



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

**PERFILES
EDUCATIVOS**

ISSN 0185-2698

Peña de la Mora, Eduardo (1989)
**“CRITICA A LA FUNDAMENTACIÓN EPISTEMOLÓGICA DE LA
MEDICIÓN DEL APRENDIZAJE ESCOLAR”**
en Perfiles Educativos, No. 45-46 pp. 27-37.

CRÍTICA A LA FUNDAMENTACIÓN EPISTEMOLÓGICA DE LA MEDICIÓN DEL APRENDIZAJE ESCOLAR

Eduardo PEÑA DE LA MORA*

INTRODUCCIÓN

La medición del aprendizaje escolar ocupa un lugar preponderante en el contexto educativo actual. Constituida por un conjunto de técnicas que pretenden cuantificar los aprendizajes que ocurren dentro de los límites escolares, la medición del aprendizaje es formalmente auxiliar de una tarea más general, que es la evaluación del aprendizaje; no obstante, en la práctica educativa la medición ha desplazado casi por completo a esta última.¹

En términos generales, la función que formalmente se atribuye a la evaluación del aprendizaje es la de contribuir positivamente al proceso de enseñanza-aprendizaje mediante la determinación de los elementos que obstaculizan o facilitan la marcha del proceso y el cumplimiento de los objetivos propuestos. Sin embargo, en la práctica educativa generalizada, la evaluación del aprendizaje no contribuye positivamente al proceso; su acción se reduce a la medición del aprendizaje o a la mera asignación de una nota al finalizar un periodo escolar.

El desplazamiento de la evaluación del aprendizaje (o la reducción de su acción) encuentra explicación en las relaciones internas que se desarrollan en la escuela y en las funciones que está desempeñando en la sociedad. Las relaciones y funciones de la escuela se conciben de diferente manera, según las diversas perspectivas de pensamiento.

Desde una perspectiva meritocrática, por ejemplo, la sociedad asigna a la escuela las funciones de selección y diferenciación de sus integrantes. Dichas funciones se llevan a cabo a través de la gratificación, el reforzamiento de comportamientos y la concesión de prerrogativas, acreditaciones y certificaciones fundamentadas en el logro diferencial escolar, es decir, en el mérito. Así, se propone que la determinación de este logro se haga de manera racional y "justa"², o sea (interpretamos) de una manera neutral y objetiva, cualidades que la ciencia positivista atribuye al método que privilegia lo cuantitativo sobre lo cualitativo. Se considera, asimismo, que lo cualitativo permite juicios subjetivos; esto es, que los juicios son esencialmente producto del sujeto y no del objeto, lo que a su vez tiene como probable consecuencia la pérdida de la neutralidad. Es así como,

* Profesor investigador del Centro Interdisciplinario de Investigación y Docencia en Educación Técnica, Querétaro, Qro.

¹ Cfr. Alicia de Alba et al., "Evaluación: análisis de una noción", en *Revista Mexicana de Sociología*, Vol. XLVI, núm. 1, 1984; A. Díaz Barriga, *Didácticas y currículum. Convergencias en los programas de estudio*. México, Nuevomar, 1984, y "Tesis para una teoría de la evaluación y sus derivaciones en la docencia", en *Perfiles Educativos*, núm. 15, 1982; P. Morán Oviedo, "Propuesta de evaluación y acreditación en el proceso de enseñanza-aprendizaje desde una perspectiva grupal", en *Perfiles Educativos*, núm. 27-28, 1985.

² Talcott Parsons, "La educación como asignadora de roles y factor de selección social", en M. de Ibarrola (comp.), *Las dimensiones sociales de la educación*. México, SEP-EI Caballito, 1985, pp. 84-87.

en la búsqueda de la objetividad y la neutralidad, la evaluación del aprendizaje deviene en meto intento de cuantificación.

El presente escrito se ubica en una perspectiva diferente. Se trata de una posición que concibe la escuela como una institución que contribuye a la reproducción de las condiciones – principalmente en el nivel ideológico- de las relaciones sociales de producción dominante. Esto es, la escuela contribuye a inculcar en los estudiantes la ideología dominante y a brindar a éstos oportunidades escolares y laborales diferenciales en función de su pertenencia de clase.³ El trabajo forma parte de uno más amplio que aborda el tema de la evaluación del aprendizaje escolar como uno de los medios de que se vale la escuela (y la sociedad) para realizar su función reproductora.

El objetivo aquí es demostrar que las técnicas de medición tienen un fundamento epistemológico que no permite generar conocimientos científicos acerca del aprendizaje escolar. Si el objetivo se lograra, se derrumbaría el argumento que esgrime la perspectiva meritocrática para sostener que la medición proporciona un conocimiento objetivo y neutral acerca del aprendizaje escolar, y se concluiría que la función de selección y diferenciación de los estudiantes en la escuela no necesariamente se fundamenta en el logro escolar.

Este trabajo está dividido en cuatro partes. En la primera se demuestra que para que el aprendizaje pueda ser cuantificado en una escala numérica, su estructura cuantitativa debe tener un grado de semejanza con el sistema numérico que pretende representarlo. En la segunda parte se establece que las diversas teorías del aprendizaje carecen de una explicación que integre los aspectos cuantitativos y cualitativos del aprendizaje; por consiguiente, se pone en tela de juicio que el sistema numérico pueda representar cuantitativamente el aprendizaje. En la tercera parte se examinan en forma crítica los modelos de medición que se aplican al aprendizaje. En ésta se concluye que la carencia en la construcción teórica del objeto de la medición no puede suplirse con las soluciones técnico-operativas propuestas por los modelos examinados. Finalmente, en la cuarta parte se sugieren, a la luz del análisis realizado, explicaciones alternativas a la perspectiva meritocrática acerca de las funciones sociales de la evaluación del aprendizaje.

