**Attribute based identification in mathematic processes in a large scale assessment: The Mexican case**

GDRR, JCPM, RVL, AFCDP

**Resumen**

Los modelos de diagnóstico cognitivo (CDM) son modelos psicométricos desarrollados para identificar con detalle, a partir de la aplicación de una prueba y a lo largo del conjunto de conocimientos y habilidades cognitivas que conforman el dominio evaluado, las fortalezas y debilidades de cada sustentante y permiten trazar perfiles detallados que orienten el desarrollo de acciones de mejora focalizadas. En este trabajo se presenta la aplicación de una técnica de retrofitting usando el modelo DINA (de la Torre, 2009) sobre los datos obtenidos en una evaluación estandarizada de Matemáticas aplicada a gran escala a estudiantes de educación básica en México: PLANEA (INEE, 2017). La matriz Q fue construida a partir de una revisión curricular y entrevistas cognitivas dirigidas a expertos y alumnos. El ajuste del modelo DINA a los datos mostró ser el mejor dentro de los CDM. A partir de este trabajo, se pudo elaborar un diagnóstico cognitivo nacional, cuyas implicaciones educativas, empíricas y de apoyo pedagógico para la mejora en el aula se discuten a detalle.

Keywords: attribute mastery, cognitive diagnosis modeling, DINA model, mathematics, large scale assessment

**Introducción**

Una reciente y robusta línea de investigación en psicometría se ha enfocado en medir múltiples dimensiones o componentes de un constructo latente para proporcionar información específica y detallada acerca del desempeño de los sustentantes, a diagnóstico cognitivo. De manera general, el objetivo principal de los distintos modelos de diagnóstico cognitivo es poder identificar el estado o grado de dominio que los sustentantes evaluados presentan a lo largo de un conjunto de habilidades, conocimientos o atributos latentes para lograr la construcción de un perfil diagnóstico específico, que permita a todo agente involucrado conocer las áreas de fortaleza y mejora identificadas para orientar el desarrollo de estrategias de mejora. En el ámbito educativo, la aplicación de estos modelos permite a padres de familia, docentes, directores y otros agentes educativos, tomar mejores decisiones con respecto al diseño, uso, manejo e implementación de recursos y materiales didácticos de apoyo hacia los estudiantes. Cabe señalar que existe una amplia gama de modelos de diagnóstico cognitivo (CDM), para una revisión más extensa consultar se sugiere consultar a Rupp y Leighton, (2017), o a van der Linden, (2016).

Los CDM pueden entenderse como modelos de clases latentes restringidos que modelan las respuestas observadas como una función de variables latentes discretas que capturan las operaciones cognitivas subyacentes (Templin y Henson, 2006). En otras palabras, los CDM interpretan el desempeño de los participantes a través de una serie de atributos cognitivos binarios, a diferencia de los modelos derivados de la Teoría Clásica de los Test (TCT) o la Teoría de Respuesta al Ítem (TRI) que interpretan el desempeño de los sustentantes como reflejo de su posición a lo largo de un solo continuo que representa el dominio de una única variable o habilidad latente evaluada.

Para que una evaluación pueda ser considerada un diagnóstico cognitivo, el diseño de la misma debe considerar la integración de diversas teorías de aprendizaje, cognición y pedagogía con teorías de medición que permitan el desarrollo de evaluaciones que no se limiten a medir y evaluar, sino que sirvan como insumo para impulsar la mejora del aprendizaje de los estudiantes (Chudowsky y Pellegrino, 2003; Shepard, 2000).

La gran mayoría de los modelos de diagnóstico cognitivo requieren de la construcción de una matriz donde se identifique para cada ítem, cuáles son las habilidades o variables latentes asociadas al dominio general evaluado que se consideran “necesarias” para obtener un acierto. El arreglo resultante es conocido en la literatura como matriz Q, y su construcción requiere del trabajo conjunto de expertos en el dominio evaluado, sustentantes que tras responder cada reactivo den cuenta de los procedimientos seguidos y de expertos en psicometría que estén constantemente revisando que los atributos identificados como “requeridos” por cada ítem, hagan sentido a la luz de las respuestas observadas, (referencias).

**Acerca de los Modelos de Diagnóstico Cognitivo**

Cada uno de los CDMs hace supuestos específicos sobre cómo el dominio de los distintos atributos es ponderado para producir una respuesta correcta o incorrecta. Una distinción importante tiene que ver con si el modelo es conjuntivo o disyuntivo (Rupp, Templin y Henson, 2010). Los modelos conjuntivos asumen que se requiere el dominio de todos los atributos asociados a cada ítem para poder responder de manera exitosa. En contraste, los modelos disyuntivos asumen que la falta de dominio de un atributo puede ser compensada por el dominio de otros atributos, (por ejemplo, algunos CDM asumen que el dominio de los atributos tiene un efecto aditivo referencias).

Algunos de los CDMs más desarrollados y utilizados en la literatura son el modelo conjuntivo DINA (entrada determinística, ruidosa "y" puerta; Junker & Sijtsma, 2001; de la Torre, 2009) y su variante disyuntiva, el modelo DINO (entrada determinística, ruidosa "o" puerta; Templin y Henson, 2006), y el A-CDM (CDM aditivo; de la Torre, 2011). Según Rupp, Templin y Henson, (2010), otros CDM bien conocidos son el modelo NIDA (determinista de entrada ruidosa y; Junker y Sijtsma, 2001, Maris, 1999), el NIDO (determinista de entrada ruidosa o, Templin, Henson, y Douglas, 2006), y el R-RUM (modelo unificado de reparación reducida; Hartz, 2002). Además, se han propuesto CDM generales que respetan los supuestos de modelos específicos (véase, por ejemplo, Henson, Templin & Willse, 2009; von Davier, 2005), o bien, que funcionan como modelos generales, como es el caso del modelo G-DINA (DINA generalizada; de la Torre, 2011), el modelo de diagnóstico cognitivo log-lineal (LCDM; Henson, Templin y Willse, 2009) y el modelo de diagnóstico general (GDM; von Davier , 2005). Este último grupo de modelos, describe la probabilidad de éxito en términos de la suma de los efectos debidos a la presencia de atributos específicos y sus interacciones.

**Acerca del Modelo DINA**

El modelo DINA constituye uno de los modelos más sencillos dentro de la familia de los CDMs (Junker y Sijtsma, 2001), pues considera únicamente dos parámetros libres por ítem, con el objetivo de describir la probabilidad de que los aciertos o errores registrados no estén relacionados con el grado de dominio que los sustentantes tienen en las habilidades requeridas. Estos parámetros, conocidos como parámetro de adivinación y parámetro de desliz, refieren a la probabilidad de obtener un acierto aún sin dominar las habilidades necesarias, (es decir, de “atinarle por por puro azar” a la respuesta correcta), y la probabilidad de errar el ítem aún dominando las habilidades necesarias, (en otras palabras, de cometer un “desliz” al momento de seleccionar una respuesta).

El modelo DINA se expresa a partir de la siguiente ecuación,

De acuerdo con la Ecuación 1, el modelo DINA está compuesto por los siguientes elementos:

1.- Una variable binaria para indicar si la persona obtuvo un acierto (1) o un error (0) en el ítem . Esto implica la existencia de un vector Y compuesto por filas e columnas, donde se concentran los aciertos y errores obtenidos por cada persona en cada ítem.

2.- Un vector que contiene a la variable binaria que señala si la persona domina (1) o no (0) el subdominio evaluado en la prueba. Al conjunto de dominios evaluados suele referírseles, en el marco de los CDM, como habilidades, conocimientos o atributos contenidos en el dominio general evaluado por la prueba.

3.- Un vector que por cada ítem , contiene a la variable binaria que señala la respuesta “idealmente” registrada (en términos de acierto o error) por el sustentante en el ítem , dado lo que la matriz Q, el modelo cognitivo detrás de la prueba, ha establecido acerca de qué habilidades se requieren para obtener un acierto en cada ítem ( y lo que el vector nos dice sobre las habilidades dominadas por el participante . De esta forma, tenemos que el vector queda plasmado como el elemento determinista del modelo, definido matemáticamente como:

, donde

4.- Un parámetro de desliz , que nos dice que aunque la respuesta idealmente esperada por el participante al ítem , sea 1 (porque domina todas las habilidades requeridas por el ítem, ), el participante cometa un “desliz” y se observe aun así que . Es decir,

)

5.- Un parámetro de adivinación , que representa la probabilidad de que un examinado que no posee todas las habilidades requeridas por el ítem , (), “le atine” a la respuesta correcta y respondan correctamente el ítem (). Es decir,

El modelo DINA nos dice que para cada ítem , solo los examinados que dominen todos los atributos requeridos tendrán una probabilidad de éxito igual a , mientras que todos los demás examinados tendrán una probabilidad de éxito igual a . Es decir, que como habíamos mencionado previamente, el modelo DINA es un modelo conjuntivo que asume que es necesario el dominio de todos y cada uno de los atributos relacionados con cada ítem.

**Acerca de la Prueba de Matemáticas para Primaria (06) del PLANEA ELCE 2015**

La Prueba de Matemáticas para Primaria (06) con la que se trabajó en el presente proyecto, forma parte del Plan Nacional para las Evaluaciones de los Aprendizajes (PLANEA), diseñado y aplicado en 2015 en México por el entonces vigente Instituto Nacional para la Evaluación de la Educación (INEE), con la intención de contar con una evaluación a gran escala que permitiera valorar la eficacia del Sistema Educativo Nacional al fomentar el dominio de los aprendizajes esenciales alineados al currículo nacional, a partir del promedio de los puntajes logrados por los estudiantes evaluados a lo largo de distintas competencias o asignaturas curriculares. Esta iniciativa pretendía cubrir cuatro grandes objetivos: (1) Informar a la sociedad mexicana sobre el estado de la Educación en términos del logro de aprendizaje de los estudiantes; (2) aportar información relevante para el monitoreo, la planeación, programación y operación del sistema educativo y sus centros escolares; (3) ofrecer información pertinente, oportuna y contextualizada a las escuelas y a los docentes, que ayude a mejorar sus prácticas de enseñanza y (4) contribuir al desarrollo de directrices para la mejora educativa con información relevante sobre los resultados educativos y los contextos en que se dan.

Como antecedente directo del desarrollo del PLANEA, destaca la revisión realizada en 2013 por el mismo INEE acerca de la validez y confiabilidad de los Exámenes Nacionales del Logro Académico en Centros Escolares (ENLACE) y los Exámenes de la Calidad y el Logro Educativo (EXCALE), administrados a nivel nacional con un propósito similar (Martínez-Rizo, 2015). Como resultado de dicha revisión, el PLANEA cuenta con tres distintas modalidades para la evaluación de los aprendizajes de los estudiantes: (1) Evaluación del Logro referida al Sistema Educativo Nacional (ELSEN), (2) Evaluación del Logro referida a los Centros Escolares (ELCE) y (3) Evaluación Diagnóstica Censal (EDC).

En especial, el PLANEA-ELCE tiene como objetivo reunir información acerca de los centros escolares como unidad de análisis, a partir de la evaluación de campos formativos específicos (Matemáticas y Lenguaje y comunicación) y a lo largo de distintos niveles de desagregación (por ejemplo, región, estado, municipio, zona escolar, modalidad, tipo de organización, entre otros), (INEE, 2015). Su propósito específico es ofrecer a cada escuela elementos de retroalimentación para su mejora, identificando el logro alcanzado en el centro escolar tras los seis años que dura la educación primaria.

.

**Método:**

A partir del Enfoque Sistémico de Diseño Cognitivo (ESDC) propuesto por Embretson (1994), y de la perspectiva top-down para el diseño y validación de pruebas cognitivas (Bejar, 2002, 2010, Gorin y Embretson, 2013 y Messick, 1989b), se desarrolló el modelo metodológico del presente estudio, cuyo foco principal fue la obtención de evidencia que sustentara la validez sustantiva del constructo y corroborara la fidelidad estructural del modelo cognitivo subyacente a la prueba de Matemáticas de primaria (06) del PLANEA ELCE 2015. El modelo metodológico empleado para este proyecto estuvo conformado por cinco fases, descritas a detalle a continuación.

**Fase I: Diseño de los estudios cognitivos**

Las primeras fases de nuestro modelo metodológico contemplan la aplicación de estudios cognitivos que permitan identificar los modelos, estrategias y procesos de respuesta adheridos a cada ítem contenido en la prueba de Matemáticas de primaria (06) del PLANEA ELCE 2015. Esto obedece a la necesidad señalada por autores como Yang y Embretson, (2007) de que toda prueba de diagnóstico cognitivo diseñada con propósitos de mejora, deben ser diseñadas y validadas a partir de modelos cognitivos que permitan identificar de manera confiable los procesos de respuesta asociados a cada ítem.

Una ventaja colateral de trabajar con la prueba de Matemáticas de primaria (06) del PLANEA ELCE 2015 es que se trata de una disciplina ampliamente abordada desde el enfoque propuesto por los CDM (Brown y Burton, 1978; Chen y Macdonald, 2011; Gierl et al., 2009; Ma, Çetin y Green, 2009; Pérez-Morán, 2014; Pérez-Morán, Contreras-Roldan, Hernández, Olivares, Chan, y Díaz, 2014; Pérez-Morán, Larrazolo, Backhoff, y Guaner, 2015; Revuelta y Ponsoda, 1998; Romero, Ponsoda y Ximénez, 2008). A lo largo del presente trabajo, se utilizará el Ítem identificado con el ID PMA01 para ilustrar los procedimientos aplicados y los resultados obtenidos.

La genealogía curricular de los ítems, la evaluación de la congruencia y la alineación de los ítems y el modelamiento de los procesos de respuesta subyacentes, se realizó a partir de un análisis cognitivo-reticular con apoyo de un panel de expertos con experiencia en la enseñanza de las Matemáticas a nivel primaria.

De forma puntual, el panel de expertos estuvo conformado por tres psicólogos con experiencia en atención a estudiantes de primaria y secundaria, un especialista en análisis cognitivos y modelamiento de procesos cognitivos del pensamiento lógico-matemático, y un docente con amplia experiencia en la enseñanza de las Matemáticas en educación básica. La selección de los expertos se hizo de acuerdo con las recomendaciones de Rupp, Templin y Henson (2010) respecto a la elección de profesionales con un conocimiento profundo de los procesos de solución de problemas que utilizan los individuos en el dominio de interés, las diferentes vertientes para el desarrollo y aprendizaje de los componentes o atributos de la prueba y los contextos en los que los examinados adquieren y utilizan dichos atributos.

Una vez seleccionados los expertos, se les capacitó en la aplicación de métodos de análisis cognitivo-reticular, técnicas de pensamiento en voz alta concurrentes y retrospectivas, y en métodos de modelado matemático de sub-tareas de respuesta.

Para el análisis cognitivo-reticular de los ítems en estudio, se utilizaron la tabla de especificaciones de la prueba y su respectivo análisis reticular, acompañado de una revisión exhaustiva de los libros de texto diseñados para los alumnos y la guía del maestro del plan de estudios 2011[[1]](#footnote-1). La Figura 1 presenta como ejemplo el diagrama de la genealogía curricular correspondiente al ítem PMA01.

