**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).

Descripción de los modelos

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, 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 , sólo 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

**Método:**

A partir del Enfoque Sistémico de Diseño Cognitivo (ESDC) propuesto por Embretson (1994), y retomando la perpectiva top-down para el diseño y valdiación de pruebas cognitivas (Bejar, 2002, 2010, Gorin y Embretson, 2013 y Messick, 1989b), se desarrolló el modelo metodológico del presente estudio. El énfasis principal siempre giró en torno a la obtención de evidencia que sustentara la validez sustantiva del constructo, es decir, que corroborara la fidelidad estructural del modelo cognitivo subyacente a la prueba de Matemáticas de primaria (06) del PLANEA ELCE 2015. La Tabla 1 presenta el modelo metodológico propuesto para alcanzar los objetivos del presente estudio.

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.

**Tabla 1.** Modelo metodológico del estudio

|  |  |  |
| --- | --- | --- |
| **Fases** | **Etapas** | **Actividades** |
| **Fase I**  **Diseño de los estudios cognitivos** | 1.1. Selección de las técnicas para el análisis cognitivo | -Determinar el conjunto de ítems para su análisis, verificar sus áreas de membresía o dominio y analizar sus características particulares de diseño.  -Definir los métodos y técnicas específicas para el análisis cognitivo con base en las características de los ítems y de la población objetivo de la prueba. |
| 1.2. Diseño de los estudios cognitivos | -Adaptar el conjunto de ítems y tareas evaluativas para el piloteo de las técnicas cognitivas.  -Definir los procedimientos de operación de los estudios cognitivos.  -Determinar el tipo de herramientas, materiales e instrumentos tecnológicos requeridos para la captura de los datos del análisis cognitivo.  -Adaptar un laboratorio acorde a las necesidades específicas del estudio cognitivo. |
| 1.3. Análisis cognitivo-reticular de los ítems | -Capacitar a expertos en el análisis cognitivo-reticular de los ítems.  **-Elaborar mediante técnicas de análisis cognitivo-reticular la genealogía curricular de los ítems.\***  **-Evaluar mediante expertos la congruencia y alineación de los ítems de la prueba de Matemáticas 06 del PLANEA ELCE con sus especificaciones, la retícula, los libros de texto y los planes de estudio.\*** |
| 1.4. Análisis por expertos de los procesos subyacentes a los ítems | -Capacitar a expertos en el método de modelado matemático de sub-tareas de respuesta  -Aplicar a expertos técnicas de pensamiento en voz alta concurrentes y retrospectivas.  -**Modelar desde la perspectiva de expertos en Matemáticas los procesos posibles de respuesta subyacentes a los ítems**.\*  -Estructurar los sistemas de categorías de los protocolos verbales y modelos cognitivos para su aplicación. |
| **Fase II**  **Piloteo y aplicación de protocolos verbales** | 2.1. Piloteo de protocolos verbales | -Establecer los criterios de selección de los estudiantes participantes del estudio piloto y del grupo de participantes del estudio cognitivo.  -Seleccionar y capacitar a los participantes del estudio piloto.  -Pilotear, probar y ajustar los protocolos, sistemas de categorías y materiales del laboratorio cognitivo. |
| 2.2. Selección del grupo de participantes | -Seleccionar al grupo de participantes del estudio cognitivo.  -Confirmar el consentimiento informado y recabar datos de identificación.  -Establecer cronograma de actividades y citas con los participantes del estudio cognitivo. |
| 2.3. Aplicación en forma de los protocolos verbales | -Entrenar a los participantes del estudio cognitivo.  -Aplicar el estudio cognitivo a los participantes.  -Recopilar la información obtenida durante el estudio cognitivo.  **-Obtener reportes verbales de los procesos de respuesta de estudiantes ante los ítems de las pruebas de Matemáticas 06 del PLANEA ELCE.\*** |
| **Fase III**  **Desarrollo y definición del modelo cognitivo** | 3.1. Análisis de los datos obtenidos durante los estudios cognitivos | -Capacitar a expertos en la verificación del proceso de respuesta de los examinados en contraste con los modelos del proceso de respuesta subyacente a los ítems definidos por docentes y especialistas en Matemáticas.  -Verificar los reportes verbales de los estudiantes participantes con base en los modelos del proceso de respuesta subyacente a los ítems definidos por docentes y especialistas en Matemáticas.  **-Elaborar los modelos sintetizados de los procesos de respuesta subyacentes a los ítems.**\* |
| 3.2. Desarrollo y definición del modelo cognitivo de la prueba | -Determinar la cantidad y el tipo de relaciones entre los ítems de la prueba y los atributos u operaciones cognitivas sustantivas determinadas por los investigadores con base en la síntesis de los resultados de los estudios reticulares y cognitivos.  **-Elaborar la matriz Q de la prueba con base en las operaciones cognitivas sustantivas determinadas *a-priori*.\***  -**Analizar y evaluar el diseño de los ítems** **para identificar posible varianza irrelevante o sesgo**.\* |
| **Fase IV**  **Aplicación del análisis componencial** | 4.1 Revisión de la estructura interna bajo el modelo de *redes nomológicas* | **-Calibrar con la aplicación del modelo de la Teoría Clásica de los Test (TCT) los ítems de la prueba.\***  **-Analizar las características psicométricas de los distractores de los ítems.\***  **-Analizar la estructura interna bajo el *modelo de redes nomológicas* (aplicación del análisis factorial confirmatorio).\*** |
| 4.2 Revisión de la estructura del *modelo cognitivo* de la prueba | -Elegir y aplicar los modelos psicométricos componenciales (MPC) acordes a las características del modelo cognitivo estructurado.  **-Analizar la validez de la matriz Q planteada.\***  -Considerar posibles mejoras y reconfiguraciones en el diseño de la matriz Q. |
| 4.3 Interpretación de los resultados del diagnóstico cognitivo de los examinados | **-Asignar puntuaciones a los examinados con base en el diagnóstico cogitivo de las operaciones definidas en la matriz Q.\***  **-Explorar resultados del diagnóstico cognitivo a nivel nacional, por estado, sexo, modalidad, tipo de servicio u otra agrupación que se considere relevante para su uso a nivel escuela.\*** |
| \*Actividades fundamentales, relacionadas con los objetivos del estudio. | | |

