
Aplicación de diferentes tipos de pruebas en un curso universitario de biología

Different examination formats applied to a university biology course

P. ESPERÓN¹, M. VITAL¹, V. MARCO¹, M. MÍGUEZ²

¹Cátedra de Biología. Facultad de Química, Gral. Flores 2124.CP11200. ; ²Unidad de Enseñanza. Facultad de Ingeniería, J. Herrera y Reissig 565.

Universidad de la República, Montevideo, Uruguay. - pesperon@fq.edu.uy

Resumen

El curso de biología general se ubica en el sexto semestre del Plan de Estudios 1980 y es obligatorio para diversas carreras que se dictan en la Facultad de Química. Su evaluación consiste en una prueba final donde los estudiantes deben alcanzar una nota mínima para la aprobación del mismo. Nos propusimos realizar un cambio en la modalidad de las preguntas utilizadas en las evaluaciones para conocer cómo influyó el planteo de las mismas en la respuesta de los estudiantes. En este trabajo se realizó la comparación de los resultados obtenidos por dos grupos de estudiantes sometidos a pruebas con dos modalidades diferentes en el planteo de las preguntas. El primer grupo fue evaluado mediante pruebas que incluyeron preguntas de respuesta abierta (PA), mientras que al segundo grupo se le aplicaron pruebas que incluyeron preguntas PA y preguntas de múltiple opción en las cuales debía justificarse la respuesta (MOV). Los resultados mostraron que el número de estudiantes aprobados era mayor al ser evaluados en pruebas con preguntas abiertas. Además, se observó que la calificación obtenida por estos estudiantes era más alta respecto al grupo de estudiantes evaluados con la modalidad de preguntas combinadas (PA y MOV).

Palabras clave: evaluación, pruebas, múltiple opción con justificación corta, preguntas de respuesta larga, aprendizaje significativo.

Abstract

The General biology course is offered for various careers as part of the Faculty of Chemistry curriculum during the 6th semester. To obtain credit, students are required

a minimum qualification in the final examination. We were interested in assessing the effects of different types of questions on the answers provided by the students. This was assayed by comparing the results obtained under two different testing formats. The first group was evaluated using “long answer” questions (PA) while the second one was evaluated using both PA and “briefly explained multiple choice” questions (MOV). The results show that the number of students in the first group (PA questions) achieved in average higher marks than those in the second group. In brief, this study strongly suggest that it is easier for students to respond to “long answer” questions rather than accounting for an answer to a “multiple choice” question on the same particular point.

Key words: evaluation, test, briefly explained multiple choice, long answer questions, Meaningful Learning.

INTRODUCCIÓN

El curso de biología general de la Facultad de Química se ubica en el 6^o semestre (del Plan de Estudios 1980). Se trata de un curso teórico-práctico que contempla un amplio temario en el ámbito molecular, celular, tisular y genético, siendo un curso introductorio de materias de corte biológico que se impartirán en una etapa más avanzada de las carreras de Farmacia e Ingeniería de alimentos.

La evaluación es un instrumento que sirve al profesor para ajustar su actuación en el proceso de enseñanza y aprendizaje, orientándolo,

reforzando contenidos y realizando la adaptación curricular necesaria (MIGUEZ y LEYMONIE, 2000). Lo que se persigue con la evaluación es transformar la realidad y mejorarla en un proceso en espiral. Si interesa conocer e identificar las ideas de los alumnos es porque el objetivo general subyacente de la instrucción es lograr que el alumno comprenda los contenidos científicos que tiene que aprender y no sólo los memorice o aprenda a resolver ejercicios aplicando fórmulas cuyo significado le resulta ajeno y extraño. Debería estar también al servicio de la comprensión y, por consiguiente, de la formación, permitiendo conocer cómo se ha realizado el aprendizaje (THORNDIKE & HASPEN, 1978).

Para tomar decisiones y medidas acertadas en la educación, es fundamental que los docentes obtengan información exacta y relevante sobre el avance de sus estudiantes.

