

Estudio de validez y confiabilidad de los instrumentos evaluativos aplicados en el examen ordinario de Química

Escuela Latinoamericana de Medicina

Marilyn Díaz Vázquez, Liusmila Puente Fumero, Daylen Sosa Herrero.

RESUMEN

El perfeccionamiento del sistema de evaluación de la asignatura de Química en función de las adecuaciones del programa según el nuevo Plan D de estudio, responde a la línea de trabajo metodológico de la Escuela Latinoamericana de Medicina, "La evaluación formativa". Es por ello que este trabajo se propone como objetivo validar la calidad de los instrumentos evaluativos aplicados en el examen ordinario de Química de la ELAM, en el curso 2015-2016. Para la investigación se utilizó el método histórico-lógico, el enfoque sistémico y el análisis estadístico. La validez se valoró mediante la consulta a los expertos y la determinación y el análisis de los índices de facilidad (IF) y de discriminación (ID) de los temarios utilizados en el examen ordinario de la asignatura. Para establecer la confiabilidad de los instrumentos aplicados se calculó el alfa de Cronbach. Según los valores del índice de facilidad promedio, se comprobó que en ambos temarios predominan las preguntas fáciles y adecuadas. No hay incisos difíciles y diez de ellos se consideran muy fáciles ($IF > 92\%$). El índice de discriminación promedio, arrojó que en cada temario las preguntas se consideran entre excelentes y buenas. Solo hay un total de diez incisos por debajo de los límites para revisar y uno malo, que deben ser reelaborados. Los valores del alfa de Cronbach oscilaron en un rango entre 0,6 y 0,8 para 75% de las preguntas de ambos temarios, considerándose estas confiables y solo dos preguntas deben ser reelaboradas.

Palabras clave: instrumentos evaluativos; validación; validez y confiabilidad.

INTRODUCCIÓN

La evaluación del aprendizaje es uno de los componentes no personales del proceso de enseñanza-aprendizaje que tiene como objetivo controlar la calidad de este y de sus resultados y que incluye como uno de sus momentos fundamentales la recogida de la información y la valoración crítica de sus resultados, mediante la aplicación de diferentes instrumentos evaluativos.(1,2) En relación con los instrumentos evaluativos se sitúa en primer plano, la necesidad de confeccionarlos con un alto grado de validez y confiabilidad, de manera que la información obtenida contribuya a incrementar la calidad del proceso de enseñanza-aprendizaje.(3)

Teniendo en cuenta que la línea de trabajo metodológico propuesta por la institución para el curso 2015-2016 fue "La evaluación formativa" y la aplicación del plan de estudios D en el centro, provocó una reestructuración en el programa de la asignatura, lo que dio como resultado la adecuación del sistema de evaluación y consecuentemente los instrumentos evaluativos de la misma.

Objetivo general: Validar la calidad de los instrumentos evaluativos aplicados en el examen ordinario de Química de la ELAM, en el curso 2015-2016.

MATERIALES Y MÉTODOS

La investigación consistió en un estudio no experimental

de tipo descriptivo transversal, de acuerdo con la clasificación de Sampieri.(4) Como métodos se emplearon el método histórico-lógico, el enfoque sistémico para el diseño de la propuesta metodológica de modificaciones a los instrumentos evaluativos y el estadístico para el procesamiento de los datos. Se midieron las variables: validez, dada por la correspondencia entre lo que se pretende verificar por el instrumento y lo que realmente se evalúa; y la confiabilidad, que se relaciona con la estabilidad en los resultados obtenidos al aplicar un instrumento.(3).

Con la finalidad de elaborar un instrumento evaluativo escrito válido a partir de las sugerencias hechas por Salas Perea, (5) se diseñó el planeamiento de dicho examen por etapas. Además se definió el dominio cognitivo del instrumento partiendo de su propósito principal y se elaboró una tabla de especificidades, para seleccionar el número de preguntas a confeccionar por objetivo. Se decidió elaborar exámenes mixtos, con preguntas tipo test objetivo y preguntas de desarrollo. El banco de preguntas confeccionado se sometió a la consulta de los expertos.(3) Posteriormente se elaboraron los manuscritos, aprobados por el Departamento Docente Metodológico. La calificación se realizó según las indicaciones de la Instrucción evaluativa del Minsap.

