

**INSTITUTO TECNOLÓGICO DE COSTA RICA
ESCUELA DE MATEMÁTICA**

PROYECTO DE INVESTIGACIÓN

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

INFORME FINAL

**Dr. Luis Gerardo Meza Cascante
MSc. Evelyn Agüero Calvo
Licda. María Gabriela Roldán Villalobos**

2012

Tabla de contenidos

Sección	Página
Resumen	
1. Introducción	4
2. Metodología	6
2.1. Tipo de investigación	6
2.2. Delimitación de la investigación	6
2.3. Selección de los casos	6
2.4. Acceso al campo.	7
2.4. Técnicas para la recolección de datos	8
2.5. Procedimiento para la recolección de datos	8
2.6. Estrategias para el análisis de los datos	9
2.7. Estrategias para dar validez a la investigación	12
3. Resultados	13
3.1. Análisis del formato de las pruebas	13
3.2. Análisis de los ítems	18
3.3. Otros hallazgos	24
4. Discusión y conclusiones	
4.1. Discusión de los resultados	32
4.2. Conclusiones	35
5. Recomendaciones	38
6. Agradecimientos	39
7. Referencias	39

Título

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

Autores y direcciones

Dr. Luis Gerardo Meza Cascante.	gemeza@itcr.ac.cr
M.Sc. Evelyn Agüero Calvo.	evaguero@itcr.ac.cr
Licda. María Gabriela Roldán Villalobos.	groidan@itcr.ac.cr

Nombre completo de los y las participantes del proyecto

Dr. Luis Gerardo Meza Cascante. Investigador Coordinador.
M.Sc. Evelyn Agüero Calvo
Licda. María Gabriela Roldán Villalobos.

Resumen

La investigación asumió como objetivo general “*Determinar la calidad técnica de las pruebas escritas en matemática aplicadas en la educación secundaria de los colegios del Cantón Central de Cartago*”. Se trata de una investigación cuantitativa realizada a partir de la recolección de pruebas escritas de matemática aplicadas en los colegios participantes.

Como parámetros para la evaluación se utilizaron los establecidos en el documento “La prueba escrita” por ser el documento oficial del Ministerio de Educación Pública (MEP) en la materia. La pretensión original de realizar una tarea similar con base en planteamientos teóricos, resultó innecesaria por cuanto las directrices del MEP encajan perfectamente en el marco teórico asumido para ese propósito.

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. Una situación similar se presenta con algunas de las directrices establecidas para el diseño de los ítemes, mas también se detectaron algunas donde el cumplimiento no es satisfactorio.

La investigación permitió constatar la existencia de pruebas que incluyen ítemes de respuesta objetiva valorados con un punto y otros de desarrollo, con el mismo grado de dificultad o menor, con puntajes superiores. Además, también se encontraron conjuntos de ítemes que en la misma prueba evalúan esencialmente el mismo conocimiento, situación que puede afectar negativamente al estudiante porque al fallar en ese conocimiento particular resulta castigado reiteradamente.

Ítemes con serios problemas de contenido matemático o de redacción también fueron detectados en la investigación.

Palabras clave

Evaluación de los aprendizajes, diseño de pruebas escritas, redacción de ítemes en matemática.

1. Introducción

Este documento corresponde al “Informe Final” del proyecto de investigación “Evaluación de los aprendizajes en matemática en la educación media: características técnicas de las pruebas escritas”, que bajo el código 5402-1440-3101 fue desarrollado en la Escuela de Matemática en el periodo comprendido entre enero y diciembre del 2011.

La pertinencia de la investigación se deriva del hecho de que la evaluación de los aprendizajes resulta de importancia en los procesos educativos. Desde de esta perspectiva, la calidad técnica ocupa un lugar preponderante.

El proyecto encuentra sustento en la Ley de Promoción del Desarrollo Científico y Tecnológico, pues este cuerpo normativo establece que uno de los objetivos específicos para el desarrollo científico y tecnológico es “f. *Fomentar todas las actividades de apoyo al desarrollo científico y tecnológico sustantivo; los estudios de posgrado y la capacitación de recursos humanos, así como el mejoramiento de la enseñanza de las ciencias, las matemáticas y la educación técnica, lo mismo que la documentación e información científica y tecnológica*”.

También encuentra respaldo en el “Modelo académico del TEC” aprobado por el III CONGRESO INSTITUCIONAL, pues en la sección de justificaciones se indica lo siguiente:

El Instituto Tecnológico de Costa Rica no puede abstraerse del entorno; por el contrario debe participar de su quehacer a toda la sociedad. Es por ello que asume un compromiso con el mejoramiento de los otros niveles educativos de la sociedad con el fin de garantizar el derecho al acceso y la permanencia exitosa de todas las personas con potencial para ingresar a instituciones de excelencia académica con las características del Instituto Tecnológico de Costa Rica.

Posición que se reitera en la sección 1.1. titulada “Una permanente vocación de servicio y desarrollo social”, en la que se plantea como uno de los grandes compromisos del ITCR el siguiente:

d. El Instituto Tecnológico de Costa Rica, en lo que le corresponda, tiene un compromiso con el mejoramiento del sector educativo nacional en todos sus niveles.

El concepto de evaluación de los aprendizajes ha venido transformándose paulatinamente desde hace varios años, pasando de un enfoque centrado exclusivamente en la medición a procesos que la conciben como elemento sustantivo del quehacer educativo encaminado a mejorarlo continuamente.

No obstante, continúa siendo imprescindible contar con instrumentos adecuados, especialmente de medición, que cumplan con las condiciones técnicas exigidas. Sin estos elementos los procesos de evaluación no tienen sustento en información confiable.

La formación de nuevos profesionales en el campo de la enseñanza de la matemática, tal como hace el ITCR en la carrera “Enseñanza de la matemática asistida por computadora”, requiere de información válida de la realidad educativa costarricense. Desde esta perspectiva, el proyecto se enmarca en los esfuerzos de la Escuela de Matemática de integrar la docencia, la investigación y la extensión en el desarrollo de la carrera indicada, y de sustentar las actividades de educación continua que los profesionales en el campo de la matemática educativa que se desempeñan en la educación media demandan.

Los resultados de la investigación también son importantes para las otras universidades costarricenses que forman profesores de matemática y para el MEP como fundamento de los programas de capacitación que desarrolla.

En este contexto, la investigación abordó como problema el siguiente:

¿Qué características técnicas presentan las pruebas escritas en matemática aplicadas en los Colegios del Cantón Central de Cartago?

El objetivo general de la investigación fue “Determinar la calidad técnica de las pruebas escritas en matemática aplicadas en la educación secundaria de los colegios del Cantón Central de Cartago” y los objetivos específicos fueron los siguientes:

1. Establecer el nivel de aplicación de las directrices establecidas por el MEP para el diseño de pruebas.
2. Identificar el nivel de aplicación de los lineamientos teóricos en el diseño de las pruebas.
3. Establecer el nivel aplicación de las directrices establecidas por el MEP para el diseño de ítemes en matemática.
4. Identificar el nivel de aplicación de los lineamientos teóricos en el diseño de los ítemes en matemática.

Los resultados permiten conocer, desde un enfoque cuantitativo, elementos de las características técnicas de las pruebas escritas de matemática aplicadas en los colegios públicos del Cantón Central de Cartago.

2. Metodología

2.1. Tipo de investigación

La investigación es de tipo exploratoria-descriptiva. Es exploratoria porque se incursiona en una temática sobre la cual no se conocían antecedentes concretos, al menos en cuanto a la región geográfica en la que se desarrolló el estudio.

A la vez es descriptiva porque permite identificar el nivel, medido a través de porcentajes, de cumplimiento de las directrices y de especificaciones normativas de tipo técnico en materia de diseño de pruebas escritas para la evaluación de los aprendizajes en matemática en la educación media. Por tanto, la investigación supera el ámbito exclusivamente exploratorio.

2.2. Delimitación de la investigación

En la investigación no se consideraron los siguientes aspectos:

- a. Si las pruebas escritas fueron elaboradas mediante la aplicación correcta de la tabla de especificaciones tal como se establece en el documento “La prueba Escrita” (MEP, 2011).
- b. La pertinencia de los ítems, según la “Taxonomía de Bloom”, en cuanto al objetivo que evaluaban.
- c. La capacidad de cada ítem para discriminar entre los alumnos que alcanzan el objetivo buscado y los que no lo hacen.
- d. Las estrategias que utilizaron los docentes para establecer los distractores en los ítems de selección única.
- e. Cuáles ítems pueden resolverse probando casos particulares con el uso de la calculadora.

2.3 Selección de los casos

La investigación se realizó exclusivamente en los Colegios públicos del Cantón Central de Cartago, que atendieron positiva y oportunamente la solicitud de participación en el proyecto.

La selección de los colegios se realizó con base en una lista facilitada por el Comité de Examen de Admisión del ITCR, instancia que por la naturaleza de las funciones que realiza dispone de información actualizada de los colegios del país.

De acuerdo con el listado, existían doce colegios potenciales de ser incluidos en la investigación. No obstante, por tratarse de una institución que funciona con una modalidad diferente a las otras, se decidió no incluir al IPEC - ARABELA JIMENEZ DE VOLIO.

2.4. Acceso al campo: aceptación de los colegios de participar en la investigación

De acuerdo con Rodríguez, Gil y García (1996) el acceso al campo se entiende como un proceso por el cual el investigador va accediendo a la información fundamental para su estudio.