La evaluación del curso de biología general se ha realizado tradicionalmente por medio de un escrito final con ocho preguntas abiertas (PA). Los antecedentes mostraban que las notas obtenidas por los alumnos en estas pruebas no eran, en promedio, demasiado altas (4/12) y que el porcentaje de estudiantes no aprobados era del 41%. Por otra parte, ocurrió un incremento en la población estudiantil extraordinario (en un 107%). Al momento de la evaluación final este aumento de estudiantes repercutiría en el tiempo de corrección y en el número de docentes que debían dedicarse a esa tarea. Con base en todos estos antecedentes, se incorporó un cambio en la modalidad de las preguntas en la prueba de evaluación final, utilizando una combinación de dos formatos: preguntas múltiple opción y preguntas abiertas. Dentro de la modalidad múltiple opción, se planteó un nuevo estilo de pregunta, en la cual se exige una breve justificación (de no más de cuarenta palabras) del por qué la opción es correcta o no, siendo dicha justificación definitoria en el momento de la calificación. A esta nueva modalidad se le denominó pregunta múltiple opción V (MOV) (HAYNIE, 1994). Este tipo de preguntas presenta una serie de características:

- No implican respuestas de larga redacción por lo que se puede aumentar el número de temas a evaluar, sin que aumente la duración de la prueba.
- Si el planteo de los distractores es correcto y éstos están estratégicamente diseñados, pueden detectarse errores conceptuales.
- Por tratarse de una variedad de múltiple opción los puntajes son más objetivos y pueden usarse para obtener una muestra representativa de un área de conocimiento. Una desventaja general de las preguntas objetivas es que no brinda información de cómo el estudiante elige la opción incorrecta y el azar puede constituir un elemento de distorsión en la medición de los conocimientos adquiridos por los alumnos. Este nuevo estilo de pregunta buscó evitar este inconveniente. Una ventaja adicional es que brinda la oportunidad de conocer el modo en que los estudiantes expresan sus conocimientos sin que implique redacciones extensas.
- Aunque su preparación y diseño resulta más costoso (materiales, tiempo y recursos humanos), la corrección resulta más rápida, lo que permite evaluar mayor cantidad de estudiantes en el mismo tiempo.

En este trabajo, se presenta un análisis comparativo del rendimiento académico obtenido por dos grupos de estudiantes al aplicarles dos modalidades diferentes de preguntas en la evaluación final del curso de biología general.

METODOLOGÍA DE LA INVESTIGACIÓN

Se evaluó una población de 522 estudiantes divididos en dos grupos. Para ambos grupos, el contenido temático del curso, el número de clases teóricas y prácticas, así como el grupo docente fue el mismo. Las pruebas tuvieron siempre una duración de tres horas.

Se analizaron los resultados de las instancias de evaluación durante el período 1999-2000, para el cual se utilizaron pruebas escritas con la modalidad PA y MOV combinadas. Para valorar si hubo una mejora en la calidad de los aprendizajes de este grupo a través del indicador del resultado de la evaluación, se tomó como referencia los resultados obtenidos en pruebas realizadas en el período 1997-1998, donde fue aplicada la modalidad de preguntas de tipo PA.

Los grupos de alumnos y el tipo de prueba aplicada a cada uno se detalla en el siguiente cuadro:

	Período	Nº de alumnos	Tipo de prueba
Grupo 1	1997-1998	170	PA
Grupo 2	1999-2000	352	PA + MOV

Las pruebas aplicadas al primer grupo de estudiantes tenían ocho PA, todas calificadas con el mismo valor (6 puntos).

Las pruebas para el segundo grupo incluían 2 PA (20 puntos c/u), y 12 MOV (5 puntos c/u), con 5 opciones por pregunta.

Para las pruebas tipo PA se establecieron pautas de corrección de modo de disminuir la variabilidad introducida por la corrección de docentes diferentes, uniformizando así los resultados siendo finalmente homogeneizados, en el caso de ser necesario, por el docente responsable del curso. Es importante reiterar aquí que el grupo de docentes del curso fue el mismo durante toda la experiencia.

Con respecto a las preguntas de respuesta corta MOV, se tomó como criterio para considerarlas correctamente contestadas cuando el estudiante:

- señaló las opciones correctas y dio una explicación breve del por qué de su elección, invalidando en la misma a los distractores.
- señaló las opciones (in)correctas, y explicó brevemente la falsedad de los distractores infiriendo la veracidad de la/s opción/es restante/s.
- señaló las opciones correctas e incorrectas, y fundamentó su elección en cada ítem por separado.
- dio una explicación general del tema al cual hace referencia la pregunta, en la cual incluyó claramente la respuesta a todos los ítems (THORNDIKE & HASPEN, 1978).

