



Revista Digital: Matemática, Educación e Internet

ISSN: 1659-0643

revistadigitalmatematica@itcr.ac.cr

Instituto Tecnológico de Costa Rica
Costa Rica

Meza C., Luis G.; Agüero C., Evelyn

Evaluación de los aprendizajes en la educación media: características técnicas de las pruebas escritas en matemática

Revista Digital: Matemática, Educación e Internet, vol. 14, núm. 2, marzo-agosto, 2014, pp. 1-12

Instituto Tecnológico de Costa Rica
Cartago, Costa Rica

Disponible en: <https://www.redalyc.org/articulo.oa?id=607973000004>

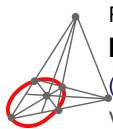
- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto



Revista digital

Matemática, Educación e Internet

(<http://www.tec-digital.itcr.ac.cr/revistamatematica/>).

Vol 14, No 2. Marzo – Agosto 2014.

ISSN 1659 -0643

Evaluación de los aprendizajes en la educación media: características técnicas de las pruebas escritas en matemática

Luis G. Meza C.
gemeza@itcr.ac.cr
Escuela de Matemática
Instituto Tecnológico de
Costa Rica

Evelyn Agüero C.
evaguero@itcr.ac.cr
Escuela de Matemática
Instituto Tecnológico de
Costa Rica

Recibido: Enero 29, 2013

Aceptado: Setiembre 15, 2013

Resumen. El artículo reporta resultados del proyecto de investigación "Evaluación de los aprendizajes en matemática en la educación media: características técnicas de las pruebas escritas", desarrollado en la Escuela de Matemática del Instituto Tecnológico de Costa Rica (ITCR) en el año 2011, bajo el código 5402-1440-3101. El estudio, de tipo exploratorio y descriptivo, utilizó como parámetros para la evaluación los establecidos en la obra La prueba escrita, por ser el documento oficial del Ministerio de Educación Pública (MEP) en la materia. Los hallazgos indican que existe, en términos generales, un alto grado de cumplimiento de las directrices emitidas por el MEP para la formulación de las pruebas. Situación similar se observa con algunas de las directrices establecidas para el diseño de los ítems, aunque se detectaron algunas donde el cumplimiento no es satisfactorio.

Palabras clave: Evaluación de los aprendizajes, diseño de pruebas escritas, redacción de ítems en matemática.

Abstract. This article reports the results of the research project "Assessment of learning in mathematics in secondary education: technical characteristics of the written tests" developed by the School of Mathematics at the Costa Rica Institute of Technology (TEC) in 2011, under research project code 5402-1440-3101. For the evaluation, we used the parameters of the study entitled La prueba escrita, as the official document approved by the Ministry of Public Education (MEP). The findings indicate that there is, in general, a high degree of compliance with the guidelines issued by the MEP for the formulation of written tests. A similar result is observed with some of the guidelines for the design of the items, but nevertheless, it was detected that some guidelines were not in satisfactory compliance.

KeyWords: Evaluation of learning, design of written tests, item wording in mathematics.

1.1 Introducción

Aunque criticados y mal vistos por los expertos en evaluación, por el alumnado e incluso por muchos profesores, se siguen realizando exámenes y que, en muchos casos, son usados como el instrumento de evaluación por excelencia (Giné y Parcerisa, 2010, p. 14). Elaborar pruebas escritas de alta calidad técnica es un asunto que compete a todos los docentes, máxime si se considera que a partir de los resultados obtenidos por los estudiantes, se toman decisiones que conllevan consecuencias sociales (Torres, Barrantes, Gutiérrez, Leitón, Marchena, Mora y Solís, 2011, p. 4).

El MEP emitió el documento titulado La prueba escrita, dirigido a los docentes de los diferentes niveles, ciclos y modalidades del sistema educativo público costarricense con el propósito de orientar y unificar criterios respecto a la elaboración y aplicación de las pruebas escritas (Torres et al, 2011, p. 4). La existencia de estas directrices no garantiza su observancia en la práctica de aula, pues, tal como plantea Carr (2002), la práctica educativa no es una especie de conducta robótica que el docente lleva a cabo de manera completamente inconsciente o mecánica.