En un primer momento, indican estos autores, supone simplemente un permiso que posibilita entrar en la institución para realizar la investigación. No obstante, tal como señalan estos mismos autores, el acceso al campo presenta características mucho más delicadas que conseguir un permiso de ingreso a la institución. En efecto, el acceso al campo puede estar limitado por la baja disposición de las personas a informar sobre ciertos temas o por la desaprobación de las personas de la forma en que un investigador enfoca ciertos problemas.

En consecuencia tenemos, siguiendo a Rodríguez, Gil y García (1996), que el acceso al campo significa la posibilidad de recoger un tipo de información que los participantes en un proceso educativo sólo proporcionan a aquellos en quienes confían y que ocultan a los demás.

Para gestionar la participación del colegio en la investigación se procedió, en primer lugar, a concertar por la vía telefónica una cita con el Director de la institución.

En cada una de las reuniones realizadas con los Directores de los colegios, la que estuvo a cargo de alguno de los investigadores participantes, se hizo entrega de los documentos que se muestran en el Anexo No. 1. El primero consiste en una carta de presentación que incluye la invitación para que el colegio participe en la investigación y el compromiso de manejo confidencial de la información suministrada y en la publicación de los resultados generados por el estudio. El segundo presenta los objetivos de la investigación.

Dado que las gestiones de participación de los colegios en este tipo de investigaciones suele requerir de mucho tiempo e inversión de recursos en transporte, se adoptó como medida operativa para decidir en que momento se daría por concluida la etapa la siguiente: tener recolectados al menos 50 ejemplares de pruebas escritas.

Tal como suele suceder en este tipo de investigaciones, en algunos de los colegios en los que se logró la autorización inicial no se concretó la participación de manera efectiva posteriormente.

Producto de los esfuerzos realizados en esta etapa, y en el marco de la política operativa indicada, se concretó la participación de las siguientes instituciones:

1. Liceo de Dulce Nombre
2. Liceo Fernando Volio
3. Liceo Vicente Lachner
4. Liceo Daniel Oduber Quirós
5. Liceo Francisca Carrasco
6. Colegio Danilo Jiménez Veiga
7. Colegio Científico Costarricense (sede ITCR)

No obstante, posteriormente se decidió no incluir al Colegio Científico Costarricense en la fase de análisis de los datos porque una revisión preliminar de las pruebas permitió establecer que correspondían a un formato, y a una temática, totalmente diferente al resto de colegios participantes en el estudio. Se trataba, entonces, de un caso particular para el cual no era ni posible ni adecuado aplicar la metodología establecida para el análisis de los datos.

Consecuentemente, la investigación se realizó con la participación de las seis instituciones restantes. Se valoró que los seis colegios que aportaron las pruebas era una cantidad adecuada para los propósitos de la investigación.

2.5. Técnicas para la recolección de datos

La investigación, por estar fundamentada en el análisis documental, no necesitó de instrumentos específicos para la recolección de la información, pues se trata de un estudio que tenía como insumo esencial documentos escritos y concretos: pruebas escritas de matemática.

La codificación de los datos se realizó mediante el uso de “listas de cotejo” diseñadas por los investigadores a partir del documento titulado “La prueba escrita”, versión a febrero de 2011, publicado por el Ministerio de Educación Pública en su página WEB.

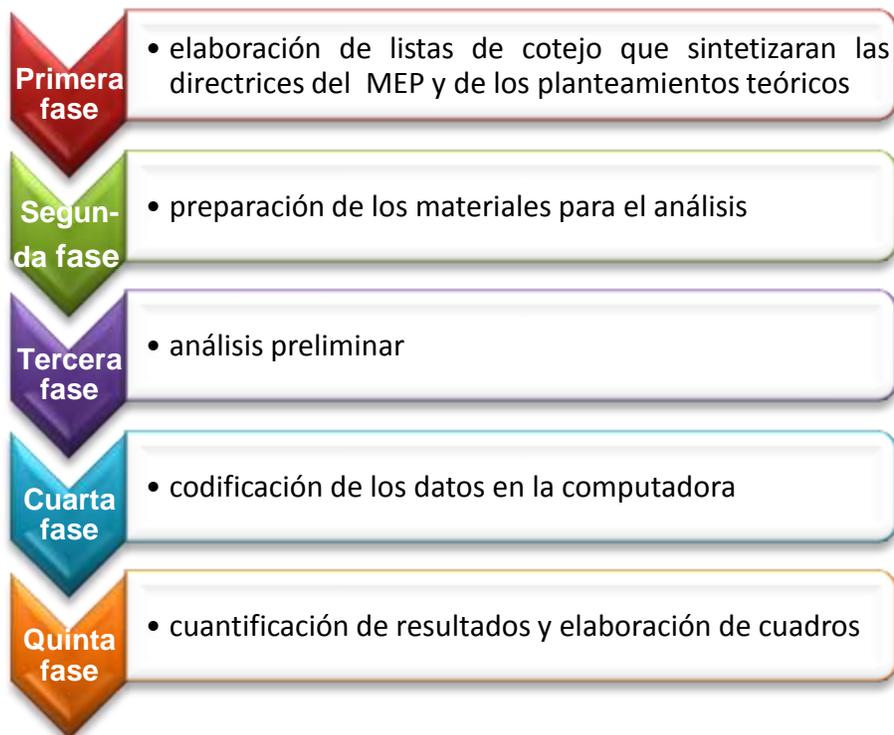
2.6. Procedimiento para la recolección de datos

Una vez lograda la autorización del Director del Colegio para que la institución participara en la investigación, se concertó, para cada caso, la forma en la cual se tendría acceso a las copias de los exámenes solicitados. Esta fase se realizó de manera distinta para unos y otros colegios, tal como se indica en el siguiente cuadro:

Colegio	Modalidad de entrega de los exámenes
Liceo de Dulce Nombre	Recogidos por los investigadores en la ocasión en que se realizó la reunión con el Director
Liceo Fernando Volio	Recogidos por los investigadores en fecha acordada en la reunión con el Director
Liceo Vicente Lachner	Recogidos por los investigadores en fecha acordada en la reunión con el Director
Liceo Daniel Oduber	Fueron enviados por el Colegio a la Escuela de Matemática del ITCR
Liceo Francisca Carrasco	Recogidos por los investigadores en la ocasión en que se realizó la reunión con el Director
Liceo Danilo Jiménez Veiga	Uno de los profesores del colegio los llevó a la Escuela de Matemática del ITCR.

2.7. Estrategias para el análisis de los datos

Una vez superada la etapa de la recolección de las pruebas escritas, se procedió con la etapa de análisis de los datos. Para tal efecto se organizó un proceso compuesto de las siguientes fases secuencialmente ejecutadas, según se muestra en el siguiente diagrama:



Primera fase: elaboración de listas de cotejo que sintetizaran las directrices del MEP y de los planteamientos teóricos

Se procedió a la elaboración de un conjunto de listas de cotejo que sintetizaran tanto las directrices del MEP como los planteamientos teóricos, contra las cuales contrastar las pruebas escritas recolectadas y los ítemes que las integraban, respectivamente. Un ejemplo de las listas de cotejo se muestra en el Anexo No. 2.

La síntesis de las directrices del MEP se realizó a partir del documento “La prueba escrita” cuyo autora es Rocío Torres Arias, Jefa del Departamento de Evaluación de los Aprendizajes del Ministerio de Educación Pública, con la colaboración de los asesores nacionales de evaluación, versión a febrero de 2011, descargada la página WEB del MEP en ese mismo mes, por ser el documento oficial en la materia.

La síntesis de las posturas teóricas se realizó con base en el libro “Aprendamos a elaborar exámenes escritos” de la Editorial EUNED cuya autora es Rose Mary Hernández Poveda, edición de 1993, por ser una obra de amplio uso en Costa Rica y responder a la realidad educativa nacional.

Segunda fase: preparación de los materiales para el análisis

La segunda fase consistió en la preparación de los materiales para ser sometidos al análisis posterior. Operativamente, esta etapa contempló las siguientes acciones:

Codificación de las pruebas

Cada una de las pruebas fue codificada según lo previsto en el “Libro de códigos” del proyecto de investigación. El código de las pruebas estuvo compuesto por un prefijo formado de dos letras mayúsculas, seguido de un número de tres dígitos el cual indicaba el número de prueba de la institución. Los prefijos establecidos para cada institución se indican en el siguiente cuadro:

Institución	Código
Liceo de Dulce Nombre	DN
Liceo Fernando Volio	FV
Liceo Vicente Lachner	VL
Liceo Daniel Oduber Quirós	DO
Liceo Francisca Carrasco	FC
Colegio Danilo Jiménez Veiga	JV

Por ejemplo, el código VL005 significa que se trata de la quinta prueba recabada en el Liceo Vicente Lachner.

Reproducción de las pruebas

Una vez que las pruebas fueron codificadas se procedió a fotocopiarlas por triplicado con el fin de contar con cuatro ejemplares de cada una.

El propósito de esta acción fue dotar a cada investigador/a de un ejemplar de cada prueba, para que pudiera realizar la etapa de análisis individual de las características de la misma, y la copia restante se utilizó para separar los ítemes que la integraban de manera que se facilitara su análisis de manera individual.

Los ítemes no fueron reproducidos en triplicado con el fin de ahorrar recursos y porque era factible su análisis por parte de cada investigador a partir de un conjunto común de fichas.

Separación de los ítemes en fichas, codificación y clasificación por categorías

La copia de las pruebas destinada al análisis de los ítemes, fue separada en los ítemes que la integraban, procediéndose posteriormente a pegar el texto de cada uno en fichas de cartulina con el propósito de ganar agilidad al manipularlos. En el Anexo No. 3 se muestra un ejemplar.