La necesidad del isomorfismo en la medición

Las diferentes definiciones de medición coinciden en señalar que ésta es la asignación de números para representar cantidades del atributo (en este caso el aprendizaje escolar) siguiendo reglas predeterminadas. De aquí se infiere que si se desea generar una escala numérica que represente cuantitativamente el atributo que se intenta medir es necesario que exista una correspondencia estrecha entre las características de los números –esto es, la matemática- y las características del objeto a medir –el aprendizaje escolar-. Esta correspondencia se conoce como “isomorfismo”, y es una condición necesaria para que la medición sea válida.

Las matemáticas son una ciencia formal abstracta producto de una práctica social –como el resto de las ciencias- que ha determinado sus propiedades; es decir, las características de su comportamiento y las reglas con que se relacionan sus abstracciones.⁴ Para los fines del presente análisis, las propiedades de las matemáticas son tres: la identidad, la ordinalidad y la aditividad. La identidad significa que un número sólo puede ser igual a sí mismo y diferente a los demás; la ordinalidad se define como la propiedad que permite ordenar un conjunto de números de manera tal que signifiquen magnitudes crecientes o decrecientes; la aditividad es la propiedad que determina

³ Evidencia contundente a este respecto se presenta en: H. Gintis, “Educación, tecnología y características de la productividad del trabajador”, en C. Biasuto, Educación y clase obrera. México, Nueva Imagen, 1978, y en H. Gintis y S. Bowles, La instrucción escolar en la América capitalista. 5ª. Edición México, Siglo XXI, 1981.

⁴ Cfr. G. Benedito, “El problema de la medida en psicología”, en N.A. Braunstein et. al., Psicología, ideología y ciencias. 5ª. Edición México, Siglo XXI, 1979.

como válida la operación de suma; si esta operación es válida, el resto de las operaciones aritméticas son válidas: la resta, la multiplicación y la división.

Para que la medición sea válida, decíamos, tiene que existir una correspondencia entre las características del atributo o cualidad a medir y las propiedades de las matemáticas. Cuanto más estrecha sea dicha correspondencia o isomorfismo, tanto mayor será la posibilidad de expresar la cantidad del atributo a medir; si existe isomorfía completa, las operaciones aritmética de las mediciones realizadas tendrán validez.⁵

Cabe señalar que, al realizar una medición, la isomorfía (o correspondencia) que se busca se encuentra en un determinado nivel de abstracción en el proceso de conocimiento. Es decir, la matemática es un medio abstracto que sirve para expresar la cantidad del objeto, y no directamente el objeto mismo. En la medición se busca la isomorfía precisamente entre las abstracciones descritas.

Debido al precario desarrollo de los aspectos cuantitativos de los conceptos en las ciencias sociales, generalmente no se tiene una estrecha isomorfía entre éstos y las matemáticas; por consiguiente, la validez de las escalas numéricas que se generan con la medición tiene grandes restricciones. No obstante, es muy usual que no se tomen en consideración dichas restricciones y que, por lo tanto, la interpretación o los resultados derivados de las mediciones carezcan de validez.

Para ejemplificar esto último cabe hacer notar que es práctica común en el medio educativo efectuar diversas “mediciones” del aprendizaje en cada sujeto durante un periodo escolar. Con los resultados de estas mediciones se efectúan adiciones, se calculan medidas, se hacen ponderaciones o se realizan otras operaciones aritméticas. Posteriormente se emite un juicio acerca del aprendizaje del alumno asignando una nota o una calificación. Sin embargo, no se comprueba previamente que exista isomorfía completa entre el aprendizaje y el sistema numérico; es decir, que el aprendizaje tenga las propiedades correspondientes a la identidad, la ordinalidad y la aditividad. Más concretamente, se desconoce si dos aprendizajes sobre temas diferentes pueden expresarse cuantitativamente con el mismo número (identidad); tampoco se tiene la certeza de que los aprendizajes de un conjunto puedan ser ordenados en función de su cantidad o magnitud de aprendizaje (ordinalidad); se ignora si dos aprendizajes “sencillos” pueden equivaler cuantitativamente a un tercer aprendizaje “mas complejo” (aditividad). Con tales incertidumbres, aun la operación aritmética más sencilla resulta tener dudosa validez, y así también los resultados, conclusiones o calificaciones derivadas de dichos procedimientos, aun cuando se expresen numéricamente y, en ocasiones, sean producto de complejos tratamientos estadísticos.

Sí, como se propone frecuentemente, la evaluación del aprendizaje requiere de la medición, entonces es necesario develar la estructura, propiedades y características del concepto aprendizaje y comparar éstas con las correspondientes a las matemáticas. Con base en esta comparación se puede determinar el grado de isomorfía y, consecuentemente, los límites de la interpretación de las mediciones del aprendizaje.

Teorías del aprendizaje y su estructura

Tanto la calidad como la cantidad son características a través de las cuales conocemos un objeto (fenómeno o proceso). En las etapas iniciales del conocimiento del objeto se aprecian estas dos características por separado; sin embargo, a medida que se profundiza en el conocimiento se

⁵ Para una discusión más profunda consultar E. Peña, Medición educativa (mimeo), 1985, y J. P. Guilford, Psychometric Methods, 2ª. Edición, Nueva York, McGraw-Hill, 1954.

descubren las interrelaciones y dependencias que se establecen entre ellas.⁶Hegel expresó teóricamente la interdependencia de las características cuantitativa y cualitativa de los objetos a través de “la ley del paso de los cambios cuantitativos a los cambios cualitativos”. Engels, a su vez, mostró la vigencia de la ley tanto en los fenómenos de la naturaleza como en los sociales.⁷

