

WILLIAM D. HEDGES

Los exámenes y la evaluación en la enseñanza de las ciencias

ETAPAS HISTORICAS

Las tentativas por medir lo que los estudiantes aprenden en la "escuela" pueden agruparse en más o menos cuatro grandes períodos. El primero va de la antigüedad hasta 1800 D.C. El segundo, comprende aproximadamente el siglo XIX e incluye el desarrollo de las técnicas estadísticas que hicieron posible la estandarización de los exámenes. El tercer período data de 1910 y concluye aproximadamente en 1960. En la fecha de prepararse este trabajo, se advierte el comienzo de un cuarto período en los exámenes y en la evaluación que acaso culmine en el desarrollo de nuevos instrumentos y técnicas de evaluación.

Si bien no se sabe cuándo se instituyó el primer sistema de exámenes escritos, si se sabe que la China había desarrollado ya un complejo sistema de exámenes nacionales en 2200 A.C. (1).

Acaso la constancia más lejana de un examen oral sea la que figura en el Antiguo Testamento (Jueces, 12:5-7). Scheidemann (2) describe la prueba a que fueron sometidos los efrateos por los gallarditas en los pasos del Jordán. Si el efrateo no podía pronunciar la palabra "Shibólet" y la pronunciaba "Sibólet", los gallarditas le daban muerte. Indudablemente, hoy diríamos que éste era un examen de criterio o una pregunta de examen.

Son muchos los casos conocidos de distintos exámenes de habilidad administrados por los grupos tribales primitivos a los jóvenes en el marco de los ritos preparatorios del ingreso a la edad adulta. Estos

• El Dr. Hedges es Jefe del Departamento de Educación Elemental de la Universidad de Florida, Gainesville, Florida, Estados Unidos.

• Documento de trabajo entregado en el II Seminario sobre Educación en Medicina Veterinaria en la América Latina, celebrado en Belo Horizonte, Brasil, del 21 al 31 de agosto de 1972.

exámenes, entre otras cosas, exigían que el joven soportara el hambre y la sed, ponían a prueba su valentía y medían su habilidad en la caza y la lucha (3).

Se sabe de los primeros exámenes objetivos para comparar los progresos de los jóvenes en Grecia. Por ejemplo, “la prueba de la evolución de un joven espartano consistía en averiguar si podía o no soportar el hambre, el frío y el esfuerzo físico; la prueba del joven ateniense radicaba en determinar si podía correr, saltar, correr carreras, declamar, repetir literatura y debatir” (4).

El examen universitario escrito, al parecer, solo comenzó a ser utilizado por las universidades hace menos de tres siglos. Probablemente el primer examen escrito en una universidad tuvo lugar en 1702, cuando fue instituido en Cambridge, Inglaterra” (5).

En resumen, el hombre ha reconocido las diferencias entre individuos por miles de años y, por ende, los ha sometido a distintos exámenes, tanto orales como escritos. Sin embargo, en fecha tan reciente como a principios del siglo XIX, los instrumentos con que podía efectuarse dicha medición eran escasamente avanzados y poco confiables y no se prestaban mucho a los fines comparativos. La aparición de este aspecto comparativo es lo que realmente nos ha llevado a algunas dificultades, como veremos.

El primer examen de interés histórico en nuestro país fue el examen de Boston de 1954 (6). Fue éste un examen escrito compuesto de 154 preguntas sobre temas de filosofía natural, geografía, historia y astronomía. El examen fue administrado a 530 estudiantes de 14 años, en 23 escuelas primarias de Boston. Dicho examen se había hecho necesario por la creciente dificultad que tenía la Comisión de Escuelas de Boston de examinar oralmente a todos los alumnos a medida que aumentaba la matrícula.

La importancia del examen de Boston radica en su efecto sobre Horace Mann, el educador norteamericano que tanto influyó en la estructuración de la educación en los Estados Unidos.

“Nos aventuramos a predecir —ha dicho Mann— que la modalidad de exámenes con preguntas impresas y respuestas escritas constituirá una nueva era en la historia de nuestras escuelas” (7). El más temprano indicio de una medición estandarizada de los resultados de la enseñanza tiene menos de 100 años de antigüedad. El profesor E. L. Thorndike informó de la preparación de un libro de escalas para la escritura, elaborado por un maestro inglés, el Rev. George Gisher, de la Greenwich Hospital School, en 1864 (8).