Para medir la validez de construcción (3) de los temarios aplicados a partir de los índices de facilidad (IF) y los índices de discriminación (ID), se siguió la metodología descrita por el estudio realizado a los exámenes diagnósticos de la asignatura.(6) Para una prueba adecuada deben

confeccionarse preguntas que tengan índices repartidos entre 30% y 70%.(7) Una pregunta con índice de discriminación mayor de 0,35 se evalúa de excelente, con este índice entre 0,25-0,34 es buena, de 0,15-0,24 está en el límite para revisar y menor de 0,15 es mala.(8)

Se determinó una muestra de 42 estudiantes por cada temario, (9) seleccionados por muestreo aleatorio simple. Para el procesamiento de los datos se utilizó la aplicación Excel de la familia Office para Windows XP. Se tomó 27% de los de mayor puntuación, denominado grupo fuerte e igual por ciento de los de menor calificación, denominado grupo débil.(10,11) La confiabilidad de los instrumentos se determinó a través del coeficiente alfa de Cronbach que puede tomar valores entre 0 y 1 y mientras más cercano sea a uno, mayor es la confiabilidad del instrumento. Para

este estudio se asume el límite mínimo de 0,65, según lo establecido para la Educación Superior para Profesionales de la Salud en Cuba.(12)

RESULTADOS

Según los valores del índice de facilidad promedio, se comprobó que en ambos temarios predominan las preguntas fáciles. En el temario 1 todas se consideran fáciles, al igual que la 1 y 2 la del temario 2; mientras que la 3 y la 5 resultaron adecuadas.(Tabla 1)

No hay incisos difíciles en ninguno de los dos temarios y diez de ellos se consideran muy fáciles (IF>92%): del temario 1 el 1i) y los 2b)c)d)f)g); así como el 1a) y los 2a) b)f) del temario 2.(Tabla 2)

Tabla 1. Valores de los índice de facilidad (IF) y de discriminación (ID) promedios para cada pregunta de los temarios 1 y 2 del examen ordinario de Química, curso 2015-1016.

Preguntas	Temario 1		Alfa de Cronbach	Temario 2		Alfa de Cronbach
	IF (promedio)	ID (promedio)		IF (promedio)	ID (promedio)	
1	83,57	0,36	0,025	76,43	0,34	0,284
2	86,43	0,27	0,785	82,09	0,36	0,614
3	79,28	0,37	0,834	66,78	0,61	0,781
5	71,43	0,57	0,796	65,47	0,69	0,761

Tabla 2. Valores de los índice de facilidad (IF) y de discriminación (ID) para cada inciso de los temarios 1 y 2 del examen ordinario de Química, curso 2015-1016.

	Pregunta 1				Pregunta 2				Pregunta 3				Pregunta 5			
	T1		T2		T1		T2		T1		T2		T1		T2	
	IF	ID	IF	ID												
a	85,7	0,28	92,9	0,14	78,6	0,43	92,9	0,14	78,6	0,43	71,4	0,43	67,9	0,64	53,6	0,93
b	85,7	0,28	78,6	0,43	92,9	0,14	92,9	0,14	71,4	0,47	75	0,5	78,6	0,43	57,1	0,86
c	85,7	0,28	82,1	0,36	92,9	0,14	75	0,5	85,7	0,29	64,3	0,71	78,6	0,43	53,6	0,93
d	85,7	0,28	78,6	0,43	92,9	0,14	83,1	0,35	71,4	0,47	75	0,36	64,3	0,71	78,6	0,43
e	75	0,5	57,1	0,28	82,1	0,36	78,6	0,43	85,7	0,29	71,4	0,57	67,9	0,64	78,6	0,43
f	82,4	0,71	82,1	0,36	92,9	0,14	92,9	0,14	82,1	0,36	64,3	0,71	71,4	0,57	71,4	0,57
g	89,3	0,21	85,7	0,14	96,4	0,07	67,9	0,64	71,4	0,47	57,1	0,86				
h	75	0,5	64,3	0,28	82,1	0,36	83,1	0,35	89,3	0,21	57,1	0,71				
i	92,9	0,14	75	0,5	75	0,5	83,1	0,35	85,7	0,29	67,9	0,64				
j	78,6	0,42	67,9	0,5	78,6	0,43	71,4	0,57	71,4	0,47	64,3	0,57				

El índice de discriminación promedio, arrojó que se consideran excelentes todas las preguntas del temario 1, así como las 2, 3 y 5 del temario 2; y buena la 1 del temario 2.(Tabla 1).

Hay en total diez incisos que deben ser reelaborados, los muy fáciles según los IF. De ellos, nueve están inmediatamente por debajo de los límites para revisar y el 2g) del temario 1 se considera malo.(Tabla 2)

Los valores del alfa de Cronbach oscilaron en un rango entre 0,6 y 0,8 para 75% de las preguntas de ambos temarios, considerándose estas confiables. Solo deben ser reelaboradas las preguntas 1 en cada temario pues si no

resultan confiables al tener valores de alfa de Cronbach por debajo de 0,3.

DISCUSIÓN

Asumir la evaluación formativa implica transformaciones en la planificación, la organización, la ejecución y en la valoración de los resultados que esta aporta, así como en la preparación de los profesores y los estudiantes. Evitar los instrumentos y los procedimientos que contribuyan a la fragmentación en esferas y aspectos demasiado puntuales debe ser una prioridad en la práctica actual de dicha evaluación.

