

10

ALFREDO SILVA

Catedrático de Educación
Universidad de Puerto Rico

LA EVALUACIÓN DEL APROVECHAMIENTO

Ponencia presentada en el Seminario sobre Evaluación bajo los auspicios de la Oficina de Evaluación y Orientación de las Escuelas Elemental y Superior de la Universidad de Puerto Rico, el 7 de marzo de 1955.

UNA PÍLDORA DE FILOSOFÍA PEDAGÓGICA

Los sistemas escolares públicos de América, incluyendo las universidades y escuelas adscritas a éstas, existen para traspasar la herencia social democrática, la parte de la cultura greco-romana o cultura de Occidente que sea factible, de las generaciones pasadas a las del presente y las del porvenir, con vistas a que estas generaciones disfruten de ella plenamente y la enriquezcan. La misión es no sólo de la universidad y escuelas a ella adscritas sino de todo el sistema escolar vigente. El caudal cultural de Occidente no es posible integrarlo en igual medida en todas las personas que constituyen la sociedad puertorriqueña. De ese caudal cultural debemos seleccionar

juiciosamente, no sólo para los distintos componentes o núcleos de alumnos del sistema escolar y la Universidad sino que, de lo seleccionado para cada componente —la escuela elemental, por ejemplo—, se hará otra selección a base de *qué* y *cuánto* para determinados grados de desarrollo, y dentro de éstos, para las distintas habilidades y aptitudes. Y serán los entendidos los que harán la selección en razonable proporción, tanto de la cultura de Occidente, como de ésta se ha definido por los eruditos, como de esa modalidad o variación que es la cultura puertorriqueña. Hecha la selección justamente proporcionada, el sistema de instrucción pública y la Universidad, valiéndose de todos los recursos económicos, filosóficos, sociológicos y pedagógicos fomentarán, no sólo la difusión de la cultura greco-romana o de Occidente, sino de su variación, la puertorriqueña, y no sólo fomentarán el enriquecimiento de la de Occidente por medio de la investigación creadora y la creación artística, sino que también se procederá de igual manera con los valores puertorriqueños. Ignorar o menospreciar los valores sociales y los valores pedagógicos de la cultura puertorriqueña, equivaldría a una amputación o poda de una rama en el árbol de la cultura matriz, la Occidental.

Una dosis sobre el arte de desarrollar planes de estudio.

El *qué* y *cuánto* de la cultura Occidental y de su variación, la cultura puertorriqueña, que deberán incluirse en el plan de estudios del sistema escolar tendrán que ser determinados por una comisión compuesta de especialistas en programas de estudio, en filosofía de la educación, en administración y supervisión escolar, en metodología, en psicología pedagógica, en psicología del niño, en psicología de la adolescencia, en sicometría, en sociología educativa, en las distintas áreas del programa de estudios de la escuela elemental y de la superior, y por último, maestros de las distintas divisiones del sistema escolar, todos de reconocida competencia. Este comité de timón, con la cooperación de otros comités, desarrollaría el plan de estudios que es labor incesante de construcción y reconstrucción.

El plan de estudios así desarrollado tendría que sustentar una filosofía de la educación para el sistema escolar con un fin último de la educación y fines generales que condujeran a aquél. Para cada disciplina —la aritmética, digamos— o grupo de disciplinas afines —las ciencias sociales, por ejemplo— también se redactarían unos fines generales, y además una serie de objetivos de tres clases por lo menos; 1) Objetivos de largo o mediano alcance; 2) Objetivos de alcance intermedio; y 3) Objetivos de alcance próximo o inmediato. Todo mentor que jurara un cargo en el sistema escolar público estaría moral y legalmente obligado a conocer y sostener la filosofía de la educación democrática que adoptara por procedimiento democrático dicho sistema escolar.

Es, pues, a los desarrolladores de los programas escolares, con la participación democrática del magisterio, y con alguna participación del alumnado y los padres, a quienes corresponde fijar las metas del plan de estudios que el magisterio deberá realizar. Solamente si los programas de estudio de las distintas disciplinas no contuvieran dichos fines y objetivos es que el maestro se verá obligado a determinar, con la cooperación de colegas de responsabilidad análoga, cuáles deberán ser esos fines generales y cuáles los objetivos específicos de distintos alcances que le orientarán en la enseñanza de su disciplina. Los tres métodos usados son: el de análisis curricular, el de conferencia, y el de cuestionario y entrevista.