Por otra parte, la calidad y la adecuación de las pruebas escritas dependen directamente de los ítems que lo conforman, quedando el todo en función de las partes (Rosado, 2005, p. 330). El estudio que aquí se presenta evalúa la calidad técnica de las pruebas escritas en matemática aplicadas en la educación secundaria de los colegios del Cantón Central de Cartago, Costa Rica, para determinar el nivel de cumplimiento de las directrices indicadas y la calidad técnica de los ítems incluidos en ellas. Los resultados son útiles para el MEP en los procesos de capacitación de educadores en servicio y para las universidades formadoras de docentes en matemática para la educación media.

1.2 Materiales y métodos

La investigación se desarrolló en los colegios públicos del cantón Central de Cartago que atendieron positivamente la invitación a participar en el proyecto y que se indican a continuación: Liceo de Dulce Nombre, Liceo Fernando Volio, Liceo Vicente Lachner, Liceo Daniel Oduber Quirós, Liceo Francisca Carrasco y Colegio Danilo Jiménez Veiga.

El estudio se realizó mediante el análisis documental de pruebas escritas de matemática. La muestra, de tipo no probabilística, constituida por setenta exámenes, se integró con ejemplares de cada institución participante, siguiendo dos criterios: al menos una prueba de cada nivel educativo y que se tratara de exámenes que habían sido efectivamente aplicados en un plazo no superior a los dos años previos.

La recolecta de los exámenes la realizaron directamente los investigadores en la mayoría de los casos, y, en menor grado, mediante el envío de los documentos por los colegios a la Escuela de Matemática del ITCR. La codificación de los datos se hizo utilizando listas de cotejo. La validación de estos instrumentos la realizó el propio equipo de investigadores mediante un procedimiento consistente en el

diseño individual de las guías y la posterior puesta en común.

El análisis de la información se inició con la codificación de las pruebas y de los ítems que las integran. El código de las pruebas estuvo compuesto por un prefijo formado de dos letras mayúsculas, uno distinto para cada colegio, seguido de un número de tres dígitos que indicaba el número de prueba de la institución correspondiente.

Con base en las listas de cotejo se realizó una primera evaluación de las pruebas escritas y de sus ítems. Esta fase se realizó porque una vez diseñadas las listas de cotejo se detectó que algunas de las características exigidas a las pruebas, o a los ítems en su caso, parecían de cumplimiento pleno. Por ejemplo, se comprobó que la totalidad de las pruebas incluían el nombre de la institución o un espacio para el nombre del estudiante.

Además, se descubrió que no era factible establecer el cumplimiento de algunos de los lineamientos a partir exclusivamente del texto de las pruebas, pues remitían a la fase de formulación de los exámenes o a las actividades de aula. Por ejemplo, este es el caso del lineamiento que plantea: considerar las vivencias del aula y el tiempo invertido en el desarrollo de cada objetivo específico, contenido procedimental o competencia, según corresponda. Cuatro fueron los criterios para los que sí fue posible establecer su cumplimiento o no a partir de las pruebas recabadas, a saber:

- Utilizar ítems objetivos y de desarrollo. Al menos uno de cada tipo.
- Garantizar la presentación nítida de la prueba: unificar tamaño y tipo de letra; así como la calidad de los dibujos, imágenes, esquemas o representaciones gráficas que se utilicen en ella.
- No utilizar tablas, cuadros u hojas de respuesta para consignar las respuestas de los ítems.
- No incluir ítems opcionales.

La validez de la codificación y del análisis de los datos se realizó mediante la triangulación de investigadores. Operativamente el procedimiento consistió en que cada investigador realizara las tareas de forma individual, con una puesta en común posterior en forma colegiada.

1.3 Resultados

Sobre el formato de las pruebas

El 100% de los exámenes cotejados cumplen con el requerimiento de contar con al menos un ítem de respuesta objetiva y otro de desarrollo. Se encontró que un 24.6% de los exámenes analizados no cumple la condición de tener una presentación nítida. Un 6.15% de las pruebas contenían preguntas opcionales y un 9.23% incluía cuadro de respuestas; ambas situaciones se separan de lo establecido por las directrices del MEP.

Parte administrativa

En el documento La prueba escrita (Torres, Barrantes, Gutiérrez, Leitón, Marchena, Mora y Solís, 2011) se establece que el formato de la prueba debe constar de dos partes: una parte administrativa y otra técnica.