Estas pautas de corrección fueron explicitadas en forma escrita en el encabezado de la prueba.

Los distractores fueron seleccionados a partir de un banco de preguntas pertenecientes al archivo de la cátedra sobre exámenes, controles de exoneración, y controles de práctico, en atención a que presentaron un buen índice de discriminación (KEHOE (a), 1995).

Otra característica del planteo de las pruebas es que los temas evaluados como PA en un año, fueron evaluados como preguntas MOV el año siguiente y viceversa. Esto permite comparar la modalidad de la evaluación independientemente de los contenidos, pudiendo estudiar los resultados obtenidos en cada evaluación analizando las fortalezas y debilidades de cada una de ellas (FRARY, 1995; KEHOE (b), 1995). A modo de ejemplo:

Modalidad PA: *“En la gametogénesis masculina indique cuántas células viables se obtendrán a partir de cada espermatogonia, así como el contenido genético que tendrán las mismas”.*

Modalidad MOV: *Luego de la gametogénesis en el hombre se obtienen:*

- Cuatro células haploides igualmente viables.
- Una célula haploide viable y tres que no servirán para fecundar.
- Cuatro células diploides igualmente viables.
- Una célula diploide viable y tres que no servirán para fecundar.
- Una célula haploide viable y tres diploides que no servirán para fecundar.

Opinión de los estudiantes respecto a las pruebas de evaluación

Los resultados de un trabajo realizado con estudiantes de inmunología de la Facultad de Química, mostraron la importancia de atender a la opinión de los estudiantes para retroalimentar y mejorar la práctica docente. Motivados por estos antecedentes y luego de realizadas las pruebas, se planteó a los estudiantes una encuesta con el objetivo de conocer su grado de afinidad frente a la nueva modalidad de evaluación. La misma consistió en las siguientes dos preguntas:

- «¿Usted prefiere que la prueba final de la materia contenga solamente preguntas PA, solamente preguntas MOV, o ambas?».
- “¿Por qué?».

RESULTADOS

Calificaciones obtenidas

La escala de calificación utilizada se extiende entre 0 y 12. Se considera al puntaje de 3 (que corresponde al 60% del puntaje otorgado a la prueba) como la calificación de aprobación en ambos tipos de prueba.

A continuación se presenta la escala de calificaciones utilizada, siguiendo el criterio adoptado en la Facultad de Química:

Calificación	Nota	Calificación	Nota
D.D.D	0	B.B.MB	7
D.D.R	1	MB.MB.B	8
R.R.D	2	MB.MB.MB	9
R.R.R	3	MB.MB.S	10
R.R.B	4	S.S.MB	11
B.B.R	5	S.S.S	12
B.B.B	6		

D: deficiente, R: regular, B: bueno, MB: muy bueno, S: sobresaliente.

Las notas (n) obtenidas por los estudiantes de ambos grupos fueron graficadas para visualizar el porcentaje de alumnos que aprobaron la prueba. Esto permitió realizar la comparación entre ambos grupos considerando la población de estudiantes que aprobó en dichas instancias (véanse gráficas I y II).



Gráfica I. Distribución porcentual de la nota obtenida por los estudiantes que realizaron las evaluaciones que incluyeron únicamente PA.



Gráfica II. Distribución porcentual de la nota obtenida por los estudiantes que realizaron las evaluaciones que incluyeron PA y MOV.

Los grupos 1 y 2 permitieron visualizar que el número de estudiantes que reprobaron era mayor cuando las evaluaciones contenían ambas modalidades de preguntas. Por otra parte, a partir del total de estudiantes que aprobaron, se ordenaron en tres categorías según la nota de aprobación obtenida (véanse tablas I y II). El criterio de categorías adoptado fue: un grupo que aprobó con calificación entre R y B, un segundo grupo entre B y MB, y un tercer grupo entre MB y S (3=n<6; 6=n<9; y 9=n<12, respectivamente) (gráficas III y IV).