En nuestro país el primer examen estandarizado fue elaborado 30

años después, en 1894, por J. M. Rice (9). El Dr. Rice tuvo la idea de tratar de poner fin a la polémica entre los docentes respecto de los méritos relativos de distintos cursos y métodos de estudio. Como primer tema eligió la ortografía y elaboró un examen de ortografía de 50 palabras. Sin embargo, Rice se adelantó a su época. Es evidente que el examen estandarizado no existía siquiera en 1900.

Cuando el siglo XIX tocaba a su fin, algunos avances ajenos a la esfera de la educación posibilitaron el rápido desarrollo de la medición educativa poco después del fin del siglo. Gran importancia tuvieron los trabajos de Galton sobre los métodos estadísticos. Otro avance significativo, que por ligero margen precedió al interés en los exámenes para evaluar el progreso escolar y el desarrollo de éstos, fue la investigación sobre la inteligencia y la evolución de las pruebas psicológicas. Por ende, alrededor del 1900 se habían echado ya los cimientos del tercer período de evolución de los exámenes. El líder más destacado de este movimiento de medición del proceso educativo con posterioridad a 1900 fue Thorndike, quien en 1904 publicó su primer trabajo sobre la medición mental y social. Otros autores siguieron rápidamente sus pasos.

Ya para la Primera Guerra Mundial se habían desarrollado las técnicas estadísticas básicas y se había determinado claramente la necesidad de contar con una medición del efecto de la actividad escolar en los estudiantes. En el decenio de 1920 se habían elaborado tantos exámenes, en tantas esferas, que Monroe publicó una pequeña bibliografía de pruebas estandarizadas para las escuelas secundarias. En 1920 se había ganado la "batalla" de la medición del proceso educativo por medio de los exámenes objetivos. Con esto no se quiere decir que los exámenes fueran muy buenos. Sin embargo, con posterioridad a 1920, se observó un gran interés en la calidad de los exámenes, cuestión que todavía no hemos resuelto.

En el decenio de 1930 surgió una nueva tendencia en la medición del proceso educativo en la cual influyó Ralph W. Tyler por conducto de sus trabajos en el "Eight-year Study". Tyler puso el acento en el concepto de la evaluación, por oposición con los exámenes y se interesó en " . . . la medición de un conjunto más amplio de objetivos (esto es, un mayor número de aspectos de las características y tendencias de respuesta de las personas) que los previstos en los exámenes de evaluación del progreso escolar corriente" (10).

Este movimiento influyó de diversas formas sobre los planes de estudio; por ejemplo, si el maestro deseaba elaborar un buen examen debía tener conciencia de lo que estaba tratando de medir, es decir, de conocer a conciencia a sus alumnos. Por lo menos, debía tener con-

ciencia de sus objetivos. Los objetivos debían ser específicos al comportamiento, en vez de vagos y generalizados.

Con posterioridad a la Segunda Guerra Mundial podemos decir que el movimiento, hasta fecha relativamente reciente, pasó por un período durante el cual los exámenes se administraban en conjunto y sucesivamente al mismo sujeto o grupo de sujetos. Una de las circunstancias más significativas ocurrida recientemente en la esfera de la medición del proceso educativo ha sido la publicación de la obra de Bloom, *The Taxonomy of Educational Objectives. Handbook I: Cognitiva Domain*. Esta taxonomía clasifica las preguntas de los exámenes según el grado de complejidad de los procesos mentales necesarios para darles respuesta.

En resumen, el medio siglo que va de 1910 a 1960 fue testigo de una gran explosión en la esfera de los exámenes estandarizados y los elaborados por los docentes. Sin embargo, los exámenes estandarizados en las ciencias y en casi todas las demás esferas del conocimiento alcanzaron un elevado nivel de eficiencia en cuanto al objeto de la medición, si bien incorporaron un concepto que entraña un gran desaprovechamiento de los recursos humanos (que aún conservan).