Es indudable que tanto la calidad como la cantidad son características que nos permiten conocer el aprendizaje, y que éstas están relacionadas entre sí. Ahora bien, ¿cómo conciben esta relación las diversas teorías del aprendizaje? En otras palabras, ¿cuál es la explicación del cómo, el qué y el cuánto se aprende? Las teorías del aprendizaje que han tenido mayor influencia en la educación contemporánea, como son el conductismo clásico, el neoconductismo, el cognoscitivismo y la aproximación piagetana, se encuentran en el nivel de descripción cualitativa del fenómeno del aprendizaje, mas no se ocupan de la cuantificación. Su preocupación se centra en la explicación del fenómeno aprendizaje a través de factores, relaciones y condiciones; es decir, en la dimensión cualitativa del fenómeno. Sin embargo, no conceptualizan la dimensión cuantitativa del aprendizaje ni la interrelación que se establece con la cualitativa. Esto, no obstante, no implica la ausencia de algún aspecto de cuantificación en las teorías.

El aprendizaje en el conductismo clásico se explica como el condicionamiento de una respuesta a un estímulo determinado. Tanto el condicionamiento de la respuesta como su permanencia se deben a ejercicios de repetición. Pueden cuantificarse las repeticiones, la intensidad del estímulo y la intensidad de la respuesta; sin embargo, no existe una construcción que permita asignar números de una escala para significar cantidades de aprendizaje. Por ejemplo, considerando el experimento de Pavlov, no existiría fundamento alguno para cuantificar mediante la asignación de un número el aprendizaje implícito en la conducta condicionada de salvación del perro, como tampoco para cuantificar el aprendizaje implicado en la conducta de rechazo a objetos peludos, en el caso del hijo de Watson. ¿Qué justificaría el que esos dos aprendizajes fuesen considerados cuantitativamente iguales?, ¿cuál de esos dos aprendizajes sería mayor?, ¿qué interrelación habría en el aprendizaje de una reacción condicionada a un estímulo y su cantidad? El conductismo clásico no ofrece respuesta a estas interrogantes.

En la corriente neoconductista, más propiamente, en la skinneriana, el aprendizaje es un cambio de conducta más o menos permanente que se explica a través del condicionamiento operante; es decir, el reforzamiento de un comportamiento que se manifiesta en las consecuencias del comportamiento mismo.⁸ En esta concepción, el comportamiento que se desea se aprenda se subdivide en comportamientos más simples. Cuando el que aprende manifiesta dichos comportamientos, recibe reforzamientos de manera inmediata. En esta corriente, sin embargo, no existen elementos suficientes que justifiquen la asignación de números que signifiquen cantidades de aprendizaje. Sería difícil aceptar que, por ejemplo, la conducta operante aprendida por un palomo, consisten en girar dos veces sobre sus patas para recibir alimento, fuese “mayor” que la conducta de levantar una pata para recibir recompensa semejante. Si esto último pudiese ser aceptado, quedaría por determinar –si se desease alcanzar una mayor isomorfía con el sistema numérico- en qué medida es mayor un aprendizaje respecto del otro.

Podría arguirse que las diferentes taxonomías del aprendizaje presentan una base para la medición. Al desarrollar su taxonomía del dominio cognoscitivo, Bloom y colaboradores señalan que “Una taxonomía ha de estar constituida de tal manera que el orden de los términos corresponda a algún orden ‘real’ de los fenómenos representados por aquéllos”.⁹ Es decir, la taxonomía debería

⁶ Véase A.P. Sheptulin, Método dialéctico del conocimiento. México, Cartago, 1984, p. 100.

⁷ F. Engels, Ddialéctica de la naturaleza. México, Grijalbo, 1981.

⁸ B.F. Skinner, Science and Human Behavior. Nueva York, 1963, p. 65.

⁹ B.S. Bloom y cols., Taxonomía de los objetivos de la educación. La clasificación de las metas educacionales. Buenos Aires, El Ateneo, 1979, pp. 17-18.

tener una relación ordinal respecto de “alguna de sus propiedades esenciales”.¹⁰ El propio Bloom reconoce que no tuvo éxito su búsqueda de alguna fundamentación teórica para realizar dicho ordenamiento.¹¹ No obstante, aun sin fundamentación teórica proponen una taxonomía basada en el ordenamiento en clases de comportamientos, que va de los más simples a los más complejos.

Para justificar esta dimensión de simplicidad complejidad argumentan que la integración de varios comportamientos simples podría constituir un comportamiento más complejo. Las clases que forman su taxonomía presentan la aditividad creciente de comportamientos simples; esto es, los comportamientos pertenecientes a una clase presuponen la integración de los correspondientes a la clase anterior con otros comportamientos simples.¹² Por ejemplo, un comportamiento perteneciente a la clase de “aplicación” implica la integración de los comportamientos de la clase de “comprensión” con otros comportamientos.

Consecuentemente, el aprendizaje de un comportamiento complejo equivaldría cuantitativamente a los aprendizajes de los comportamientos simples que lo integran. En esta concepción se considera que el conocimiento es de naturaleza aditiva y que basta con externar los comportamientos fraccionados para integrar un comportamiento más complejo. El planteamiento es erróneo porque se ignora que la realidad es un todo estructurado que se crea y desarrolla. Pretender que el conocimiento de esta realidad se expresará a través de comportamientos que reflejen los elementos que constituyen el todo, sus interrelaciones, su génesis y su estructura¹³ es un error, ya que los comportamientos simples no constituyen por sí solos un comportamiento complejo.