La parte administrativa debe consignar los datos que permitan la identificación de la prueba e incluir las instrucciones generales que orienten al estudiante sobre el procedimiento que debe seguir para su resolución. El análisis de la parte administrativa arrojó los resultados que se muestran en la Tabla 1.1

	Características	% de cumplimiento
1	Nombre de la institución	100.00
2	Nombre de la asignatura	100.00
3	Nombre del docente	73.85
4	Periodo	100.00
5	Nivel	98.46
6	Nombre del estudiante	100.00
7	Firma del padre de familia	64.60
8	Duración de la prueba	96.92
9	Puntuación total	100.00
10	Valor porcentual	87.70
11	Puntuación obtenida	100.00
12	Calificación obtenida	100.00
13	Porcentaje obtenido	92.30
14	Sección	93.85
15	Fecha	78.46
16	Instrucciones generales	100.00

Tabla 1.1: Cumplimiento de las características de la parte administrativa

Se observa que la mayoría de las pruebas escritas cotejadas cumplen con los requisitos administrativos solicitados por el MEP, siendo la firma del padre de familia, el nombre del docente y la fecha los elementos faltantes en al menos un 20% de las pruebas.

Parte técnica

De acuerdo con lo establecido en el documento La prueba escrita, referencia adoptada para el análisis, los exámenes deben estar conformados por dos tipos de ítems: los objetivos (Selección única, Respuesta corta, Correspondencia o apareamiento, e Identificación) y los de desarrollo (Respuesta restringida, Resolución de ejercicios, Resolución de problemas, Resolución de casos y Ensayo).

El 100% de las pruebas analizadas cumple con este lineamiento. La Tabla 1.2 muestra el porcentaje de cada tipo de ítem presente en las pruebas.

Tipo de ítem utilizado	% de uso en las pruebas
Ítems objetivos	
Selección única	92.31
Respuesta corta	86.15
Apareamiento	18.46
Identificación	6.15
Ítems de desarrollo	

Respuesta restringida	9.23
Resolución de ejercicios	73.84
Resolución de problemas	27.70
Resolución de casos	0.00
Ensayo	0.00

Tabla 1.2: Tipo de ítem utilizado en las pruebas cotejadas

Se desprende que los ítems de "Selección única" y los de "Respuesta corta" son los más utilizados en la categoría de ítems objetivos, con valores porcentuales muy altos. Por su parte, los ítems de desarrollo que alcanzan los mayores valores porcentuales de presencia en las pruebas son los de "Resolución de ejercicios" y de "Resolución de problemas", aunque esta segunda modalidad tiene un valor bastante pequeño en comparación con el primero. Los ítems de "Respuesta restringida" también aparecen en algunas pruebas, pero con una participación escasa y limitada exclusivamente a uno de los colegios participantes.

Análisis de los ítems

La evaluación individual de los ítems inició con el estudio de la validez de contenido, bajo la premisa que si presentaba errores conceptuales de tipo matemático sería clasificado como técnicamente mal elaborado. Se procedió de esa manera porque una carencia de ese tipo lo descalifica completamente para la evaluación de aprendizajes.

Los ítems de "Identificación" fueron desechados del análisis por su escasa representación en la muestra.

Ítems de selección única. Las pruebas recolectadas aportaron un total de 923 ítems de "Selección única", de los cuales 873 (94,58%) no contenían errores de contenido matemático, mientras que 50 (5,42 %) sí los tenían. Por tanto, se analizaron 873 ítems de "Selección única" con los resultados mostrados en la Tabla 1.3.

		Cumple criterio	No cumple criterio	% cumple criterio	% no cumple criterio
1	No más tres conocimientos previos en Tercer Ciclo y Educación Diversificada.	873	0	100,00	0,00
2	Presentar la información clara y necesaria, de manera que el estudiante comprenda lo que se le solicita, sin necesidad de leer antes las opciones.	863	10	98,85	1,15
3	Aplicar las reglas gramaticales y morfosintácticas del idioma.	758	115	86,83	13,17
4	Solo una debe ser correcta.	857	16	98,17	1,83
5	La respuesta correcta debe colocarse al azar.	865	8	99,08	0,92
6	Todas deben guardar homogeneidad, en cuanto a longitud y temática.	325	548	37,23	62,77
7	Mantener la concordancia gramatical con el enunciado.	854	19	97,82	2,18

8	En su redacción no se deben utilizar expresiones como: todas las anteriores, ninguna de las anteriores, porque induce a la eliminación de la misma.	845	28	96,79	3,21
9	Se deben redactar de tal manera que no sea necesario resolver cada una de las opciones para determinar la respuesta correcta.	860	13	98,51	1,49

Tabla 1.3: Análisis de ítems de selección única

Todos los criterios, salvo el referente a la homogeneidad de las opciones, alcanzan un porcentaje de cumplimiento muy elevado que supera el 80%, siendo que la mayoría excede el 90%.