****

*Figura 1*. Diagrama de la genealogía curricular y de la alineación del ítem PMA01 de la prueba de Matemáticas de primaria del PLANEA ELCE 2015

También se aplicaron técnicas de pensamiento en voz alta concurrentes y retrospectivas con los expertos, para identificar los procesos de respuesta asociados a cada ítem y se realizaron diversas actividades para aplicar el modelado matemático de sub-tareas. Primero, mediante un proceso inductivo-deductivo los expertos identificaron y categorizaron los procesos y atributos cognitivos vinculados con cada uno de los ítems, apoyándose en los reportes verbales de su propia ejecución con los mismos, en los descriptores de los conocimientos, habilidades previas y procesos de respuesta declarados en la matriz de especificaciones de la prueba, así como en el sistema de categorías de procesos de respuesta típicos utilizados por estudiantes de educación básica en ítems de Matemáticas (Pérez-Morán, 2014). Como ejemplo, en la **Tabla 1** se pueden observar tres modelos de respuesta hipotéticos elaborados por expertos para el ítem PMA01 de PLANEA ELCE 06. Nótese que los tres modelos presentan diferentes niveles de granulación y que sólo el primero contempla la comprensión de problemas matemáticos contextualizados.

**Tabla 1**. Modelos hipotéticos del proceso de respuesta subyacentes al ítem desde la perspectiva de expertos y docentes

|  |  |
| --- | --- |
| **No.** | **Modelos hipotéticos del proceso de respuesta subyacente al ítem desde la perspectiva de especialistas y docentes** |
| 1 | a) Leer detalladamente el reactivo, b) observar la figura con sus respectivos valores, c) comprender qué es lo que se está solicitando (integrar el contexto del problema) d) determinar el tipo de cálculo a desarrollar, e) tomar los valores proporcionados en la figura para desarrollar el cálculo elegido, e) buscar entre las opciones de respuesta el valor calculado, y f) seleccionar la opción. |
| 2 | a) Leer detalladamente el ítem, b) comprender el objetivo de la tarea evaluativa, c) observar la figura d) recordar la fórmula para obtener el área de un triángulo, e) identificar las medidas que permitan calcular el área de la figura, f) determinar los valores que se utilizarán como base y altura del triángulo, g) desarrollar la fórmula para obtener el área de la figura h) obtener el resultado, y i) seleccionar la respuesta que coincida con el cálculo realizado. |
| 3 | a) Observar la imagen, b) leer la base del reactivo, c) comprender la pregunta, d) desarrollar el cálculo que determine el área de la figura, e) comparar el resultado con las opciones de respuesta, y f) seleccionar la respuesta. |

Finalmente, un equipo de expertos trabajó con los elementos de análisis del modelo para la Evaluación del Diseño Universal (EDU) propuestos por Thompson, Johnstone y Thurlow (2002), que ha demostrado ser de gran utilidad para el desarrollo de evaluaciones más accesibles para los examinados (Johnstone, 2003) y para minimizar la varianza irrelevante del constructo originada por problemas en el diseño, formato y sesgos culturales presentes en los ítems (Haladyna, Downing, y Rodríguez, 2002). Al aplicar la EDU con el fin de evaluar el diseño de los ítems, se trabajó con categorías y códigos específicos de análisis con base en los análisis propuestos por Thompson y colaboradores (2002): (a) inclusión poblacional, (b) definición precisa del constructo (c) accesibilidad e imparcialidad (d) acomodación flexible de los contenidos, (e) procedimientos e instrucciones simples, claras e intuitivas, (f) máxima legibilidad, y (g) comprensibilidad.

**Fase II: Piloteo y aplicación de protocolos verbales**

Durante la segunda fase, se pilotearon los ítems contenidos en la prueba y se obtuvieron reportes verbales de los procesos de respuesta empleados por un grupo de estudiantes, utilizando una vez más protocolos de pensamiento en voz alta con técnicas concurrentes y retrospectivas (Ericsson y Simon, 1984, 1993; Leighton, 2009; Leighton y Gierl, 2007b). Como análisis complementarios, se realizaron análisis del sendero de la vista (*eye-tracking*; Snow y Lohman, 1989; Sternberg, 1977) y se analizaron las latencias de respuesta (Fredericksen, 1980; Posner, 1978; Posner y Rogers, 1978). La implementación de estas medidas complementarias, se considera de gran ayuda para obtener información en los casos en que se presentan dificultades para evocar el reporte verbal de los procedimientos seguidos por estudiantes de corta edad, o bien, de procedimientos que suceden en cuestión de tan sólo algunos segundos y que, por tanto, no es posible su introspección (Sternberg, 1977). Todas estas técnicas complementarias, permiten una mejor verificación de la relación entre el modelo cognitivo elaborado por los expertos y los procesos cognitivos utilizados y reportados por los examinados para responder los ítems (Messick, 1989b).

Para el diseño de los protocolos verbales con estudiantes, se tomaron en cuenta los ocho pasos generales para la aplicación de las técnicas de pensamiento en voz alta recomendados por Leighton (2009) y se tomaron en cuenta las características del formato de los ítems en estudio y las características de los participantes.

Los criterios de selección para los estudiantes participantes en el piloteo y los protocolos verbales fueron elaborados a partir de las recomendaciones de Ericsson y Simon (1984, 1993), quienes proponen incorporar tanto a participantes novatos como a expertos en el dominio de interés, por lo que se tomaron como referencia variables tales como el rendimiento escolar, grado educativo y la recomendación del profesor. La selección de estudiantes expertos se realizó tomando como referencia un promedio mayor a 8.5 y el reconocimiento de sus profesores como estudiantes sobresalientes en el dominio de las Matemáticas. Los estudiantes novatos fueron seleccionados siempre y cuando tuvieran un promedio mayor a 6.0 pero menor a 8.0 y fueran referidos por sus profesores como estudiantes con bajo desempeño o no sobresalientes en el dominio de las Matemáticas. Se buscó que la muestra final quedara conformada en un 50% de estudiantes expertos y con un 50% de mujeres. En total se realizó el piloteo con 8 participantes por ítem.

Para agilizar la recolección de información y reducir en la medida de lo posible la duración total de cada piloteo, se aplicaron 25 de los 50 ítems en cada sesión, obteniendo una duración aproximada entre los 75 y 125 minutos. Durante la aplicación de los protocolos concurrentes de respuesta en voz alta se recolectaron datos del diseño de los ítems y se verificó que el proceso de respuesta utilizado por los participantes ante los ítems de la prueba estuviera representado en el modelo cognitivo elaborado por los expertos. Por su parte, los protocolos retrospectivos permitieron realizar preguntas a los participantes inmediatamente después de contestar el ítem con el fin de complementar la información obtenida en los protocolos concurrentes.

Para la captura de los datos, se montó un laboratorio cognitivo acorde a las necesidades específicas del estudio. Se utilizó el software CAMTASIA STUDIO versión 5.0.1 (TechSmith, 2008) porque permite grabar las verbalizaciones de los examinados, la imagen del ítem en la interfaz de la computadora junto con todas las acciones ocurridas en ella durante los protocolos verbales, el sendero del indicador del mouse y el tiempo de latencia de cada una de las actividades realizadas por el examinado. Además, al final de la aplicación de las técnicas de pensamiento en voz alta, se puede obtener y editar un video con todos los datos mencionados.

Una vez registrada la respuesta de los estudiantes, se capacitó a expertos en la verificación de los procesos de respuesta reportados con base en los modelos definidos previamente por el grupo de expertos conformado por docentes y especialistas en Matemáticas. Para ello, se trabajó de forma directa aplicando técnicas de análisis inductivo-deductivo. En la Figura 2 se puede observar como ejemplo, el proceso analítico realizado en el ítem PMA01 de la prueba de Matemáticas de sexto de primaria del PLANEA ELCE 2015, donde como resultado de la verificación de los modelos hipotéticos del proceso de respuesta subyacente al ítem propuesto por el panel de expertos y los reportes verbales de los estudiantes, se desarrolló el modelo cognitivo sintetizado del ítem.



*Figura 2*. Diagrama del modelo de un proceso de respuesta erróneo subyacente al ítem PMA01 de PLANEA ELCE 06 de Matemáticas

En esta misma figura, también se pueden observar tres problemas de diseño y sesgo identificados por los expertos en el marco del EDU, correspondiendo a las categorías de: a) Problemas de inclusión poblacional, b) comprensibilidad de la base del ítem, y c) definición imprecisa del constructo, asociado este último a su vez a problemas de alineación del ítem al currículo de primaria.

**Fase III: Desarrollo y definición del modelo cognitivo**

Durante la tercera fase de nuestro marco metodológico, se buscó definir, estructurar y desarrollar el modelo cognitivo que diera cuenta de las habilidades, conocimientos y procesos cognitivos evaluados o requeridos por la prueba. Para ello, se buscó determinar la cantidad y el tipo de relaciones existentes entre los ítems y los atributos u operaciones cognitivas sustantivas determinadas por los expertos y con ello, comenzar a elaborar la matriz Q de la prueba, (Rupp, Templin, y Henson, 2010).

Mediante un análisis inductivo-deductivo realizado por diadas de expertos, se sintetizaron los modelos cognitivos asociados a los diversos procesos de respuesta para desglosarlos en operaciones cognitivas concretas que permitieran un mayor nivel explicativo del comportamiento de cada ítem. Por ejemplo, si una vez más tomamos como ejemplo el ítem PMA01, las operaciones que explican su dificultad son: Comprensión de problemas matemáticos contextualizados, Comprensión del Sistema Internacional de Unidades (SIU), Representación del modelo aritmético para calcular el área de cuadriláteros o triángulos y Aplicación de operaciones aritméticas básicas (ver Figura 3).



*Figura 3*. Diagrama del modelo de un proceso de respuesta erróneo subyacente al ítem PMA01 de la prueba de Matemáticas (06) del PLANEA ELCE 2015

**Fase IV: Aplicación del análisis componencial**

La cuarta fase de nuestro marco metodológico consideró la evaluación de la calidad técnica de la prueba en estudio, aplicando el modelo DINA con la matriz Q diseñada por los expertos en los estudios cognitivos y evaluando la posibilidad de mejorar la matriz, explorando los resultados individuales y a lo largo de diferentes tipos de agrupaciones.

En cuanto a la revisión de la calidad técnica de los ítems de Matemáticas del PLANEA ELCE 2015, se obtuvieron indicadores propios de la Teoría Clásica de los Tests (TCT) y se realizó un análisis de estructura interna mediante la aplicación de Análisis Factoriales Exploratorios (AFE) con el método de estimación de mínimos cuadrados ponderados para variables categóricas. Se seleccionó una muestra aleatoria de 5000 estudiantes para los distintos análisis. de la calidad técnica y de validación de la prueba.

Los indicadores de la TCT fueron obtenidos con la paquetería CTT del programa libre R 2.15.1. (Ihaka, R. y Gentleman, R., 1996) para ser evaluados a la luz de los estándares clásicos de calidad técnica. Los indicadores psicométricos revisados fueron el índice de dificultad, el índice de discriminación, el coeficiente de correlación punto-biserial y el coeficiente de consistencia interna para la prueba y si se elimina un ítem (**α** de Cronbach).

El primer indicador, el índice de dificultad de cada ítem señala qué proporción de la muestra evaluada obtuvo un acierto, de acuerdo a la siguiente ecuación:

,

donde es la cantidad total de aciertos registrados para este reactivo y es la cantidad total de respuestas registradas para este reactivo (aciertos y errores).

Para obtener los *índices de discriminación* altos y bajos se utilizó la siguiente ecuación:

 ,

donde es la cantidad de aciertos registrados en el ítem por el 27% de los examinados que obtuvieron las puntuaciones más altas en el examen; es la cantidad de aciertos registrados en el mismo reactivo por el 27% de examinados que obtuvieron las puntuaciones más bajas en el examen, y es la cantidad de personas en el grupo más cuantioso ( o ).

La ecuación que se utilizó para obtener el coeficiente de correlación puntual-biserial () fue:

,

Donde  es la media de las puntuaciones totales de aquellos que respondieron correctamente el ítem,  es la media de las puntuaciones totales de aquellos que respondieron incorrectamente el ítem,  es la desviación estándar de las puntuaciones totales, es la cantidad de respuestas correctas registradas en el ítem, es la cantidad de respuestas incorrectas registradas y es igual a .

Finalmente el coeficiente de consistencia interna (**α** de Cronbach) del instrumento fue computado a partir de la siguiente ecuación:

,

donde es la cantidad de ítems de la prueba,  es la varianza de las puntuaciones de la prueba y  es la sumatoria de las varianzas de los reactivos.

Para el análisis de la estructura interna se aplicó un modelo de Análisis Factorial Exploratorio usando los paquetes *nFactors* y *psych* de R (Ihaka, R. y Gentleman, R., 1996). Los indicadores de ajuste absoluto que se emplearon fueron la raíz de los residuos cuadráticos promedios estandarizados (SRMR) y la raíz del error cuadrático promedio de aproximación (RMSEA), mismos que de acuerdo a la literatura deben ser inferiores a 0,05 para poder asumir un buen ajuste de los modelos factoriales.

Para el estudio de la estructura interna de la prueba se procedió a agrupar los ítems según su pertenencia a los ejes y contenidos curriculares. En cada eje de la prueba se implementó un análisis factorial exploratorio con el fin de comprobar la dimensión dominante. Como ya se mencionó, se utilizó el método de estimación de mínimos cuadrados ponderados dado que se analizaron variables categóricas.

Fase V: Estimación del Modelo DINA

Finalmente, una vez definido el modelo cognitivo subyacente a la prueba y la matriz Q que diera cuenta de la relación entre cada ítem y las distintas operaciones cognitivas identificadas, se procedió a realizar las estimaciones correspondientes bajo el marco de modelo DINA, utilizando la librería *CDM* en R, (Ihaka, R. y Gentleman, R., 1996).

Para verificar el ajuste del modelo a los datos recolectados

Resultados del análisis psicométrico

Resultados del análisis de la precisión métrica de la prueba

En cuanto a los resultados del análisis de la precisión métrica de la prueba, se observó con la aplicación del modelo de TCT que en términos generales se cumplían moderadamente con los estándares convencionales de calidad. En lo que corresponde al índice de consistencia interna mediante el alfa de Cronbach de cada eje temático se obtuvieron valores moderados entre 0.65 y 0.79 (ver **Tabla 6**). Los valores de los índices de Discriminación **rbis** de los ítems que conforman la prueba superan el umbral establecido de 0,3 con excepción de tres de ellos: PMB01, PMB10 y PMB14. Por su parte, los valores de los índices de dificultad oscilan entre 0.12 y 0.71, colocando a la mayoría de los ítems en niveles intermedios de dificultad y dentro de los estándares preestablecidos para los propósitos del estudio.

**Tabla 6**. Alfa de Cronbach global para cada eje

|  |  |
| --- | --- |
| **Eje** | **Alfa de Cronbach (α)** |
| Espacio, forma y medida | α=0.71 |
| Sentido numérico y pensamiento algebraico | α=0.65 |
| Manejo de información | α=0.79 |
| Fuente: Elaboración propia |  |

Para el caso de los ejes *Espacio, forma y medida* y *Manejo de información* se encontró que sus índices de confiabilidad podrían considerarse en las categorías aceptable a bueno, (alfa = 0.71 y 0.79 respectivamente). Pero para el eje *Sentido numérico* y *Pensamiento algebraico* el índice de confiabilidad se encuentra en un nivel más bajo (alfa = 0.65) por lo que la calidad métrica de este eje podría considerarse como dudoso o cuestionable, esto indica una mayor variabilidad de las respuestas a los reactivos que lo integran y que, por lo tanto, es de especial preocupación para una revisión de los reactivos en vías de áreas de mejora, ya que es altamente probable que las variaciones en las puntuaciones se deban a errores aleatorios (y por lo tanto no sistemáticos) asociados al proceso de medición.