Las preguntas tipo test objetivo tienen la desventaja de requerir una elaboración muy cuidadosa, pero a su vez tienen la ventaja de ser fáciles y rápidas de calificar, en este caso las respuestas correctas obtienen todo (5) y las incorrectas, nada.(2)

A partir del análisis estadístico realizado se pudo comprobar que las preguntas más fáciles del examen se corresponden con la tipo test objetivo que responden a los objetivos generales del programa con un menor grado de dificultad en la habilidad que expresa. Corroborando así que las preguntas fáciles y muy fáciles se corresponden con las que no permiten discriminar entre los estudiantes de menor y mayor rendimiento, por lo que deben ser reelaboradas, de acuerdo con lo recogido en la bibliografía, donde se plantea que la evaluación formativa permite tener en cuenta las acciones para mejorar y ajustar la evaluación

del aprendizaje con vistas a corregir a tiempo las fallas detectadas.(13)

CONCLUSIONES

El sistema de evaluación de la asignatura Química cuenta con instrumentos evaluativos elaborados desde una concepción integradora del contenido, validado a partir del criterio de los especialistas. En estos instrumentos predominan las preguntas fáciles y adecuadas, algunas se encuentran en los límites para revisar y otras deben ser reelaboradas. De forma general los resultados obtenidos corroboran la confiabilidad y la validez de los instrumentos evaluativos aplicados en el examen ordinario de la asignatura.

REFERENCIAS BIBLIOGRÁFICAS

1. Faradales V. *Algunas consideraciones sobre la confiabilidad y la validez de los instrumentos de medición*. [acceso: febrero de 2015] Disponible en URL: <http://www.sabetodo.com/contenidos/EEuAZkuykFUreKbFjk.php>.
2. González M. *La evaluación del aprendizaje: tendencias y reflexión crítica*. *Rev Cubana Educ Med Super* 2001; 15(1):85-96 [acceso: enero de 2015] Disponible en URL: <http://www.bvs.sld.cu/revistas/ems/>.
3. Soler S, Salas R. *Calidad de los instrumentos evaluativos*. Centro Nacional de Perfeccionamiento Médico. Ciudad Habana. Cuba; 1997.
4. Hernández-Sampieri R. *Metodología de la investigación. Enfoques cuantitativo, cualitativo y mixto*. Software de la 4ta Edición McCraw-Hill Interamericana Editores. S.A.C.A. México, 2006.
5. Salas R. *La evaluación en la educación superior contemporánea*. Biblioteca de Medicina. Volumen XXIV. UMSA, La Paz, Bolivia; 1998.
6. Gui Sin Y. "Perfeccionamiento del examen diagnóstico de Química: un estudio de confiabilidad y validez". *ELAM*. 2014.
7. Gilbert J. *Guía metodológica para el personal de salud*. OMS. OPS; 1981.
8. Gilbert J. *Educational handbook for health personnel*. Revised Ed. Geneva. Ed. No. 35/1981.
9. Gil-Coya R. *Análisis multivariado aplicado a un experimento pedagógico sobre trabajo extradocente en Matemática*. [Tesis] Universidad de la Habana; 1996.
10. Escudero BE. *The level of difficulty and discrimination power of the basic knowledge and skills examinations*. *Revista electrónica de investigación educativa*. Vol 2. No. 1/2000.
11. Ebel R, Hall L. *Measuring educational. Achievement prentice*. 2000.
12. "Diplomado en Educación Superior para profesionales de la Salud". Instituto Superior de Medicina Militar. Ciudad de La Habana; 2002.
13. López-Pastor VM. *El papel de la evaluación formativa en la evaluación por competencias: aportaciones de la red de evaluación formativa y compartida en docencia universitaria*. *Rev. Docencia Universitaria*. Enero-Abril 2011;9 (1) 159 – 173

SUMMARY

Study of validity and reliability of the evaluation instruments applied in the ordinary chemistry exam

The improvement of the evaluation system of the subject of Chemistry in function of the adaptations of the program according to the new Plan D of program of study, responds to the methodological work line of the Latin American School of Medicine, "Formative evaluation". That is why this paper aims to validate the quality of evaluation instruments applied in the Chemistry regular exam of ELAM, in the academic year 2015-2016. For the investigation, the historical-logical method, the systemic approach and the statistical analysis were used. Validation was assessed by consulting experts and determining and analyzing the indices of ease (IF) and discrimination (ID) of the syllabuses used in the regular examination of the subject. In order to establish the reliability of the applied instruments, Cronbach's alpha was calculated. According to the values of the average ease index, it was verified that in both areas the questions are easy and adequate. There are no difficult items and ten of them are considered very easy (IF > 92%). The index of average discrimination, showed that in each questions book are considered between excellent and good. There are here only ten totals below the limits for review and one bad, which must be worked out again. Cronbach's alpha values ranged from 0.6 to 0.8 for 75% of the questions in both topics, and were considered reliable and only two questions had to be worked out again.

Key words: evaluative instruments; validation; exam and reliability.