PROPÓSITO Y SIGNIFICACIÓN

Establecidos los fines y objetivos en los programas de estudio, el problema se circunscribiría a realizar dichas finalidades en las personas de sus alumnos. Entra ahora la metodología en acción, asesorada por la psicología pedagógica y por la supervisión. La metodología sugerirá medios y recursos y materiales de instrucción para realizar los fines adoptados. La labor escolar se hará significativa para el mentor porque abriga propósitos. Quien no los tiene es barco sin timón, sin rumbo y a la deriva. Es ciego dando palos a ciegas. No sabe

qué hacer ni por qué hace lo que hace. En forma análoga al alumno que ejecuta sus actividades sin propósito suyo alguno, porque su maestro se lo impone, así el maestro que no abraza propósitos espontáneamente sentidos vivirá confundido, indeciso, disgustado y sin experimentar el calor del entusiasmo que provocan las vivencias intensamente sentidas.

✧ METAS, MEDICIÓN Y EVALUACIÓN

Si se acepta que lo primero que cualquier ser necesita es saber a dónde va antes de arribar a alguna meta, lo segundo sería cuánto ha recorrido de la meta y cuánto le falta por alcanzarla. Algo semejante ocurre al maestro. Determinados cuáles son los fines generales de la educación y cuáles los objetivos de la disciplina, para el grado o año que enseña, el deber ineludible del educador, luego de dirigir eficazmente el aprendizaje, será evaluar continua y sistemáticamente durante el año escolar el grado en que los objetivos van realizándose. Quien no se sienta arrastrado por fines y objetivos específicos e intensos no podrá evaluar porque carecerá de los puntos de referencia que constituyen los valores pedagógicos aceptados. De lo cual se desprende que todo educador necesita tener un conocimiento siquiera mediano de la filosofía de los valores pedagógicos, eje de la filosofía de la educación.

MEDICIÓN Y EVALUACIÓN

Evaluar es fijar el valor de alguna cosa. Es valorar, apreciar, estimar el valor, el mérito, la condición, etc., de uno o más entes, a la vez que se tiene como punto de referencia algún ideal de la educación. “Evaluar”, dice Lindquist, “significa un proceso de apreciar que envuelve la aceptación de valores específicos y el uso de una variedad de instrumentos de observación, incluyendo instrumentos de medición, como bases para emitir los juicios de valor”. Medir, en pedagogía, es determinar el cuánto, la cantidad, las unidades en que uno o más objetivos pedagógicos se han realizado. Monroe diferen-

cia la medición de la evaluación diciendo que la medición subraya los aspectos aislados del aprovechamiento en la materia de enseñanza o destrezas y habilidades específicas, mientras que la evaluación va enderezada a los cambios amplios de la personalidad y a los principales objetivos del programa escolar. Siendo la evaluación un concepto más amplio, incluye la medición como una de sus técnicas. Se mide, para evaluar mejor; se evalúa para estimar los cambios ya hechos, o que deberán hacerse, en las personas de los alumnos, en los programas escolares, en la organización escolar, y aún en la metodología, si una sensata interpretación de los resultados lo justificase.

¿Con qué se evalúa y se mide? Wrightstone ofrece la siguiente lista de instrumentos de observación con vistas a evaluar: exámenes de aprovechamiento, de actitud, de personalidad y de carácter; escalas de estimación ("rating scales"), cuestionarios, escalas para juzgar productos, entrevistas, técnicas de observación controlada, técnicas sociométricas, apuntes anecdóticos, informes taquigráficos, grabaciones, técnicas proyectivas, etc. Claro está que algunos de estos instrumentos miden aprovechamiento; otros no.

¿Qué conceptos de lógica, ética, biología, psicología, pedagogía y filosofía se usan para designar los objetivos del programa escolar y los objetivos de las asignaturas enseñadas cuya realización habrá de ser evaluada? Entre otros, se me ocurren los siguientes: Crecimiento, desarrollo por madurez, desarrollo por ejercicio, aprendizaje, leyes, principios, teorías, hipótesis, reglas de procedimiento, conceptos, datos y hechos, destrezas o habilidades intelectuales, destrezas motrices, hábitos generales, hábitos de estudio, aprendizajes concomitantes, valores, sentido de los valores, actitudes, actitudes valorativas, intereses, sentimientos, emoción estética, gustos (literario, musical, artístico, etc.), nivel de aspiración, comprensiones, adaptabilidad social, ideales, aptitudes o talentos especiales, inteligencia general, inclinaciones, aversiones, apreciaciones; y la serie numerosa de nombres simbolizadores de rasgos físicos, morales, vocacionales, estéticos, etc., de los alumnos, como por ejemplo, personalidad, carácter, cooperación, ajuste emocional, concien-