Ahora bien, en cuanto al análisis de la calidad métrica del eje temático *Espacio, forma y medida*, en la **Tabla 7** se muestran los índices de dificultad y de discriminación por ítem y el alfa de Cronbach si se elimina algún ítem. Observese que el ítem PMA08 fue respondido correctamente por el 71% de los alumnos evaluados, indicando que es un ítem relativamente fácil; por su parte el caso opuesto se ilustra con el ítem PMB01 respondido acertadamente por el 12% de los alumnos evaluados, indicando que el ítem es relativamente difícil de ser contestado correctamente. En promedio se encuentra que el índice de dificultad para el eje *Espacio, forma y medida* es de aproximadamente 0.48, indicando que en conjunto los reactivos tienen una dificultad promedio de ser respondidos acertadamente.

**Tabla 7**. Índices de la calidad técnica de *Espacio, forma y medida* de acuerdo a la TCT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **No.** | **Reactivo** | **Dificultad** | **Discriminación rbis** | **Alfa si se elimina** |
| r27 | PMB02 | 0.42 | 0.34 | 0.70 |
| r29 | PMB04 | 0.48 | 0.32 | 0.70 |
| r07 | PMA07 | 0.46 | 0.34 | 0.70 |
| r08 | PMA08 | 0.71 | 0.41 | 0.70 |
| r01 | PMA01 | 0.22 | 0.32 | 0.70 |
| r26 | PMB01 | 0.12 | 0.27 | 0.71 |
| r03 | PMA03 | 0.33 | 0.50 | 0.69 |
| r30 | PMB05 | 0.45 | 0.35 | 0.70 |
| r31 | PMB06 | 0.55 | 0.36 | 0.70 |
| r32 | PMB07 | 0.54 | 0.46 | 0.69 |
| r33 | PMB08 | 0.56 | 0.46 | 0.69 |
| r34 | PMB09 | 0.57 | 0.47 | 0.69 |
| r04 | PMA04 | 0.56 | 0.35 | 0.70 |
| r05 | PMA05 | 0.67 | 0.44 | 0.69 |
| r06 | PMA06 | 0.59 | 0.38 | 0.70 |
| r02 | PMA02 | 0.37 | 0.38 | 0.70 |
| r28 | PMB03 | 0.51 | 0.31 | 0.70 |
| Fuente: Elaboración propia | | |  |  |

Otro aspecto importante a considerar respecto a la calidad técnica de los reactivos de una prueba es que el ítem permita discriminar o diferenciar de forma efectiva entre los sujetos con alto y bajo desempeño en la prueba, es decir, un ítem será eficaz en la medida en que los sujetos de alto desempeño tiendan a responder correctamente a un criterio y los sujetos de bajo desempeño probablemente lo realicen de forma incorrecta. Nuevamente en la **Tabla 7**, se muestran los valores índices de discriminación a través de la correlación biserial (rbis), este estadístico fue propuesto por Pearson (1909) como una aproximación al coeficiente producto-momento y que dentro de sus bondades es independiente del índice de dificultad del reactivo. Por lo tanto, la interpretación más intuitiva del índice de discriminación es que cuando los valores tienden a ser altos o se acercan al valor de 1.0 se considera que ese reactivo discrimina o separa en mucho mejor medida a los sujetos de alto y bajo desempeño; en la medida que la discriminación de un reactivo tiende a cero, se consideraría que dicho reactivo no podría separar entre sujetos de alto y bajo desempeño, teniendo que ser revisado, posiblemente modificado o bien eliminado de la prueba. Para el caso del eje *Espacio, forma y medida* el reactivo 03 (PMA03) presenta la discriminación más alta rbis= 0.50. Por su parte, el promedio de la discriminación de los reactivos es superior a 0.3, excepto por el reactivo 26 (PMB01) el cual presenta un índice rbis= 0.27, el cual debería de interpretarse con cuidado.

Finalmente, la última columna de la tabla 7, representa al valor de la confiabilidad o consistencia interna del eje *Espacio, forma y medida* si se eliminara algún reactivo, dicho estimador es útil para conocer el grado de impacto que tienen cada uno de los reactivos si potencialmente se eliminaran del eje, esto ocurre cuando en algunas ocasiones algún reactivo probablemente no esté tan relacionado con una escala y bajo ciertas condiciones es posible eliminarlo de la misma, en vías de que la consistencia interna de las puntuaciones de los demás reactivos en la escala se incremente, generalmente nunca es recomendable eliminar reactivos sin un sustento teórico y estadístico robusto, que justifique su eliminación. En el caso del eje *Espacio, forma y medida* observamos que no existe un cambio considerable en la confiabilidad general (alfa = 0.71), indicando que en general los reactivos tienden a presentar la misma consistencia en sus puntuaciones, de hecho, caso opuesto ocurriría si se eliminara algún reactivo, por ejemplo, el reactivo 32 (PMB07) la confiabilidad del eje disminuiría (alfa = 0.69), lo cual le restaría consistencia interna general al eje.

A continuación, se presentan y describen los índices en la **Tabla 8**, que reflejan la calidad técnica y que corresponden al eje *Manejo de información* de la prueba de Matemáticas de primaria (06) del PLANEA ELCE 2015. Con lo que respecta a la dificultad se observa un promedio de 0.46 en el eje, resaltando que el reactivo 13 (PMA13) es respondido por el 61% de los sustentantes mientras que el reactivo 35 (PMB10) solo es contestado por el 30% de los alumnos.

Con lo que corresponde a la discriminación en el eje *Manejo de información* se observa en la antepenúltima columna de la **Tabla 8**, a nivel general se obtiene una rbis = 0.38, lo cual es relativamente aceptable; por su parte el reactivo con mejor discriminación (PMA11) presenta una una rbis = 0.50. Se observa también que los reactivos 35 y 39 (PMB10 y PMB14) presentan valores rbis = 0.25 y 0.26 respectivamente, lo que indicaría especial interpretación de los resultados asociados a los mismos, debido a que no cumplen con el estándar mínimo aceptable de discriminación, por lo cual se sugiere tomar con cautela a los mismos.

**Tabla 8**. Índices de la calidad técnica de Manejo de información según TCT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **No.** | **Ítem** | **Dificultad** | **Discriminación rbis** | **Alfa si se elimina** |
| r14 | PMA14 | 0.59 | 0.48 | 0.61 |
| r35 | PMB10 | 0.30 | 0.25 | 0.65 |
| r39 | PMB14 | 0.54 | 0.26 | 0.64 |
| r09 | PMA09 | 0.29 | 0.34 | 0.64 |
| r10 | PMA10 | 0.37 | 0.35 | 0.63 |
| r11 | PMA11 | 0.57 | 0.50 | 0.61 |
| r12 | PMA12 | 0.52 | 0.49 | 0.61 |
| r13 | PMA13 | 0.61 | 0.47 | 0.61 |
| r36 | PMB11 | 0.38 | 0.33 | 0.63 |
| r37 | PMB12 | 0.45 | 0.35 | 0.63 |
| r38 | PMB13 | 0.48 | 0.39 | 0.63 |
| Fuente: Elaboración propia | | |  |  |

Respecto al cambio en el valor de la confiabilidad del eje *Manejo de información* el cual se presenta en la última columna de la **Tabla 8**, se observa de manera general que no es recomendable eliminar ningún reactivo. Hipotéticamente, si se considerara retirar algún reactivo, esto iría en detrimento de la consistencia interna de los reactivos que se conservaran. Se recalca que se debería de tener especial cuidado/atención en las conclusiones o decisiones que se tomaran con este eje debido a que se encuentra en el umbral de la categoría dudosa o cuestionable y por lo tanto se recomienda revisar a detalle los elementos que lo integran o bien las variables que inciden directa o indirectamente, como una sugerencia de revisión en detalle a futuro.

Por su parte, como se puede observar en la tabla 9, se presentan y describen los índices, que reflejan la calidad técnica y que corresponden al eje *Sentido numérico y pensamiento algebraico* de la prueba de Matemáticas de primaria (06) del PLANEA ELCE 2015. En términos de la dificultad se observa un promedio de 0.42 en el eje, resaltando que el reactivo 19 (PMA19) es respondido por el 66% de los sustentantes mientras que el reactivo 40 (PMB15) solo es contestado por el 28% de los alumnos.

Cuando observamos el comportamiento de la discriminación en el eje *Sentido numérico y pensamiento algebraico*, a nivel general se obtiene una rbis = 0.47, lo cual es aceptable. Por ejemplo, existe un reactivo (PMA17) con una muy buena discriminación rbis = 0.60. Por su parte los reactivos restantes cumplen el criterio mínimo de discriminación (rbis ≥ 0.30), excepto por el reactivo 43 (PMB18), el cual se encuentra en el límite inferior de este umbral de calidad técnica.

Finalmente, al reportar y describir al cambio en el valor de la confiabilidad del eje *Sentido numérico y pensamiento algebraico* el cual se presenta en la última columna de la **Tabla 9**, se observa de manera general que no es recomendable eliminar ningún reactivo, ya que, si se considerara retirar algún reactivo, esto iría en detrimento de la buena consistencia interna general que ya se tiene (Alpha = 0.79).

**Tabla 9**. Índices de la calidad técnica de *Sentido numérico y pensamiento algebraico* según TCT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **No.** | **Ítem** | **Dificultad** | **Discriminación rbis** | **Alfa si se elimina** |
| r20 | PMA20 | 0.38 | 0.45 | 0.78 |
| r21 | PMA21 | 0.48 | 0.49 | 0.78 |
| r22 | PMA22 | 0.59 | 0.51 | 0.78 |
| r25 | PMA25 | 0.68 | 0.44 | 0.78 |
| r40 | PMB15 | 0.28 | 0.50 | 0.78 |
| r41 | PMB16 | 0.30 | 0.53 | 0.78 |
| r44 | PMB19 | 0.35 | 0.57 | 0.77 |
| r15 | PMA15 | 0.29 | 0.58 | 0.77 |
| r18 | PMA18 | 0.35 | 0.35 | 0.78 |
| r23 | PMA23 | 0.52 | 0.47 | 0.78 |
| r24 | PMA24 | 0.50 | 0.51 | 0.78 |
| r16 | PMA16 | 0.33 | 0.39 | 0.78 |
| r17 | PMA17 | 0.44 | 0.60 | 0.77 |
| r19 | PMA19 | 0.66 | 0.48 | 0.78 |
| r42 | PMB17 | 0.33 | 0.38 | 0.78 |
| r43 | PMB18 | 0.30 | 0.30 | 0.79 |
| r45 | PMB20 | 0.38 | 0.45 | 0.78 |
| r46 | PMB21 | 0.48 | 0.44 | 0.78 |
| Fuente: Elaboración propia | | |  |  |

Resultados del análisis de distractores

Resultados del Análisis de la estructura interna

El análisis factorial implementado para comprobar la dimensión latente en la prueba permitió determinar un factor dominante para todos los ítems sin separar por dimensión. Este factor estaría asociado al constructo habilidad matemática medida por la prueba mediante sus tres ejes y contenidos. Los valores de las cargas factoriales de los ítems superan el umbral usualmente aceptado de 0,3 con excepción de siete ítems (ver **Tabla 10**).

**Tabla 10**. Modelo de la estructura interna unidimensional de la prueba de Matemáticas de primaria (06) del PLANEA ELCE 2015

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **No.** | **Ítem** | **Carga factorial estandarizada** |  | **No.** | **Ítem** | **Carga factorial estandarizada** |
| r01 | PMA01 | 0.26 |  | r26 | PMB01 | 0.20 |
| r02 | PMA02 | 0.33 |  | r27 | PMB02 | 0.31 |
| r03 | PMA03 | 0.48 |  | r28 | PMB03 | 0.27 |
| r04 | PMA04 | 0.30 |  | r29 | PMB04 | 0.28 |
| r05 | PMA05 | 0.40 |  | r30 | PMB05 | 0.32 |
| r06 | PMA06 | 0.32 |  | r31 | PMB06 | 0.33 |
| r07 | PMA07 | 0.31 |  | r32 | PMB07 | 0.45 |
| r08 | PMA08 | 0.37 |  | r33 | PMB08 | 0.47 |
| r09 | PMA09 | 0.35 |  | r34 | PMB09 | 0.47 |
| r10 | PMA10 | 0.33 |  | r35 | PMB10 | 0.24 |
| r11 | PMA11 | 0.45 |  | r36 | PMB11 | 0.33 |
| r12 | PMA12 | 0.46 |  | r37 | PMB12 | 0.34 |
| r13 | PMA13 | 0.44 |  | r38 | PMB13 | 0.36 |
| r14 | PMA14 | 0.47 |  | r39 | PMB14 | 0.29 |
| r15 | PMA15 | 0.49 |  | r40 | PMB15 | 0.40 |
| r16 | PMA16 | 0.33 |  | r41 | PMB16 | 0.45 |
| r17 | PMA17 | 0.54 |  | r42 | PMB17 | 0.33 |
| r18 | PMA18 | 0.29 |  | r43 | PMB18 | 0.25 |
| r19 | PMA19 | 0.44 |  | r44 | PMB19 | 0.49 |
| r20 | PMA20 | 0.38 |  | r45 | PMB20 | 0.39 |
| r21 | PMA21 | 0.45 |  | r46 | PMB21 | 0.38 |
| r22 | PMA22 | 0.46 |  |
| r23 | PMA23 | 0.42 |  |
| r24 | PMA24 | 0.44 |  |
| r25 | PMA25 | 0.39 |  |
| Fuente: Elaboración propia | | |  |

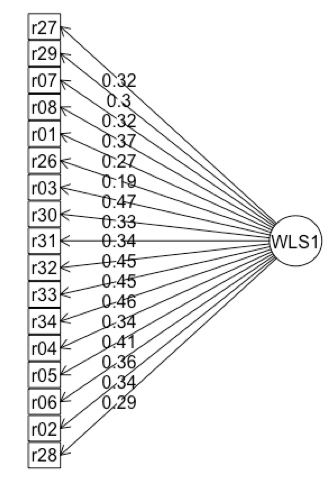
Adicionalmente, para evaluar el ajuste de la estructura factorial, se presentan dos estimadores, tanto el Standarized Root Mean Square Residual (SRMR) y el Root Mean Squared Error of Aproximation (RMSEA) resultaron con valores inferiores a 0.05 oscilando entre el 0.02 y 0.046, los cuales se pueden observar en la **Tabla 11**. Tales índices evidencian que por cada eje hay una dimensión dominante que permite asociarla con el eje teórico intencionado por parte de los desarrolladores de la prueba. Los resultados del ajuste absoluto de cada modelo unidimensional nos permiten conocer en qué medida el modelo factorial unidimensional teórico para cada eje representa a los reactivos que integran cada uno, y dado que están por debajo del criterio sugerido 0.05 se puede validar que se respetó un buen criterio de calidad técnica para el ajuste de cada uno de los modelos factoriales.