En cuanto al criterio de homogeneidad de las opciones, se hizo un estudio adicional de aquellos en cuya formulación se encontró homogeneidad por parejas (dos opciones son del mismo tipo y las otras dos son ambas de la misma naturaleza). El 22,42 % de los ítems presentaron tal condición. De esta manera, si se aceptara que la homogeneidad por parejas fuera válida, el porcentaje total de ítems que cumpliría el criterio se incrementaría al 57,63% el cual, en todo caso, sigue siendo bajo.

Ítems de respuesta corta. Bajo esta categoría se agruparon los ítems titulados como "Respuesta corta" y otros con denominaciones como "Completar", "Respuesta breve" y similares. Las pruebas analizadas aportaron un total por 788 ítems de esta categoría, ninguno de ellos con errores de contenido matemático.

Del total de 788 ítems, 657 (83,37%) correspondían efectivamente a Respuesta corta, mientras que 131 (16,62 %) eran de desarrollo. Por tanto, se analizaron 657 ítems de "Respuesta corta", pues el resto tenían que ser asumidos como técnicamente incorrectos al no tener el formato propio de este tipo de ítem, con los resultados se observan en la Tabla 1.4.

		Cumple criterio	No cumple criterio	% Cumple criterio	% no cumple criterio
1	Redactarlo de manera que la respuesta sea breve y definida.	657	131	83,38	16,62
2	Cuando presenta un problema de cálculo, no debe requerir la aplicación de más de dos conocimientos previos en I y II Ciclos y de tres en Tercer Ciclo y Educación Diversificada.	657	131	83,38	16,62
3	En las asignaturas de Matemática, Química y Física, no deben requerir cálculos y procedimientos de solución similares a los planteados en la parte de desarrollo.	657	131	83,38	16,62

4	Al final del enunciado, no se debe utilizar artículos, ni los verbos "ser "o "estar ", o preposiciones que generen ambigüedad.	653	135	82,87	17,13
---	--	-----	-----	-------	-------

Tabla 1.4: Análisis de ítems de respuesta corta

El cumplimiento de los criterios es del 83.37% en todos los casos con excepción del cuarto en el cual es ligeramente menor. Los ítems que no cumplen con los criterios fueron todos aquellos que, al incumplir con el tercero, fueron descalificados en todos los demás, pues se consideró inadecuado evaluar las otras opciones en ítems que claramente no corresponden a los de "Respuesta corta "aunque hayan sido incorporados en las pruebas bajo esa denominación o equivalente.

Ítems de correspondencia o apareamiento. Del total de 12 ítems encontrados en las pruebas analizadas, 11 (91.67%) correspondían efectivamente a apareamiento, mientras que el ítem restante era en realidad de "Respuesta corta ". Por tanto, se analizaron solo 11 ítems de "Apareamiento " con los resultados que se muestran en la Tabla 1.5.

		Cumple criterio	No cumple criterio	% Cumple criterio	% no cumple criterio
1	Describir, en forma breve, en las instrucciones el contenido de cada una de las columnas.	11	0	91.66	8.34
2	Indicar en las instrucciones si cada respuesta se puede utilizar solo una, varias o ninguna vez, o si se da una combinación de las posibilidades anteriores.	5	6	41.66	58.34
3	El contenido que se presenta en ambas columnas debe estar relacionado con un tema homogéneo.	11	0	91.66	8.34
4	Cada columna debe estar debidamente identificada.	1	10	8.33	91.67
5	Cada respuesta se debe identificar con una letra, número o símbolo que sea de fácil transcripción.	10	1	83.33	16.67
6	Se debe presentar el ítem completo en una sola página.	11	0	91.66	8.34
7	Cada enunciado debe tener una única respuesta.	11	0	91.66	8.34
8	El valor máximo del ítem es de 10 puntos.	11	0	91.66	8.34

Tabla 1.5: Análisis de ítems de correspondencia

Solo dos de las características no alcanzan porcentajes superiores al 83%. Efectivamente, "Indicar en las instrucciones si cada respuesta se puede utilizar solo una, varias o ninguna vez, o si se da una combinación de las posibilidades anteriores "obtiene un 41.66% y "Cada columna debe estar debidamente identificada "con un 8.33%.