cia social, liderato, capacidad crítica, creación, responsabilidad, tolerancia, iniciativa, independencia, persistencia, etc. Algunos de estos productos de la instrucción son medibles con exámenes de aprovechamiento; otros necesitan distintos instrumentos para ser evaluados, los que anteriormente he mencionado y que no son exámenes de aprovechamiento.

¿QUÉ ES EL APROVECHAMIENTO?

El aprovechamiento es todo lo que sirve de provecho a un ser viviente, todo lo que le aprovecha o logra tras continuado esfuerzo. El aprovechamiento debería ser el conjunto de logros resultantes de la instrucción, que aprovecha a los alumnos porque constituyen avances hacia los fines y objetivos de la educación. En este concepto amplio del aprovechamiento, habría que incluir no sólo los conocimientos, destrezas, comprensiones y creaciones, sino rasgos menos mensurables como el sentido de los valores, la persistencia en el trabajo, la capacidad crítica, los gustos, etc. Walter Cook nos dice, sin embargo, que aunque se habla mucho de aprovechamiento y de exámenes de aprovechamiento, se ha prestado escasa atención a la definición de aprovechamiento. Agrega Cook que aprovechamiento se usa en el sentido de habilidad para hacer, capacidad para hacer o tendencia a hacer”.

Así definido el concepto, el sicólogo y el pedagogo de América del Norte usan el término aprovechamiento para referirse a los logros de la instrucción en tres aspectos de la personalidad: el aspecto intelectual, el motor y el social. Pero ni americanos ni europeos incluyen los logros en el desarrollo anatómico-fisiológico, ni los avances en el desarrollo emocional, ni en el volitivo, como integrantes del aprovechamiento, si bien reconocen que hoy, temperamento y voluntad y cuerpo sano son metas valiosísimas de la educación. Y ahora viene el sicómetro. Cuando éste habla de exámenes de aprovechamiento lo que corrientemente incluye son exámenes de información, de razonamiento, de habilidades intelectuales, como la comprensión de lo leído, y habilidades sicomotrices como la ortografía.

Los exámenes de diagnóstico, hechos para descubrir lo que el alumno ha aprendido correctamente, lo adquirido incorrectamente y lo no aprendido que debió serlo, si bien no se consideran exámenes de aprovechamiento, sí se les reputa complementos indispensables de aquéllos. Tampoco se les juzga así a los exámenes de personalidad, ni por supuesto, a los de talentos especiales, ni a los de inteligencia. Luego, los exámenes de aprovechamiento miden sólo una parte de los muchos rasgos de personalidad desarrollables en el aula. Si estos rasgos medidos se habrán de considerar los más valiosos dependerá de la filosofía de la educación que adoptemos, de los valores pedagógicos que sustentemos. Sin embargo, el mentor que no pueda mostrar objetivamente los logros alcanzados en los rasgos medibles de la personalidad de sus alumnos no podrá justificarse alegando que alcanzó más valiosos fines que aquéllos por ser los más intangibles: las actitudes, los valores, gustos, intereses, etc. Por todas estas razones es que los exámenes de aprovechamiento necesitan ser complementados con otras categorías de exámenes como los de diagnóstico, de inteligencia y de talentos, escalas de actitudes, exámenes de pensamiento crítico, técnicas sociométricas, inventarios de personalidad, entrevistas, apuntes anecdóticos, etc.

Antes o después de dar los exámenes de aprovechamiento debería administrársele a cada alumno uno o más exámenes de inteligencia con el fin de descubrir una de las causas más obvias del bajo o alto aprovechamiento. Si la inteligencia del alumno es alta, deberá mostrar un grado de aprovechamiento correspondiente a su grado de desarrollo mental. Si éste es bajo el maestro no pedirá peras al olmo. Claro que en ambos casos deberá el maestro evaluar el nivel de aspiración de cada alumno. La inteligencia de X podrá ser superior, pero si su nivel de aspiración es bajo se conformará con un aprovechamiento mínimo. En este caso el maestro deberá motivar a X para elevarle su nivel de aspiración, ayudándole así a lograr más alto aprovechamiento. Asimismo deberá solicitar con tacto el maestro del alumno de escasa habilidad mental, pero con ambiciones desmedidas, que haga o aprenda primero lo que esté a su alcance intelectual y luego aspire a lo que guste.

