

ELABORACION Y ANALISIS DE REACTIVOS DE OPCION MULTIPLE EN MICROBIOLOGIA Y PARASITOLOGIA*

JOSÉ SOSA-MARTÍNEZ
LUIS ISITA SÁNCHEZ
FRANCISCO GARCÍA TORRES
ROLANDO ARCE HIDALGO

La determinación de métodos de evaluación justos y adecuados para la apreciación del logro de objetivos alcanzados por los educandos constituye un problema en la docencia.

Existen diversos procedimientos para la evaluación del proceso enseñanza-aprendizaje. En ellos se utilizan diferentes tipos de reactivos (preguntas) como son: falso-verdadero (pregunta de una alternativa de respuesta), de respuesta breve, de opción múltiple, de ensayo corto y ensayo largo, de correlación de columnas (asociación de conceptos), de interpretación de gráficas o cuadros, de prueba oral, etc. (Goring, 1971; Lemus, 1974; Lafourcade, 1969).

Metas del departamento

Cualquiera que sea el procedimiento, la evaluación es parte de un sistema en donde los reactivos deben adecuarse para determinar el logro de los objetivos de aprendizaje. Así, los reactivos deben adecuarse para determinar los reactivos de evaluación dependerán estrechamente de dichos objetivos (Lafourcade, 1974).

En relación a nuestras actividades de enseñanza de la microbiología y parasitología, se examinaron las posibilidades y limitaciones de

cada uno de los tipos de reactivos mencionados en base a nuestras metas.

Específicamente, éstas son las siguientes: a) lograr la calificación justa y precisa de un número de alumnos (que generalmente varía entre 250 y 350 por semestre), en corto tiempo para que sirva de verdadera retroinformación al estudiante y que, a la vez, se pueda calificar a todos los alumnos en una sola sesión; b) explorar los diversos niveles del conocimiento; c) utilizar un sistema confiable, es decir, que permita la reproducibilidad de las calificaciones, aun al ser obtenidas por diferentes evaluadores; y d) aplicar el análisis estadístico a los resultados para lograr un juicio preciso de las características intrínsecas del reactivo, p. ej. validez, confiabilidad, grado de dificultad, grado de discriminación, entre otros.

De los tipos de reactivos que hemos probado a través de los años, encontramos que el de opción múltiple se adaptaba mejor a nuestras necesidades. En la literatura consultada se encontró ampliamente recomendado y utilizado (Lafourcade, 1969; Goring, 1971; Lindquist, 1951 (citado por de Mattos, 1963); Neumann, 1979; Foster, 1979; Kim, 1965; Medical Examination Review, 1961). Este reactivo se caracteriza (Lafourcade, 1969) por contener las siguientes partes: a) la base o enunciado que consiste en una proposición o problema que se establece en forma de pregunta o de afirmación incompleta, y b) las opciones o alternativas que dan la respuesta a la base

* Departamento de Microbiología y Parasitología de la Escuela Superior de Medicina, Instituto Politécnico Nacional.

y que se le presentan al examinando para que elija de entre ellas la que considere correcta.

Características del reactivo de opción múltiple

Las razones por las que encontramos que el reactivo de opción múltiple satisfacía nuestras necesidades se enumeran a continuación:

1. Permite calificar un número elevado de alumnos en un tiempo corto. En forma manual, si se cuenta con una o varias tarjetas maestras preparadas previamente. Cada tarjeta con 60 reactivos puede ser calificada por un profesor a una velocidad aproximada de 30 segundos. Si, por otra parte, la detección de la opción correcta puede efectuarse bajo un sistema de computación electrónica y tarjeta *Porta-Punch Hollerith* de 80 columnas, se añade rapidez y precisión (Comité de Evaluación de Aspirantes a Residencias Médicas en las Instituciones de Salud del Sector Público, 1978). El tiempo requerido por este procedimiento es de unos minutos para 300 alumnos, obteniendo, además, el análisis estadístico de reactivos, opciones, grados de dificultad y de discriminación, etc. (*vide infra*).

2. Admite un grado marcado de objetividad, es decir, hace mediciones en relación estrecha con los objetivos delineados en el programa ya que las opciones que se proponen como alternativas se pueden expresar de tal manera que no haya ambigüedades, que no influyan en el examen factores como el azar, o factores personales de los evaluadores y aún de los mismos examinandos como, ortografía, facilidad o dificultad de expresión y otros.

3. El reactivo de opción múltiple puede, si está bien diseñado (de Mattos, 1963), tener un grado alto de confiabilidad, o sea que al aplicarse a varios grupos homogéneos en su preparación, reproduce con aproximación estadística los mismos resultados (Goring, 1971; Magnusson, 1969).

4. De igual manera, este tipo de reactivo puede tener la validez que el evaluador le asigne. Si consideramos que por validez se entiende el grado de precisión con que se desea

evaluar la profundidad y amplitud de un conocimiento (Magnusson, 1969; Goring, 1971), se deberá utilizar un reactivo ajustado al objetivo, al método de enseñanza, al campo de enseñanza, a los libros de texto, a los apoyos audiovisuales, etc.