En la mayor parte de los exámenes estandarizados y en los elaborados por los propios docentes, se observó la ascendencia del concepto de la selección por sobre el del desarrollo. El objeto de la evaluación pasó a ser (y sigue siendo) primordialmente la calificación y clasificación de los estudiantes con el objeto de excluir del sistema escolar a los ineptos. Los efectos sobre los estudiantes de todos los niveles han sido devastadores. Por un lado, unos pocos dijeron que los exámenes debían facilitar el diagnóstico de las necesidades de los estudiantes con mira a ayudarlos a aprender; por el otro, hubo quienes dijeron que los exámenes, desde el punto de vista comparativo, tenían por objeto excluir a aquellos que mostraran una aptitud insuficiente en un determinado momento. Cuando se evalúa a los estudiantes con el sistema normativo o de la curva normal, se les pone en una situación competitiva con los demás, se asegura el fracaso de muchos de ellos, se prescinde de cualquier criterio de excelencia y se alienta en los estudiantes la tendencia a no colaborar con sus pares.

“Las formas como comúnmente se utilizan los exámenes y otros elementos de evaluación en las escuelas facilitan poco el perfeccionamiento de la enseñanza y del aprendizaje y rara vez sirven para garantizar que todos (o casi todos), aprendan lo que el sistema escolar considera que son las tareas y los objetivos importantes del proceso educativo” (11).

Actualmente se perfila un concepto más amplio de la evaluación,

atinadamente enunciado por Bloom, Hastings y Madaus (12), quienes consideran la evaluación como:

1. Método para adquirir y elaborar las comprobaciones que son necesarias a fin de perfeccionar tanto el aprendizaje del estudiante como la enseñanza.
2. Una gran variedad de comprobaciones que van más allá de los exámenes escritos de rutina.
3. Instrumento para señalar con más claridad los fines y objetivos más importantes de la educación; y para juzgar el progreso alcanzado por los estudiantes en relación con dichos objetivos.
4. Sistema de control efectivo mediante el cual puede verificarse, en cada etapa, el proceso de enseñanza y aprendizaje, si éste es o no eficiente y, de no serlo, cuáles serán los cambios que será menester instituir con el objeto de preservar su eficiencia antes de que sea demasiado tarde.
5. La evaluación final como instrumento de la práctica docente para comprobar si los procedimientos sustitutivos son igualmente eficientes o no en la realización de un conjunto de fines educativos.

Los docentes contemporáneos tienen escasas dudas acerca de la necesidad de incorporar los exámenes y la evaluación en un proceso eficaz de enseñanza y aprendizaje. Por ello, es bien sabido por muchos de nosotros que los hábitos de estudio del estudiante se modifican, para bien o para mal, en función de la orientación que los pedagogos den a sus exámenes. Este factor se manifiesta en la tendencia de los estudiantes a individualizar el estilo de las pruebas de cada docente con el objeto de saber cómo prepararse para ellas.

Si los hábitos de estudio del estudiante se ven afectados por las inclinaciones que el docente manifiesta en sus exámenes, de ello se desprende que si el docente pone el acento en lo trivial, los estudiantes se inclinarán a estudiar trivialidades. Además, la enseñanza sin ninguna modalidad de evaluación de los resultados nos parece ahora virtualmente inconcebible. Esta evaluación es importante para los estudiantes; para los docentes, que pueden así percatarse de su grado de eficiencia pedagógica; para los educadores, cuando tratan de modificar su estrategia pedagógica, y para quienes elaboran los planes de estudio, en lo que hace a revisar y actualizar sus programas.

Lamentablemente se ha producido una gran confusión en cuanto a la eficacia de los exámenes en todos los niveles, desde la escuela elemental hasta el instituto de estudios superiores. Esta confusión se deriva, al

menos en parte, de la ambigüedad de los fines de los exámenes. Se deriva, además, de una confusión entre los procesos discretamente distintos de medición y evaluación. Una cosa son los resultados numéricos obtenidos en un examen válido y confiable (medición), y otra los juicios subjetivos que el docente asigna al significado de estos resultados numéricos (evaluación), en cuanto éstos se vinculan con el estudiante, su propia enseñanza, etc.