Aun suponiendo –sin conceder– que el conocimiento fuese de naturaleza aditiva, sólo se podría establecer una relación cuantitativa en un nivel de escala intervalar¹⁴ entre aprendizajes que impliquen comportamientos simples integrados para formar comportamientos complejos. Quedarían sin relación cuantitativa intervalar los aprendizajes correspondientes a comportamientos no vinculados. En las matemáticas, por ejemplo, no se podría cuantificar lógicamente el aprendizaje que representa resolver una ecuación de segundo grado comparado con el que representa aplicar una regla de tres simple (esto es, qué tanto mayor o qué tanto menor).

Según Bloom y colaboradores, las clases en su taxonomía representan un ordenamiento en función de la dimensión simplicidad-complejidad. Por su parte, Krathwohl señala que también éstas representan un ordenamiento en la dimensión concreción-abstracción.¹⁵ Suponiendo –de nuevo sin conceder– que efectivamente se den los ordenamientos en dichas dimensiones, aun faltaría justificar teóricamente las razones por las cuales dichas dimensiones podrían servir de base para la medición cuantitativa, en este caso, de nivel ordinal.¹⁶

¹⁰ Idem.

¹¹ Idem.

¹² Idem.

¹³ Karel Kosik, *Dialéctica de lo concreto*. México, Grijalbo, 1967, pp. 53-65.

¹⁴ Una escala de nivel intervalar es aquella que se forma en un proceso de medición donde la cualidad que se mide se expresa cuantitativamente sólo en los intervalos de la escala –no en los valores absolutos de la misma-. Por ejemplo, si se lograse formar una escala de nivel intervalar en la medición del aprendizaje, el valor cero de la escala no expresaría la ausencia del aprendizaje; como el valor de 40 no correspondería al doble de aprendizaje que un valor de 20 o a la mitad de un valor de 80. Las diferencias (intervalos) en la escala sí representarían diferencias cuantitativas del aprendizaje. Esto es, entre los valores 40 y 60 existiría la misma diferencia de aprendizaje que entre 73 y 93. En estas dos diferencias existiría el doble de aprendizaje que en la diferencia entre 70 y 80.

¹⁵ Consúltese D.K. Krathwohl, “Taxonomía de los objetivos educativos”, en R.M. Gagné et al., *Especificación de los objetivos de la educación*. México, Guajardo, 1969, p.28.

¹⁶ En una escala de nivel ordinal sólo existe una isomorfía parcial con el sistema numérico. Más específicamente, sólo se mantienen las propiedades de identidad y ordinalidad. Por ejemplo, en una escala de aprendizaje, un valor de 67 indica mayor aprendizaje que uno de 60 y menor que uno de 70. Sin embargo, no indica cuantitativamente qué tanto mayor o qué tanto menor. Otro ejemplo es el de una escala formada con el

Se puede preguntar por qué un comportamiento complejo abstracto representa un nivel cuantitativo de aprendizaje mayor que un comportamiento menos complejo o menos abstracto. Se podría preguntar, asimismo, por qué no utilizar como base de la medición del aprendizaje dimensiones tales como la utilidad-inutilidad social de los conocimientos implicados en el aprendizaje, o el grado de desintegración del objeto de estudio, o bien el nivel de esfuerzo por lograr el aprendizaje, etc. Estos cuestionamientos no tienen respuesta en las diferentes teorías del aprendizaje examinadas aquí.

Las teorías cognoscitivas conciben el aprendizaje como una transformación de estructuras cognoscitivas al incorporar nuevos elementos a éstas. Además, intentan explicar la forma en que se construyen y organizan las estructuras y la interacción que hay con la información que se incorpora. En esta corriente no existe conceptualización alguna del aprendizaje que permita la cuantificación de su magnitud; no se tiene el fundamento conceptual como para cuantificar. Por ejemplo, un aprendizaje por insight, en una situación dada, que pudiese transferirse a otra situación semejante, o un aprendizaje que requiriese transformaciones significativas de las estructuras cognoscitivas comparado con otro que requiriese transformaciones modestas.

De manera semejante a las nociones cognoscitivistas, la noción piagetiana del aprendizaje sostiene la existencia de estructuras o esquemas cognoscitivos que interactúan con la información proporcionada por el objeto de conocimiento. Sin embargo, a diferencia de las demás teorías, Piaget concibe la interacción de una manera dialéctica: el ser humano modifica la información que proporciona el objeto para adaptarla a sus esquemas cognoscitivos (proceso de “asimilación”) y, a su vez, estos esquemas son modificados para adecuarlos a la información que se incorpora (proceso de “acomodación”). El proceso cognoscitivo no termina aquí, ya que las actividades que desarrolla el sujeto para aproximarse al objeto le permiten transformar profunda y constantemente sus esquemas cognoscitivos.¹⁷

Nuevamente, esta concepción se centra en los aspectos cualitativos y no aporta los elementos teóricos necesarios para la cuantificación del aprendizaje. Podría arguirse que un aprendizaje que no transformase los esquemas cognoscitivos de manera profunda sería de menor magnitud que otro que sí lo hiciese. No obstante, todavía tendrían que resolverse interrogantes tales como: ¿qué diferencia de magnitud hay entre uno y otro aprendizaje? O ¿las transformaciones profundas de dos esquemas cognoscitivos diferentes implican la misma magnitud de lenguajes?, etcétera.