Si, a pesar de su inteligencia mediana y su aceptable nivel de aspiración, X logra un bajo aprovechamiento se deberá buscar la causa en otros factores que no son los dos mencionados, tal vez uno o más de los siguientes: inhabilidad especial en una o más de las disciplinas, es decir ineptitud debida a aprendizajes deficientes o falta de aprendizajes imprescindibles para dominar la asignatura. Cualquiera de estas dos causas podría ser descubierta con un buen examen de diagnóstico. Otras causas podrían ser hábitos defectuosos de estudio, ausencias excesivas del maestro y el alumno, relaciones antagónicas entre mentor y discípulo, mala salud corporal o mental del escolar, visión o audición defectuosa, trastorno emocional por traumatismo o frustración y, por consiguiente, por conflicto mental, enfermedad orgánica, etc. Cada una de estas causas requiere terapia distinta. El examen médico precede al examen psicológico o siquiátrico. Antes o después del examen de aprovechamiento los alumnos deberían tomar un examen de aptitud o talento especial con vistas a descubrir en ellas aquellas habilidades excepcionales que en su día podrían brindar al alumno dinero, prestigio, gloria o felicidad. Verificado el descubrimiento, el maestro vendrá obligado, dentro de sus limitaciones, a hacer todo lo posible a su alcance para desarrollar los talentos descubiertos hasta el máximo asequible.

En cuanto a los exámenes de diagnóstico, éstos deberán administrarse inmediatamente después de terminada la enseñanza de una unidad de la asignatura, o de alguna división del programa de estudios, o libro de texto si el programa no lo especifica. La unidad, o división del programa, o del libro de texto deberá ser analizada cuidadosamente para descubrir sus elementos más valiosos, desde el punto de vista de la filosofía educativa, del sistema escolar o de la escuela a la cual es responsable el maestro. Para hacer el examen de diagnóstico una muestra adecuada de los elementos más valiosos de la unidad, o la división del programa, se prepararán en forma de ejercicios apropiados. Ejemplo: Sobre la materia "suma de quebrados", el examen deberá contener ejercicios de cálculo y de razonamiento con quebrados de la mayor utilidad social, a base

de la frecuencia de uso, para la sociedad puertorriqueña, según estudios que se hayan efectuado en nuestro medio social. Si en la Isla no se hubieren efectuado investigaciones de esa clase ni el programa de aritmética las resume, será deber del maestro recurrir a las publicaciones de los especialistas en la enseñanza de la aritmética para asesorarse sobre el asunto. Hay tanto que enseñar dentro de los límites de cualquier disciplina que si no se evalúa y se selecciona la materia a base de algún criterio filosófico o sociológico, el maestro se arriesgará a seleccionar mal, es decir, a seleccionar lo socialmente inútil y, por consiguiente, a no enseñar ni evaluar lo útil. Ejemplos de ejercicios aceptables: $\frac{1}{2} + \frac{1}{4}$; $1\frac{1}{4} + \frac{3}{4}$; $2\frac{3}{8} - 3\frac{1}{4}$; etc. Ejemplos de ejercicios inaceptables en la escuela elemental: $1/11 - 1/13$; $5 \frac{2}{17} \times 9 \frac{4}{19}$; etc. Podrían ser aceptables en un curso de aritmética avanzada pero no en un primer curso. Algo semejante podría aconsejarse en relación con una prueba sobre el uso de los signos de puntuación, o sobre las habilidades que el programa de lectura tiene como objetivos a realizar para determinado grado. Contestado y corregido el examen de diagnóstico, se hace un censo o conteo, en papel cuadriculado, de lo no aprendido que debió serlo, y de lo erróneamente aprendido, en el caso de todos y cada uno de los alumnos. Este conteo deberá servir de base para el tratamiento (o terapia) correctivo individual sobre la unidad o división. Quien no proceda así en su labor remediadora estará dando palos a ciegas, engañándose consciente o inconscientemente al creer que sus alumnos están realizando los objetivos de las asignaturas, cuando la verdad es que no lo están.

