner y Hatch, 1989).

[24] El proceso de cuantificación, o definición del constructo rendimiento en la Evaluación de la Reforma de Medias es el siguiente: R= *f* (E). [El rendimiento es función del tipo de enseñanza; donde E= Variables aptitudinales, actitudinales, personales, familiares, contextuales y didácticas, es decir, R=*f* (Ap+Ac+Pe+Fa+Co+Di).] En el texto citado se afirma que esta es la parte aditiva del modelo, complementada por la multiplicativa y el error aleatorio de medida. (Page y cols. 1988, pág. 26 y sgs.).

pretensión de abolir la pluralidad y la diferencia, la discusión y el debate político en aras de su particular concepción *indiscutible*.

c) Fragmentación y descontextualización del aprendizaje y del conocimiento

El enfoque psicométrico mantiene una visión fragmentada del aprendizaje y del conocimiento, prácticamente desde sus orígenes, pero desde los años cincuenta, especialmente influida por el dominio de la psicología conductista y, posteriormente, por la combinación de ciertas teorías "cognitivistas". Este sentido de fragmentación puede ser detectado principalmente en la separación entre los ámbitos cognitivos, afectivos y conativos[25] y en su descontextualización.

Como ha señalado Raven, la separación artificial de los distintos ámbitos y su tratamiento como categorías diferentes distorsionan y oscurecen la riqueza y complejidad del aprendizaje humano (Raven, 1992). Por ejemplo, la medición de competencias debería ir aparejada con un conocimiento más profundo de los valores, intereses y preocupaciones de un alumno/a (pág. 89). Por tanto, resulta muy poco significativo, psicológica y educativamente, evaluar cierta habilidad sin, a su vez, reconocer el "placer" que el alumno/a experimenta por hacer algo, o su determinación para hacerlo. "Las habilidades no pueden evaluarse con independencia de los valores". En el dominio de una habilidad está implicado el valor que el alumno/a otorga a dicha destreza. La significatividad psicológica del aprendizaje es en realidad parte de la relevancia particular que dicho aprendizaje posee para un alumno/a (Pérez Gómez, 1991).

En conexión con lo anterior, las competencias o los aprendizajes en general no pueden ser, tampoco, sustraídos a los contextos sociales y culturales que determinan y modulan su relevancia y su valor (Brown, Collins y Duguid, 1989; Suchman, 1987; Bruner, 1990; Streibel, 1989). Nuestro sentido de realidad es una construcción socio-cultural (y no sólo individual) en la que se interrelacionan cognición, afectividad, conatividad y la acción práctica misma. La coherencia de nuestras acciones o el lenguaje no son ni resultado de nuestras predisposiciones individuales, ni, únicamente, procesos de codificación y transmisión de representaciones y símbolos, sino interacciones situadas, compromisos con la acción dentro de un contexto socio-cultural.

d) Neutralidad valorativa de la técnica

Este principio afirma que una vez construido un test o una prueba, de acuerdo a ciertos estándares aceptados, dicho test o prueba puede ser denominado instrumento científico; es decir, un instrumento neutral valorativamente y que sólo ha de ser juzgado en razón de su mérito científico (Berlak, 1992).

Sin embargo, es posible hacerle a dicho principio dos tipos de alegaciones complementarias. Según la primera, 1a complejidad y la precisión de una investigación (una evaluación o un instrumento) no garantizan su imparcialidad" (House, 1990, pág. 25). Según la segunda, cada instrumento, cada test o cada prueba implica una cierta "teoría" sobre la naturaleza del objeto evaluado. Dicho esto último de otra manera: la definición de un constructo es la definición sustantiva del « objeto" al que hace referencia, una forma de concebirlo y de entender sus cualidades y características.

Con respecto a la primera cuestión, a nadie se le escapa que los instrumentos técnicos están inmersos en un clima intelectual e ideológico en el que adquieren un sentido particular además de su propia legitimidad social. El clima en el que se desarrollan las evaluaciones afecta seriamente a la manera en que son llevadas a cabo, al uso de los instrumentos y las técnicas elegidos y a las conclusiones e interpretaciones que se extraigan de sus resultados. En la complejidad socio-cultural y económica en que se desenvuelven la mayoría de los procesos de evaluación, las interpretaciones finales pueden depender de aquellos grupos que, por su posición y jerarquía, hagan prevalecer sus interpretaciones como reflejo y confirmación de sus necesidades de legitimación (Gould, 1984; Fetterman, 1982; Palumbo, 1987; Madaus, 1988b).