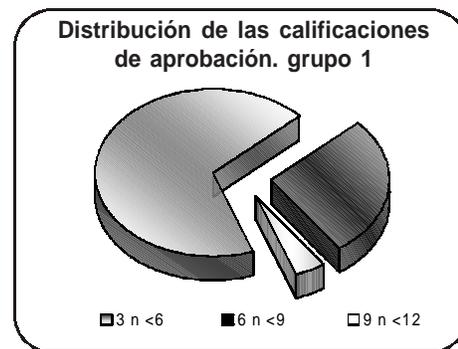
El análisis de estas gráficas mostró que las calificaciones obtenidas por los estudiantes en las pruebas que tenían únicamente preguntas abiertas eran comparativamente mayores frente a las logradas en las pruebas que incluían las dos modalidades de preguntas.

Estudiantes grupo 1	
Notas	%
3n<6	68.6
6n<	27.2
9n<12	4.1

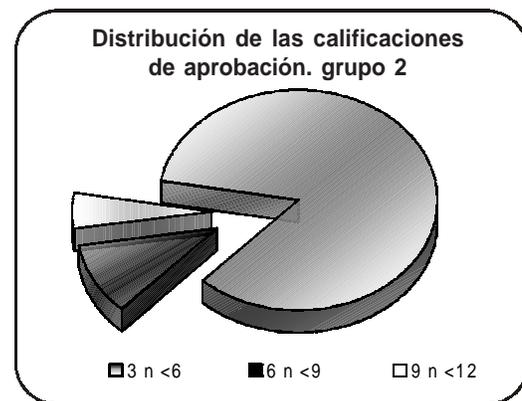
Tabla 1. Notas de los estudiantes que aprobaron las evaluaciones que incluyeron únicamente PA.

Estudiantes grupo 2	
Notas	%
3n<6	84.3
6n<	10.3
9n<12	5.4

Tabla 2. Notas de los estudiantes que aprobaron las evaluaciones que incluyeron PA y MOV.



Gráfica III. Distribución porcentual de la nota obtenida por los estudiantes que aprobaron las evaluaciones que incluyeron PA.



Gráfica IV. Distribución porcentual de la nota obtenida por los estudiantes que aprobaron las evaluaciones que incluyeron PA y MOV.

Encuesta de opinión

Se realizó una encuesta breve, anónima y sin carácter obligatorio al total de los alumnos del grupo 2.

La encuesta fue contestada por un alto porcentaje de alumnos (92%) y las respuestas mostraron que los estudiantes preferían las pruebas cuyas preguntas requerían una respuesta de desarrollo (PA) (véase tabla III).

Tabla III. Respuestas de los estudiantes frente a la encuesta.

Opciones	Preferencia (%)
Solamente PA	85
Solamente MOV	2
PA y MOV	13

Frente a la razón por la cual prefieren un determinado tipo de pregunta, se encontró como respuestas más frecuentes las siguientes: "Las PA son más fáciles de contestar"; o "Las MOV no muestran lo que en realidad estudié".

DISCUSIÓN Y CONCLUSIONES

El aprendizaje es un proceso constructivo interno por el cual la información presentada a un individuo ha de ser reconstruida por éste mediante una experiencia interna. No es una sustitución de unos conocimientos por otros. El conocimiento se explicaría como el proceso de interacción entre los esquemas previos del sujeto (que no es considerado como «tábulas» raras) y la información que le proporciona el medio que le rodea. En el proceso de dicha interacción es donde se produce la construcción de conocimientos por parte del sujeto, mediante la atribución de significado a la información con la que se va enfrentando. Para que haya un aprendizaje significativo es, entonces, necesario que el aprendiz pueda relacionar el material de aprendizaje con la estructura de conocimientos que ya dispone. Se necesita activar una idea o conocimiento previo que sirva para organizar

la situación y darle sentido. Resulta necesario distinguir entre un conocimiento que implica resistencia al cambio conceptual y el que simplemente supone un conocimiento incompleto que se mejora. Esto, en general, no se produce de manera espontánea, se recibe ayuda del entorno. Es necesario acudir al conocimiento previo, aunque esté equivocado, como puente para introducir el nuevo conocimiento.

Sin embargo debe tenerse en cuenta que no deben ser excluidas otras vías, como la asociación, que pueden conducir también a un aprendizaje si se adecuan a los objetivos de aprendizaje que se persiguen y al tipo de contenidos que se ha de aprender.