CLASES DE EXÁMENES DE APROVECHAMIENTO

Según que el examen de aprovechamiento sea de varias asignaturas o de una sola es que recibe su nombre. Si es de una sola asignatura se denomina examen de asignatura (*subject test*). Si incluye pruebas de varias asignaturas se le llama examen de aprovechamiento general. Un ejemplo de este último es el 'Modern School Achievement Test', con un total de

diez pruebas de asignatura. Otro ejemplo es el "Standard Achievement Test", dado por la Srta. Constance Towner, el que suscribe y otros, bajo la dirección del Dr. Harold O. Rugg, a 70,000 alumnos de las escuelas elementales de Puerto Rico allá por el 1925, durante el primer *survey* de nuestro sistema escolar. Tal vez ustedes recuerden el Test Hispanoamericano, desarrollado en nuestro Colegio de Pedagogía por el Dr. Fred C. Walters y la Sra. Mercedes Chiqués de Walsh y otros. La Sra. Rosa Navarro de Hayden preparó uno de ciencia elemental, ejemplo de un examen de asignatura. Para la escuela superior está el "Sones-Harry", entre otros, que es de aprovechamiento general. Exámenes de asignatura en lenguas modernas, ciencias naturales, ciencias sociales, etc., son innumerables. Los buenos exámenes de aprovechamiento tienen un manual con instrucciones para darlo, corregirlo e interpretar los resultados siempre de igual manera. Los resultados se expresan corrientemente en calificaciones de grado ("g-scores"), o calificaciones de edad, llamadas estas últimas edad educativa ("age score" o "educational age"). Los resultados de exámenes de la escuela superior y colegio se expresan percentiles. Los exámenes estandarizados contienen formas impresas para conservar los resultados, y en algunos, para dibujar el perfil pedagógico de cada examinando. Disponer de un perfil educacional para cada alumno de la escuela que asesora es uno de los ideales del sicólogo escolar.

Si se sigue un procedimiento científico al confeccionarlos, entonces deberán denominarse exámenes *estandarizados* de aprovechamiento. El opuesto a éste es el examen no estandarizado que es de dos clases: el de viejo tipo o de discusión, y el objetivo o de nuevo tipo, hecho por el maestro. Las tres variedades: estandarizados, de nuevo tipo, y de discusión pueden usarse siempre que se conozcan sus ventajas y limitaciones, que sería prolijo considerar aquí.

Los tests estandarizados tienen limitaciones para el magisterio puertorriqueño, además de las que tienen para el magisterio de los Estados Unidos. La mayor parte de los que aquí podrían usarse están en inglés, con normas estadounidenses.

No tienen normas para Puerto Rico ni han sido traducidos para nosotros. En inglés no pueden administrarse en Puerto Rico porque su validez y fidelidad bajarían al darse aquí bajo las mismas condiciones que allá. Lo ideal sería que en la Isla estandarizáramos nuestros exámenes. Un sustituto pobre sería traducir los tests estandarizados norteamericanos para re-estandarizarlos en Puerto Rico, o por lo menos, establecer normas puertorriqueñas para los traducidos de los Estados Unidos. Claro que los exámenes estandarizados de lectura inglesa no se pueden traducir para fines de medición. Lo más que podría hacerse es usar los de cuarto o quinto grado de allá, para los séptimos u octavos grados de acá, y establecer normas puertorriqueñas para ellos.

Consideremos los exámenes de discusión. Para numerosos sicómetros están desacreditados debido a los resultados de ciertas investigaciones que se han hecho. En una de ellas se repartieron fotocopias de un examen contestado en una clase de geometría, examen de un alumno de escuela superior, para que lo corrigieran profesores expertos de geometría y lo devolvieran al investigador. Starch y Elliot recibieron 116 copias del examen de geometría debidamente corregidas. A pesar de que se trataba de un examen de geometría, que se presta a la corrección objetiva, los resultados fueron los siguientes: De las 116 calificaciones o notas asignadas 2 excedían de 90, 1 era inferior a 30, 20 eran de 80 o más y 20 de 60 o menos. 47 le dieron nota de aprobado o más, mientras que 69 lo ficharon de suspenso. Cuando el éxito o el fracaso escolar de la niñez y la adolescencia depende del subjetivismo caprichosamente cambiante de sus mentores, ¿qué garantía de protección tienen los escolares para que se les haga justicia?