**Tabla 11**. Ajustes absolutos del modelo unidimensional por eje de la prueba de Matemáticas de primaria (06) del PLANEA ELCE 2015

|  |  |  |
| --- | --- | --- |
| **Eje** | **SRMR** | **RMSEA** |
| Espacio, forma y medida | 0.040 | 0.046 |
| Manejo de información | 0.020 | 0.029 |
| Sentido numérico y pensamiento algebraico | 0.020 | 0.030 |
| Fuente: Elaboración propia |  |  |

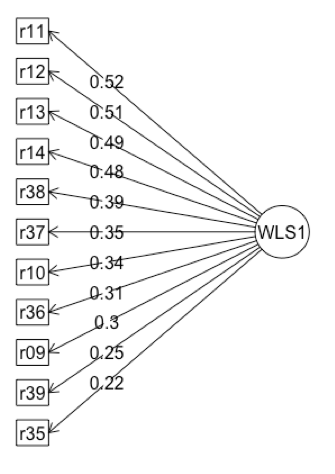
Ahora bien, para tener una mirada complementaria al comportamiento factorial de los reactivos en cada eje, se presentan los diagramas en las figuras 1 a 3 para cada uno de los modelos de la prueba de Matemáticas de primaria (06) del PLANEA ELCE 2015. Los gráficos muestran las respectivas cargas factoriales obtenidas por eje.

En el caso del primer eje *Espacio, forma y medida*, esbozado en la Figura 5, se observa que la mayoría de los reactivos presentan una carga factorial aceptable a buena (≥0.30), solamente tres de ellos, los reactivos 01, 26 y 28, presentan cargas factoriales (0.27, 0.19 y 0.29, respectivamente) menores al criterio sugerido, por lo que se sugiere prestar especial énfasis en las características particulares e inherentes a estos tres reactivos, con el objetivo de revisarse y potencialmente someterse a mejora en sus características técnicas.



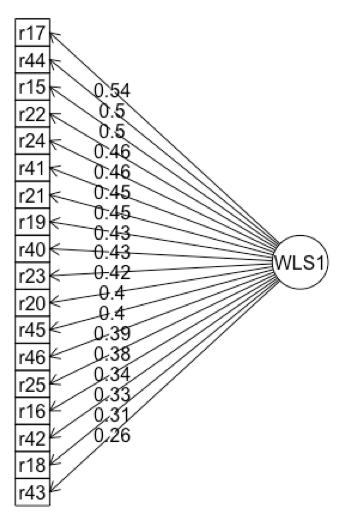
*Figura 5.* Modelo gráfico de la estructura interna unidimensional del eje Espacio, forma y medida de la prueba de Matemáticas de primaria (06) del PLANEA ELCE 2015

Para el caso del segundo eje “Manejo de información”, la mayoría de los reactivos presentan una carga factorial mayor o igual a la mínima recomendada (0.30), se puede observar en la Figura 6, que solo los reactivos 39 y 35 con cargas factoriales 0.25 y 0.22 respectivamente, no cumplen con este criterio técnico, por lo que con la misma lógica del eje anterior se sugiere prestar atención y tener cuidado acerca de las decisiones y conclusiones asociadas a dichos reactivos.



*Figura 6.* Modelo gráfico de la estructura interna unidimensional del eje Manejo de información de la prueba de Matemáticas de primaria (06) del PLANEA ELCE 2015

Finalmente, en la Figura 7 se presenta el tercer eje “Sentido numérico y pensamiento algebraico”, en donde nuevamente la mayoría de los reactivos cumple el criterio técnico de carga factorial (≥0.30), solamente el reactivo 43 con una carga factorial de 0.26, bajo las sugerencias antes hechas se recomienda prestar atención.



*Figura 7.* Modelo gráfico de la estructura interna unidimensional del eje Sentido numérico y pensamiento algebraico de la prueba de Matemáticas de primaria (06) del PLANEA ELCE 2015

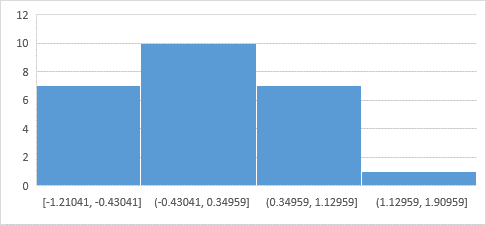
Resultados Teoría de Respuesta al Ítem (TRI)

Los resultados del análisis de los reactivos mediante la óptica de la TRI, se presentan en la **Tabla 12**, en donde se coloca el parámetro dificultad, el cual fue calibrado bajo un modelo de Rasch. Dicho parámetro se interpreta como una medición continua de la cantidad de rasgo latente en una escala estandarizada (con media 0 y desviación estándar 1) que se requiere para poder acertar con una probabilidad p=0.50 cada reactivo. Por lo tanto, por ejemplo, el reactivo 24 (PMA24) que presenta una dificultad cercana a cero (β=0.02), requiere un nivel de habilidad latente (θ) promedio. Otro caso se ilustra con el reactivo 26 (PMB01) el cual presenta una dificultad β=2.15, lo que indica, es que se necesita más cantidad de habilidad latente (θ) y por lo tanto tienen una probabilidad de acierto menor para aquellos sustentantes que tengan un nivel de habilidad latente baja.

**Tabla 12**. Ajustes absolutos del modelo unidimensional por eje de la prueba de Matemáticas de primaria (06) del PLANEA ELCE 2015

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **No.** | **Ítem** | **Dificultad** |  | **No.** | **Ítem** | **Dificultad** |
| r01 | PMA01 | 1.42669 |  | r26 | PMB01 | 2.15816 |
| r02 | PMA02 | 0.49278 |  | r27 | PMB02 | 0.07363 |
| r03 | PMA03 | 0.69758 |  | r28 | PMB03 | -0.19353 |
| r04 | PMA04 | -0.37634 |  | r29 | PMB04 | -0.2288 |
| r05 | PMA05 | -0.9039 |  | r30 | PMB05 | -0.09398 |
| r06 | PMA06 | -0.51266 |  | r31 | PMB06 | -0.54563 |
| r07 | PMA07 | -0.21105 |  | r32 | PMB07 | -0.71395 |
| r08 | PMA08 | -1.21041 |  | r33 | PMB08 | -0.52764 |
| r09 | PMA09 | 0.84757 |  | r34 | PMB09 | -0.76852 |
| r10 | PMA10 | 0.37918 |  | r35 | PMB10 | 0.70304 |
| r11 | PMA11 | -0.41153 |  | r36 | PMB11 | 0.26823 |
| r12 | PMA12 | -0.06558 |  | r37 | PMB12 | -0.12717 |
| r13 | PMA13 | -0.5199 |  | r38 | PMB13 | -0.16598 |
| r14 | PMA14 | -0.48669 |  | r39 | PMB14 | -0.58977 |
| r15 | PMA15 | 0.89901 |  | r40 | PMB15 | 0.81615 |
| r16 | PMA16 | 0.53155 |  | r41 | PMB16 | 0.69606 |
| r17 | PMA17 | 0.34281 |  | r42 | PMB17 | 0.52643 |
| r18 | PMA18 | 0.48815 |  | r43 | PMB18 | 0.65836 |
| r19 | PMA19 | -0.69305 |  | r44 | PMB19 | 0.39225 |
| r20 | PMA20 | 0.26237 |  | r45 | PMB20 | -0.02094 |
| r21 | PMA21 | -0.04339 |  | r46 | PMB21 | -0.35727 |
| r22 | PMA22 | -0.40305 |  |
| r23 | PMA23 | -0.08483 |  |
| r24 | PMA24 | 0.02239 |  |
| r25 | PMA25 | -0.93435 |  |
| Fuente: Elaboración propia | | |  |

Podemos observar en la Figura 8 que, en términos de la distribución de dificultades de los reactivos, predominan reactivos con dificultad promedio, pero solamente el reactivo 26 (PMB01) con un nivel de dificultad relativamente difícil (β=2.15).



*Figura 8*. Distribución de frecuencias del nivel de dificultad de la prueba de Matemáticas de primaria (06) del PLANEA ELCE 2015.

Resultados de la aplicación de los métodos cognitivos

Resultados del análisis por expertos del diseño de la prueba

Con respecto al análisis cognitivo-reticular aplicado a la prueba PLANEA ELCE 2015 en su formato Matemáticas 06, se utilizó el Modelo de EDU (Thompson, Johnstone y Thurlow, 2002) para lograr una detección y sistematización minuciosa de los errores presentados por los ítems. Este modelo abarca seis elementos que forman parte de la evaluación del diseño de los ítems de una prueba: 1) inclusión poblacional, 2) definición precisa del constructo, 3) accesibilidad e imparcialidad, 4) acomodación flexible de los contenidos, 5) de comprensibilidad, y 6) procedimientos e instrucciones simples, claras e intuitivas. El panel de expertos asignó 12 categorías en total a los elementos del EDU, mismas que se le relacionan negativamente. Lo anterior, debido a que cada categoría hace alusión a deficiencias en el abordaje de los elementos del EDU. Cabe señalar que cada ítem puede encontrarse en más de una categoría e incluso presentar más de un error de diseño en la misma categoría (véase apéndice 10).

Los resultados del análisis cognitivo arrojaron que, de los 50 ítems que componen la prueba, 46 ítems presentaron al menos un error de diseño, esta cifra corresponde al 92% de los reactivos. En total se detectaron 101 errores, donde la moda corresponde a 2 errores por ítem y los valores extremos corresponden entre cero y 6 errores Respecto a los errores detectados con base en el Modelo EDU, 9 ítems presentaron al menos un problema relacionado con el elemento *Inclusión poblacional*; es decir, el 18% total de la prueba. 34 de 50 de los ítems presentaron problemas relacionados con el elemento *Definición precisa del constructo*, correspondiente al 68%; 3 de 50 ítems presentaron problemas relacionados con el elemento *Accesibilidad e imparcialidad*, que atañe al 6% de la prueba; el elemento *Acomodación flexible* de los contenidos no presentó problemas relacionados, esta situación se explica con mayor detalle en el apartado *Problemas de alineación* del presente documento; 8 de 50 ítems manifestaron problemas relacionados con el elemento *Procedimientos e instrucciones claras, simples e intuitivas*, correspondiente al 16% de la prueba; por último, 10 de los 50 ítems presentaron problemas relacionados con el elemento *Comprensibilidad*, que corresponde al 20% de la prueba (véase **Tabla 13**).

**Tabla 13.** Problemas de diseño acorde al modelo EDU

|  |  |  |
| --- | --- | --- |
| **Elementos del EDU** | **Cantidad de ítems** | **Porcentaje de ítems** |
| Inclusión poblacional | 9 | 18% |
| Definición precisa del constructo | 34 | 68% |
| Accesibilidad e imparcialidad | 3 | 6% |
| Acomodación flexible de los contenidos | 0 | 0% |
| Procedimientos e instrucciones claras, simples e intuitivas | 8 | 16% |
| Comprensibilidad | 10 | 20% |
| Fuente: Elaboración propia | | |

Otra forma de sistematizar los errores detectados por los expertos a través del análisis cognitivo, es la clasificación de los errores en relación a su alineación, sesgo y diseño (véase tabla 14).

**Tabla 14.** Problemas de los ítems relacionados con alineación, sesgo y diseño

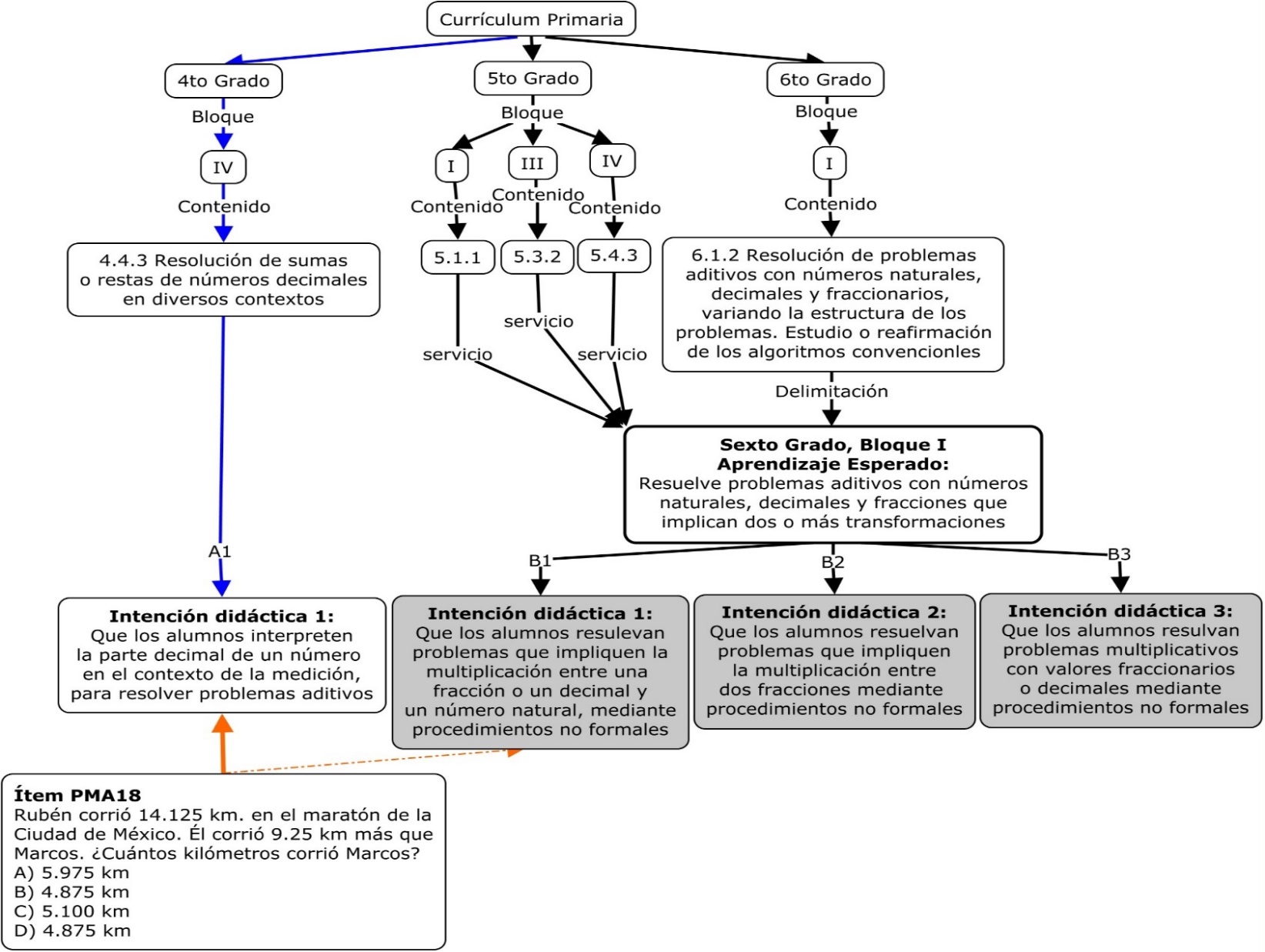
|  |  |  |
| --- | --- | --- |
| **Problemas** | **Cantidad de ítems** | **Porcentaje de ítems** |
| Alineación | 34 | 68% |
| Sesgo | 12 | 24% |
| Diseño | 31 | 62% |
| Fuente: Elaboración propia | | |

Problemas de alineación

Durante el análisis cognitivo los expertos detectaron que 34 ítems presentan *Problemas de alineación*, esto equivale al 68% del total de la prueba. Se contemplan 4 etiquetas para ubicar a los ítems en esta clasificación: (1) inadecuada alineación del ítem a las indicaciones de su respectiva especificación, (2) inadecuada alineación de la especificación a la retícula, (3) inadecuado cuidado del proceso de respuesta al ítem, y/o (4) inadecuada alineación curricular, esa última etiqueta refiere que presentan un grado de dificultad menor o mayor por no corresponder con el grado escolar; o bien, pueden no coincidir con el contenido a evaluar.