Un claro ejemplo del reconocimiento de estas cuestiones lo encontramos en los requisitos necesarios para validar los tests. Según Messick (1989), la validez de un test está compuesta por el cruce entre la validez de las interpretaciones y de los *usos* a los que se destine, por un lado, y las evidencias y las *consecuencias*, por el otro. Sin entrar en pormenores más profundos, esto quiere decir que para validar nuestros instrumentos no podemos sustraernos a las implicaciones valorativas que les subyacen y a las consecuencias sociales que de su empleo se derivan. Messick (1989) recalca que la validez es un concepto unitario, en el que no se pueden separar los distintos ámbitos.

[25] Los componentes conativos de la conducta y la acción humana hacen referencia a la determinación, la persistencia y la voluntad.

La segunda cuestión introduce una problemática mucho más interesante. En primer lugar, nos señala que los datos, las mediciones, los instrumentos, "los constructos", los resultados, no son algo estático. Por el contrario, son utilizados activamente para comprobar y determinar, en la dinámica social en la que se encuentran inmersos los sujetos, los fines y los objetivos de la acción y los intereses políticos en juego. En este sentido pueden representar, y de hecho representan, una *ideología* de la realidad. Las evaluaciones legitiman realidades y percepciones.

Es decir, los constructos que sustentan los instrumentos que empleamos definen concepciones sobre la institución escolar, el conocimiento, los valores educativos, las relaciones interpersonales, el aprendizaje y los procesos de enseñanza. Los instrumentos no son neutrales porque imponen concepciones educativas particulares y legitiman ideologías políticas sobre la escolaridad. Al aceptar su neutralidad estamos aceptando subrepticiamente unas concepciones y unas ideologías que deberían concurrir en el terreno público de la discusión y el debate (Angulo, 1992b).

e) Necesidad de control centralizado y dominio experto

Incluso en sus variantes más ligeras, los sistemas tecnológicos exigen para su desarrollo y aplicación un fuerte grado de centralización en su gestión (Winner, 1986). Centralización que se amplifica si de la tecnología de los tests se trata, y si es dicha tecnología la que da cuerpo, fundamentalmente, a los sistemas de evaluación educativa. Como señala Berlak (1992, pág. 18), los "procedimientos de test son formas de 'vigilancia' cuyo uso se centra en la imposición de relaciones de poder... La razón de ser de todo proceso de evaluación en educación (...) se encuentra en modelar las acciones educativas, ejerciendo el control sobre profesores/as y/o estudiantes".

Pero lo que aquí me interesa señalar, además de lo que se afirmó anteriormente, es que la centralización supone acaparar el poder de *determinación* y de *interpretación* de lo que pueda ser considerado valioso y aceptable. En la evaluación, la legitimidad procede de las comunidades de expertos relacionados íntimamente con la Administración y el poder político. Las evaluaciones además de objetivar a los individuos y los procesos sociales, a través de la asignación de un lugar numérico –llamado rendimiento– en un orden preestablecido, sitúan la decisión y la comprensión fuera del escrutinio y la crítica pública. La apropiación por los expertos del significado del conocimiento y de las valoraciones, es algo más penetrante que su apropiación del poder público, es, sin duda, la apropiación de la democracia, de la creación y decisión política a la que los ciudadanos/as tienen derecho.

El control unidireccional del poder que los sistemas de evaluación ejercen, articulados a través del enfoque psicométrico, es una de las estructuras más avanzadas de control social que la tecnocracia educativa adoptará y está adoptando en el mundo occidental.

La necesidad de un enfoque educativo de evaluación

Como ya quedó claro al principio de este trabajo, la evaluación del sistema educativo es una necesidad prioritaria pero prescindible si dominan las funciones primarias y secundarias y si es el enfoque psicométrico el adoptado. Antes expuse algunas estrategias alternativas con respecto a las funciones. Aquí quisiera ofrecer esquemáticamente otras con respecto al enfoque que dicha evaluación debería adoptar. Este nuevo enfoque está justificado en los trabajos de ciertos psicómetras e investigadores descontentos con el rumbo de sus disciplinas, con la tecnología al uso y con su empleo, y comprometidos con el replanteamiento político de algunos procedimientos de evaluación (Irvine, Miles y Evans, 1979; Secada, 1989; Berlak y cols., 1992).