Queda claro, entonces, que en las concepciones teóricas del aprendizaje revisadas aquí no se establecen las interrelaciones y dependencias mutuas entre calidad y cantidad; por consiguiente, no se devela la estructura cuantitativa del aprendizaje. Dado el estado en que se encuentra la construcción teórica del objeto de conocimiento, se concluye que no existen los fundamentos suficientes para cuantificar la magnitud de aprendizaje. A lo sumo, en algunos casos existiría justificación suficiente para establecer relaciones limitadas de ordinalidad entre ciertos aprendizajes. No obstante, para estas escasas ocurrencias se formaría una escala de nivel ordinal cuya isomorfía con el sistema numérico sería parcial, por lo tanto, carecería de validez el uso de cualquier operación aritmética. Por ello, aun en estos limitados casos son inválidas las operaciones (o transformaciones) que frecuentemente se realizan en el ámbito escolar, como son, entre otras, el cálculo de promedios (medias aritméticas), ponderaciones de calificaciones y agregados de medición (sumas de calificaciones).

orden de llegada a la meta de los competidores de una carrera pedestre, el que obtuvo el tercer lugar corrió más rápido que el que obtuvo el décimo lugar, y más lento que el que obtuvo el segundo; sin embargo, no se sabe el tiempo que emplean los competidores en el recorrido ni las diferencias de tiempo.

¹⁷ Cfr. E. Ruíz Larraguivel, “Reflexiones en torno a las teorías del aprendizaje”, en *Perfiles Educativos*, núm. 2 (nueva época), 1983, y M. Pansza, “Una aproximación a la epistemología genética de Jean Piaget”, en *Perfiles Educativos*, núm. 18, 1982.

Las técnicas de medición y la cuantificación del aprendizaje

La ausencia de una construcción teórica que explique integradamente las características cualitativas y cuantitativas del aprendizaje impide establecer, entre otras cosas, la estructura cuantitativa del aprendizaje y su relación, por ejemplo, con el desempeño de un estudiante en un examen. La corriente dominante de la medición educativa en nuestro medio pretende suplir dicha carencia teórica con la aplicación de técnicas de medición basadas en complejos modelos matemáticos.

Las técnicas de medición han sido desarrolladas con la pretensión de lograr un alto grado de generalidad, aislándolas casi por completo del objeto y la intención de la medición. Así, se aplican en cualquier contexto, con diversos propósitos y para medir cualquier atributo-sea éste una característica o una habilidad- que posean los seres humanos o hasta “animales de menor nivel”.¹⁸ No se consideran, al desarrollar estas técnicas, las características propias de los atributos, las características de los seres que poseen dichos atributos, la relación existente entre ellas, la intención ni el contexto de la medición. Además, donde se aplican con mayor frecuencia estas técnicas es en la medición del aprendizaje escolar.

Uno de los modelos más utilizados es el llamado lineal o sumativo. Este modelo establece que existe una relación lineal recta entre el atributo medido en los sujetos y el puntaje total obtenido en el examen¹⁹ (o cuestionario). Dicho puntaje se calcula mediante la suma de los resultados obtenidos en cada uno de los reactivos²⁰ del examen. En el caso que nos ocupa, el modelo indica que existe una relación lineal recta entre el aprendizaje y el puntaje total del examen, lo que significa que el puntaje total de un examen sería una medida de nivel de escala intervalar del aprendizaje. Se elige un modelo y se termina así el engorroso problema de determinar la estructura cuantitativa del aprendizaje. En la práctica de la medición educativa, los maestros no suelen explicitar (o elegir conscientemente) modelo alguno, pero por lo general subyace el modelo lineal.

Antes de aceptar la relación propuesta por el modelo, se debe cuestionar la validez de los supuestos que lo fundamentan. Los supuestos básicos del modelo son tres. El primero supone que la relación entre el aprendizaje y la probabilidad de responder correctamente un reactivo cualquiera está representada geoméricamente por una línea ascendente, de tal suerte que una persona con un nivel bajo de aprendizaje tiene poca probabilidad de contestar correctamente un reactivo, mientras que una persona con un nivel alto de aprendizaje sí. El modelo establece que a mayor aprendizaje corresponde siempre una mayor (o igual) probabilidad de contestar correctamente el reactivo. La línea que representa la relación descrita se llama “línea de trazo monótonica”.²¹ El supuesto no implica que las diversas líneas de trazo correspondientes a los reactivos tengan la misma forma; sólo que son monótonicas. Aunque intuitivamente parezca aceptable, no es posible determinar con certeza la validez de este supuesto, porque se desconoce la estructura cuantitativa (y por consiguiente la escala de medición) del aprendizaje.

El segundo supuesto establece que la suma de las líneas de trazo de los reactivos de un conjunto –como, por ejemplo, los de una examen- tiene como resultado una relación aproximadamente rectilínea entre el aprendizaje (atributo) y el puntaje total en el examen. En los escritos de quienes proponen este modelo, las justificaciones de los supuestos están ausentes.²² Se

¹⁸ J. C. Nunnally, *Psychometric Theory*. Nueva York, McGraw-Hill, 1967, p. 72.

¹⁹ Dada la naturaleza probabilística del modelo, es más preciso referirse al “valor esperado” del puntaje total.

²⁰ Utilizamos aquí el término reactivo como sinónimo de ítem o de pregunta en un cuestionario o examen.

²¹ La línea de trazo (trece líneas) se califica como monótonica cuando se mantiene una relación ascendente. La forma de esta línea se asemeja a una “s”.

²² El modelo de los supuestos es claramente expuesto por J.C. Nunnally, *op.cit.*, pp. 70-74 y 236.

esgrime únicamente el argumento de que la suma de las líneas de trazo de reactivos con diferentes grados de dificultad tiende a compensar sus formas y a resultar en una línea aproximadamente recta. Dado que se desconocen las formas de las líneas de trazo de los reactivos, no es posible comprobar la validez del supuesto. Éste, en sí mismo, define casi en su totalidad las características del modelo, el cual, por consiguiente, nunca trasciende el supuesto. Esto es, se parte del supuesto de la existencia de una relación rectilínea entre el aprendizaje y los puntajes del examen y se llega, en el modelo, a la misma relación lineal.