FUNCIONES ESPECIFICAS DE LOS EXAMENES

Es sumamente importante que se adviertan con claridad las funciones que deberán desempeñar los exámenes. Si bien las decisiones pertinentes corresponden esencialmente a los docentes y a quienes elaboran los planes de estudio, ellos por lo menos deberían tener una idea clara de las funciones que sus exámenes han de cumplir. Seguidamente se indicarán someramente algunas de sus principales funciones.

Primero, cuando las preguntas de los exámenes u otras medidas de evaluación giran en torno a objetivos esenciales y claramente definidos, es decir, en torno a los elementos más significativos del plan de estudios, los estudiantes tienden a concentrar el aprendizaje en estos objetivos.

Segundo, una evaluación preliminar cuidadosa puede servir para que el estudiante y el profesor sepan dónde y cómo deben comenzar. Evidentemente esto resulta particularmente importante en el caso de los niños pequeños, pero ciertamente es útil cuando se trata de individualizar los programas en la enseñanza superior. Es común, sin embargo, que empecemos “por el comienzo”, es decir, partiendo de la hipótesis de un igual grado de ignorancia (o conocimiento) de todos los estudiantes.

Tercero, los exámenes *a posteriori* pueden comprobar el grado de asimilación o la falta de ésta y servir de base para evaluar factores tales como la eficiencia pedagógica, el progreso del estudiante y la modificación del plan de estudios. Cuando se combinan las pruebas *a posteriori* con los exámenes preliminares, se pueden observar indicios de progreso o cambio en los estudiantes. No debe pasarse por alto la orientación para el aprendizaje complementario que puede obtenerse mediante el uso de procedimientos eficientes de exámenes *a posteriori*.

Cuarto, los exámenes provisionales o formativos cumplen una función de vigilancia, gracias a la cual el estudiante y el docente ajustan sus esfuerzos con miras a aumentar su eficiencia. Esta verificación frecuente es el reverso de los exámenes acumulativos, menos frecuentes, en los que lo que interesa son las “notas”, que para lo que sirven primordial-

mente es para producir tensión en algunos estudiantes. Se ha comprobado que cuando hay un grado considerable de ansiedad la motivación auténtica es poca o nula.

Quinto, la elaboración y evolución de los exámenes u otras medidas de criterio afines suelen perfeccionar y reforzar la idea que tiene el docente de sus objetivos pedagógicos. Cuando el docente formula sus objetivos primero, elabora sus exámenes en segundo término y, a renglón seguido, prepara su material y su estrategia pedagógica, aumenta radicalmente su eficiencia y equilibra, al propio tiempo, el programa. Compárese la metodología que se ha descrito con la secuencia más tradicional que va de la selección de los materiales primero, la elección (más o menos aleatoria) de las preguntas de los exámenes en segundo término, y por último, cuando se requiere el “imaginarse”, más o menos, los objetivos. Cuando tenemos un conocimiento —por somero que éste sea— de la pedagogía, un claro concepto de los objetivos que queremos alcanzar sirve para aumentar nuestra eficiencia.

Sexto, no debe pasarse por alto ni minimizarse la función de aprendizaje que se deriva de las pruebas frecuentes. Los teóricos del proceso de aprendizaje sostienen que el recuerdo de lo aprendido refuerza la retención. El descubrimiento de las respuestas correctas a las preguntas objetivas cuidadosamente formuladas de un examen va más allá del reconocimiento y entraña un grado considerable de recordación.

En todo lo antes expuesto es de suma importancia el concepto de la comprobación del resultado. Una sencilla analogía física servirá para aclarar este punto. Por ejemplo, el tirador sólo puede modificar su puntería cuando conoce la ubicación del impacto de los disparos anteriores. De tener esto presente, al planificar y elaborar nuestros exámenes y procedimientos de examen podremos evitar algunos de los errores más comunes que han difamado a los exámenes. Seguidamente examinaremos algunos de los defectos más comunes de los exámenes elaborados por los propios docentes.