Por defecto, todos los ítems que presenten problemas de Alineación curricular, por estar mayormente relacionados a un contenido de otro grado escolar, presentan también problemas en la clasificación Dificultades en la adaptación de los contenidos; ya que esta clasificación incluye todos los ítems que no estén adecuados a la edad del sustentante. Sin embargo, el error en cuanto a la adaptación de los contenidos no se puntuó dentro de los errores de diseño detectados en Matemáticas 06 PLANEA ELCE 2015 con base en el Modelo EDU, en consideración a que el error se puntuó en la clasificación Alineación curricular previamente (ver apéndice 10).

Los ítems ubicados en esta clasificación fueron: PMA01, PMA03, PMA04, PMA05, PMA06, PMA07, PMA08, PMA09, PMA10, PMA13, PMA15, PMA16, PMA17, PMA18, PMA19, PMA20, PMA21, PMA22, PMA23, PMA24, PMB02, PMB05, PMB06, PMB08, PMB11, PMB12, PMB13, PMB14, PMB17, PMB18, PMB20, PMB21, PMB22, PMB23. Un ejemplo de los problemas que se manifestaron en esta clasificación es el caso del ítem PMA18, que se describe a continuación; en su correspondiente tabla de especificación se señala que el contenido a evaluar corresponde al bloque I de sexto grado y que dicho ítem evaluará el contenido *Resolver problemas aditivos con números decimales. La incógnita es el resultado*. Sin embargo, los modelos de enseñanza en el libro de textos *Desafíos matemáticos: Sexto grado* no corresponden con lo evaluado por el ítem (ver Figura 9).



*Figura 9.* Extracto del diagrama de análisis reticular del ítem PMA18

Tras el análisis realizado por los expertos, se encontró relación directa del ítem y su especificación con el contenido *4.4.3. Resolución de sumas o restas con números decimales en diferentes contextos* y con el modelo de enseñanza que le corresponde en el libro de texto *Desafíos matemáticos: Cuarto grado* en el bloque IV; mismo de donde se desprende la intención didáctica *Que los alumnos interpreten la parte decimal de un número en el contexto de la medición, para resolver problemas aditivos* (ver Figura 10)*.*

|  |
| --- |
| **A1: Intención didáctica 1**  **Grado:** 4to  **Eje:** Sentido numérico y pensamiento algebraico  **Tema:** Problemas aditivos  **Modelo del proceso de enseñanza:** a)Hacer parejas, b) leer el problema 1 que se muestra en la consigna 1, c) observar el contenido de la tabla e identificar los valores que se encuentran representados en decimales, d) leer las preguntas, e) comprender lo que solicita cada pregunta, f) elegir que valores y operaciones aritméticas serán utilizados para obtener el resultado en base a lo que solicite cada pregunta (sumas y restas), g) realizar las operaciones aritméticas necesarias con ayuda de la tabla, h) escribir el resultado que correspondiente a cada pregunta; i) leer el problema 2 de la consigna 1, j) observar la imagen de la cancha con sus respectivos datos, k) leer las preguntas del problema, l) comprender lo que solicita cada pregunta con base a los datos que se encuentra a en la imagen, m) determinar qué valores y tipo de cálculo matemático serán utilizados según lo que solicite cada pregunta, n) realizar las operaciones aritméticas necesarias con ayuda de la tabla, y ñ) escribir el resultado que correspondiente a cada pregunta. |

*Figura 10.* Extracto del modelo de enseñanza en los libros de texto del ítem PMA18

En la **Tabla 15** se pueden observar aquellos elementos de la especificación de los ítems de la prueba que los expertos consideraron como información útil para el diagnóstico de los procesos cognitivos subyacentes a los ítems. Nótese que el último nivel granular descriptivo son los rubros: *Especificació*n, *Conocimientos y habilidades previos requeridos para contestar el reactivo correctamente* y *Procesos cognitivos involucrados en la contestación correcta del ítem*. Cabe señalar que dichos rubros de los formatos de las especificaciones son cumplimentados por los desarrolladores de la prueba tomando como base el *Aprendizaje esperado* definido por un comité de expertos. A simple vista puede notarse que el rubro *Especificación* se encuentra sobre-simplificado a diferencia de los rubros relacionados con la descripción de los conocimientos previos, y los procesos cognitivos involucrados requeridos para contestar correctamente el ítem. Los expertos consideraron que la relación entre los elementos de la especificación es ambigua y presenta problemas de congruencia. En especial se consideró que la descripción del rubro *Especificación* no podría ser utilizada para ofrecer información diagnóstica de calidad a los evaluados debido a que no toma en cuenta la diversidad de operaciones cognitivas sustantivas que se requieren realmente para contestar correctamente el ítem PMA01.

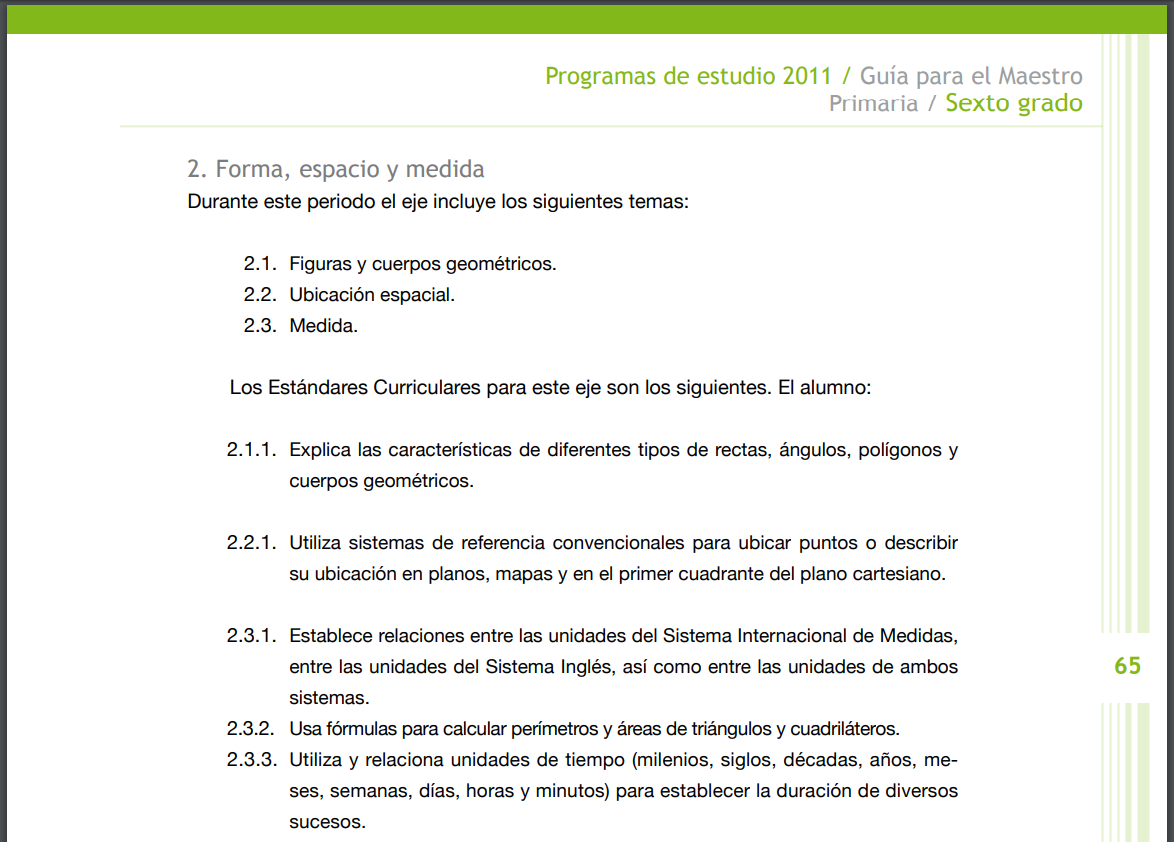
**Tabla 15**. Elementos de ubicación y de menor granulo de la especificación del Ítem PMA01 de la prueba de Matemáticas de primaria (06) del PLANEA ELCE 2015

|  |  |
| --- | --- |
| **Ubicación señalada en la retícula** | **Relación con otros contenidos según la retícula** |
| **Grado:** 5to.  **Eje:** Forma, espacio y medida.  **Tema:** Medida.  **Aprendizaje esperado:** Calcula el perímetro y el área de triángulos y cuadriláteros.  **Especificación:** Calcular el área de triángulos. | **-Contenidos fuente indicados en retícula:** 5.2.6 Construcción y uso de una fórmula para calcular el área de paralelogramos (rombo y romboide).  -**Contenidos que reciben servicio indicados en retícula:** 6.5.4. Armado y desarmado de figuras en otras diferentes. Análisis y comparación del área y el perímetro de la figura original y la que se obtuvo. |
| **Conocimientos y habilidades previos requeridos para contestar el reactivo correctamente según su especificación:** Para resolver correctamente el reactivo el alumno debe tener el concepto de base y altura de un triángulo, identificar estas dos medidas con respecto a cualesquiera de sus lados, saber la fórmula que resuelve correctamente y haber desarrollado la habilidad de multiplicar y dividir números naturales y decimales. | |
| **Procesos cognitivos involucrados en la contestación correcta del ítem según su especificación:** el alumno debe leer la base del reactivo, comprender la pregunta, analizar la figura dada para identificar los datos que resuelven el problema, conocer y comprender la fórmula para calcular el área, tener un manejo de técnica adecuada para operar con números naturales y decimales y haber desarrollado adecuadamente el sentido numérico para validar los resultados. El dominio cognitivo al que alude el procedimiento que realice el alumno se refiere a la resolución de problemas complejos. También se avanza en el proceso cognitivo que permite la construcción y uso de una fórmula para calcular el área de paralelogramos (rombo y romboide); el contenido culmina con la construcción completa del aprendizaje esperado que está referido a la construcción y uso de una fórmula para calcular el área del triángulo y el trapecio (SEP, 2011). | |

De igual forma, cuando los expertos revisaron la retícula para rastrear la genealogía curricular del ítem, se encontraron con dificultades para interpretar las indicaciones de los elementos que señalan las interrelaciones. En la Figura 11 se puede observar que del rubro *Aprendizaje esperado* se granula el contenido *5.3.6 Calcula el perímetro y el área de triángulos cuadriláteros* que recibe servicio del contenido 5.26 y aporta servicio al contenido 6.5.4. Sin embargo, las líneas que interrelacionan los diferentes elementos mencionados no indican la dirección de la relación por lo que los expertos tuvieron que deducirla por la numeración de los contenidos. Lo mencionado anteriormente pudiera ser considerado como irrelevante, pero si pensamos en sistematizar un diagnóstico y esquema de la genealogía curricular asociada al ítem para retroalimentar la mejora de la operación del currículo, sería importante tener claridad en dicho aspecto. Otro problema que también se consideró importante a señalar es que en la especificación del ítem no se encuentra descrito el contenido origen que se sistematiza.

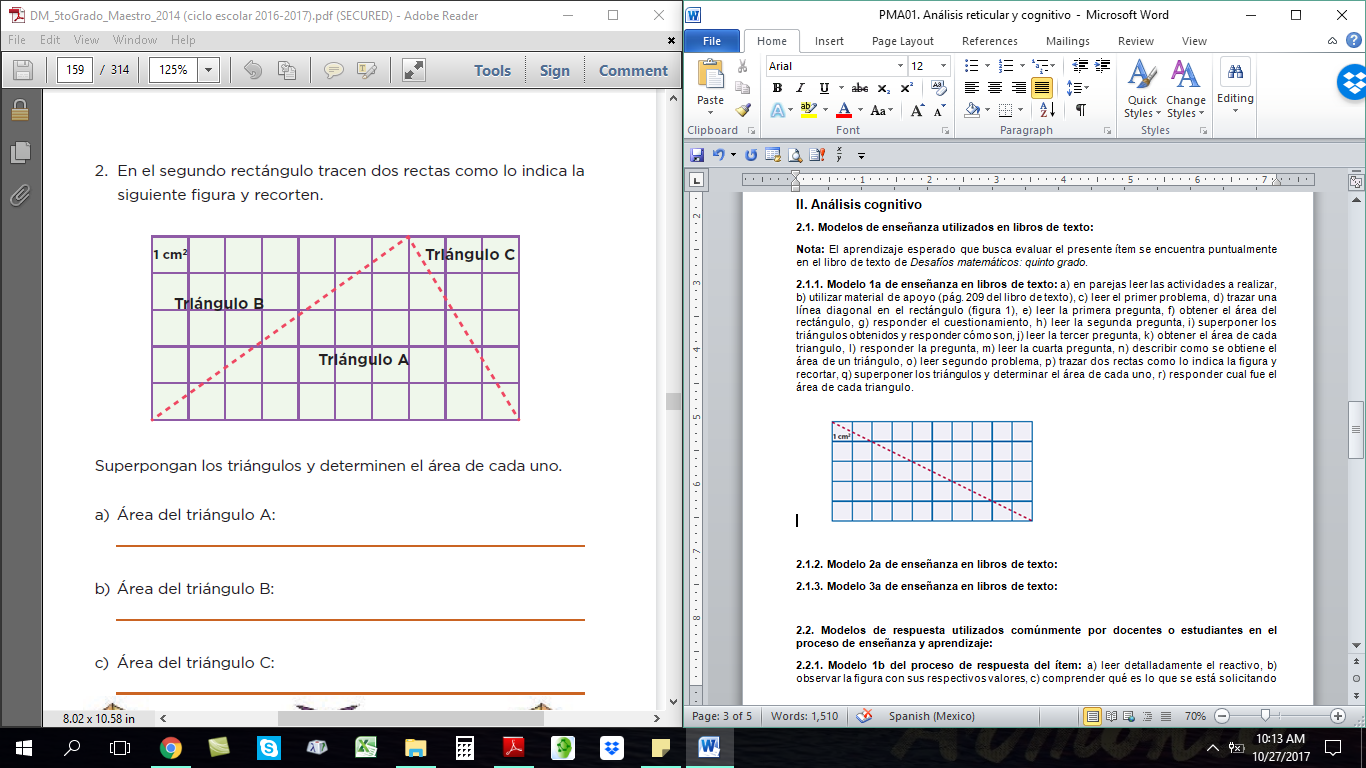
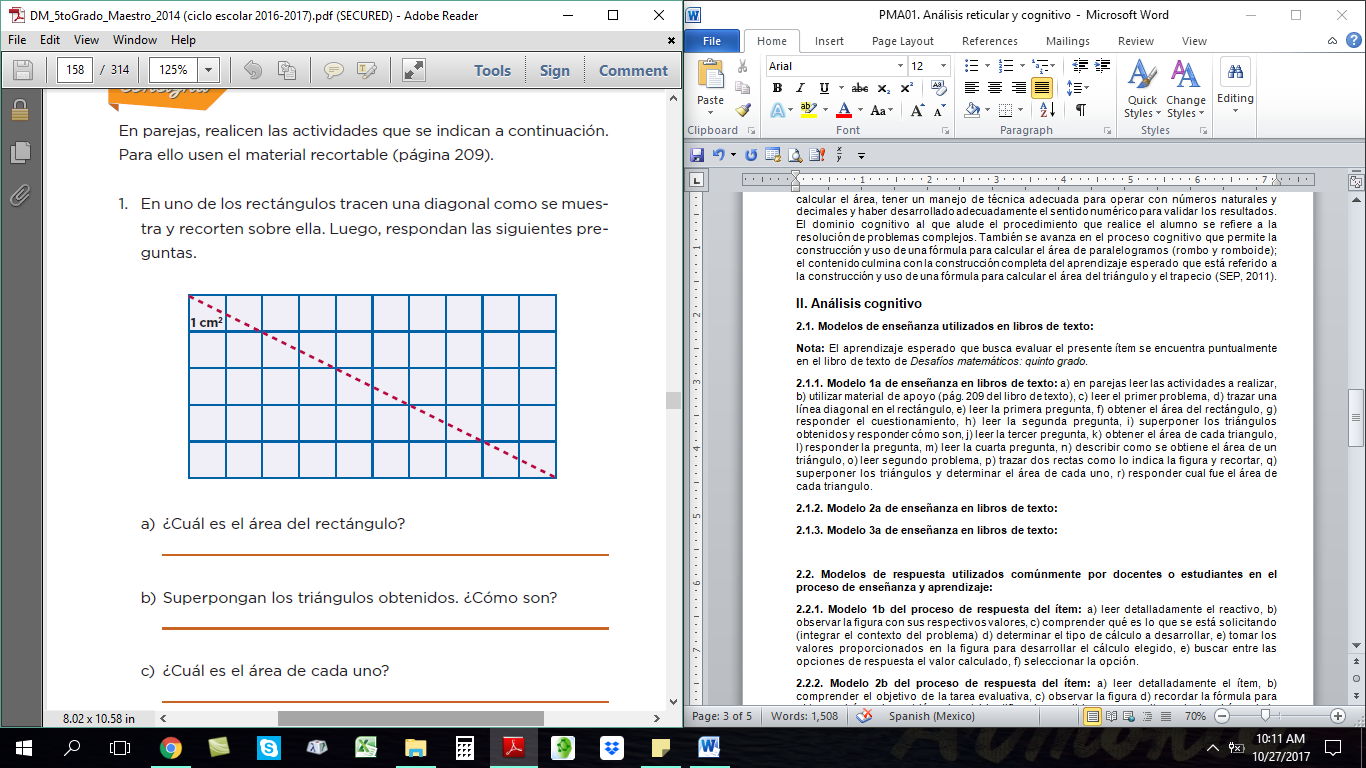
*Figura 11*. Secciones de la retícula de PLANEA ELCE de Matemáticas (06) relacionadas con la especificación del ítem PMA01