Oigan ahora este segundo escándalo que nos relata el eminente sicómetra Ben D. Wood. Había un grupo de correctores expertos ("readers") evaluando papeles de exámenes de historia. Un corrector, luego de corregir varios papeles de exámenes, preparó lo que él consideró una clave modelo del examen de discusión con el fin de usarla para corregir los restantes papeles. Dicha clave se traspapeló y cayó en manos de otro co-

rector que creyó que era el examen de un alumno. Corrigió la clave modelo, siempre con la falsa creencia de que evaluaba el examen de un alumno. Le asignó una nota de fracaso. Como era la costumbre de que cada papel fuera corregido por varios de los correctores, así se hizo, con el resultado de que la clave modelo de un papel de historia redactada por un corrector experto recibió calificaciones que variaron entre 40 y 90.

El viejo examen de discusión, sin embargo, recibe reconocimiento en algunos tratados de sicometría. Odell, por ejemplo, cree que debe usarse. En todo caso la corrección de este examen deberá objetivarse lo más que sea posible. Debe recordarse, sin embargo, que en determinado tiempo el examen de discusión cubre solamente una cuarta parte de los detalles de información que el nuevo tipo objetivo de exámenes incluye en igual tiempo, según las investigaciones de Ruch. También es más trabajoso para corregir. Gates y otros afirman que el principal valor del examen de discusión radica en su uso como una base para apreciar destrezas tales como resumir, organizar, evaluar o aplicar información, más bien que para recordar detalles. En las páginas 558 y 559 de la tercera edición de *Educational Psychology*, Gates nos ofrece siete excelentes reglas para aumentar la validez y fidelidad de la evaluación del examen de discusión, reglas que el tiempo me impide transcribir. Las mismas reglas podrían aplicarse para evaluar otros productos de la instrucción, tales como la calidad de la caligrafía, la composición escrita, la costura, el dibujo natural, etc. El capítulo 13 de *Educational Measurement* por Lindquist y otros, lo dedica Stalnaker totalmente al examen de discusión.

En cuanto al examen objetivo de nuevo tipo cierto-falso, pareo, etc., todo maestro que salga de nuestro Colegio de Pedagogía deberá conocer la técnica sobre cómo hacerlos, corregirlos e interpretarlos. Para esto último deberá saber, por lo menos, convertir las notas brutas en notas derivadas. Su popularidad entre los que han tomado cursos de sicometría es notable. El tratado clásico sobre este tipo de examen es *The Objective or New Type Examination*, por M. Ruch. Si el lector sólo dispone de tiempo para leer un capítulo, éste deberá ser

el número 7 de *Educational Measurements*, por Lindquist y otros. Sería pueril que trate aquí de sintetizar la técnica de hacerlos, corregirlos e interpretarlos. Deberá recordarse que, entre los muchos propósitos que existen para dar exámenes de aprovechamiento, los cuatro siguientes deberán ser recordados: 1) Para descubrir y evaluar el grado de progreso de cada alumno hacia los objetivos que persigue la enseñanza de las asignaturas; 2) Para provocar en los alumnos el deseo de estudiar más y mejor; 3) Para desarrollar un interés permanente, o por lo menos, duradero, en la asignatura, y 4) Para dar notas. La materia de enseñanza, símbolo de la cultura, sugerida en los programas de estudio, es para ser dominada de la manera más eficaz posible y por los métodos de aprendizaje y enseñanza que estén más en armonía con la psicología pedagógica y la mejor filosofía de la educación.

El nuevo tipo de examen requiere planeo antes de empezar a hacerlo. Parte importante del planeo consiste en determinar de antemano los porcentajes del contenido del examen. Por contenido del examen se entiende, según Vaughn, "la materia de enseñanza del examen o del ejercicio individual del examen y el tipo de habilidad que dicha materia requiere". Ejemplo de planeo en el Examen de Aprovechamiento de Ciencia Premédica de la Asociación de Colegios Médicos Americanos:

1.—Objetivos de la instrucción

a) Comprensión, interpretación de material científico	35	35
b) Aplicación de conceptos, principios, etc.	30	36
c) Recuerdo de conceptos básicos	35	29
TOTAL	100	100