Una ventaja colateral de trabajar con la prueba de Matemáticas de primaria (06) del PLANEA ELCE 2015 es que se trata de una disciplina que ha sido 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 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) con 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, de las diferentes vertientes para el desarrollo y aprendizaje de los componentes o atributos de la prueba y de 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 un análisis reticular, acompañado de una revisión exhaustiva de los libros de texto del alumno y del maestro, así como 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

Finalmente, se aplicaron técnicas de pensamiento en voz altas (concurrentes y retrospectivas) con los expertos para identificar los procesos de respuesta asociados a cada ítem. Al aplicar el *modelado matemático de sub-tareas de respuesta* se realizaron diversas actividades: primero, mediante un proceso inductivo-deductivo los expertos identificaron y categorizaron los procesos y atributos cognitivos subyacentes a los ítems, apoyándose en lo reportado durante su propia ejecución ante los ítems, en los descriptores de los rubros de los conocimientos, habilidades previos y procesos de respuesta declarados en las especificaciones de los ítems de la prueba, así como en el sistema de categorías de procesos de respuesta típicos utilizados por estudiantes de educación básica ante ítems de Matemáticas (Pérez-Morán, 2014).

En la **Tabla 4** se puede observar tres modelos de respuesta hipotéticos del ítem PMA01 de PLANEA ELCE 06 elaborados por expertos. Nótese que los tres modelos del proceso presentan diferentes niveles de granulación, sin embargo, el primer modelo contempla un proceso que fue muy poco tomado en cuenta como importante para contestar correctamente el ítem: la comprensión de problemas matemáticos contextualizados.

**Tabla 4**. 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. |

Se obtuvieron reportes verbales de los procesos de respuesta empleados por un grupo de estudiantes evaluados con los ítems de la prueba, mediante la aplicación de protocolos de pensamiento en voz alta con técnicas concurrentes y retrospectivas (Ericsson y Simon, 1984, 1993; Leighton, 2009; Leighton y Gierl, 2007b). Además, se realizaron análisis del sendero de la vista (*eye-tracking*; Snow y Lohman, 1989; Sternberg, 1977), y se analizó también las latencias de respuesta (Fredericksen, 1980; Posner, 1978; Posner y Rogers, 1978).