Otros hallazgos. De manera complementaria al análisis cuantitativo realizado tanto a las pruebas como a los ítems, se efectuó un estudio de carácter más cualitativo. Este proceso produjo los hallazgos que se enuncian a continuación.

Ítems de "Selección única "de igual o mayor dificultad que algunos de desarrollo. Dentro de los hallazgos de la investigación está la detección de ítems de desarrollo de un nivel de dificultad similar al de ítems objetivos de "Selección única "o "Respuesta corta ", estos últimos con una asignación de un punto.

Por ejemplo, en una de las pruebas se encontraron los siguientes ítems en la sección de desarrollo:

- $-3 - 7 - 5 - 4 - 2$ (5 puntos)
- $-3 \cdot 2 - 4 \cdot 7$ (4 puntos)
- $-36 \div -2 \div -9$ (4 puntos)

En la misma prueba, pero ahora en la parte de selección única, se localizaron ítems de una dificultad similar cuyo valor era de un punto cada uno:

- $800 \div 50 \div 4$
- $-10 \cdot -2 \cdot -3$
- $-12 + 14 + -21 + -41 + 18$

Situaciones similares fueron encontradas en otras siete pruebas.

Ítems de "Falso o Verdadero "encubiertos como de "Selección única "o de "Respuesta corta". Los ítems de "Falso o verdadero "no están autorizados por las directrices emanadas del MEP. No obstante, el análisis realizado de las pruebas recabadas permitió revelar la existencia de algunos ítems de "Selección única "que son, esencialmente, ítems de falso o verdadero.

Para ejemplificar estos hallazgos se muestran dos ejemplos en la Tabla 1.6.

Para la función f dada por $f(x) = \frac{1}{x-2}$ considere las siguientes proposiciones

A . $f(3) > f(1)$

B. $f(1) > f(0)$

De ellas son verdaderas:

() solo la A

() ambas

() solo la B

() ninguna

De las siguientes afirmaciones es verdadera:

- a) dos rectas son concurrentes si se intersecan en un punto.
- b) dos rectas son concurrentes solamente cuando forman ángulo recto.
- c) dos rectas son concurrentes si están en el mismo plano y nunca se intersecan.
- d) dos rectas son concurrentes si están en planos diferentes y nunca se intersecan.

Tabla 1.6: Ítems de selección única equivalentes a falso o verdadero

Este tipo de ítems fue encontrado en 10 de las pruebas estudiadas.

Ítems carentes de pertinencia. La investigación enfatizó en los aspectos formales exigidos a las pruebas y a los ítems por parte del MEP. No obstante, el proceso de análisis mostró la existencia de ítems cuyo contenido motiva al cuestionamiento de la pertinencia de su inclusión en pruebas escritas cuyos resultados son base para decidir sobre la promoción del estudiante al siguiente nivel escolar.

En la Tabla 1.7 se muestran algunos ejemplos de estos ítems.

Cuando se denota una figura con una letra griega (ϵ, β, ϕ) se refiere a

- A. Un punto o un plano
- B. Un ángulo o un plano
- C. Un rayo o un segmento
- D. Una recta o un plano

El ángulo con vértice B es el siguiente:

- A. $\angle BCA$
- B. $\angle CBA$
- C. $\angle ACB$
- D. $\angle CAB$

¿Cuántos ángulos internos posee un triángulo cualesquiera?

- A. 1
- B. 6
- C. 4
- D. 3

Tabla 1.7: Ítems de pertinencia cuestionable (Elaboración propia de los autores, 2012).

1.4 Discusión de los resultados

La mayoría de las pruebas escritas cotejadas cumplen con los requisitos administrativos solicitados por el MEP. Ninguno de los tres elementos ausentes en al menos el 20% de las pruebas, a saber, la firma

del padre de familia, el nombre del docente y la fecha, constituye una fuente potencial de invalidez de la prueba o de pérdida de confiabilidad, pues corresponden a cuestiones complementarias de poca significación.