Los ítemes fueron codificados mediante un código cuyo prefijo correspondía al código asignado a la prueba de donde provenía al cual se le agregó un código según el tipo de ítem de que trataba más un número de tres cifras para indicar el número correspondiente del mismo.

Los códigos de los ítemes se muestran en el siguiente cuadro:

Tipo de ítem	Código
Falso o Verdadero	FV
Selección única	SU
Selección múltiple	SM
Pareo (Asocie)	PA
Completar	CO
Identificación	ID
Desarrollo	DE

Por ejemplo, el código FC001SU003 corresponde al ítem número tres, del tipo selección única, del examen número uno, del Colegio Francisca Carrasco.

Producto de estas acciones resultó fácil agrupar los ítemes por categorías según el tipo al que correspondían (selección única, "Respuesta corta", etc.)

Tercera fase: análisis preliminar

Con base en las listas de cotejo diseñadas en la primera etapa, se procedió a una primera evaluación de las pruebas escritas en cuanto tales y de los ítemes que las integraban en particular.

Esta fase se incorporó porque una vez diseñadas las listas de cotejo se detectó que algunas de las características exigidas a las pruebas, y a los ítemes, parecían de cumplimiento pleno. Por ejemplo, era factible esperar que la totalidad, o la casi totalidad, de las pruebas incluyeran el nombre de la institución o un espacio para el nombre del estudiante. De igual manera en el caso particular de los ítemes era esperable que la totalidad, o la casi totalidad, de los ítemes de selección única, por ejemplo, incluyeran solo cuatro opciones tal como exige el MEP.

Consecuentemente, se detectó que algunas de las características que deben reunir las pruebas y los ítemes se cumplían en la casi totalidad de las pruebas, lo que se podría determinar con una revisión preliminar haciéndose innecesaria una codificación formal. Procediéndose de esta manera se lograba un ahorro importante de tiempo y de esfuerzos innecesarios.

Con este propósito cada investigador realizó una revisión preliminar en forma individual, sin codificación formal, de las pruebas y de los ítemes, utilizando tanto las listas de cotejo de las directrices del MEP como las que contenían las cuestiones teóricas. Este procedimiento culminó con una fase de “puesta de acuerdo” entre los tres investigadores en una reunión plenaria.

Un hallazgo inesperado de esta etapa fue que las listas de cotejo de las directrices del MEP y de las cuestiones teóricas podrían considerarse como equivalentes, lo cual hacía innecesaria la codificación formal de ambos tipos. Ante esta circunstancia, se optó por utilizar las “listas de cotejo” generadas a partir de las directrices del MEP dado que son de acatamiento obligatorio para los docentes por haber sido establecidas por su patrono.

Cuarta fase: codificación de los datos en la computadora

Al tratarse de condiciones que las pruebas, o los ítemes en su caso, deben cumplir, la evaluación se realizó de manera estricta bajo la directriz operativa se cumple o no se cumple. Por tanto, la codificación para cada una de las características evaluadas fue 1 para el caso en que se cumpliera la característica y 0 para el caso contrario.

La única excepción a este procedimiento general fue en la evaluación de la característica de la homogeneidad de las opciones de los ítemes de selección única, en el que se utilizó un tercer código, a saber 0.5, con la finalidad de identificar a los ítemes de “Selección única” que presentaban homogeneidad en las opciones por parejas (característica detectada en la tercera fase del análisis).

Quinta fase: cuantificación de resultados y elaboración de cuadros

Con los resultados obtenidos en la fase de codificación se procedió a la cuantificación de los resultados obtenidos, y a la elaboración de cuadros sinópticos para facilitar la presentación de los datos.

2.8. Estrategias para dar validez a la investigación

Por tratarse de una investigación en la cual los datos provenían específicamente de documentos impresos, correspondiendo el análisis a la emisión de juicios en cuanto al ajuste de los mismos a determinados criterios establecidos a partir de directrices emitidas por el MEP (o en el marco de elementos teóricos obtenidos de obras concretas), se utilizó como estrategia para darle validez a la investigación la triangulación por investigadores.

Consecuentemente, la estrategia utilizada consistió en que cada uno de los tres investigadores realizara de manera individual la evaluación de las pruebas o de los ítems, según el caso, asignando los códigos previamente definidos, y realizando sesiones plenarias de trabajo posteriores con el fin de generar consensos.

Como procedimiento operativo se estableció que de haber coincidencia inicial en la valoración se aceptaba la clasificación sin más trámite. Para los casos en los que no hubiera coincidencia cada investigador expondría los argumentos que sustentaban su escogencia con el fin de volver a realizar una nueva evaluación. De persistir las diferencias en la clasificación se resolvería por votación. No obstante, en ningún caso fue necesario realizar votación y en aquellos casos en que se presentaron diferencias en las clasificaciones iniciales se logró consenso rápidamente.

Por tanto, las clasificaciones obtenidas de las pruebas y de los ítems que finalmente sirvieron para la obtención de las conclusiones, contaron con el consenso de los tres investigadores.

3. Resultados

3.1. Análisis del formato de las pruebas

Según el documento “La prueba escrita” elaborado por el Ministerio de Educación Pública, versión de febrero de 2011, para la construcción de pruebas escritas deben ser considerados 18 lineamientos.

No obstante, un análisis de los mismos, y en el marco de las limitaciones de la investigación, permitió establecer que la información aportada directamente por los ejemplares de las pruebas escritas era pertinente para evaluar solo lo indicado en el siguiente cuadro:

Lineamientos generales de las pruebas escritas evaluables a partir de las copias recolectadas

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.

El análisis realizado permitió constatar que el 100% de los exámenes cotejados cumplen con el requerimiento de contar con al menos un ítem de respuesta objetiva y otro de desarrollo.

Con respecto a la presentación nítida de la prueba se determinó que un 24.6% de los exámenes analizados no cumple con esa condición.

También se encontró que en un 6.15% de las pruebas habían preguntas opcionales, y que un 9.23% de las mismas incluye cuadro de respuestas, situaciones que se separan de lo establecido en las directrices del MEP.

Con respecto al formato de la prueba, el mismo documento plantea que esta debe constar de dos partes: la parte administrativa y la parte técnica.

Parte administrativa

En la parte administrativa se deben consignar los datos que permitan identificar la prueba; así como las instrucciones generales las cuales sirven para orientar al estudiante acerca del procedimiento que debe seguir para su resolución.

Para el análisis de la parte administrativa se elaboró la siguiente lista de cotejo para verificar en cada prueba escrita los siguientes elementos:

	Características
1	Nombre de la institución
2	Nombre de la asignatura
3	Nombre del docente
4	Periodo
5	Nivel
6	Nombre del estudiante
7	Firma del padre de familia
8	Duración de la prueba
9	Puntuación total
10	Valor porcentual

11	Puntuación obtenida
12	Calificación obtenida
13	Porcentaje obtenido
14	Sección
15	Fecha
16	Instrucciones generales

Según el análisis realizado, los resultados obtenidos fueron los siguientes:

	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

Esta información se muestra en el Gráfico No. 1.

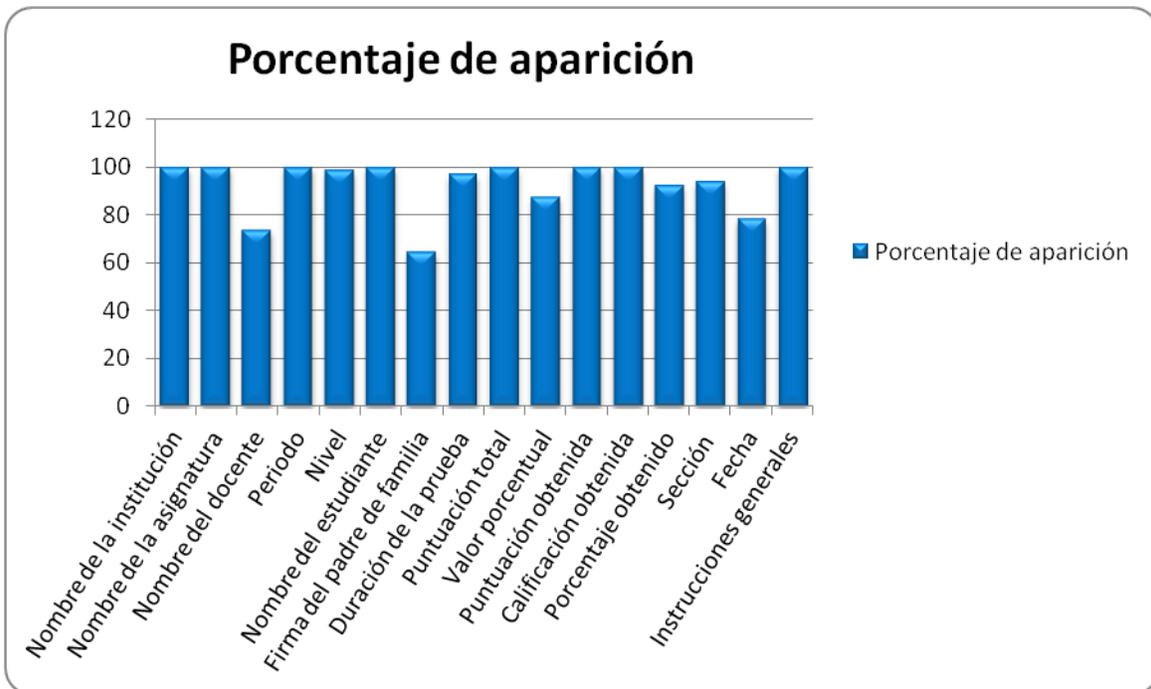


Gráfico No. 1

Por lo tanto, se observa que la mayoría de las pruebas escritas cotejadas cumplen con los requisitos administrativos solicitados por el MEP, siendo particularmente la *firma del padre de familia, el nombre del docente y la fecha* los datos 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 tal como se indicó anteriormente, la prueba escrita debe estar conformada por dos tipos de ítems: los objetivos y los de desarrollo.