DEFICIENCIAS COMUNES DE LOS EXAMENES ESCRITOS PREPARADOS POR LOS DOCENTES

Un defecto importante es la insistencia en verificar que el alumno haya adquirido un conocimiento de bajo nivel, esencialmente de memorización. Un estudio realizado por el autor y otros colegas respecto a los exámenes de mitad del año en física, química, biología y ciencias del primer año en un centenar de escuelas secundarias del Estado de Virginia (Estados Unidos) puso de manifiesto que se da preferencia a la memorización (13).

En particular, el 78% de las preguntas de los exámenes se refería al recuerdo de información objetiva (el nivel 1,00 de Bloom). De este 78%, el 64% de las preguntas (o sea, el 50% de las 14,000 preguntas aproximadamente) sólo requería la forma más sencilla de recordar (el nivel 1,10 de la taxonomía de Bloom). Si se parte de la hipótesis de que los maestros incluyen en sus exámenes de mediado de año los puntos que consideran más importantes, y si se parte de la hipótesis, ya mencionada, de que los hábitos de estudio de los estudiantes se ven afectados por los hábitos que pone de manifiesto el docente en la preparación de sus exámenes, cabe pensar que la situación de la enseñanza de las ciencias en las escuelas de segunda enseñanza de Virginia en ese punto era, por lo menos, alarmante. Si se realizaran estudios similares en las escuelas y "colleges" de nuestros institutos de estudios superiores, quizás se hallarían situaciones análogas.

Otro gran defecto de los exámenes preparados por los maestros radica en el desconocimiento que éstos tienen de las virtudes, limitaciones y funciones de los distintos tipos de pruebas, preguntas y otros instrumentos y procedimientos de pruebas. Al formular las preguntas de los exámenes escritos no tienen presente las funciones específicas de los de tipo analítico, las pruebas de opción múltiple, los exámenes de lo verdadero y de lo falso, los de correspondencia, los de complementación, los de solución de problemas y los de situación o aplicación. Son frecuentes las polémicas ciegas entre docentes sobre si se debe o no usar un determinado tipo de examen. Estas polémicas no tienen razón de ser porque el criterio de utilización debe depender del objetivo del docente y de si un determinado tipo de pregunta conduce o no a dicho objetivo.

Un tercer defecto importante es el desconocimiento básico de los principios de redacción de preguntas y de preparación de exámenes en general. Su planificación, preparación, administración y análisis requieren cuidado y conocimiento.

Un cuarto error consiste en que rara vez se emplean procedimientos de análisis de preguntas sencillas con el objeto de reunir preguntas acertadas y precisas en torno de objetivos específicos. Se observa que las técnicas estadísticas elementales utilizadas para analizar y perfeccionar los exámenes son ineficientes. Este uso limitado va acompañado de una falta general de comprensión de los conceptos básicos de la medición, a saber, validez, confiabilidad, correlación, error típico, etc.

Otro defecto importante no se deriva tanto de las fallas de los exámenes mismos, sino de la utilización de los resultados. Uno de los "flagelos" más importantes y comunes de este movimiento de los exámenes ha sido la aplicación ciega del concepto normativo a la evaluación acumulativa de los estudiantes. En el método de calificaciones utilizado corriente-

mente se aplica el concepto de la curva normal, lo que significa que, por definición, el 50% de los estudiantes de cada clase recibirá la evaluación de deficiente, sea cual fuere su nivel de rendimiento. La recíproca también es cierta. Además, la aplicación ciega de este concepto normativo ha llevado, con el correr del tiempo, a un gran temor, tirantez y hostilidad por parte de los estudiantes, dando lugar a una intensa competencia desvinculada de niveles particulares de excelencia. Los estudiantes han querido tener un nivel superior al de los demás, sea cual éste fuere, y se ha comprobado una considerable variación de calidad entre las instituciones de enseñanza, los “colleges” de estas instituciones y los departamentos de estos “colleges”, por no mencionar las diferencias entre los docentes de un departamento.

Por otra parte, si los docentes precisaran lo que los estudiantes deben aprender y comprobaran luego si éstos han adquirido o no dicho conocimiento, el éxito vendría a ser sinónimo del hecho de alcanzar el nivel de asimilación previsto para el objetivo prescrito y diferiría de lo que los demás estudiantes pudieran o no hacer. Este es un concepto fundamental en el aprendizaje por asimilación conforme se formula en la obra de Block, Bloom y otros, circunstancia que nos excusa de extendernos más a este respecto en este trabajo.