El tercer supuesto se refiere a que la totalidad de los reactivos tiende a medir un solo aspecto o cualidad, en este caso, el aprendizaje de un tema en particular. Esto es, la respuesta a los reactivos del examen sólo depende del aprendizaje específico que se desea medir. Este supuesto refleja claramente la posición positivista, que concibe la conducta aislada de su contexto histórico social. Se supone que la conducta de responder a un examen es únicamente producto del grado de aprendizaje adquirido por el estudiante en un tema específico. El supuesto pretende ignorar la interrelación múltiple de la conducta humana y los factores del contexto histórico social donde se desarrolla el acto educativo. Por ejemplo, se intenta dejar de lado los aprendizajes previos del sujeto, las estrategias cognoscitivas desarrolladas, los aprendizajes no planeados o no sujetos a medición, el proceso grupal de aprendizaje, los intereses del sujeto en el aprendizaje, los rechazos o acercamientos a una situación del examen,²³ los juegos de poder involucrados en la relación maestro-alumno, etc. En términos de la función reproductora de la escuela, se pretende ignorar los efectos de la socialización, la certificación de conocimientos, la acreditación del saber, etc. Por la pretensión de ignorar los factores arriba mencionados y su interrelación con el aprendizaje de un tema específico, creemos que no es válido el supuesto de que el aprendizaje que se pretende medir es el único factor que interviene en el desempeño de un examen.

Pensamos, además, que el modelo lineal o sumativo carece de sustento que valide su aplicación. Insistimos en la carencia de una teorización que considere de manera integrada los aspectos cuantitativos y cualitativos del aprendizaje escolar, y en que ésta no puede ser cubierta con modelos técnico operativos.

Existen otros modelos de medición que, de manera semejante al modelo lineal, pretenden establecer la relación entre el grado de atributo que posee una persona y las respuestas que externa en un examen o cuestionario. Aunque estos modelos, llamados de características o rasgos latentes,²⁴ no son nuevos, es muy reciente su difusión y popularización. Buscan mayor generalidad y precisión en la relación descrita que el modelo lineal, sin embargo, guardan similitudes con los supuestos que lo fundamentan.

Las líneas de trazo para cada reactivo propuestas por los modelos de características latentes no son, a diferencia de las propuestas por el modelo sumativo, cualquier línea monotónica, sino líneas de naturaleza logarítmica. Estas líneas son conocidas como curvas logísticas. En un examen donde se pretende medir el conocimiento, cada curva logística representa la relación entre el aprendizaje y la probabilidad de contestar correctamente el reactivo correspondiente. La relación estará determinada por la dificultad del reactivo, por su discriminación y por un factor conocido como parámetro de adivinación del mismo reactivo.²⁵

²³ En relación con las afectividades que se generan en los exámenes, se sugiere consultar Julio Aray, "El examen, la neurosis de examen y el examen como factor neurotizante", en Sadismo en la enseñanza. Caracas, Monte Ávila, 1980.

²⁴ R.K. Hambleton y L.L. Cook, "Latent Trait Models and Their Use in the Analysis of Educational Test Data", En Journal of Educational Measurement, V. 14, núm.2, 1977.

²⁵ En realidad, sólo las curvas logísticas de tres parámetros son las que incluyen los tres factores mencionados. Las curvas logísticas de dos y un parámetro constituyen casos particulares de aquella. En las curvas de dos parámetros sólo se consideran los factores de discriminación y dificultad del reactivo, y en las de un parámetro sólo la dificultad. Las curvas normal-ojiva, también utilizadas en el medio educativo, son un caso particular de

Dado que en la realidad se desconoce el nivel de aprendizaje (o del conocimiento) de los sujetos que responden a un reactivo, no es posible determinar la validez del uso de curvas logísticas en un modelo sin una teoría de la estructura cuantitativa del aprendizaje que le dé fundamento. El modelo pretende superar el desconocimiento del grado de aprendizaje a través de un procedimiento técnico que se basa en los mismos resultados que arroja el examen. Primeramente, para cada sujeto se estima, de manera inicial, su nivel de aprendizaje a través de los puntajes obtenidos en su examen. Enseguida, a partir de las estimaciones de los niveles de aprendizaje de los estudiantes, se ajustan, con procesos recurrentes, las curvas logísticas a los datos que arrojan los reactivos del examen.²⁶ En palabras de Hambleton y Cook:

La escala de aprendizaje se ‘estira’ y ‘comprime’ en algunas secciones de manera tal que se maximice el ‘ajuste’ entre las respuestas a los reactivos, las curvas logísticas y los puntajes de habilidad (en este caso, de aprendizaje).²⁷

Para lograr las maximizaciones propuestas en los ajustes se recurre a complejos algoritmos manejados a través de medios computacionales. Mediante este procedimiento se estiman los parámetros que definen matemáticamente las curvas lógicas de cada reactivo y el modelo matemático que establece la relación entre el aprendizaje de los estudiantes y su desempeño en el examen.

De nuevo, la carencia de una concepción teórica que explique integradamente los aspectos cuantitativos y cualitativos del aprendizaje pretende subsanarse con un complejo modelo técnico. Aun cuando este modelo incorpora algunas características de los reactivos que constituyen una prueba, que el modelo lineal no incluye, no considera las características propias del aprendizaje humano y la intención de la medición del mismo; tampoco se explica por qué se parte de la idea de que el comportamiento del ser humano se relaciona precisamente en forma logarítmica con el aprendizaje. En los mismos términos en que está planteado el modelo se puede preguntar también por qué no se parte de otra forma de relación diferente a la logarítmica.