En ese documento se establece cuáles son los ítems objetivos y los ítems de desarrollo que pueden ser utilizados, los que se indican en el siguiente cuadro.

Ítems objetivos	Ítems de desarrollo
Selección única	Respuesta restringida
Respuesta corta	Resolución de ejercicios
Correspondencia o apareamiento	Resolución de problemas
Identificación	Resolución de casos
	Ensayo

Según el análisis realizado para determinar el tipo de ítems utilizado en las pruebas cotejadas, se tienen los siguientes datos:

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

Esta información se muestra en los Gráficos No. 2 y No. 3.



Gráfico No. 2

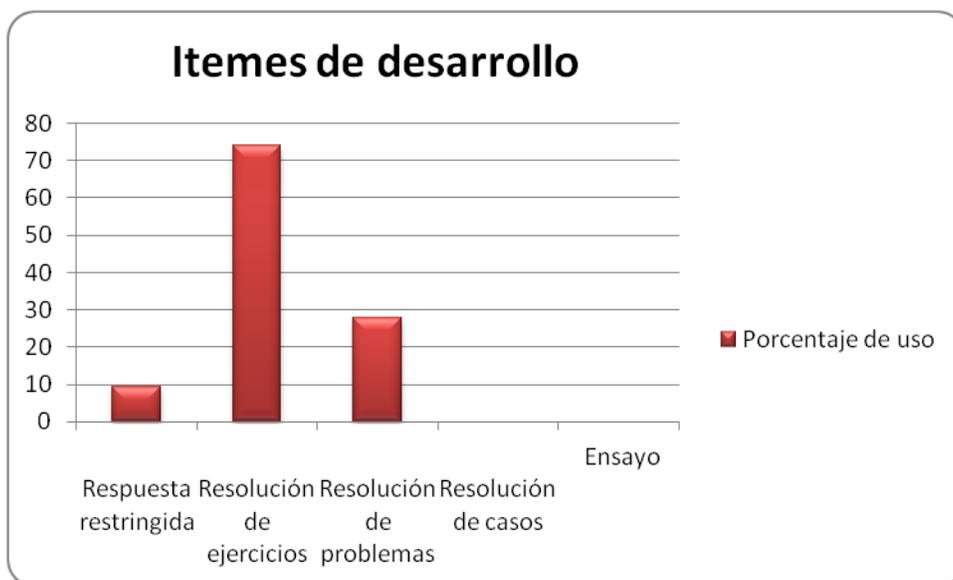


Gráfico No. 3

De esta información se desprende que los ítemes de “Selección única” y los de “Respuesta corta” son los más utilizados dentro de la modalidad de ítemes 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 ítemes indicados. Efectivamente, los ítemes de “Apareamiento” están presentes solo en el 18.46% de las pruebas y los de “Identificación” en un 6.16%.

Por su parte, los ítemes de desarrollo que alcanzan los mayores valores porcentuales de presencia en las pruebas cotejadas 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 ítemes de “Respuesta restringida” también aparecen en algunas pruebas, pero con una participación escasa y limitada exclusivamente a uno de los colegios participantes.

3.2. Análisis de los ítemes

La evaluación de los ítemes se inició con el estudio de la validez de contenido, asumiéndose que si presentaba errores conceptuales de tipo matemático sobre la materia a evaluar se clasificaba como técnicamente mal elaborado. Se consideró en estos casos que una carencia de ese tipo lo descalificaba completamente.

Además, dado que la cantidad de ítemes de “Identificación” encontrada en las pruebas fue muy pequeña (solo el 6.15% de las pruebas los incluían) se resolvió no analizarlos.

Ítemes de selección única

Las pruebas recolectadas permitieron contar con un total de 923 ítemes de “Selección única”, según la distribución por colegio indicada en la siguiente tabla:

Institución	Cantidad	Porcentaje
Liceo Dulce Nombre	112	12,13
Liceo Fernando Volio	110	11,91
Liceo Vicente Lachner	123	13,32
Liceo Daniel Oduber	165	17,87
Liceo Francisca Carrasco	133	14,40
Liceo Danilo Jiménez Veiga	280	30,33
TOTAL	923	100

Del total de 923 ítemes, 873 (94,58%) no contenían errores de contenido matemático, mientras que 50 (5,42 %) si tenían este tipo de característica. Por tanto, se analizaron 873 ítemes de “Selección única” con los siguientes resultados:

		Cumple criterio	No cumple criterio	TOTAL	% cumple criterio	% no cumple criterio
1	No más tres conocimientos previos en Tercer Ciclo y Educación Diversificada.	873	0	923	94,58	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	923	93,50	1,08
3	Aplicar las reglas gramaticales y morfosintácticas del idioma.	758	115	923	82,12	12,46
4	Solo una debe ser correcta.	857	16	923	92,85	1,73
5	La respuesta correcta debe colocarse al azar.	865	8	923	93,72	0,87
6	Todas deben guardar homogeneidad, en cuanto a longitud y temática.	325	548	923	35,21	59,37

7	Mantener la concordancia gramatical con el enunciado.	854	19	923	92,52	2,06
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	923	91,55	3,03
9	Se deben redactar de tal manera que no sea necesario resolver cada una de las opciones para determinar la respuesta correcta.	860	13	923	93,17	1,41

Esta información se muestra en el Gráfico No. 4.

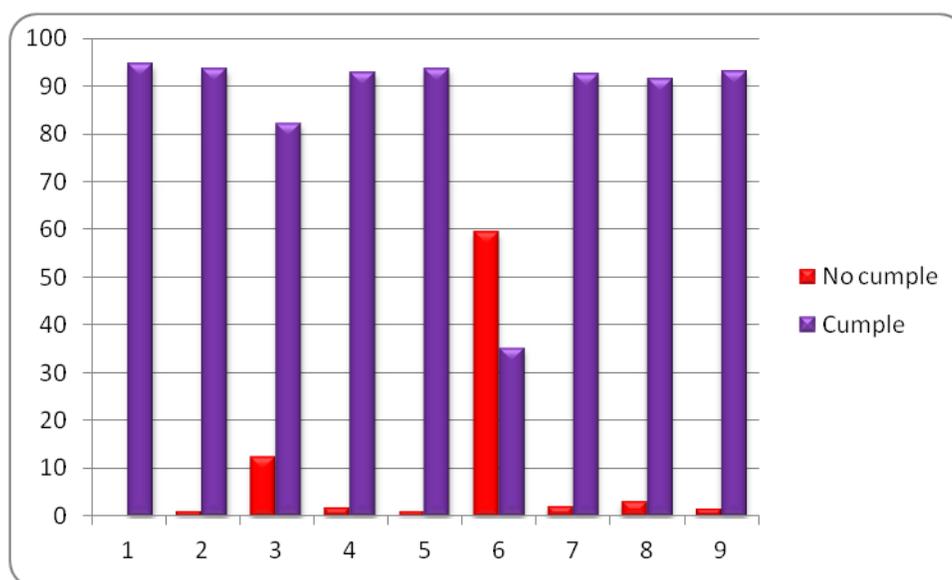


Gráfico No. 4

Como puede observarse, todos los criterios, salvo el referente a la homogeneidad de las opciones, tienen un porcentaje muy alto de cumplimiento, superándose el 80% en todos ellos, siendo que la amplia 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. En este caso el 22,42 % de los ítems presentaron homogeneidad por parejas.

De esta manera, si se aceptara que la homogeneidad por parejas fuera válida el porcentaje total de ítems que cumplirían el criterio se incrementaría al 57,63% el cual, en todo caso, sigue siendo bajo.

Ítemes de “Respuesta corta”

Bajo esta categoría se agruparon los ítemes titulados como de “Respuesta corta” propiamente dicha, y todos aquellos con denominaciones como “Completar”, “Respuesta breve” y similares.

Las pruebas analizadas estaban compuestas en total por 788 ítemes de “Respuesta corta”, de acuerdo con la siguiente distribución:

Institución	Cantidad	Porcentaje
Liceo Dulce Nombre	35	4,44
Liceo Fernando Volio	127	16,12
Liceo Vicente Lachner	98	12,44
Liceo Daniel Oduber	231	29,31
Liceo Francisca Carrasco	107	13,58
Liceo Jiménez Veiga	190	24,11
Total	788	100,00

Del total de 788 ítemes, 657 (83.37%) correspondían efectivamente a “Respuesta corta”, mientras que 131 (16.62 %) eran de desarrollo. Por tanto, se analizaron 657 ítemes de “Respuesta corta” con los siguientes resultados:

		Cumple criterio	No cumple criterio	TOTAL	% Cumple criterio	% no cumple criterio
1	Redactarlo de manera que la respuesta sea breve y definida.	657	131	788	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	788	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	788	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	788	82,87	17,13

Esta información se muestra en el Gráfico No. 5.