2.—Campo de la materia de enseñanza

a) Ciencias biológicas	40	39
b) Química	40	40
c) Física	20	21
TOTAL	100	100

Luego de terminarse el planeo, se procede a redactar los ejercicios que deberán ser de varios tipos: 2, 3 ó 4 tipos diferentes. La materia de enseñanza que se seleccione para cada ejercicio será la que aparezca más relacionada con los objetivos de la asignatura. Como los exámenes miden con más exactitud mientras más largos y objetivos son, cualquier examen de nuevo tipo debería incluir no menos de 75 ejercicios. Los llamados “quizzes” de 15, 20, 25 ó 30 preguntas no son fidedignos porque miden con escasa exactitud. Otorgarán una medida al alumno que en un número considerable de casos será un engaño, un fraude y no más. Luego de dar a los alumnos por lo menos tres pruebas de veinticinco ejercicios se podrían sumar los tres resultados, y proceder luego a convertir dichos resultados en notas derivadas por fórmulas conocidas.

Como un fin de la instrucción para todas las disciplinas es capacitar para razonar con la materia de la asignatura, no menos de una tercera parte de los ejercicios de los exámenes en las asignaturas de contenido se prepararán en forma de problemas y razonarse y aplicación de principios, leyes, teorías, hipótesis y reglas de procedimiento. Otra tercera parte será para descubrir la habilidad para comprender e interpretar la materia enseñada, y la restante tercera parte intentará descubrir el recuerdo de conceptos básicos.

El examen de pareo y el de respuestas múltiples, conocido como “best answer”, si están bien hechos, se adaptan para la medición de los procesos racionales. El problema, en términos de la materia de enseñanza se redacta, y debajo de aquél se sitúa la contestación correcta, que será una de entre 3 a 5 respuestas. Cada alternativa contendrá algún grado de plausibilidad o “atracción”, pero la alternativa correcta será la que tenga más plausibilidad. El pareo se presta para que el alumno aplique reglas a ejemplos, leyes científicas a casos que las ilustren. Una ventaja del tipo cierto-falso es que estimula y desarrolla la habilidad crítica. En la vida diaria el hombre instruido aprecia la veracidad o falsedad de las afirmaciones que lee o escucha aunque por urbanidad, timidez o temor no emita un juicio sobre ellas.

Al redactar exámenes objetivos se tendrá cuidado de que todos los ejercicios de cualquier subexamen sean del mismo tipo. Ejemplos: el primer subexamen podría ser todo de "cierto-falso"; el segundo, todo de 3 alternativas, el tercero, todo de pareo; y el cuarto, todo de llenar blancos. En los subexámenes hechos de ejercicios de 2, 3, 4 ó 5 alternativas, éstos se corregirán usando la fórmula para neutralizar la adivinación. Resultado = No. correctas — No. incorrectas sobre N-1. *Ejemplo:* En el cierto-falso el resultado será el número de ejercicios incorrectamente contestados menos el número de ejercicios correctamente contestado. Esto significa que del total se descuentan los ejercicios incorrectamente contestados a dos puntos cada uno y los ejercicios no contestados a un punto. No hay tiempo para explicar las otras fórmulas. Para mayor exactitud y rapidez en la corrección de estos tipos de exámenes se usa una clave preparada por el autor o los autores del examen. Para ciertos exámenes estandarizados hay máquinas correctoras.

La creación literaria sólo puede evaluarse motivando al alumno para que escriba composiciones, narraciones, descripciones, discursos, poemas, apreciaciones y crítica sobre temas de su experiencia. Cualquiera de estas producciones literarias que fuera escrita por todos los alumnos sobre el mismo tema podría ser evaluada guiándose por las siete reglas de Gates que aparecen en *Educational Psychology*. Hace años Thorndike y Hillegas estandarizaron dos escalas para evaluar composición literaria. Para evaluar las composiciones literarias las siete reglas de Gates podrían adaptarse, si bien ha habido tanteos para inventar más mejores instrumentos de evaluación. Las escalas de Ayres, Gray, Leamer y Thorndike están en el mercado para evaluar la calidad de la escritura. Los carteles de Freeman y de West pueden usarse para diagnosticar los elementos deficientes de la caligrafía. Las escalas de Ayres, Buckingham y Iowa existen para medir deletreo u ortografía inglesa. El Negociado de Publicaciones del Teachers College, Columbia University, vende exámenes de pensamiento crítico en los estudios sociales, cuyo autor es Wrightstone.