Se ha comprobado, además, como ya se dijo y como lo han señalado Cronbach, Bloom, Block y otros autores, que se gana la mitad de la batalla cuando se enuncia con claridad y precisión lo que debe aprenderse en términos de comportamiento observables, en un determinado nivel de competencia y en condiciones claramente establecidas. En esas circunstancias los exámenes dejan de ser juegos de adivinación y pasan a ser realmente instrumentos auxiliares del estudiante y de los docentes. Se elimina así una fuente importante de celos y desconfianzas, pues el estudiante ya no se ve forzado a tratar de adivinar las intenciones del profesor.

Lamentablemente, en esta etapa suele plantearse otro problema. Los estudiantes comienzan a desempeñarse demasiado bien, es decir, hay muchos “sobresalientes” y ello puede dar lugar a que la ira de los órganos directivos de las instituciones de enseñanza se descargue sobre los profesores. Esto es interesante (y hasta irónico) cuando se recuerda que presumiblemente una, por lo menos, de las funciones de la enseñanza es el aprendizaje; en ese caso, la afirmación de que los estudiantes están aprendiendo demasiado bien, se transforma en un aspecto reprehensible de la enseñanza eficiente. Sin embargo, ésta es la realidad que tiene ante sí el profesor que, luego de haber con exactitud informado a sus estudiantes que deben alcanzar un determinado grado de compe-

tencia, se ve obligado a producir 27 “sobresalientes” y tres “distinguidos” en una clase de 30 alumnos. Una de las soluciones que en los últimos años se ha dado a este problema ha sido la transición al sistema del cual se utilizan solamente las notas de “aprobado” e “insuficiente”.

Aunque en este trabajo se ha destacado hasta ahora la utilización de los exámenes escritos, no se quiere con ello sugerir que éstos sean los instrumentos únicos o necesariamente más útiles. Actualmente se están ampliando las investigaciones destinadas a vincular este tipo de exámenes con las situaciones que entrañan comportamientos observables, de modo que puedan utilizarse indistintamente (14). Además, hay muchos otros elementos de un programa que, hasta la fecha, sólo pueden ser evaluados mediante la observación directa del desempeño del estudiante en el contexto de una situación dada. Esto nos lleva a preguntarnos, ¿Qué debemos medir?

Solemos tender a destacar las dimensiones congoscitivas del aprendizaje con menoscabo de lo actitudinal. Sin embargo, sabemos que el éxito que llegamos a tener en nuestras esferas de actividad depende tanto de nuestros sentimientos, creencias, opiniones, valores, actitudes y aptitudes en las relaciones humanas como de nuestra pericia técnica. ¿Saben ustedes de alguien que haya sido despedido “porque no podía trabajar con nadie más, ni siquiera consigo mismo”? Con esto sólo quiero decir que se están haciendo avances en la evaluación del ámbito afectivo, a pesar de la renuencia cultural que se observa en los Estados Unidos a penetrar en lo que, por lo común, se supone es del dominio privado.

Además de incorporar las dimensiones del ámbito afectivo en el “objeto de la medición”, es preciso que el nivel de ésta en el ámbito congoscitivo sobrepase el de la memorización o el recuerdo intenso con mira a insistir cada vez más en la aptitud de aplicación, análisis, síntesis y evaluación, conforme a lo sugerido por Bloom y otros autores.

EL PERFECCIONAMIENTO DE NUESTROS PROGRAMAS DE EXAMENES

Existen diversas posibilidades de perfeccionar nuestros programas de exámenes y de evaluación. Su éxito, sin embargo, dependerá de que los docentes sepan formular claramente los objetivos y tengan, por lo menos, cierto conocimiento de los principios sobre los cuales se basan los exámenes y la evaluación.

El segundo paso, sería elaborar los procedimientos de evaluación, incluidas las combinaciones de preguntas en torno de objetivos. Si

se obtiene la cooperación y coordinación de los profesores de los diversos establecimientos de enseñanza interesados en asignaturas similares, no será necesario que cada profesor elabore por sí solo dichos procedimientos.