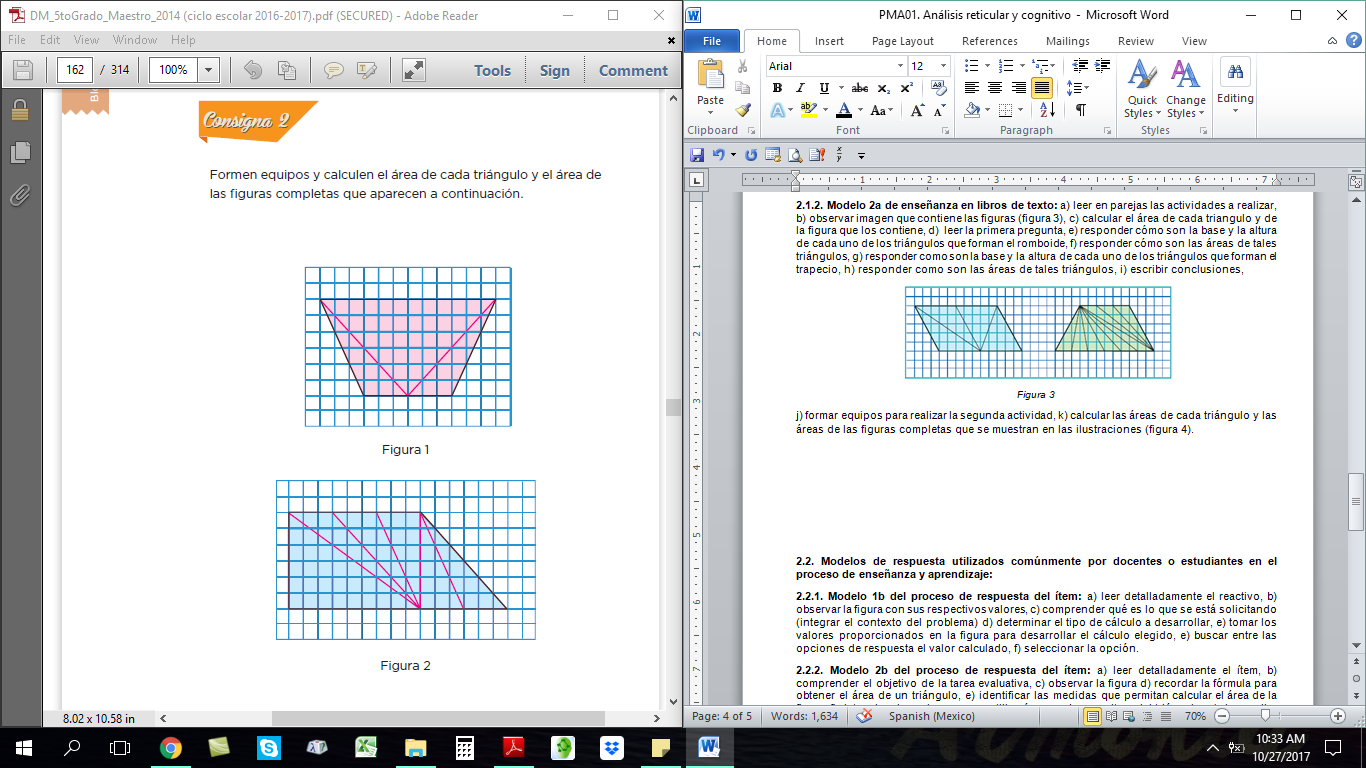
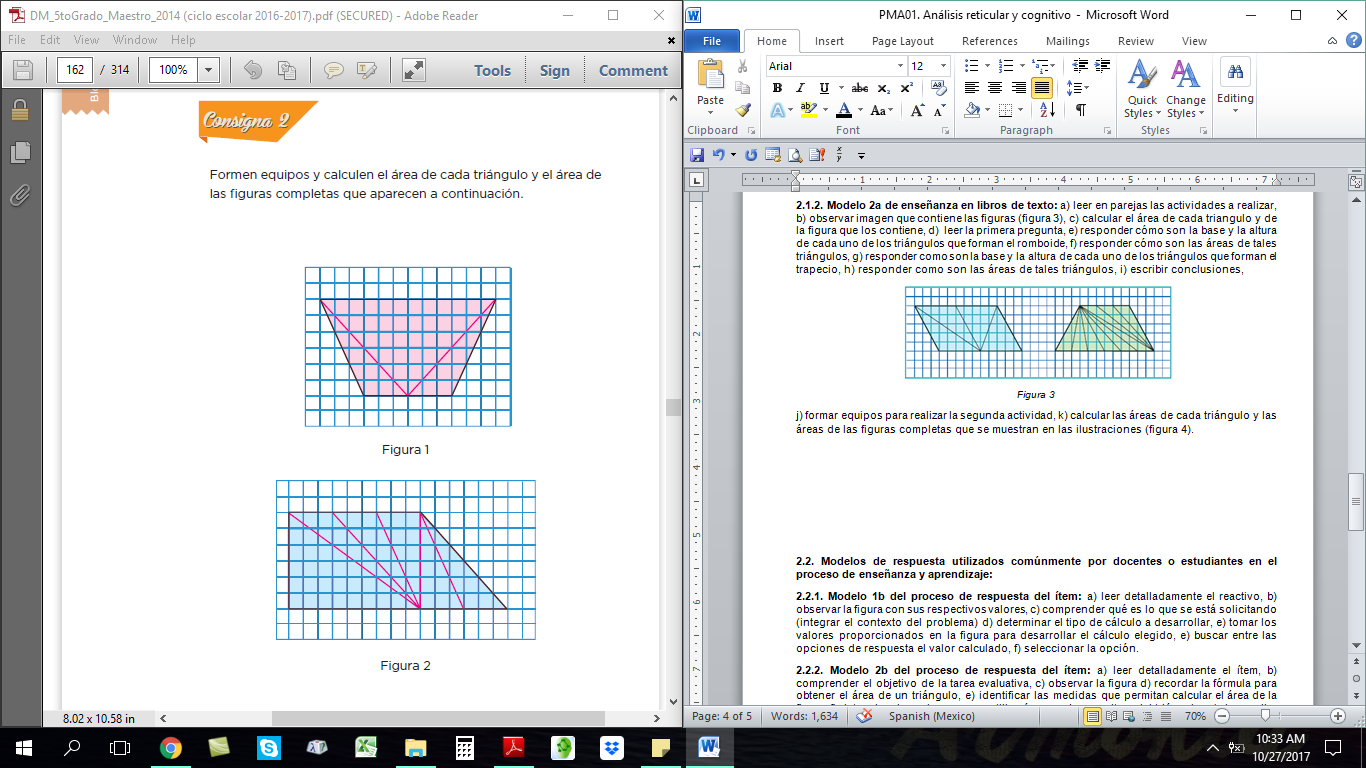
Consecutivamente los expertos analizaron los documentos curriculares oficiales asociados con los ítems en estudio. En el caso del ítem PMA01 se encontró el estándar curricular en la *Guía del maestro* del Plan de estudio 2011 asociado con el *Aprendizaje esperado* descrito en la especificación del ítem (ver Figura 12). Nótese que el estándar curricular presenta una descripción general diferente al *Aprendizaje esperado* en especial a lo referente con el uso de fórmulas.



*Figura 12*. Estándar curricular al que se alinea la especificación del ítem PMA01 de la prueba de Matemáticas de primaria del PLANEA ELCE 2015

De igual forma, los expertos analizaron los libros de texto del maestro como los del alumno para poder diseñar la genealogía curricular el ítem. Se encontró que el ítem se encuentra asociado especialmente con las intenciones didácticas 1 y 2 del contenido 5.3.6. del libro de texto del alumno. En la Figura 13 puede observarse de forma resumida los modelos de enseñanza de las intenciones didácticas en las que se encontró mayor alineación con el ítem PMA01 y su especificación.



**Modelo del proceso de enseñanza de la intención didáctica 1:** a) En parejas leer las actividades a realizar, b) utilizar material de apoyo (pág. 209 del libro de texto), c) leer el primer problema, d) trazar una línea diagonal en el rectángulo (figura 1), e) leer la primera pregunta, f) obtener el área del rectángulo, g) responder el cuestionamiento, h) leer la segunda pregunta, i) superponer los triángulos obtenidos y responder cómo son, j) leer la tercer pregunta, k) obtener el área de cada triangulo, l) responder la pregunta, m) leer la cuarta pregunta, n) describir como se obtiene el área de un triángulo, o) leer segundo problema, p) trazar dos rectas como lo indica la figura y recortar, q) superponer los triángulos y determinar el área de cada uno, y r) responder cuál fue el área de cada triangulo.

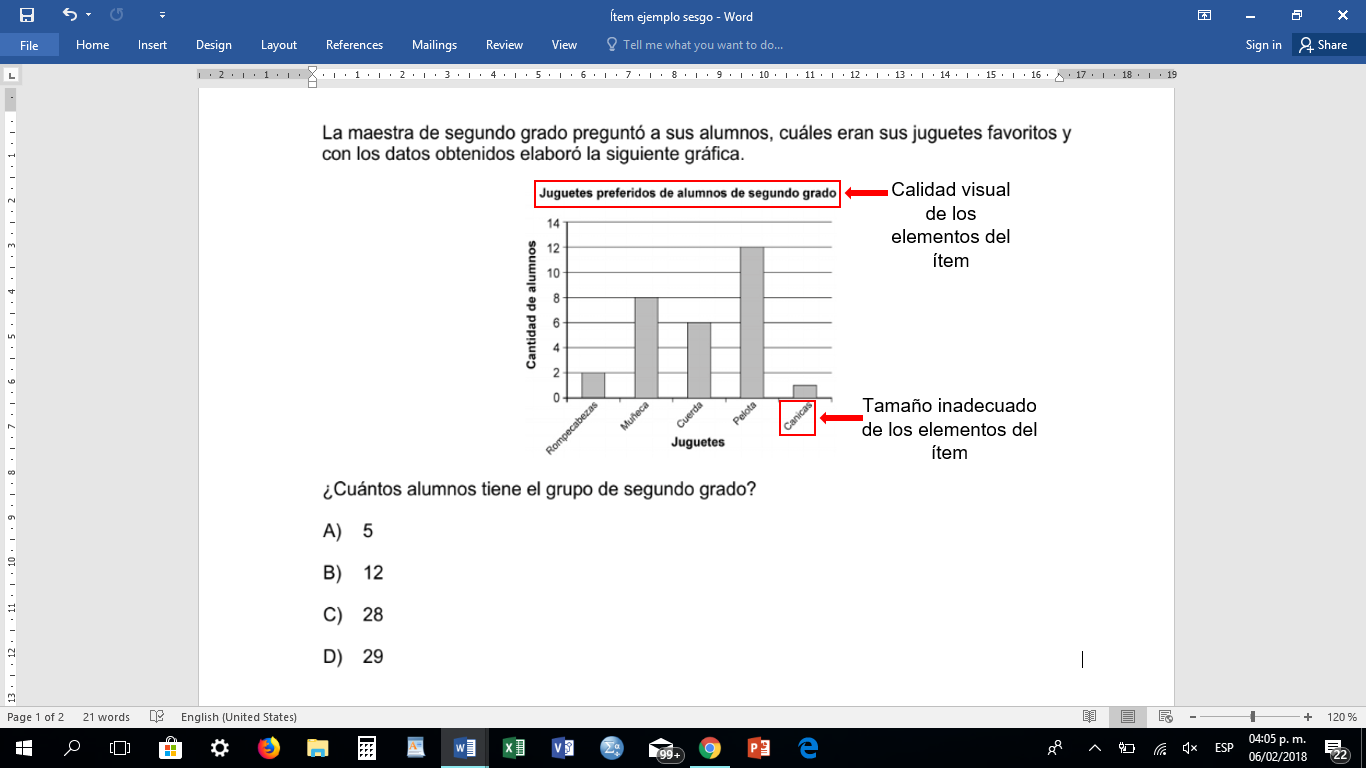
**Modelo del proceso de enseñanza de la intención didáctica 2:** a) Leer en parejas las actividades a realizar, b) observar imagen que contienen las figuras, c) calcular el área de cada triangulo y de la figura que los contiene, d) leer la primera pregunta, e) responder cómo son la base y la altura de cada uno de los triángulos que forman el romboide, f) responder cómo son las áreas de tales triángulos, g) responder cómo son la base y la altura de cada uno de los triángulos que forman el trapecio, h) responder cómo son las áreas de tales triángulos, i) escribir conclusiones, j) formar equipos para realizar la segunda actividad, y k) calcular las áreas de cada triángulo y las áreas de las figuras completas que se muestran en las ilustraciones.