5. En relación con el inciso anterior, es pertinente manifestar que a este reactivo se le puede variar su grado de facilidad casi con la precisión requerida y, más aún, el evaluador puede, de acuerdo a sus necesidades, hacer el reactivo más fácil o más difícil, lo que bastaría con leves modificaciones a una o varias de las opciones que contenga. Así, si es de aceptarse de la Tabla Analítica de Reactivos proporcionada por la computadora que el grado de facilidad promedio es cuando lo contesta acertadamente más del 35 por ciento pero no más del 65 por ciento, entonces se puede redactar este reactivo de manera que su grado de facilidad quede dentro de los límites deseados. En este momento consideramos pertinente exponer que, si bien la mayoría de los autores utiliza el término "grado de dificultad", en el presente ensayo le llamamos "grado de facilidad", de acuerdo con Nérci (1973). Agregaremos que para nosotros es más útil llamarlo "grado de facilidad" ya que en los datos de la computadora se expresa que a mayor dificultad corresponde un índice menor, y viceversa, lo cual parecería antitético.

6. Más aún, si los reactivos de opción múltiple pueden ser analizados por métodos electrónicos, se puede modificar de igual manera su grado de discriminación, o sea aquella cualidad del reactivo de poder diferenciar entre alumnos de alto rendimiento escolar y alumnos de bajo aprovechamiento.

7. El reactivo de opción múltiple es muy versátil en su aplicación. Puede utilizarse tanto para resolver propuestas como para la interpretación de gráficas, discusión de casos clínicos, planteamiento de problemas de diversa índole, diferenciación de definiciones y aseveraciones, solución de problemas matemáticos, etc.

8. El reactivo se presta bien a la exploración de los distintos niveles del área cognos-

citiva (Bloom, 1975; Schyfter, 1976). Así, permite determinar conocimiento, comprensión, aplicación y aún análisis, que son niveles más explorados en una asignatura básica a nivel de pregrado.

9. El alumno no pierde el tiempo en redactar y escribir las respuestas, como sucede con otros tipos de reactivos, por lo que dedica casi todo el tiempo del examen para pensar. Además de lo anterior, el reactivo requiere poco tiempo del alumno para ser contestado; un examen conteniendo 60 reactivos suele ser contestado entre 50 y 70 minutos por un 75 por ciento de los alumnos examinados y casi por un 100 por ciento a los 90 minutos.

10. Facilita la organización y manejo de un banco de reactivos, conteniendo el número que se considere necesario de manera que se eviten repeticiones en los diversos exámenes, por una parte y, por otra, la probable memorización por los alumnos. Más aún, este banco posibilita la selección de reactivos de acuerdo con el énfasis que se desee poner según lo cubierto del programa y, además, se hace impersonal con respecto al docente ya que el departamento puede seleccionar los reactivos, ya sea a través del jefe o de la academia de profesores.

11. Si el alumno desea revisión de examen es fácil mostrarle sus errores y aciertos en forma rápida y precisa y, además, se le puede informar sobre las áreas específicas en donde tiene fallas o debilidades en su aprendizaje, si es que las áreas del programa en evaluación se distribuyen en determinados reactivos, p. ej., los reactivos del número 11 al 15 pueden cubrir el tema de estafilococos e infecciones por estafilococos.

12. Si la evaluación se hace, como ya lo hemos mencionado, con 60 reactivos en una tarjeta perforable y si en un examen, p. ej., de inmunología, los últimos 10 reactivos (del 51 al 60) corresponden a la exploración de conocimientos del tema "Hipersensibilidad", entonces al revisar los resultados de estos reactivos podremos hacer comparación entre los distintos grupos de alumnos que corresponden a un período lectivo. Por ende, puede ser fácil-

mente investigar, al hacer un examen departamental (los mismos reactivos a todos los grupos de diferentes profesores), si un grupo en especial exhibió alguna deficiencia significativa, la que puede ser discutida con el profesor del grupo correspondiente.

13. De igual manera, se podrán hacer comparaciones con respecto a sexo, edad, grupos, temas del programa, y múltiples factores que, en forma natural o experimental, se introduzcan en el proceso enseñanza-aprendizaje como son: conferencias, películas, textos, monografías, conferencias de integración, etc., cuya influencia se podrá determinar si se introducen como variable única en algunos grupos.

14. En comparación con el tipo denominado "falso-verdadero", el reactivo de opción múltiple reduce considerablemente el acierto aleatorio. Así, en nuestra experiencia hemos observado que los aciertos por azar, por lo general, no pasan de 20 por ciento en comparación con la cifra de 50 por ciento en "falso-verdadero".

Preparación de reactivos de opción múltiple

A continuación se exponen los principios básicos para la preparación de reactivos de opción múltiple.

La base debe ser explícita y sencilla en su redacción y expresada en lenguaje bien conocido por los examinados, ya que una palabra poco usual puede constituir un factor de desorientación. De igual manera, debe estar de acuerdo con el objetivo que se desee evaluar. Tendrá una longitud suficiente para el planteamiento correcto de la proposición. No debe contener palabras que señalen la respuesta ni que, por otra parte, sirvan de señuelos que desvíen demasiado la atención del alumno. Es pertinente enfatizar el hecho aceptado por diversos autores (Goring, 1971; Lafourcade, 1969; Lemus, 1974) que es preferible que la base no contenga negaciones y mucho menos dobles negaciones, ya que el alumno dedicará parte de su atención a resolver la construcción gramatical y su significado. De igual manera, se procurará no hacer propuestas que

incluyan la expresión "excepto", como: "Todas las siguientes sustancias son linfocinas, excepto:", ya que en un momento crítico, como por lo general lo es un examen, puede introducir un elemento de confusión.