Una evaluación educativa del sistema educativo debería englobar la acepción amplia de evaluación. Esto significa al menos lo siguiente:

1. Evaluar el sistema educativo debería significar evaluar todo el sistema. Es decir, recoger información tanto sobre el papel de la Administración como sobre los aprendizajes, valores, motivaciones y circunstancias socio-económicas y culturales de los alumnos/as. Una evaluación educativa del sistema no debería omitir la responsabilidad de la Administración en la calidad educativa, de la misma manera que no suele omitir la responsabilidad de los centros y los docentes.
2. La información que genere la evaluación del sistema educativo debería ayudarnos a comprender con más profundidad los logros y dificultades del mismo, las circunstancias y determinantes de la desigualdad (racial, sexual y socio-cultural). Asimismo debería potenciar aquellos procesos

educativos que propicien la recreación crítica del conocimiento en la escuela, las relaciones democráticas, y desarrollen los valores públicos de libertad, igualdad y solidaridad.

3. Los instrumentos y estrategias de recogida de información no deberían restringirse a las pruebas nacionales, los tests u otras técnicas por el estilo, sino desarrollar estrategias alternativas como los estudios de casos y los grupos de discusión, u otras.

4. Los instrumentos y estrategias deberían ser sensibles a las complejidades de los fenómenos sociales y educativos, a la ecología de la escuela y a las necesidades de los directamente interesados: profesores/as, padres/madres, alumnos/as y Administración. Además se necesita desarrollar instrumentos y estrategias que no separen artificialmente los aspectos cognitivos, afectivos, conativos, interactivos y situacionales del aprendizaje.

5. Deberían potenciarse instrumentos y estrategias a través de los cuales los distintos grupos sociales implicados pudieran expresar libremente sus opiniones, sus críticas, sus incertidumbres y sus problemas.

6. La validez de los instrumentos y estrategias debería tener en cuenta los usos a los que serán destinados y limitar aquellos que conllevasen consecuencias negativas: sociales, personales, culturales y educativas. En general, no deberían utilizarse todos aquellos instrumentos y estrategias que atentasen y menoscabasen la democracia, la justicia, la solidaridad y la igualdad.

7. En lugar de obligar a adaptar la enseñanza a los requisitos estructurales y técnicos de las pruebas, como ocurre con las criteriales y de competencia mínima, se tendrían que someter aquéllas a la naturaleza compleja de los aprendizajes, los conocimientos escolares, los procesos de enseñanza y el *curriculum*.

8. La evaluación del sistema educativo debería hacer explícito el sentido de escolaridad, educación y enseñanza que subyace a sus procedimientos, sus criterios y sus orientaciones, aceptando que la pluralidad de puntos de vista y de valoraciones es su terreno social de desarrollo, no un obstáculo que neutralizar.

9. La evaluación del sistema educativo no es un acto puntual, sino una dinámica de conocimiento social y público sobre la calidad de los procesos e intervenciones educativos tanto de reforma actual como de cualquier otra iniciativa educativa, sea o no de la Administración.

10. La evaluación del sistema educativo no puede ser una “responsabilidad” exclusiva de una institución tecnocrática, central (nacional) o periférica (autonómica), sino el compromiso de todos los sectores y de todos los grupos relacionados, directa e indirectamente, con el sistema educativo. La evaluación de este sistema ha de ser un proceso descentralizado en el que ni los expertos, ni los gestores políticos tengan poder de decisión y elección exclusivo y último.

11. Antes de emprender cualquier macro-proceso de evaluación, resulta necesario negociar y recoger un amplio consenso social sobre la misma. Un consenso que no puede remitirse únicamente a los Consejos Escolares Nacional y Autonómicos, sino que debería plantearse directamente con colectivos de docentes (Movimientos de Renovación), de padres/madres y de alumnos/as, así como con colectivos amplios de especialistas. En este debate nadie puede arrogarse la posesión de la verdad, y menos que nadie los expertos y la Administración.

BIBLIOGRAFÍA

Airasian, P. W. (1987): "State Mandated Testing and Educational Reform: Context and Consequences", *American Journal of Education*, Vol. 95, n.º 3, págs. 393-412.

Althusser, L. (1975): *Escritos 1968-1970*. Barcelona. Laia.