*Figura 13.* Modelos de enseñanza de las Intenciones didácticas 1 y 2 del contenido 5.3.6 de 5to de primaria

Problemas de sesgo

Durante el análisis cognitivo, los expertos detectaron que 12 ítems presentaron problemas de sesgo, correspondiente al 24% total de la prueba. Estos ítems son: PMA01, PMA02, PMA07, PMA08, PMA14, PMA19, PMB01, PMB03, PMB04, PMB10, PMB14, PMB24. Los problemas detectados están relacionados con (1) tamaño inadecuado de los elementos del ítem, (2) calidad visual de los elementos del ítem, (3) uso de conceptos con nivel bajo de inclusión de nivel socioeconómico y/o (4) uso de conceptos asociados con culturas o lugares particulares.

Un inadecuado manejo de los elementos mencionados en el párrafo anterior ocasiona que los sustentantes no puedan participar o no participen por igual en la realización del ítem debido a la presencia de una discapacidad, algún problema del lenguaje o por pertenecer a subgrupos poblacionales (de acuerdo a su género, raza, etnia y/o nivel socioeconómico). Un ejemplo de lo que contempla esta clasificación son los errores detectados en el ítem PMA14 (Figura 14), ya que cuenta con letras pequeñas y borrosas, dificultando la tarea de contestar el ítem a las personas con problemas de visión. Durante el laboratorio cognitivo los sustentantes tuvieron que acercarse a la prueba para leer las letras de la gráfica; por unanimidad manifestaron que las letras eran muy pequeñas. Además, perdieron tiempo descifrando las palabras e inclinaron la cabeza para leer la señalización “Cantidad de alumnos”.



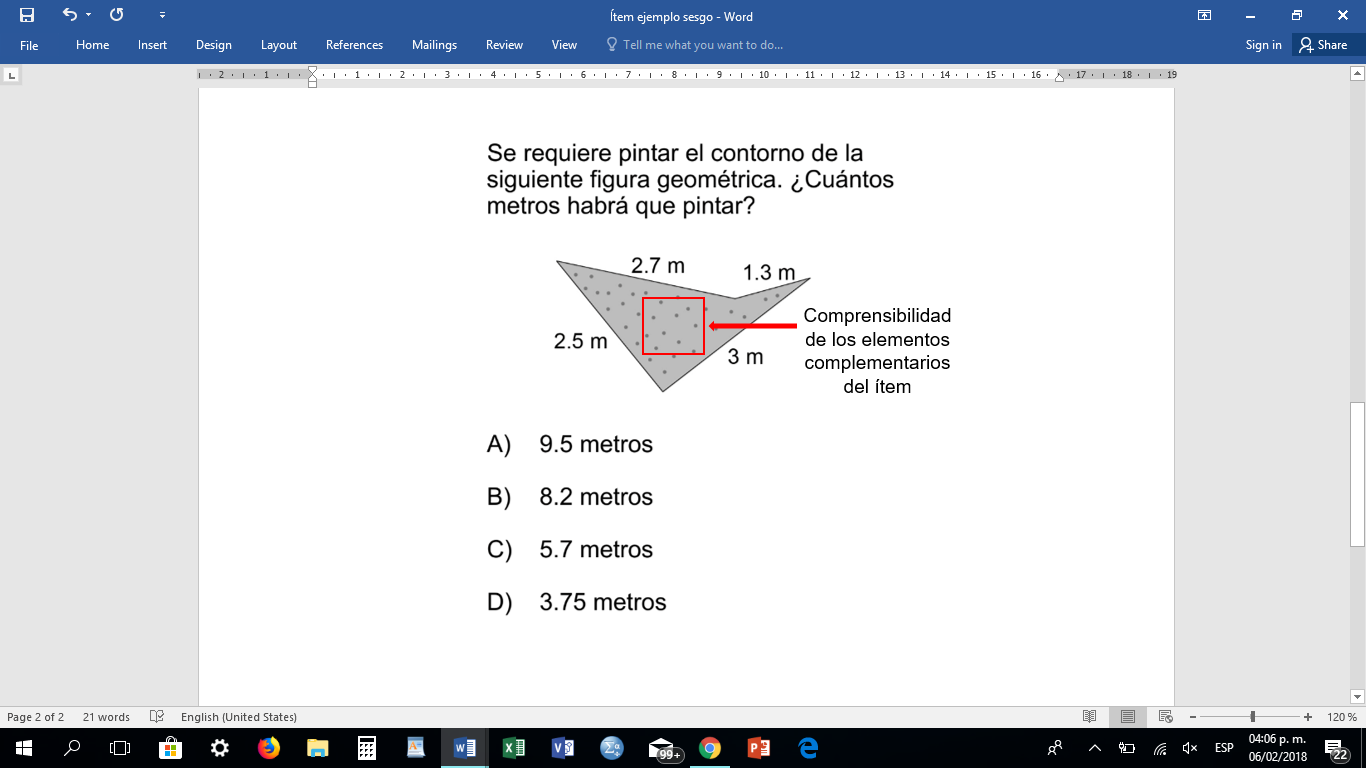
*Figura 14.* Ítem PMA14

Problemas de diseño

Durante el análisis cognitivo se detectaron 31 ítems en esta clasificación, correspondiente al 62% del total de la prueba. Los ítems que entraron en esta clasificación presentaron problemas relacionados con (1) claridad de las instrucciones por falta de comprensibilidad total o parcial de las instrucciones, (2) comprensibilidad de la base del ítem por presentar una base del ítem poco clara o ambigua, uso ambiguo de términos y conceptos en la base del reactivo o por desorganización lógica de las ideas de la base del ítem y (3) comprensibilidad de los elementos complementarios del ítem por ocasionar procesos de distracción en los sustentantes.

Los ítems detectados en esta clasificación son los siguientes: PMA01, PMA02, PMA03, PMA06, PMA08, PMA09, PMA11, PMA12, PMA14, PMA15, PMA16, PMA17, PMA21, PMA22, PMA23, PMA24, PMA25, PMB01, PMB02, PMB03, PMB04, PMB05, PMB08, PMB10, PMB12, PMB13, PMB14, PMB15, PMB22, PMB24, PMB25.

En particular, un claro ejemplo de la clasificación *Problemas de diseño*, es el presentado por el ítem PMB08 (Figura 15), ya que este cuenta con un polígono irregular donde se incluyen elementos que hacen énfasis en el relleno (área) de la misma, como el sombreado y el punteado que presenta. Durante el laboratorio cognitivo los sustentantes manifestaron que el sombreado de la figura y los puntos (que señalaron como abolladuras, puntos, marcas) los hacían pensar que se les solicitaba obtener el área; por lo que tuvieron que regresar a las indicaciones hasta en más de una ocasión para asegurar la tarea solicitada. Algunos sustentantes invirtieron tiempo en encontrarle forma a la figura (pájaro, nave, entre otras), perdiendo tiempo valioso para la realización de la prueba.



*Figura 15.* Ítem PMB08

Resultados del modelamiento matemático de sub-tareas de respuesta

Como resultado de la primera fase de aplicación del método de *modelado matemático de sub-tareas de respuesta* (Embretson, 1983) apoyado con el *análisis de expertos* (Rupp, Templin, y Henson, 2010) se obtuvieron diferentes productos que aportan a la definición integral del modelo cognitivo de los ítems de la prueba de Matemáticas 06 de PLANEA ELCE 2015. Un primer producto de la aplicación de dichos métodos fue el modelo del proceso de respuesta (mencionado anteriormente como sub-tareas) de cada uno de los ítems de la prueba. Dichos modelos se encuentran desarrollados bajo un esquema lineal en términos de los pasos a seguir para responder correctamente los ítems de la prueba. Por ejemplo, como se observó con el ítem PMA01, los expertos definieron que la secuencia de pasos a seguir para contestarlo correctamente es la siguiente: (a) leer detalladamente las indicaciones del ítem, (b) comprender el contexto del problema, (c) comprender el objetivo de la tarea evaluativa, (d) deducir de elementos del contexto el cálculo del área del triángulo, (e) comprender el objetivo de la tarea evaluativa, (f) recordar y representar la fórmula para calcular el área del triángulo, (g) seleccionar y sustituir los valores de la fórmula, (h) aplicar operaciones aritméticas básicas, (i) reconocer el resultado dentro de las opciones de respuesta y (j) seleccionar la opción de respuesta (ver **Tabla 16**).

**Tabla 16.** Esquema del modelo cognitivo de procesos de respuesta del ítem PMA01

|  |  |
| --- | --- |
| **Ítem PMA01** | **Proceso de respuesta** |
|  | Leer detalladamente las indicaciones del ítem |
| Comprender el contexto del problema |
| Comprender el objetivo de la tarea evaluativa |
| Deducir de elementos del contexto el cálculo del área del triángulo |
| Comprender el objetivo de la tarea evaluativa |
| Recordar y representar la fórmula para calcular el área del triángulo |
| Seleccionar y sustituir los valores de la fórmula |
| Aplicar operaciones aritméticas básicas |
| Reconocer el resultado dentro de las opciones de respuesta |
| Seleccionar la opción de respuesta |

Resultados de los protocolos verbales

Durante la aplicación de los reportes verbales, los expertos pudieron constatar que algunas de las hipótesis establecidas durante el análisis inicial de los ítems de la prueba Matemáticas 06 PLANEA ELCE 2015, se confirmaron, pues los sustentantes manifestaron problemas al tratar de resolver las tareas evaluativas. Además, mientras los sustentantes resolvían la prueba con la técnica de pensamiento en voz alta, los expertos detectaron otros problemas de diseño que no habían sido detectados durante el análisis de los ítems; algunos de ellos referentes a problemas de sesgo y/o de diseño.

Un protocolo verbal y concurrente aplicado durante la resolución del ítem PMA01 es un claro ejemplo de lo expresado en este apartado (ver apéndice 12); pues durante el análisis de los ítems por expertos, se consideró que los sustentantes podrían manifestar problemas para comprender la palabra “banderola”. Esta hipótesis quedó respaldada, ya que los sustentantes presentaron problemas para pronunciar dicha palabra y en su mayoría manifestaron desconocer el término; tal fue el caso del sustentante de alto rendimiento P007:

“Leonor va a hacer una…, una bander-ola para su equipo de voleibol…, con las medidas que indica en el dibujo…, 31, 60, 62. ¿Cuánta tela ocupará para el bander-ol…, la…, para la bander-ola…?

En contraste, gran parte de los sustentantes utilizaron la imagen complementaria del ítem para suponer que era una “bandera”. Sin embargo, y típicamente, las banderas presentan medidas distintas a las señaladas en el ítem, son rectangulares y las medidas de base y altura son diferentes para poder obtener el área, por lo que algunos sustentantes se confundieron al seleccionar la base y altura de la figura. Este problema se presentó incluso en sustentantes de alto rendimiento, como es el caso del sustentante P007:

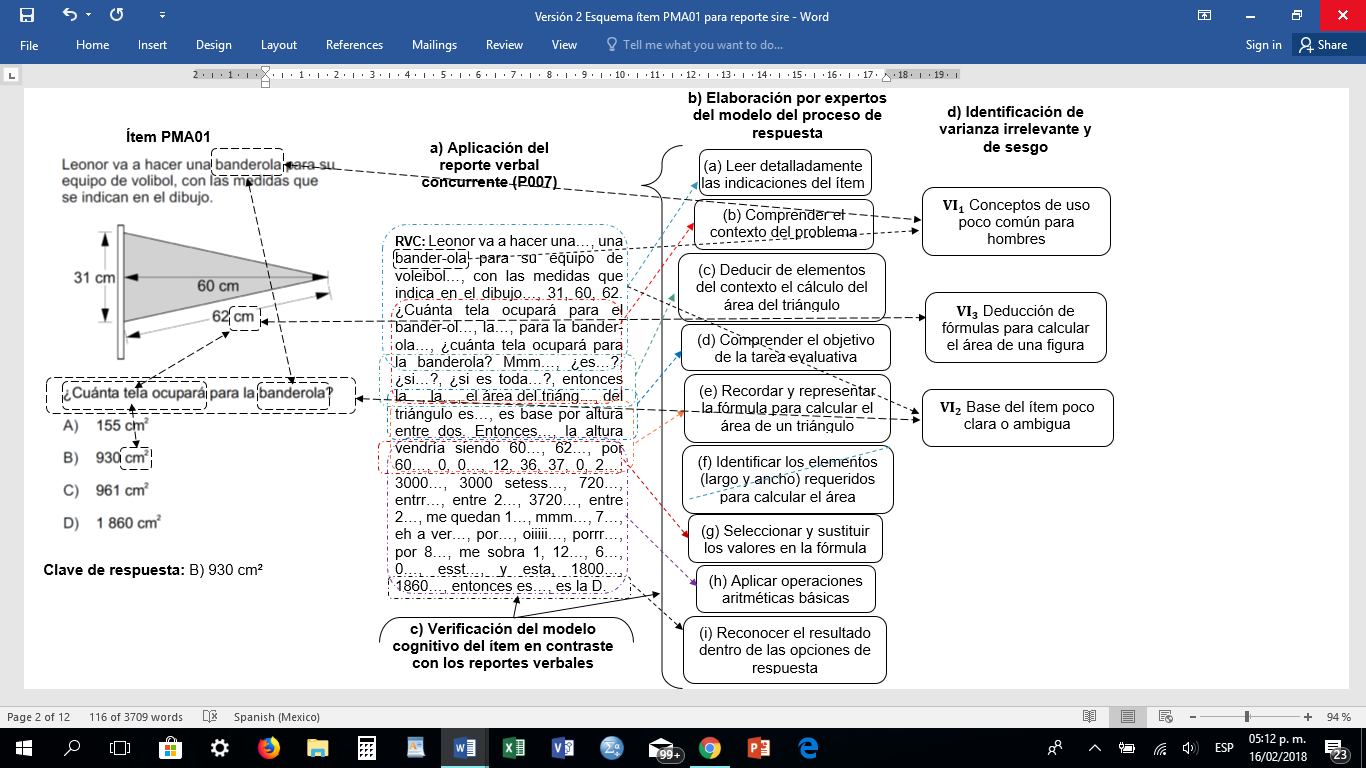
“Leonor va a hacer una…, una bander-ola para su equipo de voleibol…, con las medidas que indica en el dibujo…, 31, 60, 62. ¿Cuánta tela ocupará para el bander-ol…, la…, para la bander-ola…?, ¿cuánta tela ocupará para la banderola? Mmm…, ¿es…?, ¿si…?, ¿si es toda…?, entonces la…, la…, el área del triáng…, del triángulo es…, es base por altura entre dos. Entonces…, la altura vendría siendo 60…, 62…, por 60…, 0, 0…, 12, 36, 37, 0, 2…, 3000…, 3000 setess…, 720…, entrr…, entre 2…, 3720…, entre 2…, me quedan 1…, mmm…, 7…, eh a ver…, por…, oiiiii…, porrr…, por 8…, me sobra 1, 12…, 6…, 0…, esst…, y esta, 1800…, 1860…, entonces es…, es la D.” (ver apéndice 12).