Son varios los supuestos de los que se parte en el modelo de características latentes; sin embargo, en este trabajo sólo discutiremos los dos que consideramos más importantes. En primer lugar, el supuesto de que las características latentes que subyacen en el desempeño de los estudiantes en un examen son múltiples y ortogonales; es decir, es un conjunto de características que no se relacionan entre sí. El supuesto de la ortogonalidad es inaceptable, ya que sería difícil, si no imposible, encontrar un conjunto de características del ser humano que no se relacionen entre sí, más aún si dicho conjunto de características influye en un determinado comportamiento, como es el desempeño de un examen.²⁸

El segundo supuesto establece que los reactivos de un examen sólo miden una característica latente. Este supuesto es idéntico al formulado para el modelo lineal, por lo que cuestionamos ahora

las curvas logísticas de dos parámetros. Para mayor profundización en el tema consultar R.K. Hambleton y L.L.Cook, op.cit., y F.M. Lord, “Practical Applications of Item Characteristic Curve Theory”, en *Journal of Educational Measurement*, V. 14, núm. 2, 1977.

²⁶ R.K. Hambleton y L.L. Cook, op.cit., pp. 87-89.

²⁷ Idem, p. 84.

²⁸ La forma en que algunos autores justifican el supuesto de la ortogonalidad es concibiendo las características latentes como factores producto de una transformación estadística llamada “análisis factorial ortogonal”. Sin embargo, ésta es una forma artificial de lograr factores no correlacionados entre sí. Frecuentemente dichos factores agrupan variables de tal manera que no existe una clara interpretación del significado de éstos. Además, no existe la seguridad de que en esta factorización aparezca un factor con la habilidad que se desea medir. Si en los casos de interés de este trabajo no aparece el factor aprendizaje, la utilización del modelo de características latentes pierde sentido. Revisar las críticas al análisis factorial de Pablo González Casanova, en *Las categorías del desarrollo económico y la investigación en las ciencias sociales*. México, UNAM, 1967, p. 23.

su validez en los mismos términos que lo hicimos anteriormente. Se pretende ignorar la interrelación múltiple de la conducta humana con factores dentro del contexto histórico social. Esta interrelación no nos permite aislar el efecto de una sola característica para medirla a través de una prueba. Algunos de los que proponen este modelo expresan sus dudas acerca de la validez del supuesto, como Hambleton y Cook, quienes, parafrasando a Lord, señalan que...” el supuesto acerca de la naturaleza unidimensional de un conjunto de reactivos no es estrictamente verdadera para la mayoría de las pruebas. Sin embargo (...) en algunos casos constituye una aproximación aceptable”.²⁹ Los procesos y productos del uso tanto del modelo lineal como de los modelos de características latentes no alcanzan el estatus científico; son, a lo sumo, en términos de Gloria Benedito,... “abstracciones razonadas con precisiones cuantitativas, producto de una práctica ideológica o bien materia prima para una práctica científica”.³⁰

Conclusiones

Si la tendencia dominante actual en el desarrollo teórico del concepto aprendizaje no permite realizar cuantificaciones válidas, cabe entonces preguntar ¿por qué se continúa intentado cuantificar el aprendizaje?, ¿por qué se pretende suplir la construcción teórica del concepto aprendizaje con procedimientos técnico operativos?; más aún, ¿por qué la medición ha desplazado a la evaluación del aprendizaje? Sólo podremos encontrar respuesta a estas y otras interrogantes relacionadas a través del análisis de las funciones que la evaluación y la medición tienen en el contexto escuela-sociedad, y de las interrelaciones que éstas mantienen con otras funciones escolares. Para llegar a este análisis es necesario determinar la génesis de las funciones e interrelaciones y su evolución histórica.

Estos análisis y la búsqueda de respuestas a las interrogantes planteadas quedan fuera de los objetivos que nos habíamos propuesto en este escrito. No obstante, aquí se esboza lo que constituiría dicho análisis.

Desde la perspectiva funcionalista se concibe a la sociedad regida por principios meritocráticos; en ella, las diversas funciones debe ser realizadas de acuerdo con las capacidades y conocimientos de las personas; las responsabilidades y recompensas se conciben asociadas a las funciones. Lo anterior constituye la base de la estratificación en la sociedad.³¹ Es precisamente la escuela la institución donde la sociedad ha depositado la atribución de acreditar y certificar la capacidad y el conocimiento. A través de estas funciones, la escuela contribuye eficazmente a la estratificación de la sociedad: algunos sectores son promovidos a niveles elevados de escolaridad y alcanzan las mejores posiciones en la sociedad; otros grupos son inducidos a adquirir menos escolaridad y a conformarse con posiciones de más bajo estatus y menores recompensas.

Las funciones de acreditación y certificación se apoyan en la evaluación del aprendizaje escolar. Algunos autores sostienen que el concepto de evaluación surgió estrechamente vinculado al desarrollo industrial de Estados Unidos y, particularmente, al desarrollo de los conceptos tayloristas de la administración científica del trabajo.³² Se puede añadir que no sólo el concepto de evaluación surge con esos vínculos, sino también una nueva conceptualización eficientista de la escuela, que pretende dar satisfacción a las demandas del desarrollo industrial. Así se ha desarrollado, desde

²⁹ Lord, citado por Hambleton y Cook, op.cit., p. 77.

³⁰ Gloria Benedito, “El problema de la medida”, en N.A. Braunstein et al., op.cit., p. 168. Cabe aclarar que la frase citada la dirige Benedito a la aplicación del método cuantitativo en psicología; sin embargo, consideramos que se adecua perfectamente la caso de la medición del aprendizaje.

³¹ Consúltese K. Davis y W.E. Moore, “Some Principles of Stratification”, en *American Sociological Review*, V. 10, 1945, pp. 242-244, y Tlacott Parsons, op.cit.