Álvarez Méndez, J. M. (1985): *Didáctica, Curriculo y Evaluación: Ensayos sobre Cuestiones Didácticas*. Barcelona. Alamex.

Álvaro Page, M. y cols. (1988): *Evaluación Externa de la Reforma Experimental de las Enseñanzas Medias*. *Vol. 1*. Madrid. C.I.D.E., Ministerio de Educación y Ciencia.

— (1992): "El Proyecto de la OCDE. Indicadores Educativos". *Cuadernos de Pedagogía*, n.º 199, págs. 13-17.

— **e Izquierdo, M. C.** (1992): 'Para Saber Más". *Cuadernos de Pedagogía*, n.º 199, págs. 25-26.

Angulo, J. F. (11990): *Innovación y Evaluación Educativa*. Málaga. EAC, Universidad de Málaga.

— (1992a): "Descentralización y Evaluación en el Sistema Educativo Español: Algunas Claves para el Pesimismo". *Escuela Popular*, n.º 3, págs. 23-87.

— (1992b): "Objetividad y Valoración en la Investigación Educativa. Hacia una Orientación Emancipadora". *Educación y Sociedad*, n.º 10, págs. 91-129.

— (1992c): "El Caballo de Troya. Calidad de Enseñanza y Tecnocracia". *Cuadernos de Pedagogía*, n.º 206, págs. 62-67.

— (1993): "La Participación Local en el Sistema Sanitario: Claves y Compromisos", en Gallo Vallejo, F. J. e. a. (Comp.) (1993). *Manual del Residente de Medicina Familiar y Comunitaria*. Madrid. Beecham; págs. 372-381.

—, **Contreras, J. y Santos, M. A.** (19911): "Evaluación educativa y participación democrática". *Cuadernos de Pedagogía*, n.º 195, págs. 74-79.

Baudelot, C. y Establet, R. (1990): *El Nivel educativo sube. Refutación de una antigua idea sobre la pretendida decadencia de nuestras escuelas*. Madrid. Morata.

Bell, D. (1976): *El Advenimiento de la Sociedad Post-Industrial*. Madrid. Alianza.

Beltrán Llavador, F. (1991): *Política y Reformas Curriculares*. Valencia. Servei de Publicacions. Universitat de Valéncia.

Berlak, H. (1992): "The Need for a New Science of Assessment", en H. Berlak, F. M. Newmann, E. Adams, D. A. Archbald, y cols. (Comp.) (1992). *Toward a New Science of Educational Testing and Assessment*. Albany, N. Y. State University of New York; págs. 1-21.

—, **Newmann, F. M., Adams, E., Archbald, D. A., y cols.** (1992): *Toward a New Science of Educational Testing and Assessment*. Albany. N. Y. State University of New York.

Boli, J., Ramírez, F. O. y Meyer, M. J. (1990): "Explicación de los Orígenes y el Desarrollo de la Educación de Masas", en P. G. Altbach y G. P. Kelly. (Comp.) (1990). *Nuevos enfoques en educación comparada*. Madrid. Mondadori; págs. 123-152.

Broadfoot, P. (11983): "Evaluation and the Social Order in Advanced Industrial Societies: the Educational Dilemma". *International Review of Applied psychology*, Vol. 32, págs. 307-325.

— (Comp.) (1984): *Selection, Certification & Control. Social Issues in Educational Assessment*. Londres. The Falmer Press.

— (1986) "Assessment Policy and Inequality: The United Kingdom Experience". *British Journal of sociology of Education*, Vol. 7, NY 2, págs. 205-224.

—, **Murphy, R. y Torrance, H.** (Comps.) (1990): *Changing Educational Assessment. International Perspectives and Trends*. Londres. Routledge.

Browm, J., Collins, A. y Dugid, P. (1989): "Situated Cognition and the Culture of Learning". *Educational Researcher*, Vol. 18, n.º 1, págs. 32-42.

Bruner, J. (1990): *Actos de Significado. Más Allá de la Revolución Cognitiva*. Madrid. Alianza.

Cole, N. S. (1990): "Conceptions of Educational Achievement". *Educational Researcher*, Vol. 19, n.º 3, págs. 2-7.

Conner, C. (1991): *Assessment and Testing in the Primary School*. Basingstoke, Hampshire. The Falmer Press.

Consejo Escolar del Estado (1992): *Informe sobre el Estado y Situación del sistema Educativo*. Madrid. M.E.C./C.E.E.