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 tomó en cuenta las características del formato de los ítems en estudio y las características propias de los participantes.

Los criterios de selección de los participantes de 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 tanto, se estableció como criterios de selección el uso de variables tales como el rendimiento escolar, grado educativo y la recomendación del profesor. Además, se estableció como restricción que el 50% de los participantes tenían que ser mujeres.

La selección de estudiantes expertos se realizó tomando como punto de referencia un promedio mayor a 8.5 y el reconocimiento de sus profesores como estudiantes sobresalientes en el dominio de las Matemáticas. Por su parte, los estudiantes novatos consideraban un promedio mayor a 6.0 pero menor a 8.0 y que fueran referidos por los 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. 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; también 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* (ver **Apéndice 4**).

*L*a matriz Q señalando las operaciones cognitivas sustantivas subyacentes a la ejecución en cada ítem contenido en la prueba, fue elaborada a partir del método de *modelado matemático de sub-tareas de respuesta* propuesto por Embretson (1983), en conjunto con el *análisis de expertos* en el contenido evaluado (Rupp, Templin, y Henson, 2010) y los resultados obtenidos de los protocolos verbales. Para ello, 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.

La implementación de estas técnicas complementarias a las técnicas cognitivas, se considera de gran ayuda para obtener información sobre 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 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.

Como primera actividad de la Etapa 3.1 de la Fase III, se capacitó a expertos en la verificación del proceso de respuesta de los examinados con base en los modelos del proceso de respuesta subyacente a los ítems definidos por docentes y especialistas en Matemáticas. Para dicha actividad se trabajó de forma directa con datos reales aplicando técnicas de análisis inductivo-deductivo. En la Figura 3 se puede observar el proceso analítico realizado en el ítem PMA01 de la prueba de Matemáticas de sexto de primaria del PLANEA ELCE 2015. Nótese que, como resultado de la verificación inductiva-deductiva de los modelos hipotéticos del proceso de respuesta subyacente a los ítems propuestos por los expertos con los reportes verbales de los estudiantes, se desarrolló el modelo cognitivo sintetizado del ítem.



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

Además de elaborar los modelos sintetizados de los procesos de respuesta subyacentes a los ítems, los expertos analizaron y evaluaron el diseño de los ítems para identificar posible varianza irrelevante o problemas de sesgo. Al igual que el proceso anaílitico, en la Figura 3 se pueden observar tres de los principales problemas de diseño y sesgo en el ítem PMA01 identificados por los expertos. En el EDU, los problemas de diseño y sesgo señalados corresponden 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* (ver **Apéndices 7 y 8**).

Etapa 3.2. Desarrollo y definición del modelo cognitivo de la prueba

Para la última etapa de la Fase III se buscó en especial determinar la cantidad y el tipo de relaciones entre los ítems de la prueba y los atributos u operaciones cognitivas sustantivas determinadas por los expertos, y con base en ello, se elaboró la matriz Q de la prueba. De igual forma mediante un análisis inductivo-deductivo y por pares de expertos se sintetizaron los modelos cognitivos del proceso de respuesta en operaciones cognitivas de menor granulo que tuvieran un nivel explicativo alto del comportamiento de los ítems. De tal forma que, si tomamos el ejemplo del ítem PMA01, las operaciones que explican mayormente su dificultad son: 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 Figura 4). Dicho procedimiento se realizó para cada uno de los ítems de la prueba analizada (ver **Apéndice 7**).



*Figura 4*. 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

En la Fase IV se llevaron a cabo las actividades requeridas para revisar las características de calidad técnica de la prueba en estudio, aplicando el modelo cognitivo-componencial DINA tomando en cuenta la matriz Q diseñada por los expertos en los estudios cognitivos, evaluando mejoras del diseño de dicha matriz Q, y explorando los resultados por individuo y por diferentes tipos de agrupaciones de los datos con base en las operaciones cognitivas obtenidas en los estudios cognitivos. Para el logro de dichas actividades se definieron tres etapas: Etapa 4.1 Revisión de la estructura interna bajo el modelo de redes nomológicas; etapa 4.2 Revisión de la estructura del modelo cognitivo de la prueba; y etapa 4.3 Interpretación de los resultados de los examinados.