Además de los problemas previamente previstos por los expertos para el ítem PMA01 y manifestados por los sustentantes, se resaltaron otros que no habían sido considerados. Un ejemplo es el caso del sustentante P016, que durante el protocolo retrospectivo comentó que al ver la imagen complementaria no comprendió qué era la barrita (señalando el bordón), ni que tenía que hacer con ella y las flechitas que la señalaban; o el caso del sustentante P008 que mencionó que no sabía por qué le habían dado el dato de 62 cm en la figura complementaria del ítem.

Resultados del análisis de los protocolos verbales con base en los modelos del proceso de respuesta

El análisis de los protocolos verbales es una parte muy importante del proyecto, ya que conjunta los esfuerzos de las diferentes etapas del análisis cognitivo de los ítems. Respecto al trabajo de los expertos, posteriormente al análisis de los protocolos verbales, se confirmaron o adecuaron los modelos de los procesos de respuesta de cada ítem, se comprobaron algunos de los problemas hipotéticos que los sustentantes podían presentar al responder los ítems, se detectaron otros errores de diseño o sesgo que no se habían detectado en los análisis previos, y se diseñó con mayor precisión el esquema que explica la varianza irrelevante que determina el comportamiento del ítem.

En especial, respecto al análisis de los modelos hipotéticos de proceso de respuesta, los expertos se encargaron aplicar técnicas de análisis inductivo-deductivo para verificar el proceso de respuesta mediante los modelos del proceso de respuesta subyacente a los ítems, mismos que fueron previamente determinados por docentes y especialistas en Matemáticas. En la Figura 16, correspondiente al ítem PMA01, se observa el resultado de la labor de los expertos, ya que se presenta el modelo cognitivo sintetizado del ítem, donde se pueden observar tres de los principales problemas de diseño y sesgo que presenta acorde con el modelo EDU. Los problemas señalados se relacionan con las categorías: a) *Problemas de inclusión poblacional,* b) *definición imprecisa del constructo* (referente a la alineación del mismo), y c) *comprensibilidad de la base del ítem*.



*Figura 16*. Diagrama del modelo de un proceso de respuesta erróneo subyacente al ítem PMA01

Definición de las operaciones cognitivas que explican la dificultad de los ítems

El análisis descrito en el apartado anterior *Resultados del análisis de los protocolos verbales con base en los modelos del proceso de respuesta* fue el preámbulo para culminar esta etapa, ya que, por medio del trabajo de análisis realizado por los expertos, se determinó las relaciones entre los ítems (tipos de relación y la magnitud de las relaciones) y los atributos u operaciones cognitivas sustantivas determinadas por los expertos. A través del análisis por pares e inductivo-deductivo se realizó una síntesis de los modelos cognitivos del proceso de respuesta en operaciones cognitivas específicas, que poseyeran un nivel explicativo agudo del comportamiento de los ítems. Un referente son las operaciones que explican primordialmente la dificultad del ítem PMA01: O1 Comprensión de problemas matemáticos contextualizados, O2 Comprensión del Sistema Internacional de Unidades (SIU), O12 Representación del modelo aritmético para calcular el área de cuadriláteros o triángulos, y O5 Aplicación de operaciones aritméticas básicas (ver **Tabla 17**).

**Tabla 17.** Operaciones cognitivas del eje Espacio, forma y medida (Extracto de la tabla original)

|  |  |  |
| --- | --- | --- |
| **Eje** | **No.** | **Operaciones cognitivas** |
| Espacio, forma y medida | O1 | Comprensión de problemas matemáticos contextualizados |
| O2 | Comprensión del Sistema Internacional de Unidades (SIU) |
| O3 | Operación de valores posicionales con números naturales y decimales |
| O4 | Ubicación de una coordenada en el primer cuadrante del plano cartesiano |
| O5 | Aplicación de operaciones aritméticas básicas |
| O6 | Definición de tecnicismos del lenguaje formal de la geometría (p. e., los tipos de ángulos) |
| O7 | Representación viso-espacial de figuras geométricas |
| O8 | Identificación de las características geométricas de los cuadriláteros |
| O9 | Identificación gráfica de tipos de líneas rectas (paralelas, perpendiculares y decantes) |
| O10 | Representación del modelo aritmético de la conversión de unidades de medida |
| O11 | Representación del modelo aritmético para calcular el perímetro de una figura geométrica (triángulo o cuadrilátero) |
| O12 | Representación del modelo aritmético para calcular el área de cuadriláteros o triángulos |
| O13 | Deducción de fórmulas para calcular el área mediante descomposición de figuras geométricas |

De la anterior tabla, emerge la Matriz Q, que es un modelo sintetizado de las operaciones cognitivas específicas de cada uno de los ítems, extraídas mediante el análisis de los expertos a los procesos de respuesta hipotéticos y empíricos. Dentro de la Matriz Q, pueden observarse los ítems alineados al eje que les corresponde; en la parte superior de la tabla se encuentra el código de la operación cognitiva, dentro de la tabla se puntúa con el número 1 si en el ítem interviene la operación cognitiva y con el número 0 si no la presenta. Utilizando como ejemplo el ítem PMA01, pueden observarse en la Tabla 18 las operaciones cognitivas que le corresponden (fila señalada en azul claro); en este caso O1, O2, O5, O12. Posteriormente basta con volver a la **Tabla 17.** *Operaciones cognitivas del eje Espacio, forma y medida* e identificar las operaciones de acuerdo a los códigos adjudicados al ítem en la Matriz Q. Esta Matriz permite tener un acceso ágil a la información relevante del ítem, pues de primera mano puede intuirse la dificultad del mismo, parte del comportamiento del ítem, someterla a otros análisis, entre otras cosas.

**Tabla 18.** Matriz Q (Extracto de la tabla original)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Código de ítem | **O1** | **O2** | **O3** | **O4** | **O5** | **O6** | **O7** | **O8** | **O9** | **O10** | **O11** | **O12** | **O13** |
| PMA05 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| PMB09 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| PMB07 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| PMA06 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| PMA04 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| PMA03 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| PMB06 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| PMB08 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| PMB05 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| PMA01 | **1** | **1** | 0 | 0 | **1** | 0 | 0 | 0 | 0 | 0 | 0 | **1** | 0 |
| PMB01 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| PMB04 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| PMA07 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| PMB02 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| PMA08 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| PMB03 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| PMA02 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Resultados del diagnóstico cognitivo

Resultados de los análisis con el modelo DINA de los parámetros de adivinación, desliz y de las probabilidades de dominio de las operaciones cognitivas por parte de los examinados

En la **Tabla 19** se muestra los resultados de las probabilidades de dominio de las operaciones cognitivas por parte de los examinados a nivel nacional. Para los tres ejes, se puede observar que las probabilidades oscilan entre 0.49 y 0.68. De manera más específica se encuentra que el promedio de la probabilidad de dominio por cada eje temático es 0.59, 0.63 y 0.58 respectivamente. Estas probabilidades de tamaño moderado permiten evidenciar que de manera sistemática, todas las operaciones requerirá un plan de mejoramiento en el dominio por parte de los examinados.

**Tabla 19.** Probabilidades de dominio de las operaciones cognitivas por parte de los examinados en los tres ejes temáticos del PLANEA ELCE (06) 2015 de Matemáticas

|  |  |  |
| --- | --- | --- |
| **Eje** | **Clave** | **Probabilidad de dominio** |
| I. Espacio, forma y medida | O1 | 0.60 |
| O2 | 0.63 |
| O3 | 0.63 |
| O4 | 0.51 |
| O5 | 0.67 |
| O6 | 0.50 |
| O7 | 0.57 |
| O8 | 0.57 |
| O9 | 0.59 |
| O10 | 0.50 |
| O11 | 0.61 |
| O12 | 0.68 |
| O13 | 0.57 |
| II. Manejo de información | O1 | 0.68 |
| O2 | 0.59 |
| O3 | 0.65 |
| O4 | 0.59 |
| O5 | 0.68 |
| O6 | 0.63 |
| O7 | 0.63 |
| O8 | 0.66 |
| O9 | 0.49 |
| O10 | 0.68 |
| III. Sentido numérico y pensamiento algebraico | O1 | 0.58 |
| O2 | 0.59 |
| O3 | 0.65 |
| O4 | 0.52 |
| O5 | 0.57 |
| O6 | 0.59 |
| O7 | 0.65 |
| O8 | 0.57 |
| O9 | 0.51 |
| O10 | 0.61 |
| O11 | 0.63 |
| O12 | 0.57 |
| O13 | 0.50 |

Con lo que respecta a los análisis mediante el modelo DINA, se presenta en la **Tabla 20** los parámetros de adivinación y desliz para el eje *Espacio, forma y medida*, para el parámetro adivinación el intervalo cubre los valores de 0.08 a 0.61 con un valor promedio de 0.33. A su vez, el parámetro deslíz, presenta valores comprendidos entre 0.01 y 0.79 con un valor promedio de 0.25. Por ejemplo, los ítems r31, r25 y r39 tienen valores muy altos en el parámetro de adivinación. Esto indica que los examinados tienen una alta probabilidad de dar una respuesta correcta a ítems, aunque no tengan dominio de la operación. Otros ítems como el r18, r03 y r09 tienen valores muy altos en el parámetro de desliz, indicando que las personas con el dominio de la operación cognitiva del ítem, presentan una baja probabilidad de responder correctamente el ítem.

**Tabla 20.** Parámetros de adivinación y deslíz del eje *Espacio, forma y medida* según el modelo DINA

|  |  |  |  |
| --- | --- | --- | --- |
|  | Item | Adivinación | Deslíz |
| 1 | r27 | 0.30 | 0.30 |
| 2 | r29 | 0.28 | 0.32 |
| 3 | r07 | 0.35 | 0.32 |
| 4 | r08 | 0.61 | 0.06 |
| 5 | r01 | 0.12 | 0.64 |
| 6 | r26 | 0.08 | 0.79 |
| 7 | r03 | 0.07 | 0.43 |
| 8 | r30 | 0.33 | 0.23 |
| 9 | r31 | 0.47 | 0.02 |
| 10 | r32 | 0.34 | 0.02 |
| 11 | r33 | 0.36 | 0.09 |
| 12 | r34 | 0.37 | 0.01 |
| 13 | r04 | 0.37 | 0.22 |
| 14 | r05 | 0.51 | 0.08 |
| 15 | r06 | 0.48 | 0.14 |
| 16 | r02 | 0.24 | 0.35 |
| 17 | r28 | 0.41 | 0.26 |

Por su parte, en la **Tabla 21,** se presentan los parámetros de adivinación y desliz del modelo DINA para el eje *Manejo de información*. Para el parámetro adivinación los valores oscilan entre 0.18 y 0.45 con un valor promedio de 0.30. De igual manera, el parámetro deslíz, comprende valores entre cero y 0.48 con un valor promedio de 0.22.

**Tabla 21**. Parámetros de adivinación y deslíz del eje *Manejo de información* según el modelo DINA

|  |  |  |  |
| --- | --- | --- | --- |
|  | Item | Adivinación | Deslíz |
| 1 | r14 | 0.34 | 0.02 |
| 2 | r35 | 0.22 | 0.24 |
| 3 | r39 | 0.45 | 0.25 |
| 4 | r09 | 0.18 | 0.48 |
| 5 | r10 | 0.26 | 0.35 |
| 6 | r11 | 0.37 | 0.11 |
| 7 | r12 | 0.28 | 0.05 |
| 8 | r13 | 0.38 | 0.00 |
| 9 | r36 | 0.24 | 0.37 |
| 10 | r37 | 0.35 | 0.27 |
| 11 | r38 | 0.27 | 0.27 |

Finalmente, para el eje *Sentido numérico y pensamiento algebraico*, se presnetan los valores de los parámetros adivinación y desliz del modelo DINA. Se encuentra que el parámetro Adivinación fluctúa entre 0.13 y 0.56 con un promedio de 0.28, por su parte el parámetro deslíz se encuentra entre cero y 0.53 con un promedio de 0.25.

**Tabla 22**. Parámetros de adivinación y deslíz del eje *Sentido numérico y pensamiento algebraico* según el modelo DINA

|  |  |  |  |
| --- | --- | --- | --- |
|  | Item | Adivinación | Deslíz |
| 1 | r20 | 0.27 | 0.14 |
| 2 | r21 | 0.33 | 0.04 |
| 3 | r22 | 0.45 | 0.09 |
| 4 | r25 | 0.56 | 0.08 |
| 5 | r40 | 0.13 | 0.43 |
| 6 | r41 | 0.15 | 0.38 |
| 7 | r44 | 0.17 | 0.29 |
| 8 | r15 | 0.13 | 0.05 |
| 9 | r18 | 0.25 | 0.44 |
| 10 | r23 | 0.37 | 0.17 |
| 11 | r24 | 0.30 | 0.15 |
| 12 | r16 | 0.23 | 0.45 |
| 13 | r17 | 0.23 | 0.16 |
| 14 | r19 | 0.50 | 0.00 |
| 15 | r42 | 0.23 | 0.43 |
| 16 | r43 | 0.24 | 0.53 |
| 17 | r45 | 0.20 | 0.38 |
| 18 | r46 | 0.36 | 0.21 |

De manera general, por ejemplo, los ítems r31, r25 y r39 tienen valores muy altos en el parámetro de adivinación. Esto indica que los examinados tienen una alta probabilidad de dar una respuesta correcta a ítems, aunque no tengan dominio de la operación. Otros ítems como el r18, r03 y r09 tienen valores muy altos en el parámetro de desliz, indicando que personas con el dominio de la operación cognitiva del ítem, presentan una baja probabilidad de responder correctamente el ítem.

Resultados del diagnóstico cognitivo de los estudiantes de 6to de primararia

Descriptivos del dominio de los atributos del total de la población (en el discurso hacer infasis de las fortalezas y debilidades) y descriptivos a nivel nacional por cada atributo.

Discusión: JCPM, GDRR, RVL

Conclusiones: JCPM, GDRR, RVL

Referencias: RVL

Pérez-Morán, J. C., Vázquez-Lira, R., y Rojas, G. (Coords.) (2017). Análisis de bases de datos y de validez de pruebas y de los patrones de respuesta PLANEA. Ciudad de México: INEE.

1. Libro de texto del alumno: Desafíos matemáticos. Libro para el alumno. 6° Matemáticas. Secretaría de Educación Pública. URL: <http://libros.conaliteg.gob.mx/content/restricted/libros/carrusel.jsf?idLibro=1269>

   Libro de texto del maestro: Desafíos matemáticos. Libro para del maestro. 6° Matemáticas. Secretaría de Educación Pública. URL: <http://libros.conaliteg.gob.mx/content/restricted/libros/carrusel.jsf?idLibro=1270>

   Secretaría de Educación Pública (2011). Programas de estudios 2011. Guía para el maestro. Educación Básica. Primaria. Sexto grado. México. SEP. URL: <http://edu.jalisco.gob.mx/cepse/sites/edu.jalisco.gob.mx.cepse/files/sep_2011_programas_de_estudio_2011.guia_para_el_maestrosexto_grado.pdf> [↑](#footnote-ref-1)