Cronbach, L. J. y Meehl, P. (1955): "Construct Validity in Psychological Tests". *Psychological Bulletin*, Vol. 52, Julio, págs. 281-302.

Crozier, M. J., Huntington, S. P. y Watanuki, J. (1975): *The Crisis of Democracy. Report on the Governability of Democracies to the Trilateral Commisión*. Nueva York. New York University Press.

Dale, R. (1989): *The State and Education Policy*. Milton Keynes. Open University Press.

Del Río, E., Jover, D. y Riesco, L. (1991): *Formación y Empleo. Estrategias Posibles*. Barcelona. Paidós.

Elliott, J. (1992a): "¿Son los 'Indicadores de Rendimiento', indicadores de la Calidad Educativa?" *Cuadernos de Pedagogía*, n.º 206, págs. 56-60.

— (1992b): "¿Son los 'Indicadores de Rendimiento', indicadores de la Calidad Educativa?" *Cuadernos de Pedagogía*, n.º 207, págs. 44-47.

Feito, R. (1990): "CONCAPA y CEAPA: Dos Modelos de Intervención de los Padres en la Gestión de la Enseñanza". Educación y Sociedad, n.º 9, págs. 35-57.

Foucault, M. (1981): *Vigilar y Castigar. Nacimiento de la Prisión*. Madrid. Siglo XXI.

Gardner, H. y Hatch, Th. (1989): "Multiple intelligences go to School. Educational implications of the Theory of Multiple Intelligences". *Educational Researcher*, Vol. 18, n.º 8, págs. 4-10.

Gould, S. J. (1984): *La Falsa Medida del hombre*. Barcelona. Antoni Bosch.

Granheim, M. (1990): "The education system in Norway: the Need for Evaluation", en M. Granheim, M. Kogan y U. P. Lundgren (Comps.) (1990). *Evaluation as Policymaking. Introducing Evaluation into a National Decentralised Educational System*. Londres. Jessica Kingsley Publishers; págs. 12-19.

Habermas, J. (1975): *Problemas de Legitimación en el Capitalismo Tardío*. Buenos Aires. Amorrortu.

Hargreaves, A. (1989a): The Crisis of Motivation and Assessment, en (Comp.) (1989a). *Education Policies: Controversies and Critiques*. Londres. The Falmer Press; págs. 41-63.

— (1989b): *Currículum and Assessment Reform*. Milton Keynes. Open University Press.

Hartnett, A. y Naish, M. (1990): "The Sleep of Reason breeds Monsters: the Birth of a Statutory Currículum in England and Wales". *Journal of Currículum Studies*, Vol. 22, n.º 1, págs. 1-16.

Henkel, M. (1991): Gov*erment, Evaluation and Change*. Londres. Jessica Kingsley Publisher.

House, E. R. (1978): "Evaluation as Scientific Management in U. S. School Reform". *Comparative Education Review*, Vol. 22, n.º 3, págs. 388-401.

— (1990): "Methodology and Justice", en K. A. Sirotnik (Comp.) (1990). *Evaluation and Social justice: Issues in Public Education. New Directions for Program Evaluation, Nº 45*. San Francisco. Jossey-Bass Inc., Publishers; págs. 23-36.

Irvine, J., Miles, I. y Evans, J. (Comps.) (1979): *Demystifying Social Statistics*. Londres. Pluto Press.

Karier, C. J. (1974): "Ideology and Evaluation, in Quest of Meritocracy", en M. W. Apple, M. J. Subkoviak y H. S. Lufler (Comps.) (1974). *Educational Evaluation: Analysis and Responsability*. Berkeley, California. McCutchan Publish. Corp.; págs. 279-320.

Linn, R. L. (1993): "Educational Assessment: Expanded Expectations and Challenges". *Educational Evaluation and Policy Analysis*, Vol. 115, n.º 1, págs. 1-16.

M.E.C. (1989): *Libro Blanco para la Reforma del Sistema Educativo*. Madrid. Ministerio de Educación y Ciencia.

— (1990): *Ley Orgánica de Ordenación General del Sistema Educativo*. Madrid. Ministerio de Educación y Ciencia.

MacDonald, B. (1976): *Who's Afraid of Evaluation*. C.A.R.E., University of East Anglia. Trabajo.