En cuanto a las características de las opciones, mencionaremos las siguientes:

Se procurará que una sola de ellas sea la correcta, o la más adecuada, o la más compatible. El número de opciones que se utilizan es de 4 y de 5. Aunque por una parte, el número de 5 disminuye los aciertos por azar, por otra aumenta el trabajo requerido por los docentes que se encargan de redactar los reactivos.

Es lógico asumir que los distractores deben ser suficientemente atractivos para los alumnos y así conferir al reactivo el grado de facilidad y de discriminación adecuados.

Además, al igual que la base, deberán estar redactados en forma explícita y precisa. La opción correcta deberá representar ciertamente la contestación apropiada del conocimiento que se pretende evaluar. Las opciones no deberán ser extensas en su redacción, deberán estar gramaticalmente bien expresadas y ser semejantes en su extensión.

No es conveniente que entre las opciones exista alguna como: "Ninguna de las anteriores" o "Todas las anteriores", ya que se considera que si el alumno se percata que dos o más opciones son compatibles con las palabras "ninguna" o "todas" se inclinará más a alguna de ellas, aun sin poseer el conocimiento requerido.

Creemos pertinente mencionar que es inaceptable que las preguntas (o bases) sean capciosas o ambiguas, ya que desvirtúan el objetivo de evaluar el aprendizaje de los alumnos. De igual manera, es de evitarse aquella serie de reactivos secuenciales en donde la contestación del primero determina las respuestas de los reactivos subsecuentes, pues en caso de que aquél no haya sido resuelto correctamente hace que los siguientes tampoco lo sean ("reacción en cadena").

Una vez que los reactivos han sido elaborados, es muy conveniente presentarlos para su

resolución al conjunto de profesores que imparten la materia, bajo los dos objetivos fundamentales de: a) revisar si todos los profesores cubrieron los temas evaluados en forma adecuada; y b) si hay consenso en las respuestas, ya que la condición *sine qua non* es que cada reactivo deberá tener una opción correcta y no deben existir dualidades de criterio.

Finalmente, es muy importante que antes de un examen (o a principios del periodo lectivo) se proporcionen a los alumnos instrucciones pertinentes sobre el tipo de examen a que se sujetarán; sobre el uso de la tarjeta, de las características de las opciones, de la necesidad de leer adecuadamente los reactivos, de perforar únicamente una sola opción, de la seguridad de que han seleccionado las opciones que crean correctas antes de perforar, y otras.

Análisis de reactivos de opción múltiple por medio de la computadora

En base a lo expresado se prepararon reactivos con la finalidad de evaluar el logro de objetivos de aprendizaje de nuestro curso de microbiología y parasitología, cubriendo la temática de las 4 secciones, a saber: I. Inmunología y Generalidades de Microbiología y Parasitología, II. Bacteriología, III. Virología-Micología y IV. Parasitología. Estos reactivos estaban compuestos por la base expresada ya en forma de pregunta sobre un tema específico o en la descripción de un breve caso clínico, o de una gráfica o gráficas sobre las que se formularon los reactivos pertinentes. Cada base tenía 5 opciones (*a, b, c, d y e*), de las que una de ellas únicamente era la correcta y las otras 4 estaban en la categoría de distractores.

Desde un principio se comprendió que lo expuesto anteriormente como índices de facilidad y de discriminación deberían caracterizar a estos reactivos. Sin embargo, también se reconoció que para determinar la calidad de los mismos en forma real era necesario apli-

carlos previamente en las pruebas de evaluación de los alumnos y, de las contestaciones obtenidas hacer los debidos cálculos matemáticos para definir los parámetros señalados anteriormente, lo cual comprendía una labor de extraordinario consumo de tiempo. Además, era necesario fijar si las opciones consideradas como distractores llenaban el objetivo desde el punto de vista de semejanza con la contestación correcta, pero sin llegar a confundirse con ésta; ni tampoco alejarse tanto en el contenido conceptual como para no llegar a ser distractor. Así, si un 50 por ciento de los alumnos contestan correctamente la opción verdadera, las contestaciones del otro 50 por ciento de los alumnos deberán repartirse entre los cuatro distractores y, de preferencia, en forma parigual. En otras palabras, si uno solo de los distractores recibe un 40 por ciento de contestaciones, entonces se confunde con la opción verdadera; mientras que si, por otro lado, otro distractor recibe únicamente un 5 por ciento o menos de contestaciones, reflejará que se aleja mucho de la opción verdadera y, consecuentemente, no llena el objetivo de ser distractor.

De acuerdo con lo anterior, se vio la necesidad de utilizar un procedimiento electrónico apropiado para hacer en forma rápida y precisa los cálculos que permitieran seleccionar los buenos reactivos, o mejorar aquéllos con algún defecto en sus opciones o en sus grados de facilidad y discriminación. Al respecto, mencionaremos que el Departamento de Estadística de esta escuela nos proporcionó todas las facilidades para llevar a cabo los cálculos correspondientes por medio de una computadora con la que se resolvió la problemática de tiempo y precisión en los estudios estadísticos.