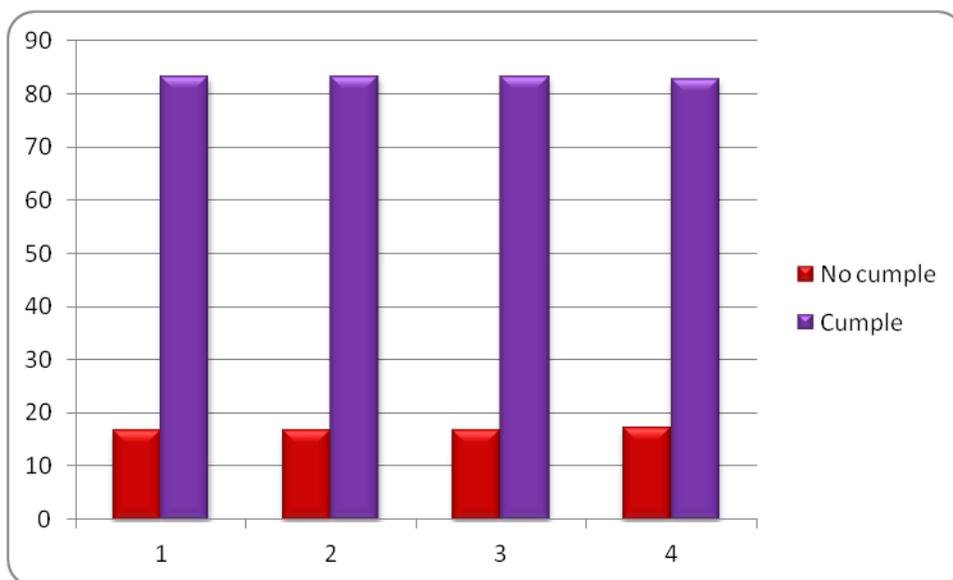


Gráfico No. 5

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”.

Ítems de correspondencia o apareamiento

Este tipo de ítem consiste en la presentación de dos columnas paralelas llamadas enunciados y respuestas cuyo contenido se relaciona entre sí.

Las pruebas analizadas estaban compuestas en total por 12 ítems de correspondencia o apareamiento, de acuerdo con la siguiente distribución:

Institución	Cantidad	Porcentaje
Liceo Dulce Nombre	5	41.67
Liceo Fernando Volio	0	0.00
Liceo Vicente Lachner	3	25.00
Liceo Daniel Oduber	4	33.33
Liceo Francisca Carrasco	0	0.00
Liceo Jiménez Veiga	0	0.00
Total	12	100,00

Del total de 12 ítems, 11 (91.67%) correspondían efectivamente a apareamiento, mientras que el ítem restante era en realidad de “Respuesta corta”. Por tanto, se analizaron 11 ítems de “Apareamiento” con los siguientes resultados:

		Cumple criterio	No cumple criterio	TOTAL	% Cumple criterio	% no cumple criterio
1	Describir, en forma breve, en las instrucciones el contenido de cada una de las columnas.	11	0	12	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	12	41.66	58.34
3	El contenido que se presenta en ambas columnas debe estar relacionado con un tema homogéneo.	11	0	12	91.66	8.34
4	Cada columna debe estar debidamente identificada.	1	10	12	8.33	91.67
5	Cada respuesta se debe identificar con una letra, número o símbolo que sea de fácil transcripción.	10	1	12	83.33	16.67
6	Se debe presentar el ítem completo en una sola página.	11	0	12	91.66	8.34
7	Cada enunciado debe tener una única respuesta.	11	0	12	91.66	8.34
8	El valor máximo del ítem es de 10 puntos.	11	0	12	91.66	8.34

Esta información se ilustra en el Gráfico No. 6.

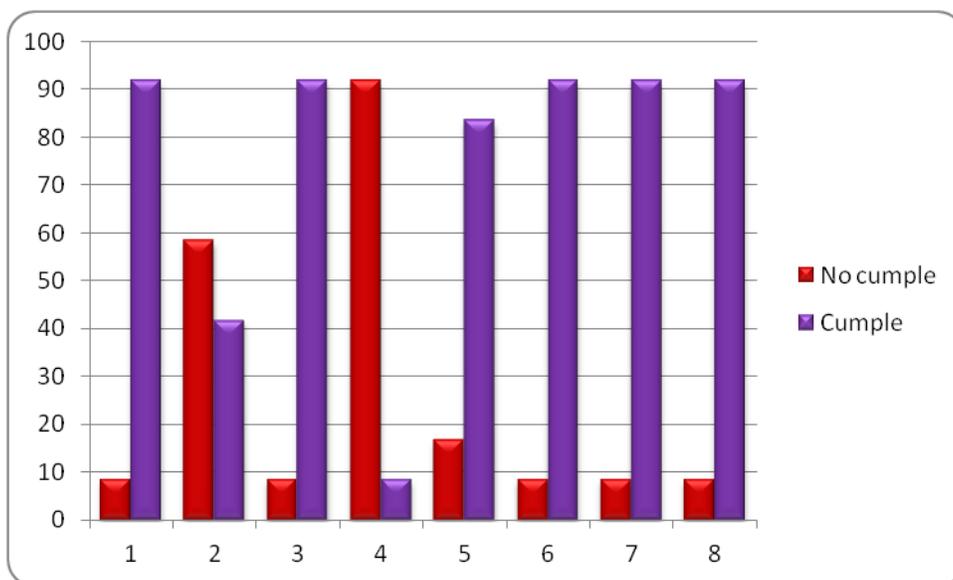


Gráfico No. 6

Como puede observarse 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%.

3.3. Otros hallazgos

De manera complementaria al análisis cuantitativo realizado tanto a las pruebas como a los ítemes, se efectuó un estudio de carácter más cualitativo. De esta manera se identificaron cinco características de las pruebas que se consideran de importancia, a saber:

- Ítemes de “Selección única” de igual o mayor dificultad que algunos de desarrollo
- Ítemes de “Falso o Verdadero” presentados como de “Selección única”
- Ítemes con serios problemas de redacción
- Ítemes que evalúan reiteradamente lo mismo
- Ítemes carentes de pertinencia

3.3.1. Ítemes de “Selección única” de igual o mayor dificultad que algunos de desarrollo

En los lineamientos del MEP para redactar la base o enunciado de un ítem de selección única, se indica que cuando se presenta un problema de cálculo este 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.

Con esto en cuenta, a la hora de analizar los ítems de “Selección única” se identificaron algunos en los que la aplicación de conocimientos requeridos para resolverlos sobrepasa lo establecido por el MEP y, por lo tanto, el valor asignado de un punto no es el adecuado.

Como ejemplo de la situación descrita, se transcriben los siguientes ítems de “Selección única” detectados en diferentes exámenes de distintos colegios:

Enunciados de ítems de “Selección única” equivalentes a ítems de desarrollo
La expresión $\frac{y}{x} - \frac{y}{x(x+1)}$ es equivalente a:
Una expresión equivalente a $\frac{x-1}{(x-1)(x+1)} + \frac{2x}{(x-1)^2}$ es:
Una expresión equivalente a $\frac{x}{(x+3)(x+2)} - \frac{2}{(x+2)(x+1)}$ es:
El conjunto solución de $4^{2x} \cdot 8^x = 2^{3x+2}$ es:
El conjunto solución de $\left(\frac{16}{81}\right)^{2x-1} = \sqrt[4]{\frac{4}{9}}$ es:
Al factorizar $8a^3 - 18ab^2 - 12a^2b + 27b^3$ dos de los factores son:
La expresión $\frac{4x^2-1}{9x^2-25} \cdot \frac{3x^2-14x+15}{2x^2-5x-3}$ es equivalente a:
La expresión $\frac{xy-2y}{x^2y^2-4y^2} \cdot \frac{x^2y+4xy+4y}{xy+2y}$ es:
El cociente de $\frac{2x^2-9x-5}{3x-2} : \frac{2x^2-5x-3}{3x-2}$ es:
El valor de k para que la ecuación $9x^2 + 5kx = -4 - 2x$ tenga dos raíces reales iguales corresponde a:
El conjunto solución de $2\sqrt{x} - \sqrt{x+5} = 1$ es:
La expresión $(2x+1)(2x-1) + (2x+3)(2x-5)$ es equivalente a:
El resultado de efectuar la operación $-x(x-1) - 3(x-2)^2$ corresponde a:

Todos estos ítems requieren para su solución de un procedimiento que claramente supera lo esperado para un ítem de selección única.

Adicionalmente, en la parte de resolución de ejercicios correspondiente a la sección de desarrollo de las pruebas, fueron detectados ítems en los cuales el puntaje asignado no se corresponde con la dificultad o que, comparados con los ítems objetivos como los de “Selección única” o los de “Apareamiento”, son similares por lo cual el puntaje asignado no es proporcional para toda la prueba.