— (1988): "Evaluation", en A. Bullock (Comp.) (1988). *The Fontana Dictionary of Modern Thought*. Londres. Fontana Press; págs. 291.

Madaus, G. F. (1985a): "Testing as an Administrative Mechanism in Educational Policy: What does the Future Hold?" *Educational Horizons*, Vol. 63, Special Issue, págs. 34-39.

— (1985b): "Public Policy and the Testing Profession- You've never had it so Good?" *Educational Measurement. Issues and Practice*, Vol. 4, n.º 4, págs. 5-11.

— (1988a): "The Influence of Testing on the Curriculum", en L. N. Tanner (Comp.) (1988a) Chicago. 87th Yearbook of the National Society for the Study oí Education. The NSSE; págs. 83-121.

— (1988b): "The Distortion of Teaching and Testing: High-Stakes Testing and Instruction". *Peabody Journal of Education*, Vol. 65, n.º 3, págs. 29-46.

— (1990): *Testing as a social technology. Inaugural Annual Boisi Lecture in Education and Public Policy*. Boston. Boston College.

McCarthy, T. (1987): *La Teoría Crítica de Jürgen Habermas*. Madrid. Tecnos.

Messick, S. (1989): "Meaning and Values in Test Validation: The Science and Ethics of Assessment". *Educational Researcher*, Vol. 18, n.º 2, págs. 5-11.

Muñoz-Repiso, M. e. a. (1992): *Las Desigualdades en la Educación en España*. Madrid. Ministerio de Educación y Ciencia-C.I.D.E.

N.C.E.E. (National Commision on Excellente in Education = Comisión Nacional para la Excelencia de la Educación) (1985): "Estados Unidos. Una nación en peligro: el imperativo de una reforma educativa", *Revista de Educación*, n.º 278 (septiembre-diciembre), págs. 135-153.

O'Connor, J. (1989): *El Significado de la Crisis. Una Introducción Teórica*. Madrid. Editorial Revolución.

Palumbo, D. J. (Comp.) (1987): *The Politics of Program Evaluation*. Londres. Sage.

Pérez Gómez, A. I. (1991): "Cultura Escolar y Aprendizaje Relevante". *Educación y Sociedad*, n.º 8, págs. 59-72.

— **y Gimeno Sacristán, J.** (1992): *Documento Síntesis: Claves para Comprender la Evolución de un Proceso de Innovación. Evaluación de la Reforma del Cielo Superior de E.G.B. en Andalucía*. Depto. de Didáctica y Organización Escolar. Universidad de Málaga.

Raven, J. (1992): "A Model of Competence, Motivation, and Behavior, and a Paradigm for Assessment", en H. e. a. Berlak (Comp.) (1992). *Toward a new Science of Educational Testing & Assessment*. Nueva York, Albany. State University of New York Press; págs. 85-116.

Rivas, F. y Alcatud, F. (1989): *La Evaluación Criterial en la Educación Primaria*. Madrid. C.I.D.E.

Secada, W. G. (Comp.) (1989): *Equity in Education*. Londres. The Falmer Press.

Shapiro, H. S. (1985): "Capitalism at Risk: The Political Economy of the Educational Reports of 1983". *Educational Theory*, Vol. 35, n.º 1, págs. 57-72.

Sriven, M. (1991): *Evaluation Thesaurus*. Londres. Sage.

Stake, R. E. (1989): *Quality Control and Deceptive Packaging: Does Reform Require Accurate Portrayal of Education?* Conferencia dada en la Universidad de North Carolina en Greensboro, Octubre 23. C.I.R.C.E., University of Illinois at Urbana-Champaing.

Suchman, L. A. (1987): *Plans and Situated Actions, the Problem of Human-Machine Communication*. Cambridge. Cambridge University Press.

Walberg, H. J., Bottani, N. y Delfau, I. (1990): "OECD indicators of Educational Productivity". *Educational Researcher*, Vol. 19, n.º 4, págs. 30-33.

Weiler, H. N. (1983): "Legalization, Expertise, and Participation: Strategies of Compensatory Legitimation in Educational Policy". *Comparative Education Review*, Vol. 27, n.º 2, págs. 259-277.

Winner, L. (1986): *The Whale and the Reactor. A Search for Limits in an Age of High Technology Chicago*. The University of Chicago Press.

Wood, R. (1987): *Measurement and Assessment in Education and Psychology*. Colección de documentos 1967-1987. Londres. The Falmer Press.