Los ítems de "Selección única" y los de "Respuesta corta" son los más utilizados dentro de la modalidad de ítems objetivos, alcanzando valores muy altos en los porcentajes calculados. Las otras opciones admisibles según los lineamientos del MEP alcanzaron valores muy pequeños en comparación con los correspondientes a los indicados.

Dado que las directrices del MEP solo permiten cuatro tipos de ítems objetivos (Selección única, Respuesta corta, Apareamiento e Identificación) y considerando las dificultades técnicas que suele presentar la redacción de ítems de apareamiento (Constenla, 2011), más el hecho de que en matemática los ítems de identificación se confunden fácilmente con los de "Respuesta corta", no es de extrañar la preferencia de los profesores por los ítems de "Selección única" o de "Respuesta corta".

Los ítems de "Selección única" son utilizados ampliamente en diversidad de pruebas y, tal como indica Constenla (2011), tienen la ventaja de adecuarse tanto a aprendizajes simples como complejos, aunque, tal como señala este autor, tienen la desventaja de que el docente no obtiene información de cuáles son las dificultades que el alumno presenta frente a los ítems respondidos en forma incorrecta.

Este hecho es preocupante porque, por una parte, tal como queda evidenciado en la investigación, la asignación de puntaje en algunos de los ítems objetivos no resulta justa en comparación con algunos de desarrollo. Por otra parte, esta situación parece revelar falencias del sistema educativo costarricense en cuanto al uso de la resolución de problemas como estrategia para el aprendizaje de la matemática.

La mayoría de las características exigidas por el MEP se cumplen en los ítems examinados. No obstante, la evaluación de los ítems de "Selección única" permitió detectar que un 5,42 % presentaba errores de contenido matemático.

Aunque este porcentaje pueda parecer bajo en comparación con el de los ítems que no tienen esa carencia, la situación no deja de ser preocupante porque son ítems elaborados por profesores titulados que no deberían incurrir en ese tipo de errores.

Lo esperado era que el porcentaje de ítems con errores de contenido matemático estuviera cercano al cero por ciento, por tratarse de colegios del cantón central de una provincia, que, por su ubicación geográfica, resultan atractivos para los profesores titulados. No obstante, esta situación no resulta del

todo inesperada porque, tal como se indica en el Tercer Informe Estado de la Educación (2011), "En el 2010 la Dirección de Gestión y Evaluación de la calidad del MEP aplicó una prueba de Matemáticas a 1.733 docentes activos. Los resultados mostraron que un tercio de los profesores no domina los conocimientos matemáticos y los procesos contemplados en el programa de estudios vigente en la secundaria".

La situación relativa a la falta de homogeneidad de las opciones encontrada en los ítems de "Selección única", podría encontrar explicación en la dificultad que conlleva la selección de las opciones (aunque autores como Santamaría (2010) afirmen que su construcción es fácil), pues tal como plantea Rocha (2011) "elaborar opciones plausibles es dispendioso".

La importancia de este hecho no debe subestimarse porque la homogeneidad de las opciones resulta elemento importante porque, tal como afirma Santamaría (2010), los niveles de confiabilidad de los ítems de "Selección única" se afecta por los malos distractores.

En cuanto a los ítems de "Respuesta corta" el cumplimiento de los criterios es del 83.37% en todos los casos con excepción del cuarto ("Al final del enunciado, no se debe utilizar artículos, ni los verbos "ser" o "estar", o preposiciones que generen ambigüedad") en el cual es ligeramente menor. No obstante, estos resultados serían engañosos si no se señala que los que no cumplen con los criterios fueron aquellos que, al incumplir con el tercero de los criterios evaluados ("no deben requerir cálculos y procedimientos de solución similares a los planteados en la parte de desarrollo"), fueron descalificados en todos los demás, pues se consideró inadecuado evaluar las otras opciones en ítems que claramente no corresponden a los de "Respuesta corta".

Este punto es importante porque uno de los hallazgos más interesantes de la investigación ha consistido en develar la existencia de puntuaciones inadecuadas en los ítems objetivos claramente desventajosos en comparación con algunos de desarrollo de igual o menor dificultad.