Con el objeto de llevar a cabo las evaluaciones, se prepararon listas de 60 reactivos para cada una de las secciones mencionadas del programa. Cada lista fue aplicada a su debido tiempo, de acuerdo con la calendarización de evaluaciones del curso.

Tabla analítica de reactivos

	Muy difíciles menos de 20%	Difíciles 20 a 35%	Promedio 36 a 65%	Fáciles 66 a 80%	Muy fáciles más de 80%
Correctos					
Mejorables					
Defectuosos					
Desechables					

La lista de reactivos de cada sección fue dada individualmente a cada alumno junto con una tarjeta *Porta Punch Hollerith* de 80 columnas, en la que se especificaba la clave, nombre del alumno, materia, clase de examen, fecha y grupo. Esta tarjeta podía ser perforada con un broche en la opción que el alumno considerara como correcta. A los alumnos se les dieron instrucciones precisas sobre el uso de la tarjeta y su manejo. El tiempo proporcionado fue de 90 minutos para los 60 reactivos, el que en general, fue suficiente.

Ese mismo día, o al día siguiente, las tarjetas fueron analizadas en una computadora Burroughs 6700, con lenguajes Algol y Fortran IV.

Datos que proporciona la computadora

Los datos proporcionados por la computadora fueron los siguientes:

a) *Índice de facilidad de cada reactivo.* En las hojas de la computadora se expresa el índice de facilidad de cada reactivo, que co-

respondió al porcentaje de alumnos que acertó la opción marcada como correcta en la tarjeta maestra perforada. Con este dato y considerando que un reactivo debe tener como índice de facilidad promedio entre 36 y 65, se pudo clasificar la calidad de los reactivos y mejorar aquéllos que quedaron fuera de estos límites. (Ver Tabla Analítica de Reactivos).

b) *Distribución de contestaciones de los alumnos por opciones.* Además del índice de facilidad, también se expresa en las hojas de la computadora el número y porcentaje de alumnos que contestó como correcta cada una de las opciones de cada reactivo. Con estos datos se conocieron los distractores que llegaban a confundirse o acercarse demasiado a la contestación correcta, así como lo que no servían como distractores por haber sido escogidos por muy pocos alumnos.

c) *Tabla analítica de reactivos.* La computadora provee de un cuadro en donde los 60 reactivos los clasifica en 20 casilleros. (Ver la Tabla Analítica de Reactivos). Los encabezados de los renglones se expresan como: *Correctos, Mejorables, Defectuosos y Desechables.* Los encabezados de las columnas son: *Muy Difíciles* (menos de 20 por ciento de índice de facilidad), *Difíciles* (de 20 a 35 por ciento de índice de facilidad), *Promedio* (de 36 a 65 por ciento de índice de facilidad), *Fáciles* (de 66 a 80 por ciento de índice de facilidad) y *Muy Fáciles* (de más de 80 por ciento de índice de facilidad). Los datos anteriores permitieron saber rápidamente la clasificación de los reactivos, y en forma específica, precisar cuáles eran adecuados y cuáles habría que modificar o desechar.

d) *Contestaciones correctas e incorrectas de cada alumno por reactivo.* La computadora proporciona para cada examinando las opciones marcadas como correctas y las marcadas como incorrectas.

e) *Índice de discriminación de los reactivos.* La computadora reconoce a los alumnos buenos y a los malos. Esta nomenclatura de "alumnos buenos" y "alumnos malos" podría ser objetable y discutible, pero estaba fuera

de nuestro alcance modificarla. Preferiríamos las denominaciones de "alumnos de alto rendimiento escolar" y "alumnos de bajo aprovechamiento", respectivamente. Sin embargo, en el presente artículo continuaremos utilizando la nomenclatura que viene expresada en las hojas de la computadora, y que es la primera mencionada. Los primeros son los que obtienen un porcentaje de aciertos entre 100 y 73 (100 menos 27), correspondiendo a los alumnos malos los que aciertan en 27 por ciento o menos de los reactivos. Se considera como un reactivo correcto, es decir, con un aceptable índice de discriminación, cuando lo contestan correctamente la mayoría de los alumnos buenos y pocos alumnos malos. De igual manera, un reactivo es defectuoso si no acierta la mayoría de los alumnos buenos o si acierta una alta proporción de los alumnos malos.

En otras palabras, la computadora señala si un reactivo no cumplió su objetivo cuando es contestado correctamente por la mayoría de los alumnos malos, lo cual indicaría que el reactivo puede contestarse al azar o que es tan confuso que los alumnos buenos no escogen la opción correcta. Lo anterior casi siempre sucede con reactivos que corresponden a temas insuficientemente tratados con clase o de los que existe poca información en los textos.

f) *Histograma de distribución de aciertos de los alumnos.* Otro de los datos proporcionados por la computadora se refiere a que en forma digital y de histograma proporciona el número (frecuencia) de alumnos que contesta bien cada una de los reactivos.

En el histograma se puede observar si la curva de distribución es unimodal o bimodal, o si es aplanada en meseta.

De igual manera, se obtienen los datos en frecuencia y porcentajes acumulativos de la distribución.