En relación con el análisis estadístico de las preguntas que forman parte de nuestros exámenes, desde hace ya algún tiempo que venimos utilizando la elaboración electrónica de datos en lo que se relaciona con el análisis, la confiabilidad y la validez de las preguntas. En la elaboración de exámenes, especialmente de los exámenes escritos, se recurre cada vez más a las combinaciones de preguntas. Las facultades interesadas en perfeccionar sus programas de exámenes deben contar con una persona competente en dicha esfera que se encargue de organizar y ejecutar estos trabajos con la colaboración de cada una de ellas.

Sin embargo (y acaso éste sea el punto principal de nuestra posible orientación futura), el programa moderno de evaluación hace hincapié en los procedimientos de evaluación formativos, es decir, en la evaluación sistemática durante los procesos de elaboración de los planes de estudio, enseñanza y aprendizaje, por oposición a la evaluación acumulativa al término de un curso o programa. En la evaluación acumulativa se juzga *a posteriori* del proceso y no durante el proceso. Esta diferencia es crucial, pues modifica todo el concepto de la evaluación, ya que en lugar de considerársela como medio de eliminación y selección, se la concibe como un arbitrio, gracias al cual el estudiante puede percibir los objetivos que debe alcanzar para permanecer en el proceso educativo. En lugar de ser un proceso punitivo negativo, la evaluación pasa a ser un proceso gratificante y útil.

En resumen, en el breve tiempo que se me ha asignado, he tratado de presentar, aunque en forma somera, la historia y las principales tendencias de la evaluación en los últimos años. Si bien, mi presentación ha sido bastante esquemática, abrigo la esperanza de haber logrado bosquejar en sus mentes la idea que responderá a estas dos preguntas: ¿De dónde venimos? ¿Dónde estamos?

REFERENCIAS

- (1) Lang, Albert R.: *Modern Methods in Written Examinations*. Houghton Miffling Company, Boston, 1930, pág. 1.
- (2) Scheidemann, Norma V.: The Earliest Recorded Objective Test, *School and Society*, Vol. XXIX, junio, 1929, pág. 702.
- (3) Russell, Charles: *Standard Tests*. Ginn and Company, Boston, pág. 15.
- (4) Russell, Charles: *Standard Tests*. Ginn and Company, Boston, pág. 16.

- (5) Lang, Albert R.: *Modern Methods in Written Examinations*. Houghton Miffling Company, Boston, 1930, pág. 3.
- (6) Lang, Albert R.: *Modern Methods in Written Examinations*. Houghton Miffling Company, Boston, 1930.
- (7) Caldwell, Otis W., y Curtis, Stuart A.: *Then and Now in Education, 1845-1923*. World Book Company, Yonkers-on-Hudson, Nueva York, 1924, pág. 44.
- (8) Ayres, Leonard P.: "History and Present Status of Educational Measurements", en *Measurement of Educational Products*, Seventeenth Yearbook of the N. S. S. E., Parte II. Public Publishing Company, Bloomington, Illinois, 1918, pág. 9.
- (9) Monroe, N.: *Journal of Education and Research*, Vol. XXXVIII, pág. 339.
- (10) Monroe, N.: *Journal of Education and Research*, Vol. XXXVIII, pág. 339.
- (11) Bloom, B. S., Hastings, J. T., y Madaus, G.: *Handbook of Formative and Summative Evaluation of Student Learning*. McGraw-Hill Book Company, Nueva York, 1971, pág. 7.
- (12) Bloom, B. S.; Hastings, J. T., y Madaus, G.: *Handbook of Formative and Summative Evaluation of Student Learning*. McGraw-Hill Book Company, Nueva York, 1971.
- (13) Hedges, William D.; Lochhead, Ian C., y McDougall, Mary Ann: *An Analysis of Test Items to Evaluate the Objectives of Science Teachers in Virginia Combined and Secondary Schools*. División de Investigación Educativa de la Universidad de Virginia, Charlottesville, N. C., Estados Unidos, 1964.
- (14) Lawrence, Gordon: Measuring Teacher Competences for the Middle School, *The National Elementary Principal*, Vol. 51:60-66 (nov.), 1972.