Etapa 4.1. Revisión de la estructura interna bajo el modelo de redes nomológicas

Previo a la aplicación de los modelos componenciales, se realizó un análisis de la calidad técnica de los ítems de Matemáticas del PLANEA ELCE 2015. Dicho análisis consistió en verificar la calibración mediante la aplicación del enfoque de la Teoría Clásica de los Tests (TCT) y un análisis de la estructura interna de la prueba mediante el Análisis Factorial Exploratorio (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.

La calibración de la prueba con la TCT se implementó con la paquetería CTT del programa libre R 2.15.1. (Ihaka, R. y Gentleman, R., 1996). La finalidad del análisis psicométrico de los ítems de la prueba de matemáticas de sexto de primaria del PLANEA ELCE con la TCT es calibrarlos y estimarlos a la luz de los *estándares* clásicos de calidad técnica. Los indicadores psicométricos que puntualmente se analizaron son el *índice de dificultad*, el *índice de discriminación* (correlación biserial ítem-test) y el *coeficiente de consistencia interna* para la prueba y si se elimina un ítem (**α** de Cronbach). El procedimiento para la obtención de los indicadores psicométricos mencionados consistió de cuatro ecuaciones principales. La primera que se requirió para la obtención del *índice de dificultad* del reactivo fue la ecuación (**1**):

 (**1**)

En esta ecuación (**6**) pi es el *índice de dificultad* del reactivo,*A*i es la cantidad de aciertos en el reactivo y *N*ies la cantidad de aciertos más la cantidad de errores en el reactivo. La ecuación (**2**) que se utilizó para obtener el *índice de discriminación* (altos-bajos) fue:

 (**2**)

En esta ecuación (**7**) *D*i es el *índice de discriminación* del reactivo i, *GA*i es la cantidad de aciertos del reactivo i del 27% de examinados que obtuvieron las puntuaciones más altas en el examen,*GB*i la cantidad de aciertos del reactivo i del 27% de examinados que obtuvieron las puntuaciones más bajas en el examen, y N es la cantidad de personas en el grupo más cuantioso (*GA*io *GB*i). La ecuación (**3**) que se utilizó para obtener el *coeficiente de correlación puntual-biserial* fue:

 (**3**)

En esta ecuación (**3**)  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, n1 es la cantidad de casos que respondieron correctamente al ítem, n0 es la cantidad de casos que respondieron incorrectamente al ítem y n es igual a n1 + n0. Por último, la ecuación (**4**) que se utilizó para obtener el *coeficiente de consistencia interna* (**α** de Cronbach) del instrumento fue:

 (**4**)

En esta ecuación (**4**) *α* es el coeficiente de consistencia interna, n 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 AFE mediante la paquetería *nFactors* y *psych* del programa libre R 2.15.1. (Ihaka, R. y Gentleman, R., 1996). Los indicadores de ajuste absoluto que se emplearon corresponden a la raíz de los residuos cuadráticos promedios estandarizados (SRMR) y la raíz del error cuadrático promedio de aproximación (RMSEA). Tanto el SRMR como el RMSEA deben ser inferiores a 0,05 para obtener 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 (ver **Apéndice 1**). 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.

Etapa 4.2. Revisión de la estructura del modelo cognitivo de la prueba

El Modelo de Diagnóstico Cognitivo (MDC) seleccionado es el modelo DINA (Junker y Sijtsma, 2001). En lo que respecta a la aplicación del modelo DINA, se utilizó la librería CDM implementada en el programa libre R 2.15.1. (Ihaka, R. y Gentleman, R., 1996). Los parámetros de adivinación y desliz del modelo DINA permiten interpretar la calidad técnica de cada ítem en función de las operaciones cognitivas de los procesos de respuesta (de la Torre, 2009).

Etapa 4.3. Interpretación de los resultados de los examinados

Para el diagnóstico cognitivo por examinado y por estrato o agrupación de resultados, se optuvieron de la paquetería CDM las cadenas o vectores de probabilidades del dominio de los atributos para cada estudiante. Estas probabilidades son transformadas a ceros (0) y unos (1) para facilitar la interpretación diagnóstica del dominio de los atributos por parte de los estudiantes (ver **Apéndice 10**).

Metodología: GDRR y JCPM

Resultados: GDRR y JCPM

Discusión: JCPM, GDRR, RVL

Conclusiones: JCPM, GDRR, RVL

Referencias: RVL

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)