Ejemplos de análisis de reactivos. A continuación se discuten algunos reactivos que fueron diseñados para la evaluación del logro de objetivos. Es preciso, en este momento, enfatizar el hecho de que las cualidades de los reactivos dependen no solamente del reactivo

mismo sino del grupo de alumnos en que dicho reactivo se aplica. El grado de aprovechamiento de los alumnos, al igual que el énfasis con que el tema fue explicado u orientado por el maestro, repercute indefectiblemente en las características del reactivo.

Por lo tanto, parecería interesante exponer algunos ejemplos que encontramos en la práctica con el objeto de ilustrar más objetivamente la problemática inherente a la elaboración de estos reactivos.

Ejemplo de reactivo demasiado fácil.

Para la sección de Inmunología se preparó el siguiente reactivo:

Base:

“La reacción de tipificación de los grupos sanguíneos es una reacción de:”

<i>Opciones:</i>	<i>Núm. y por ciento de alumnos que contestaron</i>	
* a) aglutinación	211	(100%)
b) precipitación	0	(0%)
c) neutralización	0	(0%)
d) inmunofluorescencia	0	(0%)
e) fijación del complemento	0	(0%)
* Opción correcta	Índice de facilidad = 100 por ciento	

Como se podrá observar, este reactivo resultó con un índice de facilidad de 100 por ciento, ya que los 211 alumnos examinados contestaron la opción correcta. Consecuentemente, este reactivo fue calificado como “Defectuoso” y “Muy fácil” por la computadora, y además, este instrumento emitió la siguiente expresión: “Substituya las opciones de *b*, *c*, *d* y *e* por otras más próximas a la respuesta correcta, ya que resultaron muy poco atractivas para los estudiantes”.

El hecho anterior indujo a la modificación de este reactivo y con el objetivo disminuir su índice de facilidad se modificó y fue aplicado a otro grupo de 312 alumnos que no habían tenido contacto con el reactivo en cuestión. La modificación quedó como sigue:

Base:

“La reacción de tipificación de los grupos sanguíneos es una reacción en que:”

<i>Opciones:</i>	<i>Núm. y por ciento de alumnos que contestaron</i>	
a) participa el complemento	37	(12%)
b) intervienen las reagentinas	26	(8%)
* c) participan las IgG y las IgM	72	(23%)
d) ocurre el fenómeno de inmunoadherencia	84	(26%)
e) ocurre el fenómeno de hemólisis	93	(29%)
* Opción correcta	Índice de facilidad = 23 por ciento	

En este nuevo ensayo se encontró que el índice de facilidad del reactivo disminuyó a 23 por ciento y la computadora calificó al reactivo como “Defectuoso” y “Difícil”.

En otras palabras, como se habrá advertido, se modificaron tanto las opciones, sobre la misma base, que resultó demasiado difícil a los alumnos. La expresión de la computadora al respecto fue: “Reactivo defectuoso que presenta opciones que confunden a los estudiantes”. Lo anterior fue analizado por la academia de profesores del departamento, quienes concluyeron que la causa de lo anterior fue alguna de las siguientes razones:

- 1) El alumno confundió el fenómeno de tipificación (*in vitro*) con el de reac-

ción postransfusional (*in vivo*), en donde la reacción de hipersensibilidad tipo II se acompaña de hemólisis, debido a que, como estudiantes de medicina, suelen poner más atención a los fenómenos en relación con el caso clínico que a los del laboratorio.

- 2) El alumno confundió el fenómeno de inmunoadherencia con el de aglutinación, ya que pudieron pensar que en la aglutinación puede haber adherencia mediada por los elementos inmunes.
- 3) Faltó de parte de los profesores una información y diferenciación básica más adecuada con respecto al método de tipificación.

Debido a lo anterior, se acordó en el seno de la academia de profesores del departamento diferenciar este fenómeno a mayor profundidad y señalar a los alumnos del nuevo semestre lectivo la importancia del tema en medicina, con el objeto de que le pusieran mayor atención. En un examen posterior a un nuevo grupo de 396 alumnos se volvió a incluir este último reactivo exactamente en la misma forma (base y opciones) y se encontró que su índice de facilidad fue de 42 por ciento y que el número de alumnos y porcentaje correspondiente que contestó cada opción fue:

a)	29	(7%)
b)	26	(7%)
* c)	167	(42%)
d)	100	(25%)
e)	74	(19%)

* Opción correcta Índice de facilidad
= 42 por ciento

Lo anterior nos vuelve a señalar que son diversos los factores que intervienen en la calidad de un reactivo. Para este caso, la importancia dada en clase por los profesores fue fundamental en convertir el reactivo en un

elemento válido de evaluación.

Ejemplo de reactivo promedio en facilidad.
En este caso tenemos el ejemplo de un reactivo con índice de facilidad promedio (47 por ciento), pero en donde uno de los distractores se acerca mucho a la respuesta correcta:

Base:

“Una bacteria anaeróbica es la que:”

Opciones: *Núm. y por ciento de
alumnos que contestaron*

a) utiliza compuestos oxidados	72	(34%)
* b) utiliza compuestos orgánicos como aceptores de hidrógeno	97	(47%)
c) respira O ₂ y elimina CO ₂	4	(2%)
d) utiliza el O ₂ como fuente de energía	21	(10%)
e) consume alcoholes	15	(7%)
* Opción correcta	Índice de facilidad = 47 por ciento	

En este reactivo, el distractor *a* se acercó demasiado a la respuesta correcta *b*. Por otra parte, el distractor *c* no fue atractivo para los estudiantes. La computadora se refirió a este reactivo como “Mejorable” y “Promedio en dificultad”. Además, refiriéndose al distractor *a* expresó: “Cambie esta opción por otra que difiera más de la respuesta correcta pues confunde a los estudiantes”.