Por ejemplo, en una prueba se encontraron las siguientes preguntas en la parte de desarrollo:

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

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

- $800 \div 50 \div -4$
- $-10 \cdot -2 \cdot -35$
- $-12 + 14 + -21 + -41 + 18$

En este mismo sentido, en otra prueba se encontraron los siguientes ítems en la parte del desarrollo:

- $\left(\frac{4}{3} - \frac{3}{2}\right) - \left(\frac{1}{2} - 3\right)$ (4 puntos)
- $\left(\frac{1}{4} + \frac{3}{2} - 1\right) \div \frac{3}{4}$ (4 puntos)

Y en la parte de apareamiento de esa prueba se encontraron ítems similares pero con valor de un punto, como los siguientes:

- $\left(\frac{3}{2} \cdot 2 - \frac{1}{3}\right) \cdot \frac{2}{4}$
- $2 - 0.7 + \frac{1}{10}$

Y en una tercera prueba, y como parte del desarrollo, se localizaron los siguientes ítems:

- $(5 - 3) + (5 - 7 \cdot 4)$ (4 puntos)
- $-10 - (-11 - 10)$ (4 puntos)
- $-5 + 3 \cdot (-2 + 7 + 11 \cdot 5) - 3$ (5 puntos)
- $-6 + -3 + 5 + 7 - 8 + -9 + 4 + -3 + 10 + -5 + -3 - 4$ (5 puntos)

Encontrándose en esa prueba los siguientes ítems similares, ahora en la parte de apareamiento, pero con valor de un punto:

- $-7 + 6 - 3 + 2 + -5 + -8 + 11$
- $(-3 + 2 \cdot 4) + (5 \cdot -2 + 6)$
- $8 + 5 \cdot 2 \div 10 + 1$

También, en una cuarta prueba se incluyeron como parte del desarrollo los siguientes ítems:

- $(2x^2 - 3y)^2$ (3 puntos)
- $(4x - 5y)(4x + 5y)$ (3 puntos)

Mientras que en la sección de apareamiento de esa prueba, y con valor de un punto, se plantearon los siguientes:

- $(8x - 2y)(8x + 2y)$
- $(5m^3 + 2n)(5m^3 - 2n)$
- $(4m - 3n)^2$

Por otra lado, en una quinta prueba, se propusieron los siguientes ítems como parte del desarrollo:

- Utilizando los pasos necesarios factorizo el siguiente polinomio: $3m^2 - 6mn + 4m - 8n$ (6 puntos)
- Factorice el siguiente trinomio de segundo grado por el método de fórmula general: $4x^2 - 6x - 2$ (5 puntos)

Encontrándose en esa prueba los siguientes ítems similares en la parte de “Respuesta corta”:

- Determine los factores de $2x^2 - 3xy + 4x - 6y$ (5 puntos)
- Determine los factores de $25t^4 - 60t^2v^3 + 36v^6$ (5 puntos)

Y en la parte de “Selección única” de la misma prueba se detectaron ítems con procedimiento similar pero ahora con valor de un punto:

- Un factor del polinomio $x^3 + 2x^2 - 3x - 6$ es:
- Un factor de $ac + bc + ad + bd$ es:

En una sexta prueba se encontró la misma situación: ítems de “Respuesta corta” similares en dificultad a los planteados en la sección de desarrollo, algunos de los cuales eran los mismos de la prueba anterior. Entre las preguntas de “Respuesta corta” estaban los siguientes ítems:

- Determine los factores de $2x^2 - 3xy + 4x - 6y$ (5 puntos)
- Determine los factores de $8t^3 - 27$ (4 puntos)
- Determine los factores de $4m^2 - 100$ (4 puntos)
- Determine los factores de $2x^2 + x - 10$ (5 puntos)

Y entre las preguntas de desarrollo de esa prueba estaban las siguientes:

- Utilizando los pasos necesarios factorizo el siguiente polinomio: $3m^2 - 6mn + 4m - 8n$ (6 puntos)
- Factorice en forma completa el siguiente polinomio: $5t^2 + 20t + 20$ (6 puntos)

Otro problema detectado en otras pruebas fue la falta de correspondencia del puntaje (1 punto) con el nivel de dificultad de la pregunta. Por ejemplo, los siguientes ítems son tomados de la parte de “Respuesta corta”, con un valor de 1 punto cada una:

- Determine el resultado obtenido al efectuar cada una de las siguientes operaciones:
 - $3(5x^3 + 5y^2) - (3x^3 - 4y^2)$
 - $(2n + 5y^2)(n - y^2)$
 - $(3x^4y^3 + 7x^2y^2 + x^5y^3) \div (x^2y^3)$
 - $\frac{-16x^5y^7}{-4x^{11}y^2z^{11}}$
 - $\frac{-3}{4}x^{1001}w^{10001}z^7; -4x^{1001}w^{9001}$
 - $(x^{a+b} - x^{b-a})^2$

La solución de estos ítems requiere de procedimientos manifiestamente exige de la asignación de un puntaje mayor a un punto.

3.3.2. Ítems de “Falso o Verdadero” encubiertos como de “Selección única” o de “Respuesta corta”

Los ítems de “Falso o verdadero” no están autorizados en las directrices emanadas del MEP para la construcción de las pruebas escritas. No obstante, en el análisis realizado de las pruebas escritas recabadas se detectaron algunos ítems de “Selección única” que son, esencialmente, ítems de falso o verdadero.

En efecto, se detectaron 29 ítems entre “Selección única” o de “Respuesta corta” que, aunque tienen una redacción para adecuarlos al formato de ítem con el cual se le incorpora en la prueba, corresponden básicamente a ítems de “Falso o Verdadero”.

Para ejemplificar esta situación en el siguiente cuadro se incluye una muestra tomada al azar de los 29 ítems que presentan la condición indicada:

Ítemes de "Selección única" o de "Respuesta corta" equivalentes a "Falso o Verdadero"	
Para la función f dada por $f(x) = \frac{1}{x-2}$ considere las siguientes proposiciones	
<p style="text-align: center;">A. $f(3) > f(1)$ B. $f(1) > f(0)$</p>	
De ellas, son verdaderas	
<p>() solo la A () ambas () solo la B () ninguna</p>	
De las siguientes afirmaciones es verdadera	
<p>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</p>	
Conteste SI o NO de acuerdo con algunos teoremas de triángulos.	
¿en todo triángulo se cumple que la suma de las medidas de dos ángulos internos es igual a la medida de uno de sus ángulos externos? _____	

3.3.3. Ítemes con serios problemas de redacción

Se detectaron 10 ítemes con serios problemas de redacción, en unos casos por falta de precisión matemática y en otras porque no se comprende lo que plantea.

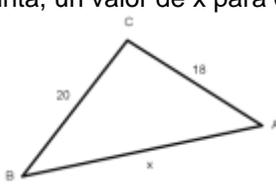
Para ejemplificar esta situación en el siguiente cuadro se incluye una muestra tomada al azar de los 10 ítemes que presentan la condición indicada:

Ítemes con serios problemas de redacción	
La recta notable de un triángulo cuyo punto de intersección es el "ortocentro" se llama	
<p>() Mediatriz () Bisectriz () Mediana () Altura</p>	
Dos rectas concurrentes forman entre sí la figura geométrica llamada	
<p>(a) recta (b) segmento (c) punto (d) rayo</p>	
Si las rectas l y m son paralelas entre sí, entonces su intersección representa la figura geométrica llamada	
<p>(a) punto (b) plano (c) recta (d) no hay intersección</p>	

3.3.4. Ítems que evalúan reiteradamente lo mismo

Se detectaron 53 conjuntos de ítems, en grupos de 2, 3, 4, 5 o 6 ítems, que evaluaban esencialmente el mismo conocimiento en la misma prueba.

Para ejemplificar esta situación en el siguiente cuadro se incluye una muestra tomada al azar de los 53 grupos ítems que presentan la condición indicada (la numeración corresponde a la utilizada en la prueba en la que aparecen):

Ítems de la misma prueba que evalúan esencialmente lo mismo	
1. La terna que corresponde a los lados de un triángulo es	
a) (4,3,9)	
b) (3,5,9)	
c) (16,30,42)	
d) (5,7,12)	
2. Considere las siguientes ternas	
I. (19,34,50)	
II. (9,10,25)	
III. (12,20,30)	
De ellas, corresponden a los lados de un triángulo	
a) Solo I	
b) Solo I y II	
c) Todas	
d) Ninguna	
3. Si la terna (10,x,23) corresponde a las medidas de los lados de un triángulo un valor posible para x puede ser	
a) 40	
b) 12	
c) 27	
d) 7	
4. De acuerdo con la figura adjunta, un valor de x para que el $\triangle ABC$ exista corresponde a	
	
a) 40	
b) 12	
c) 27	
d) 7	

3.3.5. Ítems carentes de pertinencia

La evaluación realizada en la investigación enfatizó en los aspectos formales exigidos a las pruebas y a los ítems. No obstante, en el proceso de evaluación se detectaron ítems cuyo contenido motivó a la reflexión en cuanto a la pertinencia de su inclusión en pruebas escritas con propósitos sumativos (constituyen base para definir la promoción del estudiante al siguiente nivel escolar).

Se detectaron 26 ítems cuyo contenido motiva al cuestionamiento de la utilidad de incluir este tipo de preguntas en las pruebas escritas encaminadas a la evaluación sumativa.

Para ejemplificar esta situación en el siguiente cuadro se incluye una muestra tomada al azar de los 26 ítems que presentan la condición indicada:

Ítems de pertinencia cuestionable
Quando 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
El valor de $\text{sen}\left(\frac{2\pi}{3}\right)$ corresponde a (a) $\frac{-2}{\sqrt{3}}$ (b) $\frac{\sqrt{3}}{2}$ (c) $\frac{-1}{2}$ (d) 2

4. Discusión y conclusiones

4.1. Discusión de los resultados

Las pruebas escritas continúan siendo el principal instrumento de medición que sustenta la evaluación de los aprendizajes en la educación costarricense. Aun cuando desde hace varios años se ha planteado formalmente que la educación costarricense se enfoca en las corrientes del racionalismo, el humanismo y el constructivismo, es claro que se ha avanzado poco en la discusión, y menos aun en el uso de instrumentos, de la temática de la evaluación de los aprendizajes desde una perspectiva constructivista. La realidad educativa se caracteriza, como ya se dijo, por el predominio de las pruebas escritas.

Sin asumir como necesariamente válida la situación actual, la investigación se planteó desde una perspectiva pragmática para conocer el estado actual de cumplimiento de las directrices que el MEP ha formulado para la elaboración de las pruebas escritas y de los ítemes que las integran, y del ajuste de tales pruebas a preceptos teóricos de ordinaria aceptación. Por tanto, dado que las pruebas escritas siguen siendo el instrumento de medición de mayor empleo y que existen disposiciones que regulan su formulación, era necesario conocer la concordancia entre tales pruebas y las directrices emitidas.