Por otra parte, un aprendizaje memorístico o por repetición es aquel en el que los contenidos están relacionados entre sí de un modo arbitrario, sin poseer necesariamente un significado para el educando. El aprendizaje significativo es más eficaz que el memorístico ya que se produce una retención más duradera de la información, se facilita nuevos aprendizajes relacionados y porque el mismo produce cambios profundos o significativos que persisten más allá del olvido de los detalles concretos. Sin embargo, los dos tipos de aprendizaje constituyen un continuo, por lo que el aprendizaje memorístico y el significativo no son excluyentes, sino que deben coexistir (AUSUBEL y col., 1983).

Se considera entonces que el aprendizaje es un proceso constructivo, acumulativo, autorregulado, dirigido a una meta (intencional), situado (relacionado con un contexto), cooperativo (social), con diferencias y matices individuales, en el que se elaboran conocimientos, se dan significados y se desarrollan habilidades y actitudes. Aprender implica un cambio duradero y transferible a nuevas situaciones. Si bien todo aprendizaje es cambio, no todos son de la misma calidad. En general el aprendizaje constructivo tiende a producir resultados más estables y duraderos.

La evaluación del progreso de los alumnos es uno de los aspectos primordiales del trabajo del docente ya que da una buena imagen sobre el adelanto que logra el alumno en el proceso educativo. Si bien el docente tiene muchas formas de evaluar el adelanto logrado por el alumno, la realización de pruebas escritas resulta un instrumento valioso en el momento de tener que realizar la certificación de los alumnos.

Como forma de convertir las pruebas escritas en un instrumento de evaluación eficaz, y determinar las causas que llevan a que los alumnos no tengan buen rendimiento en las mismas, se compararon los resultados obtenidos por dos grupos de estudiantes que realizaron pruebas de biología pero con modalidades diferentes en el planteo de las preguntas.

La hipótesis de la cual se partía era que, al plantear un nuevo tipo de prueba (MOV) se obtendría igual o mayor calificación final en las pruebas de evaluación, ya que para su aprobación sólo era necesaria una respuesta corta y concreta. Esta hipótesis se basaba en el hecho que los alumnos habían cursado, en los cinco semestres anteriores, materias en las cuales se evaluaba su capacidad para resolver situaciones problema, principalmente mediante el uso de fórmulas y de muy poca redacción. Por esta razón, los alumnos presentarían un cierto grado de dificultad en contestar preguntas que requerían una larga redacción en la respuesta. Este instrumento de evaluación, no sería entonces el adecuado y las calificaciones no estarían reflejando lo que en realidad el alumno aprendió.

Al contrario de lo esperado se pudo observar que, frente a un tema, los estudiantes obtenían una mayor calificación cuando se les formulaba una pregunta de tipo abierta (véanse gráficas III y IV). Esto pudo deberse a que para los estudiantes del grupo 1 era más fácil responder preguntas cuyas respuestas podían ser resultantes de un aprendizaje memorístico únicamente. Sin embargo, los estudiantes del grupo 2 debían discernir entre la opción correcta o los distractores y además justificar su elección, para que la respuesta de las preguntas MOV fueran consideradas correctas. El porcentaje de aprobación en el grupo 1 (59%) fue superior respecto al grupo 2 (39%), lo que concuerda con lo anteriormente observado. En síntesis, los resultados de este análisis permiten concluir que, si la respuesta involucraba un gran componente de aprendizaje memorístico, el porcentaje de estudiantes aprobados era superior y con una mejor calificación. Pero, al enfrentarlos a una situación donde debían jerarquizar conceptos de forma de elaborar y crear una respuesta estableciendo una interacción entre los mismos, el índice de aciertos era francamente menor. Por otra parte, los alumnos manifestaron, en la encuesta realizada, su preferencia por las preguntas abiertas frente al diseño de una prueba objetiva debido a que «son más fáciles de contestar» (tabla III).

Los resultados obtenidos en este trabajo muestran que la mayoría de los estudiantes, al enfrentarse a materias con las características que presenta biología general, cuya prueba de evaluación requiere una correcta redacción, estudian memorizando los temas y por lo tanto se desenvuelven de manera aceptable al desarrollarlos en forma no tan restringida. Por el contrario las

preguntas MOV, por su estructura, requieren de una respuesta corta y específica. Esto significa que los estudiantes deben elaborar las respuestas con sus propias palabras relacionando conceptos a partir de temas que pueden haber memorizado, mientras que, la sola memorización no es suficiente para lograr una correcta respuesta.