Ejemplo de reactivo mejorable en facilidad.
El siguiente reactivo fue calificado por la computadora como “Mejorable” y “Promedio en dificultad”.

Base:		a) sangre	20	(6%)	
		b) hígado	32	(10%)	
"Las exotoxinas se caracterizan por:"		c) cerebro	20	(6%)	
Opciones:	<i>Núm. y por ciento de alumnos que contestaron</i>	* d) corazón	222	(68%)	
		e) bazo	31	(10%)	
a) ser termoestables	35	(17%)			* Opción correcta
* b) actuar en forma específica en determinadas regiones anatómicas	134	(63%)			Índice de facilidad = 68 por ciento
c) liberarse únicamente durante la lisis bacteriana	13	(6%)			
d) ser lipopolisacáridos	21	(10%)			
e) no producir toxoides	8	(4%)			
* Opción correcta					Índice de facilidad = 63 por ciento

En este reactivo se observó que los distractores *c* y *e* fueron poco atractivos, aunque, en general, el reactivo es aceptable. Sin embargo, cabe mencionar que una vez probados los reactivos, la actividad de mejorarlos reside fundamentalmente en corregir distractores como *c* y *e*, o sea, tratar de elevar ligeramente su dificultad. En nuestra experiencia, esta labor es la que consume más tiempo, en general.

Ejemplo de reactivo correcto. Ejemplo de buen reactivo, en donde el índice de dificultad se encuentra cercano al fácil, y con distractores semejantes en su poder de distracción, es el siguiente:

Base:

"Cuando *Trypanosoma cruzi* infecta al hombre en forma crónica produce daño principalmente en:"

Opciones: *Núm. y por ciento de alumnos que contestaron*

En efecto, la computadora lo calificó como "Correcto" y "Fácil" y el único comentario emitido por la máquina fue: "Este reactivo presentó opciones adecuadas, tanto en lo que se refiere al grado de dificultad recomendable como su buen poder discriminativo".

Ejemplo de reactivo correcto. Otro reactivo algo mejor aún, de acuerdo a nuestros indicadores, fue el siguiente:

Base:

"La sintomatología del sistema músculo esquelético en la triquinosis está determinada, entre otros factores, por:"

Opciones: *Núm. y por ciento de alumnos que contestaron*

* a) número de parásitos	197	(61%)
b) hiperplasia de la miofibrilla	44	(13%)
c) necrosis de la miofibrilla	23	(7%)
d) lesión mecánica producida por el gusano adulto	40	(12%)
e) lesiones de nervios periféricos	21	(7%)

* Opción correcta Índice de facilidad = 61 por ciento

Este reactivo manifestó un índice de facilidad apropiado y, como se observa en los porcentajes de las opciones, los distractores atrajeron la atención en forma semejante. La

computadora lo calificó como "Correcto" y "Promedio en dificultad".

Sin embargo, en un análisis más profundo, la computadora señaló que la opción *b* era defectuosa pues: "No discrimina entre los alumnos buenos y malos". En nuestra opinión esta falta de discriminación podría deberse, entre otros factores, a que el término hiperplasia no está suficientemente dilucidado en alumnos que aún no han cursado la asignatura de anatomía patológica. Por ello, los alumnos malos pueden escogerlo cuando no conocen con precisión la respuesta correcta y los alumnos buenos pueden escogerla porque las miofibrillas muestran hipertrofia, la que probablemente confundieron con hiperplasia. El mismo aparato señaló que: "Las opciones *c*, *d* y *e* son adecuadas tanto por su dificultad como su buen poder discriminativo".

Ejemplo de reactivo aparentemente correcto. Un ejemplo de un reactivo aparentemente correcto sería el siguiente:

Base:

"¿Cuál de las siguientes sustancias tiene más posibilidades de ser antigénica?:"

Opciones: *Núm. por ciento de alumnos que contestaron*

- | | | |
|--|----|-------|
| a) proteína extraña al organismo, soluble en el plasma, de bajo peso molecular | 91 | (23%) |
| b) proteína propia, soluble en el plasma, de alto peso molecular | 58 | (15%) |
| c) fosfolípido, extraño al organismo, soluble | | |

en el plasma, de alto peso molecular	70	(17%)
--------------------------------------	----	-------

* d) carbohidrato, de alto peso molecular, soluble en el plasma extraño al organismo	164	(41%)
--	-----	-------

e) polipéptido, soluble en el plasma, extraño al organismo, de bajo peso molecular	17	(4%)
--	----	------

* Opción correcta Grado de facilidad = 41 por ciento

El grado de facilidad de este reactivo es adecuado y, como se observa en la distribución porcentual, hay un buen grado de aproximación entre los distractores, con excepción de *e*, que es bajo en relación con los demás. Pero lo importante de comentar con respecto a este reactivo es que la computadora, si bien calificó al reactivo como "Correcto" y de "Dificultad promedio", también señaló que: "A pesar de ser reactivo adecuado, la opción *a* fue seleccionada por muchos estudiantes malos". Aunque al parecer sería lógico pensar que fueran los estudiantes malos los que fallaran la contestación aceptando a *a* como válida, en un análisis más fino deberemos preguntarnos el porqué de este cuestionamiento, es decir: ¿Qué tiene de raro el que muchos estudiantes malos hayan escogido a *a* como la opción correcta?