Los hallazgos relacionados con la existencia de ítems objetivos con igual o mayor dificultad que otros de desarrollo pero con puntaje menor resultan preocupantes por cuanto conlleva una injusticia para los estudiantes que tienen que dedicar tiempo y esfuerzos similares en un caso y en otro, obteniendo puntajes distintos. Esta situación merece atención especial en los procesos formativos, razón por la cual deberá ser atendida por las universidades formadoras de profesores de matemática y en las acciones de capacitación que desarrolla el MEP para los docentes en servicio.

Los resultados obtenidos sugieren la necesidad de reforzar los procesos formativos de profesores de matemática o de fortalecer la capacitación profesional, de manera que se reduzca la presencia de ítems en la misma prueba que evalúan reiteradamente lo mismo. Esta situación también debe ser atendida porque atenta contra la validez y la confiabilidad de la prueba, además de que introduce elementos de injusticia en contra del estudiante: el desconocimiento de un tópico le resulta castigado de manera reiterada, sin que exista razón justa que lo explique.

Mejorar la calidad de la redacción de los ítems y fortalecer la formación matemática de los profesores de matemática que así lo requieran, son elementos que merecen la atención prioritaria del MEP en los procesos de capacitación del personal docente en servicio y de las universidades formadoras de profesores de matemática para la educación media.

1.5 Conclusión

- a.) Hay un alto porcentaje de cumplimiento de las directrices emitidas por el MEP para la elaboración de las pruebas escritas, tanto en los aspectos formales como en la construcción de los ítems.
- b.) Los ítems de "Selección única" y los de "Respuesta corta" son los más utilizados en la modalidad de ítems objetivos. Los otros tipos de ítems permitidos por los lineamientos del MEP, muestran una presencia muy poco significativa en comparación con estas dos opciones.

- c.) La "Resolución de ejercicios "y la "Resolución de problemas "son los ítems de desarrollo más utilizados, aunque la segunda opción en menor cantidad que la primera.
- d.) Los ítems de "Selección única "cumplen satisfactoriamente todos los criterios, salvo el referente a la homogeneidad de las opciones. No obstante, se encontraron ítems de esta modalidad con errores de contenido matemático.
- e.) Las pruebas incluyen ítems de "Selección única "o de "Respuesta corta", que deben tener una asignación de un punto, por así disponerlo las directrices del MEP, con similar dificultad que ítems de desarrollo con puntajes superiores. Este hallazgo revela una seria deficiencia en el diseño de las pruebas.
- f.) Existe un uso camuflado de ítems de "Falso o Verdadero ", los que no están autorizados por las directrices del MEP.
- g.) Algunas pruebas evalúan reiteradamente el mismo conocimiento, lo que atenta contra la validez y la confiabilidad de los exámenes como instrumentos de medida.

Agradecimientos

Los autores agradecen a la Vicerrectoría de Investigación y Extensión del ITCR el apoyo para el desarrollo de la investigación.

Bibliografía

- [1] Carr, W. (2002). *"Una teoría para la educación. Hacia una investigación educativa crítica."*. Madrid, España: Ediciones Morata.
- [2] Constenla, J. (2011). *"Los procedimientos de pruebas"*. Recuperado de: maestro.ucsc.cl/html/pec0006/apunteprocedpruebas2.doc
- [3] Costa Rica, Programa Estado de la Nación. (2011) . "Desempeño de la educación básica y el ciclo diversificado. Tercer Informe Estado de la Educación." Recuperado de: <http://www.estadonacion.or.cr/images/stories/informes/educacion003/docs/Parte1Capitulo3.pdf>
- [4] Giné, N. y Parcerisa, A. (2007). "Evaluación en la educación secundaria. Elementos para la reflexión y recursos para la práctica." Barcelona, España: Graó.
- [5] Rocha, M. (s.f). "Reglas para elaborar ítems de selección múltiple con única respuesta.". Recuperado de: <http://w3.icfes.gov.co:8080/seminariointernacional/ktmllite/files/uploads/MARTHA%20ROCHA%20-%20ICFES.pdf>
- [6] Rosado, M. (2005). "Análisis de ítems. Teoría clásica y teoría de respuesta al ítem", *Revista Denarius* 10(1): 321-345.
- [7] Santamaría, M. (2010). "). *¿Cómo evaluar aprendizajes en aula?"*. San José, Costa Rica: EUNED.
- [8] Torres, R., Barrantes, J., Gutiérrez, X., Leitón, O., Marchena, J. y Mora, I. (2011). "La prueba escrita.". Recuperado de: <http://www.mep.go.cr>