³² Consultar A. Díaz Barriga (1987 y 1984), op.cit.; A.de Alba et al. (1984), op.cit.; E. Peña, “Las determinantes sociales de la evaluación del aprendizaje escolar”, en *Perfiles Educativos*, núm. 43.44, 1990.

principios del presente siglo, una tecnología educativa que tiene gran semejanza con las propuestas tayloristas de la administración científica del trabajo, y que recientemente se ha visto influida por la corriente humanista en la administración.³³ El fin primario que se pretende alcanzar con la tecnología educativa es resolver a través de la técnica la problemática educativa de una manera óptima, con una aséptica neutralidad ideológica.³⁴

En la actualidad, la tecnología educativa se encuentra presente en todos los ámbitos del quehacer educativo. Uno de ellos es el de la medición del aprendizaje. Las técnicas desarrolladas intentan cuantificar el aprendizaje con la pretensión de lograr un conocimiento científico-consecuentemente objetivo y neutral- que auxilie la tarea de la evaluación. Sin embargo, como se señaló, en la práctica educativa, en vez de fungir como auxiliar, la medición del aprendizaje ha desplazado a la evaluación.

Otras perspectivas diferentes de la funcionalista, las teorías del conflicto social, plantean su concepción de las funciones que realiza la escuela. Estas teorías hacen énfasis en que, en toda sociedad, los grupos dominantes tienden a reproducir las fuerzas productivas y las relaciones sociales de producción existentes. Para que esta reproducción tenga lugar es necesario que tanto los medios materiales de producción como la fuerza de trabajo sean reproducidas. Desde esta perspectiva, se sostiene que la escuela participa en la reproducción, entre otras formas, al contribuir en la estratificación de la sociedad. Se señala que la escuela transmite la ideología dominante y participa en la selección y preparación de los individuos para que pertenezcan a grupos de estatus particulares en la sociedad. La escuela inculca en los estudiantes el respeto por el orden establecido y las normas que rigen el comportamiento social,³⁵ y lo hace de manera diferenciada al infundir los valores de ... “independencia, logro y seguridad para los miembros de las clases dominantes y de subordinación, aceptación y fatalidad para los integrantes de las clases dominadas”.³⁶

Desde este punto de vista la evaluación se concibe como un elemento que contribuye a que la escuela realice la función reproductora. Esto es, las calificaciones o notas escolares, producto de mediciones del aprendizaje escolar –y de supuestos procesos de evaluación-, reflejan algo más que el aprendizaje. Diversos estudios muestran que sólo un bajo porcentaje del significado de las calificaciones corresponde a aspectos cognoscitivos, y que el principal determinante de las calificaciones son las actitudes del maestro.³⁷ A través de la calificación el maestro recompensa la subordinación del estudiante al orden social escolar y castiga a quienes manifiestan autonomía, creatividad, independencia de juicio y autoconfianza, etc.³⁸ Aún más, las calificaciones recompensan el nivel de coeficiente intelectual “más allá de su contribución al logro verdadero”.³⁹ De esta manera, las calificaciones que pretenden ser el reflejo de la evaluación del aprendizaje, contribuyen a reproducir las relaciones jerárquicas que se encuentran en el trabajo, promueven la subordinación a

³³ Véase Eduardo Peña, *Análisis de los factores que influyen en el rendimiento escolar de los estudiantes*. (Mimeo.) 1987.

³⁴ Consúltese Alfredo Kuri y Roberto Follari, “Para una crítica de la tecnología educativa. Marco teórico e historia”, en A. de Alba et.al., *Tecnología Educativa*. Querétaro, México, Universidad Autónoma de Querétaro, 1985.

³⁵ Ver P. Bourdieu, “Cultural Reproduction and Social Reproduction.” y R. Collins, “Functional and Conflict Theories of Educational Stratification”, en J. Karabel y A. H. Halsey, *Power and ideology in Education*. Nueva York, Oxford University Press, 1977, y T.J. La belle y R. E. Verthine, “Nonformal Education and Occupational Stratification: Implications for Latin America”, en *Harvard Education Review*. V. 45, pp. 160-189.

³⁶ Carlos Ornelas, “Educación y sociedad: ¿consenso o conflicto?”, en G. González y C.A. Torres (coords.), *Sociología de la educación*.

³⁷ En términos estadísticos esto se refiere al porcentaje de varianza de las calificaciones, explicado por variables cognoscitivas. Consúltese H. Gintis, “Educación, tecnología y...”, en C. Biasutto, op.cit., p. 63, y H. Gintis y S. Bowles, op.cit., pp. 181-182.

³⁸ Evidencia contundente a este respecto puede encontrarse en H. Gintis, op.cit., pp. 63-71, y S. Bowles, op.cit., pp. 58-62 y 181-188.

³⁹ H. Gintis, op.cit., p. 66.

la autoridad jerárquica, la disciplina en las reglas establecidas y los modos de comportamiento cognoscitivos sobre los afectivos.⁴⁰ En este trabajo se señaló la ausencia de un cuerpo teórico que permitiese la cuantificación del aprendizaje y se proporcionaron los argumentos para afirmar que los modelos técnicos propuestos para la medición del aprendizaje no pueden suplir la carencia de una construcción teórica que, entre otras cosas, integre los aspectos cualitativos y cuantitativos del aprendizaje. El breve esbozo expuesto sugiere que la evaluación tiene funciones de reproducción diferentes de las que generalmente se le atribuyen. Por consiguiente, para formar el objeto de estudio de la evaluación del aprendizaje es necesario un marco teórico que considere el papel que desempeña y las interrelaciones que tiene –y ha tenido históricamente- en el contexto escuela-sociedad. Una vez que el objeto de estudio se haya formado, se estará en posibilidades de estudiar acuciantes temas escolares, como son el bajo rendimiento escolar, la alta reprobación, la frecuente deserción, etcétera.

⁴⁰ Ibid, 67-69.