El hecho es que cuando la computadora expresa que la opción *a* fue seleccionada por muchos estudiantes malos, nosotros interpretamos esta aseveración en el sentido que la proporción de estudiantes malos que indicaron esta opción como correcta fue en una proporción mucho mayor que las que se presentaron para los distractores *b*, *c* y *e*, las que también eran incorrectas.

Al discutir este problema en el seno de la academia de profesores (12 docentes) se llegó a la conclusión que fue muy probable que la razón de la alta proporción de estudiantes malos que se inclinó por el distractor *a* fue la siguiente: los alumnos fijan en la memoria la definición clásica de antígeno como: "Substancia extraña al organismo, soluble en el plasma, de alto peso molecular que induce al organismo a la formación de anticuerpos". Además, durante el curso se van formando la idea que las proteínas son las que más frecuentemente actúan como antígenos, o en otras palabras, que a este grupo químico pertenecen los antígenos más comunes y potentes. Por otra parte, los estudiantes buenos estudian con más amplitud y profundidad, y ejercen mayor reflexión sobre lo estudiado, es decir, llegan a asimilar que también los polisacáridos y otras sustancias de alto peso molecular, por lo general, son capaces de actuar como antígenos.

La anterior discusión promovió en la academia de profesores el consenso de que este punto representaba uno más de la lista que los docentes deberían señalar a los alumnos como importante de caracterizar y diferenciar.

CONCLUSIONES

En el presente ensayo nos ha parecido importante exponer la experiencia obtenida en nuestro curso de microbiología y parasitología con el reactivo de opción múltiple. Las características que describimos al principio de este trabajo nos parecieron atractivas y, en realidad, podemos decir que satisficieron nuestras necesidades.

Específicamente, confirmamos que el reactivo en cuestión tiene una gran versatilidad para evaluar varios niveles del aprendizaje y puede ser aplicado con facilidad en el estudio de propuestas, problemas, casos clínicos, interpretación de gráficas, etc. El uso de procedimientos electrónicos facilita enormemente la evaluación al mismo tiempo que proporciona datos de gran valor para el análisis de los reactivos mismos. En alguna otra ocasión pre-

sentaremos con mayor detalle la información que puede llegar a dar la computadora y la manera de utilizarla para mejoramiento de los métodos de evaluación.

Por otra parte, mencionaremos algunos aspectos que el evaluador debe tomar en consideración en la utilización de este tipo de reactivo. La elaboración de los reactivos presenta ciertas dificultades, principalmente a los que se inician en la utilización del reactivo, ya que los grados de dificultad y de discriminación, así como la validez de los distractores, se precisan después de aplicarlos a los alumnos, y esto requiere de la adquisición de práctica. En ocasiones, subjetivamente se cree haber preparado excelentes reactivos y el análisis estadístico demuestra lo contrario.

En nuestra experiencia hemos encontrado que después de una evaluación y de analizar los datos estadísticos propios de los reactivos hemos tenido que modificar un 30 por ciento con el objeto de mejorarlos. Lo anterior representa consumo de tiempo por parte de dos o tres profesores, a razón, por lo general, de un promedio 5 a 7 reactivos por hora.

Es necesario señalar, también, que es indispensable preparar un banco de reactivos para utilizarlos de acuerdo a las necesidades. Facilita el manejo de las evaluaciones en base al énfasis que se desee dar a algunos objetivos del aprendizaje.

Consideramos la posibilidad de establecer intercambio de reactivos con bancos de otras escuelas de medicina ya que en ocasiones los programas son semejantes. Sin embargo, debemos hacer hincapié que cualquier diferencia en los objetivos, ya sea en forma o en profundidad, así como en el tipo de enseñanza (tradicional o activa), e inclusive los medios de información (libros de texto, apuntes, etc.) limitarían su uso.

Es pertinente mencionar en este momento que la valoración de los logros de los objetivos consiste en una integración de diversas apreciaciones efectuadas en el estudiante con el objetivo de obtener una imagen lo más completa posible de sus cambios de conducta. Así, además de la investigación de conocimientos

adquiridos a través del reactivo de opción múltiple, en la calificación integral conferida a cada educando es de tomarse en consideración elementos de las áreas psicomotora y afectiva. En relación con esta última nos referimos a las señaladas por Nérici (1973), como son: respeto y consideración hacia sus compañeros, cooperación, altruismo, actitudes morales como veracidad, solidaridad y honorabilidad, orden en los trabajos, puntualidad, sentido de responsabilidad, perseverancia, control emocional y hábitos higiénicos.

Finalmente, al considerar las ventajas del tipo de reactivo de opción múltiple y compararlo con las limitaciones creemos que vale la pena el esfuerzo y el tiempo invertidos si la meta es lograr una evaluación justa, rápida, objetiva, práctica y científica.

RESUMEN

El contenido del presente ensayo representa una revisión de las características del reactivo de opción múltiple al ser aplicado a la evaluación del área cognoscitiva de los estudiantes del curso de microbiología y parasitología en una escuela de medicina.