Lo anterior se hace necesario porque, tal como indica Hernández (1988 p. 9), *“si nuestro maestro construye instrumentos de medición válidos y eficaces, estará garantizando, en buena medida, la preparación académica del alumno y contribuyendo, en gran medida, con su integración en la sociedad, es decir, ayudando para que cada uno de sus alumnos alcance la plena realización, intelectual y social”*.

Tampoco podemos olvidar que las pruebas escritas forman parte de la llamada evaluación formativa, y que, tal como señala Brenes (1990, p. 69), sus resultados se utilizan para decidir si el estudiante pasa o no al siguiente nivel escolar. Esto implica que los resultados obtenidos en las pruebas escritas tienen consecuencias directas para los/as estudiantes.

La interpretación de los resultados obtenidos en esta investigación debe realizarse desde la óptica de que los profesores de matemática involucrados laboran en colegios de un cantón central de provincia. Este hecho reúne la particularidad de que se debe asumir que todos, o al menos una mayoría muy significativa, cuentan con la formación profesional mínima de bachillerato universitario para el desempeño del puesto, situación que no es necesariamente cierta para las distintas regiones educativas del país en donde laboran profesores sin la titulación adecuada.

Desde esa perspectiva, resulta que algunas carencias en el diseño técnico de las pruebas o de los ítemes, aun cuando tenga una incidencia porcentual baja o muy

baja, resultan inaceptables como producto de la labor de un profesional acreditado como tal.

4.1.1. Sobre las pruebas escritas

La mayoría de las pruebas escritas cotejadas cumplen con los requisitos administrativos solicitados por el MEP. No obstante, tres elementos se muestran ausentes en al menos el 20% de las pruebas: la *firma del padre de familia*, el *nombre del docente* y *la fecha*. Ahora bien, aunque estos tres elementos deberían estar presentes en las pruebas por así disponerlo el MEP, es cierto que ninguno de ellos constituye una fuente potencial de invalidez de la prueba o de pérdida de confiabilidad de la misma, 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 (Constela, 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 Constela (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.

Por su parte, los ítems de desarrollo que alcanzan los mayores valores porcentuales de presencia en las pruebas cotejadas 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. Sobre este hallazgo es necesario resaltar que una cantidad importante de los ítems de desarrollo evaluados, clasificados en la modalidad de “Resolución de ejercicios”, en realidad presentan poca o ninguna diferencia con algunos ítems objetivos de la misma prueba.

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, 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.

4.1.2. Sobre el diseño de los ítemes

La mayoría de las características exigidas por el MEP se cumplen en los ítemes examinados. No obstante, la evaluación de los ítemes 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 ítemes que no tienen esa carencia, es preocupante porque al haber sido elaborados por profesores titulados no deberían presentar ese tipo de error. Lo esperado era que el porcentaje de ítemes con errores de contenido matemático estuviera cercano al 0%, por tratarse del cantón central de una provincia pues los colegios de estas zonas son instituciones 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”*.

Para los ítemes de “Selección única” todos los criterios, salvo el referente a la homogeneidad de las opciones, alcanzaron un porcentaje muy alto de cumplimiento, superándose el 80% en todos ellos, siendo que la amplia mayoría excede el 90%.

La situación relativa a la falta de homogeneidad de las opciones podría encontrar explicación en la dificultad que conlleva la redacción de ítemes de “Selección única” en cuanto a 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 ítemes de “Selección única” se afecta por los malos distractores.

Por esta razón en la fase de análisis de los datos se hizo un estudio adicional con los ítemes de “Selección única” en cuya formulación se encontró homogeneidad por parejas (dos opciones eran homogéneas entre sí y las otras dos entre ellas). El resultado obtenido, consistente en que el 22.42 % de los ítemes presentaron homogeneidad por parejas, sugiere un tema para un análisis posterior sobre el efecto en la confiabilidad de la prueba si se aceptara que los ítemes de “Selección única” en matemática puedan diseñarse con opciones completamente

homogéneas o al menos con homogeneidad por parejas. La investigación, por los propósitos que la caracterizan, no permitió profundizar en este punto.

En cuanto a los ítemes 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 ítemes que claramente no corresponden a los de “Respuesta corta”. El punto es importante porque uno de los hallazgos más interesantes de la investigación ha consistido en develar que existen puntuaciones inadecuadas en los ítemes objetivos claramente desventajosos en comparación con algunos de desarrollo de igual o menor dificultad.

Dos de las características de este tipo de ítem no alcanzaron 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%.

La primera de estas características si parece tener, al menos potencialmente, efecto en la confiabilidad de la prueba, en tanto que no parece ser el caso de la segunda.

4.1.3. Sobre otros hallazgos

Los hallazgos relacionados con la existencia de ítemes objetivos con igual o mayor dificultad que otros de desarrollo pero con puntaje menor, resulta preocupante 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, razón por la cual debe ser atendida en los procesos formativos de profesores de matemática y en las acciones de capacitación que desarrolle el MEP.

Los resultados obtenidos también 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 ítemes en la misma prueba que evalúan esencialmente lo mismo. Esta situación 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 ítemes, así como fortalecer la formación matemática de los profesores de matemática que lo requieran, son dos elementos

que requieren de atención prioritaria del MEP en los procesos de capacitación del personal docente activo. En este sentido los esfuerzos realizados en el segundo semestre del año 2010 mediante el convenio firmado entre CONARE y el MEP, ejecutado por medio de convenios adicionales de las universidades estatales y CONARE, son esfuerzos que deben replicarse.

4.2. Conclusiones

Lo planteado anteriormente, se concreta entonces en las siguientes conclusiones:

Sobre el análisis del 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.
- Un 24.6% de los exámenes analizados no cumplen con la condición de nitidez exigida en las directrices del MEP.
- En el 6.15% de las pruebas había preguntas opcionales, y en un 9.23% de las mismas incluye cuadro de respuestas, situaciones que se separan de lo establecido por las directrices del MEP.

Sobre la parte administrativa de las pruebas

- La mayoría de las pruebas analizadas cumplen con los requisitos administrativos solicitados por el MEP, siendo particularmente la *firma del padre de familia*, *el nombre del docente* y *la fecha* los datos faltantes en al menos un 20% de las pruebas.

Sobre la parte técnica de las pruebas

- 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 ítems indicados.
- Los ítems de “Apareamiento” están presentes solo en el 18.46% de las pruebas y los de “Identificación” en un 6.16%.
- Los ítems de desarrollo que alcanzan los mayores valores porcentuales de presencia en las pruebas cotejadas 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 ítemes de “Respuesta restringida” aparecen en algunas pruebas, con una participación escasa y limitada exclusivamente a uno de los colegios participantes.

Sobre los ítemes de “Selección única”

- Los ítemes de “Selección única” cumplen satisfactoriamente todos los criterios, salvo el referente a la homogeneidad de las opciones, con porcentajes muy alto de cumplimiento que superan el 80% en todos ellos, siendo que la amplia mayoría excede el 90%.
- Solo el 35,21% de los ítemes de “Selección única” cumple con el requisito de que sus opciones sean homogéneas.
- El 22.42 % de los ítemes de “Selección única” presentaron homogeneidad por parejas.
- Si se aceptara que la homogeneidad por parejas fuera válida en la redacción de ítemes de “Selección única”, el porcentaje total de ítemes que cumplirían el criterio de homogeneidad se incrementaría al 57,63% el cual, en todo caso, sigue siendo bajo.

Sobre los ítemes ítemes de “Respuesta corta”

- El cumplimiento de los criterios en los ítemes de “Respuesta corta” es del 83.37% en todos los casos con excepción del que indica que “Al final del enunciado, no se debe utilizar artículos, ni los verbos “ser” o “estar”, o preposiciones que generen ambigüedad”.
- Un 16.62 % de los ítemes incorporados en las pruebas bajo la modalidad de “Respuesta corta” correspondían a ítemes de desarrollo.

Sobre los ítemes de “Apareamiento”

- Solo dos de las características de los ítemes de “Apareamiento” 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%.

Sobre los ítemes de “Selección única” de igual o mayor dificultad que algunos de desarrollo

- Se identificaron ítemes de “Selección única” en los que la aplicación de conocimientos requeridos para resolverlos sobrepasa lo establecido por el MEP y, por lo tanto, el valor asignado de un punto no es el adecuado.

- En la parte de resolución de ejercicios correspondiente a la selección de desarrollo de las pruebas, fueron detectados ítemes en los cuales el puntaje asignado no se corresponde con la dificultad o que, comparados con los ítemes objetivos como los de “Selección única” o los de “Apareamiento”, son similares por lo cual el puntaje asignado no es proporcional para toda la prueba.

Sobre los ítemes de “Falso o Verdadero” encubiertos como de “Selección única” o de “Respuesta corta”

- Se encontraron 29 ítemes de “Selección única” o de “Respuesta corta” que, aunque tienen una redacción para adecuarlos al formato de ítem con el cual se le incorpora en la prueba, corresponden básicamente a ítemes de “Falso o Verdadero”. Los ítemes de “Falso o Verdadero” no están autorizados por las directrices del MEP.

Sobre los ítemes con serios problemas de redacción

- Se detectaron 10 ítemes con serios problemas de redacción, en unos casos por falta de precisión matemática y en otros porque no se comprende lo que plantea.

Sobre los ítemes que evalúan reiteradamente lo mismo

- Se detectaron 53 conjuntos de ítemes, en grupos de 2, 3, 4, 5 o 6 ítemes, que evaluaban esencialmente el mismo conocimiento en la misma prueba.