Las notas obtenidas en ambos tipos de pruebas en promedio (5 y 4 en los períodos 1997-1998 y 1999-2000, respectivamente) no llegan a la nota de 6 (calificación de bueno según la escala de calificaciones). Estas bajas calificaciones obtenidas no sólo estarían reflejando la dificultad de la prueba en sí misma, sino que además hay que considerar una serie de factores externos que pueden influir en el desempeño de los estudiantes. Entre ellos destacamos:

- Los alumnos que ingresan a la Facultad de Química, en un alto porcentaje (un 60%) provienen de un preuniversitario de corte matemático. En los primeros cinco semestres de la carrera, los estudiantes han cursado materias de corte lógico-matemático como Física, Química, Físicoquímica y Matemáticas. Como Biología General se ubica en el tercer año de la carrera, al momento de realizar este curso han transcurrido mínimo 5 años de su último contacto con la materia.
- En los 2 primeros años de la carrera no se cursan materias que impliquen lecturas ni redacciones largas.
- En la Facultad de Química la escolaridad promedio no es alta (promedio 4,5).
- Los estudiantes cursaban otras tres asignaturas simultáneamente, todas con clases teóricas y con clases prácticas de larga duración, lo que implicaba varios controles semestrales y poco tiempo para la preparación de la prueba de Biología.
- La materia Biología es, con respecto a su aprobación, menos exigente comparada con otras del mismo semestre, lo que generalmente lleva a los estudiantes a dedicarle mucho menos tiempo de estudio.
- Como material de apoyo para las evaluaciones, los alumnos tienen acceso a pruebas de años anteriores en la biblioteca de la Facultad, por lo que los estudiantes suelen focalizarse en el tipo de prueba para sus estrategias de estudio. Aunque fueron informados oportunamente del cambio de modalidad y se les proporcionó una "evaluación tipo" elaborada por la cátedra. Este es un punto que amerita profundización en el análisis, y acciones en consecuencia, ya que se detecta la práctica de los estudiantes de prepararse para un determinado tipo de pruebas, preocupándose más de éste que de sus aprendizajes.

Una razón que puede conducir a los alumnos a no intentar comprender y sólo limitarse a buscar el éxito, puede ser que sus intentos anteriores por comprender materiales potencialmente significativos hayan concluido en fracaso, por ausencia de ideas con las que puede relacionar el nuevo material. Como ya fue expuesto, el aprendizaje significativo sólo se produce cuando se relaciona información nueva con algún concepto ya existente en la estructura cognitiva del alumno que resulte relevante para el nuevo material que se pretende aprender. De esta forma se modifica la estructura cognitiva del sujeto (Pozo, 1989).

Una perspectiva futura de nuestro trabajo es promover la motivación y la comprensión para que el estudiante logre relacionar de modo sustancial los nuevos contenidos con los previos y no recurra al simple aprendizaje memorístico para alcanzar la aprobación en las pruebas.

AGRADECIMIENTOS

Este trabajo ha sido financiado por la Comisión Sectorial de Enseñanza (proyecto CSE N° 246) de la Universidad de la República, Uruguay.

BIBLIOGRAFÍA

- AUSUBEL, D.P.; NOVAK, J.D.; HANESIAN H., *Psicología educativa*, Trillas, México 1983.
- FRARY, R.B., More Multiple-Choice Item Writing Do's and Don't's. *The Educational Resources Information Center (ERIC) Digest*, 398238, 1995.
- HAYNIE, W.J., Effects of Multiple-Choice and Short-Answer Tests on Delayed Retention Learning. *Journal of Technology Education*, 6 [1], 1994.
- KEHOE, J.(a), Basic Item Analysis for Multiple-Choice Tests. *The Educational Resources Information Center (ERIC) Digest*, 398237, 1995.
- KEHOE, J.(b), Writing Multiple-Choice Test Items. *The Educational Resources Information Center (ERIC) Digest*, 398236, 1995.
- MÍGUEZ, M. & LEYMONIÉ, J., Observing, listening and exchanging ideas. *Journal of Science Education*, 1, [1], p39-42, 2000.
- POZO, J.I., *Teorías cognitivas del aprendizaje*, Morata, Madrid, 1989.
- THORNDIKE R.L. & HASPEN S., *Test y técnicas de medición en psicología y educación*, p40-85, 1978.

Received: 29.05.2003, accepted: 10.10.2003