Este tipo de reactivo fue valorado en base a las necesidades de nuestro departamento, como son: obtener una calificación justa y precisa de 250 a 350 alumnos; explorar los diversos niveles del conocimiento; utilizar un sistema confiable; y poder aplicar el análisis estadístico a los resultados.

Encontróse que este tipo de reactivo tenía una serie de características que, en general, resolvían nuestras necesidades y fue aplicado a la evaluación de logros de aprendizaje en el área cognoscitiva en la asignatura de microbiología y parasitología. Para ello se utilizó el sistema de proveer a los examinandos de una lista de 60 reactivos por evaluación y de una tarjeta perforable. Los resultados de la evaluación de los conocimientos y de las características de cada reactivo en particular fueron obtenidos por medio de una computadora, haciendo una relación de algunos da-

tos de los reactivos que se analizan a través de este aparato.

En el presente ensayo se relatan algunos ejemplos de reactivos que fueron analizados en base a los datos proporcionados por la computadora y de las experiencias obtenidas en el ajuste o modificación de los mismos al tratar de mejorarlos. Nuestra experiencia al momento es de 25 evaluaciones en grupos de 250 alumnos como promedio, utilizando 60 reactivos de opción múltiple en cada evaluación.

SUMMARY

The content of this essay represents a review of the characteristics of the multiple choice question when applied to the evaluation of students of microbiology and parasitology in one medical school.

This type of question was appraised in basis of the needs of our department such as: to obtain a fair and precise marking of 250 to 350 students to explore the different levels of knowledge, to use a reliable system, and to be able to apply statistical analysis to the results.

It was found that this type of question had characteristics which, in general, fulfilled our needs and were applied to the evaluation of learning of microbiology and parasitology courses. For this purpose, a system of providing the examinees of a list of 60 questions per evaluation and a card to be perforated was utilized. The results of knowledge evaluation and the characteristics of each particular question were obtained by means of a computer.

A discussion is also made of some examples of questions which were analyzed in basis of the data afforded by the computer and from the obtained experiences in the adjustment and modification of the questions in trying to improve them. Our actual experience is of 25 evaluations in groups of an average of 250 medical students, using 60 multiple choice questions in each evaluation.

BIBLIOGRAFIA

- BLOOM, B. S. 1975 *Taxonomía de los Objetivos de la Educación*. Librería "El Ateneo" Editorial, Buenos Aires. (Traducción de *Taxonomy of Educational Objectives*. D. McKay Company, Nueva York).
- Comité de Evaluación de Aspirantes a Residencias Médicas en las Instituciones de Salud del Sector Público. 1978 II Examen nacional de conocimientos médicos para ingresar a la residencia rotatoria y al curso universitario asociado. México, D.F.
- DE MATTOS, L. A. 1963 *Compendio de Didáctica General*. Editorial Kapelusz, Buenos Aires.
- FOSTER, J. W., JR. 1979 *Microbiología. Autoevaluación y Repaso*. Libros McGraw-Hill de México, S.A. de C.V. (Traducido de *Microbiology: Pretest Self-Assessment and Review*. 1976. Pretest Services, Inc.).
- GORINC, P. A. 1971 *Manual de Mediciones y Evaluación del Rendimiento en los Estudios*. Editorial Kapelusz, Buenos Aires.
- KIM, C. W. 1965 *Microbiology Review*. Medical Examination Publishing Company, Inc., Nueva York.
- LAFOURCADE, P. D. 1969 *Evaluación de los Aprendizajes*. Editorial Kapelusz, Buenos Aires.
- LAFOURCADE, P. D. 1974 *Planeamiento, Conducción y Evaluación en la Enseñanza Superior*. Editorial Kapelusz, Buenos Aires.
- LEMUS, L. A. 1974 *Evaluación del Rendimiento Escolar*. Editorial Kapelusz, Buenos Aires.
- MAGNUSSON, D. 1969 *Teoría de los Tests*. Editorial Trillas, México. (Traducción de *Testteori*. 1966. Almqvist & Wiksell Gebers Forlag AB. Estocolmo).
- Medical Examination Review Book*. Vol. 6. Public Health and Forensic Medicine. Medical Examination Publishing Company, Inc., Nueva York, 1961.
- NEUMANN, R. D. 1979 *Patología. Autoevaluación y Repaso*. Libros McGraw-Hill de México, S.A. de C.V. (Traducido de *Pathology: Pretest Self-Assessment and Review*. 1976. Pretest Services, Inc.).
- NÉRICI, I. G. 1973 *Hacia una Didáctica General Dinámica*. Editorial Kapelusz, Buenos Aires.
- SCHYFTER, G. 1976 *Cómo elaborar reactivos que evalúen el logro de objetivos de aprendizaje*. En *Evaluación del Aprovechamiento Escolar*, Fascículo I, 2a. versión, edit. por Comisión de Nuevos Métodos de Enseñanza, U.N.A.M.

AGRADECIMIENTO

Los autores agradecen al Lic. Ramón Villaurrutia Ayala, Jefe del Departamento de Estadística de la Escuela Superior de Medicina del Instituto Politécnico Nacional, la colaboración brindada para las evaluaciones efectuadas y su interés en llevar a efecto el presente trabajo.