Sobre los ítemes carentes de pertinencia

- Se identificaron 26 ítemes cuyo contenido motiva al cuestionamiento de la utilidad de incluir este tipo de preguntas en las pruebas escritas encaminadas a la evaluación sumativa.

5. Recomendaciones

Con base en los resultados de la investigación se plantean, muy respetuosamente, las siguientes recomendaciones:

A las universidades

- a. Realizar investigaciones similares en otras regiones educativas del país.
- b. Considerar los resultados de esta investigación en la formación de profesores/as de matemática.

Al MEP

- a. Valorar los resultados de esta investigación en el desarrollo de los programas de capacitación de los docentes en servicio.
- b. Retomar los convenios con las universidades estatales por medio de CONARE para darle continuidad a la formación matemática y pedagógica de los docentes.
- c. Mejorar los sistemas de control sobre la elaboración de las pruebas escritas con el fin de corregir las deficiencias detectadas en esta investigación.

A la Escuela de Matemática del ITCR

- a. Apoyar el desarrollo de nuevos proyectos de investigación que le den continuidad a este estudio.
- b. Desarrollar investigaciones que atiendan los siguientes problemas:
 - ¿Cuál es la eficacia del “Cuadro de balanceo” en el diseño de pruebas escritas?
 - ¿Qué características presentan los “Cuadros de balanceo” elaborados por los docentes de matemática en la educación media?
 - Desarrollar un programa de educación continua dirigido a docentes de matemática de la educación media encaminado a la atención de las deficiencias detectadas en la investigación.
 - Incorporar los resultados de la investigación en el curso “Evaluación de los aprendizajes” de la carrera EMAC.

6. Agradecimientos

Se agradece el apoyo de la Vicerrectoría de Investigación y Extensión para el desarrollo de este proyecto y de los seis colegios participantes.

7. Referencias

Brenes, F. (1990) *Principios y técnicas de evaluación II*. San José: EUNED.

Camillioni, A., Celman, S., Litwin, E. y Palou, M. (1998) *La evaluación de los aprendizajes en el debate didáctico contemporáneo*. Recuperado el 1 de abril de

2010, de http://www.farg.edu.uy/estructura/unidades_de_gestion/uap/mateval/aprend/Edith%20Litwin%20y%20Susana%20Celman.pdf

Castro, C. (1998) *Instructivo para la confección de ítem de pruebas estructuradas*. Recuperado el 10 de noviembre de 2011, de www.saidem.org.ar/docs/Preparar/Castro%20C.%20Instructivo%20para%20la%20redacci%F3n%20

Constenla, J. (s.f) *Los procedimientos de pruebas*. Recuperado el 7 de diciembre de 2011, de maestro.ucsc.cl/html/pec0006/apunte_proced_pruebas2.doc

D'Agostino, G. (2000) *Aspectos teóricos de la evaluación educacional. Una orientación para su puesta práctica*. San José: EUNED.

Hernández, R. (1990) *Principios y técnicas de evaluación IV*. San José: EUNED.

Hernández, R. (1993) *Aprendamos a elaborar exámenes escritos*. San José: EUNED.

Flores, R. (1999) *Evaluación pedagógica y cognición*. Colombia: Mc Graw Hill.

Gil, F., Rico, L. y Fernández, A. (2002) *Concepciones y creencias del profesorado de secundaria sobre evaluación en matemática*. *Revista de Investigación educativa*, 20 (1), 47-75. Recuperado el 10 de abril de 2010, de <http://revistas.um.es/rie/article/viewFile/97441/93541.2022>

Ministerio de Educación Pública (MEP). (2005) *Programas de matemática para la educación secundaria*. San José: Publicaciones del Ministerio de Educación Pública.

Programa Estado de la Nación. (2011) *Desempeño de la educación básica y el ciclo diversificado. Tercer Informe Estado de la Educación*. Recuperado el 23 de enero de 2012, de http://www.estadonacion.or.cr/images/stories/informes/educacion_003/docs/Parte_1_Capitulo_3.pdf

Rocha, M. (s.f) *Reglas para elaborar ítems de selección múltiple con única respuesta*. Recuperado el 12 de octubre de 2011, de <http://w3.icfes.gov.co:8080/seminariointernacional/ktmlite/files/uploads/MARTHA%20ROCHA%20-%20ICFES.pdf>

Rodríguez, G., Gil, J. y García, E. (1996) *Metodología de la investigación cualitativa*. Málaga: Ediciones ALJIBE.

Santamaría, M. (2005) *¿Cómo evaluar aprendizajes en aula?* San José: EUNED.

Torres, R. y otros. *La prueba escrita*. Recuperado el 12 de febrero de 2011, de www.mep.go.cr

Anexo No. 1.

Cartago, __ de __ del 2011

M.Sc. _____
Director Liceo _____

Estimado señor Director:

Somos profesores e investigadores de la Escuela de Matemática del Instituto Tecnológico de Costa Rica. En el presente año estamos desarrollando un proyecto de investigación titulado "***Evaluación de los aprendizajes en matemática en la educación media: características técnicas de las pruebas escritas***", cuyos objetivos se plantean en el documento adjunto.

Por la distinguida trayectoria de la Institución que usted dirige, estamos interesados en poder incluirla dentro de las instituciones participantes. Consecuentemente, le solicitamos nos facilite copia de exámenes escritos de matemática que hayan sido efectivamente aplicados en el año 2010 o en lo que va del 2011. Estos exámenes pueden ser de cualquiera de los niveles educativos, aunque lo ideal es contar con exámenes de cada uno de los niveles. Por tanto, nos gustaría poder contar con al menos cinco exámenes aplicados en su institución.

Nos comprometemos a tratar de manera confidencial la información que nos suministre, razón por la cual la publicación de los resultados se hará en forma general sin que se brinden detalles particulares de cada una de las instituciones participantes.

Firmemente convencidos de que los resultados de esta investigación permitirán mejorar los procesos de evaluación de los aprendizajes en el campo de la matemática en la educación media, nos permitimos reiterar la solicitud.

Agradecemos su fina atención. Atentamente,

Dr. Luis Gerardo Meza Cascante

M.Sc. Evelyn Agüero Calvo

Licda. Gabriela Roldán Villalobos

Profesores Escuela de Matemática
Instituto Tecnológico de Costa Rica

Instituto Tecnológico de Costa Rica

Escuela de Matemática

Proyecto de investigación

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

Objetivo general:

Determinar la calidad técnica de las pruebas escritas en matemática aplicadas en la educación secundaria de los colegios del Cantón Central de Cartago.

Objetivos específicos:

1. Establecer el nivel de aplicación de las directrices establecidas por el MEP para el diseño de pruebas.
2. Identificar el nivel de aplicación de los lineamientos teóricos en el diseño de las pruebas.
3. Establecer el nivel de aplicación de las directrices establecidas por el MEP para el diseño de ítems en matemática.
4. Identificar el nivel de aplicación de los lineamientos teóricos en el diseño de los ítems en matemática.

Investigadores responsables:

Dr. Luis Gerardo Meza Cascante
M.Sc. Evelyn Agüero Calvo
Licda. Gabriela Roldán Villalobos

Anexo No. 2.

**INSTITUTO TECNOLÓGICO DE COSTA RICA
ESCUELA DE MATEMÁTICA**

PROYECTO DE INVESTIGACIÓN

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

Lista de cotejo para ítemes de “Selección única” basada en el documento “La prueba escrita”

Selección única	Item	Item	item
Estar constituido por un texto, un gráfico, un esquema, una pregunta, una oración incompleta, un dibujo o imagen, entre otros, seguido de una oración introductoria o explicativa de la información suministrada.			
No más tres conocimientos previos en Tercer Ciclo y Educación Diversificada.			
Cuando se trata de un enunciado con forma de oración incompleta, el espacio por completar debe ubicarse siempre al final de la oración			
Presentar la información clara y necesaria, de manera que el estudiante comprenda lo que se le solicita, sin necesidad de leer antes las opciones.			
La(s) palabra(s) que se repite(n) al inicio de todas las opciones, deben colocarse al final del enunciado.			
Aplicar las reglas gramaticales y morfosintácticas del idioma.			
Cada ítem debe ser independiente de los demás; es decir, la respuesta de una pregunta no debe depender de la respuesta de otra anterior.			
A cada ítem se le asigna el valor de un punto.			
Pautas para redactar las opciones			
Solo una debe ser correcta.			
La respuesta correcta debe colocarse al azar.			
Todas deben guardar homogeneidad, en cuanto a longitud y temática.			
Mantener la concordancia gramatical con el enunciado.			
El número de opciones es cuatro para II y III Ciclos de Educación General Básica y Educación Diversificada.			
En su redacción no se deben utilizar expresiones como: todas las anteriores, ninguna de las anteriores, porque induce a la eliminación de la misma.			

Deben presentarse en la misma página en que se presenta el enunciado.			
Se pueden colocar en forma horizontal y/o vertical.			
Se identifican con números, paréntesis, letras (mayúsculas o minúsculas) o cualquier otro elemento que el docente considere apropiado, los mismos no deben representar una distracción para el estudiante.			
En III Ciclo de la Educación General Básica y en Educación Diversificada cada opción puede presentar una, dos o tres respuestas.			
Se deben redactar de tal manera que no sea necesario resolver cada una de las opciones para determinar la respuesta correcta.			

Anexo No. 3.

00015007

- ♣ La suma de los ángulos internos de un triángulo corresponde a
- 60°
 - 90°
 - 180°
 - 360°