COCIENTE DE APROVECHAMIENTO Y COCIENTE DE PROFICIENCIA

Los resultados de los exámenes estandarizados de aprovechamiento pueden expresarse en términos de años educacionales, grados de habilidad ("g-score"), calificaciones "standard", percentiles, calificaciones en puntos ("point-score"), etc. Si el resultado del examen estandarizado lo expresa el manual de direcciones en años educacionales, se podrían obtener dos medidas evaluativas para cada alumno: el cociente educacional ("educational quotient") y el coeficiente de proficiencia ("accomplishment quotient"). El cociente educacional se obtiene dividiendo la edad educacional por la edad cronológica y multiplicando por 100. Si la edad educacional que se obtiene es en una asignatura se llamará edad en lectura, en aritmética, etc. Dividiendo la edad en una asignatura por la edad cronológica se obtiene el cociente en la asignatura. De ese modo si Pedro tiene 12 años en habilidad para leer y 12 de edad cronológica su cociente en lectura es 100, que es normal.

Si la edad educacional se divide por la mental y el resultado se multiplica por 100, tendremos el coeficiente de proficiencia. Parecería que ésta es la medida ideal, el sueño de un sicómetra, para evaluar el aprovechamiento general de un alumno, pues con él se está comparando lo que un alumno efectúa o realiza (su edad educacional) con lo que podría lograr (su edad mental). Ejemplo: un alumno que obtuviera una edad educacional de quince años con una edad mental de 15 años estaría aprendiendo a toda capacidad. Pero si su edad educacional fuera doce años y su edad mental quince, su eficiencia en el aprendizaje sería 80%. A pesar de tal bella teoría Cureton en el *Journal of Experimental Education* de marzo de 1937, aporta datos para desacreditar el coeficiente de proficiencia.

CÓMO EVALUAR OTROS LOGROS

El desarrollo intelectual, motor y social que se da por sentado en los exámenes de aprovechamiento no es el único digno de evaluación. Existe también el aprovechamiento que es consecuencia de las experiencias de laboratorio, antes lla-

madas actividades extracurriculares, las cuales la pedagogía moderna valora en alto grado. Se necesitará una extraordinaria habilidad mental para crear un método válido y consistente para medir o evaluar los aprendizajes derivados de las experiencias de laboratorio. ¿Cuándo y bajo qué condiciones se efectúan espontáneamente, con sinceridad, estas actividades? ¿Cuándo y bajo qué condiciones se efectúan a regañadientes, por lo cual los aprendizajes concomitantes pueden ser indeseables o efímeros? La única ayuda más práctica del momento para evaluar las experiencias de laboratorio la constituye la lista de cotejo.

Hay dos experiencias o actividades extracurriculares que todos los semestres solicito de mis alumnos y que tengo que evaluar. Son éstas el estudio de un caso de un niño o un adolescente, según sea el curso de sicología, y las lecturas suplementarias.

El estudio del niño o adolescente que mis alumnos me entregan el 30 de noviembre, o de abril, lo evaluó como "aceptado" o "rechazado". Si no lo entregan o si se lo rechazo, reciben una calificación de "incompleto" y tienen un semestre para entregarme otro aceptable. Si lo acepto no les doy crédito específico en puntos, aunque por su excelencia podría significar una subida de nota de C a B, por ejemplo.

En cuanto a las lecturas suplementarias, por cada informe de un capítulo de por lo menos una carilla que me entreguen conteniendo una síntesis creadora con no menos de dos implicaciones derivadas de la síntesis, les acredito $\frac{1}{4}$ de punto hasta un máximo de 6 puntos de crédito por 24 informes. Estos 6 puntos de crédito se los agrego a su promedio a fin de año al computar la calificación final.

BIBLIOGRAFÍA

1. Cureton, E. E., "The accomplishment Quotient Technic", *Journal of Experimental Education*, March 1937, pp. 315-326.

2. Gates, Arthur, *Educational Psychology*, third rev. ed., pp. 556, Macmillan, 1948.
3. Lindquist, E. F., *Educational Measurement*, American Council on Education, 1951. Chapters 10 and 13.
4. Monroe, Walter S., *Encyclopedia of Educational Research*, Macmillan, 1950. "Evaluation".
5. Ross, C. C., *Measurement in Today's Schools*, Prentice Hall, 1954.
6. Ruch, G. M., *The Objective or New Type Examinations* Scott, Foresman, 1929.
7. Skinner, Charles E., *Educational Psychology*. Prentice-Hall, 1951, Chapters 22 and 23.