

MODELOS DE DIAGNÓSTICO COGNITIVO: FUNDAMENTOS, DIDÁCTICA Y APLICACIONES EN CIENCIAS DEL COMPORTAMIENTO

COGNITIVE DIAGNOSIS MODELS (CDMS): THEORY, DIDACTICS AND APPLICATION IN BEHAVIORAL SCIENCES

Recibido: 14 de noviembre de 2016 | Aceptado: 25 de marzo de 2017

DOI: 10.22199/S07187475.2017.0001.00005

DIEGO ARMANDO **LUNA** BAZALDUA ¹,

¹· LYNCH SCHOOL OF EDUCATION, BOSTON COLLEGE, Boston, United States of America

RESUMEN

Los modelos de diagnóstico cognitivo representan una alternativa multidimensional en psicometría cuando se asume que las variables observables y latentes son categóricas. El presente artículo hace una revisión de la literatura sobre los modelos originales, así como aplicaciones y avances recientes en investigación psicométrica en este tema. Además, se incluye una revisión de software estadístico utilizado para la estimación de los modelos y ejemplos de su aplicación en ciencias del comportamiento y de la salud. Ejemplos didácticos con datos simulados demuestran la capacidad de los modelos para clasificar a las personas de acuerdo a sus atributos latentes.

PALABRAS CLAVE: Modelos de diagnóstico cognitivo, Matriz Q, Psicometría, Clasificación, Variables latentes.

ABSTRACT

CDMs represent a multidimensional choice in psychometrics when observable and latent variables are believed to be categorical. The study presents a thorough revision of the literature about original models, as well as its applications and current developments in psychometric research. Additionally, a revision of statistical software to estimate models and examples of applications in behavioral sciences and health care is also presented. Didactic examples based on simulated data show the capability of models to classify people according to their latent attributes.

KEY WORDS: CDMs, psychometrics, latent variables.

INTRODUCCIÓN

El presente texto tiene por objetivo dar una presentación general sobre los modelos de diagnóstico cognitivo¹ (CDM, por su nomenclatura en inglés) como un área de investigación en psicometría, con el fin de familiarizar al investigador aplicado en ciencias sociales y del comportamiento con esta aproximación que puede serle útil en su investigación. En la primera sección de este manuscrito se presenta un panorama general en torno a los CMD, seguida de una sección donde se presenta un estudio de simulación utilizando R (R Core Team, 2012), software libre de licencia. En el apéndice se incluye un tutorial acerca del uso de dicho software estadístico para la estimación de este tipo de modelos.

Es importante enfatizar que CDM es un término genérico que se refiere un conjunto de modelos psicométricos que representan un área de creciente investigación, sobresalen de otras aproximaciones psicométricas por su capacidad de integrar una evaluación estandarizada dentro de un modelo psicométrico multidimensional ligado a teoría cognitiva (Rupp, 2007; Rupp & Templin, 2008).

Desde el punto de vista matemático, los CDM se enfocan en analizar patrones de respuestas a reactivos en una prueba utilizando variables categóricas latentes de los examinados con especificaciones en torno a la relación entre las variables latentes requeridas para contestar cada reactivo (Templin & Henson, 2006). Así, el propósito principal de los CDM es la clasificación de los examinados en clases latentes dicotómicas o politómicas – usualmente referidas como atributos, habilidades, o estados de conocimiento– determinadas por vectores de indicadores binarios para cada atributo latente en cada

reactivo (Chiu & Douglas, 2013; de la Torre & Douglas, 2004).

Desde el punto de vista psicológico, los CDMs permiten contrastar la ejecución de los participantes en una prueba estandarizada a partir de un conjunto predefinido de atributos psicológicos asumidos como requeridos para contestar correctamente cada reactivo, de modo tal que los examinados puedan obtener retroalimentación de acuerdo a su desempeño y se puedan estimar parámetros para cada reactivo en la prueba (Rupp, Templin, & Henson, 2010). Dependiendo del objetivo de la evaluación, los CDM se pueden utilizar para determinar si el examinado ha dominado un conjunto de habilidades (Rupp & Templin, 2008; Huo & de la Torre, 2014) o si presenta un conjunto de signos o síntomas relacionados con la presencia de un trastorno clínico (Templin & Henson, 2006).

Los CDM resultan particularmente útiles cuando las respuestas observadas en una prueba o inventario se califican como dicotómicas (por ejemplo, reactivos de opción múltiple calificados como como correctos o incorrectos, listas de cotejo que indican la presencia o ausencia de determinados signos y síntomas, etcétera), y cuando existe una teoría psicológica que asume a los constructos latentes como categóricos (por ejemplo, habilidades dominadas o no dominadas, o un rasgo de personalidad definido como presente o ausente). A pesar de que existen CDM que permiten la definición de los atributos latentes como politómicos, la mayoría de la investigación en este tema se ha centrado en el desarrollo de modelos para atributos latentes dicotómicos (Rupp, Templin, & Henson, 2010).

En los siguientes párrafos, la nomenclatura para los examinados queda denotada con subíndices $i = 1, \dots, N$, para los reactivos con $j = 1, \dots, J$, y para el número de atributos latentes con $k = 1, \dots,$

1. Otros nombres encontrados en la literatura para estos modelos son: modelos de clasificación diagnóstica, modelos de clases latentes de múltiple clasificación, modelos psicométricos cognitivos, modelos de respuesta latente, modelos de clases latentes restringidas, y modelos de teoría de respuesta al reactivo estructurados (Rupp & Templin, 2008).

K , siguiendo los estándares de nomenclatura en CDM (Rupp, Templin, & Henson, 2010). Los reactivos son expresados en estos modelos con la letra Y , de modo tal que una respuesta correcta del examinado i en el reactivo j se denota como $Y_{ij} = 1$ y una respuesta incorrecta como $Y_{ij} = 0$. Los atributos latentes se expresan con la letra griega α , y se puede expresar si el examinado i posee el atributo latente k como $\alpha_{ik} = 1$ o si no lo posee como $\alpha_{ik} = 0$.

Idealmente, una teoría cognitiva debería ser utilizada dentro de la planeación de la construcción de la prueba o inventario psicológico, de modo tal que la teoría defina que atributos latentes son requeridos para cada reactivo y describa el proceso a través del cual las habilidades están ligadas para producir una respuesta correcta (Henson, Templin, & Willse, 2009). La teoría cognitiva queda especificada en la matriz Q (de la Torre, 2009), misma que es un elemento común en los distintos modelos contenidos bajo el término genérico CDM. La matriz Q especifica en modo binario los atributos latentes medidos por la prueba que son requeridos para producir una respuesta correcta a cada reactivo. Las dimensiones de la matriz Q se definen a partir del número total de reactivos J en la prueba como filas y el número total de atributos latentes K a medir como columnas, de modo tal que la celda q_{jk} en la matriz específica si el atributo latente α_k es medido en el reactivo Y_j (Tatsuoka, 1990).

La inclusión de la matriz Q en los CDM resulta un número fijo de posibles clases latentes en las que se puede clasificar a los examinados y define dicha clasificación a partir de patrones ideales de respuesta y desviaciones con respecto a dichos patrones ideales (Chiu & Douglas, 2013). Por ejemplo, si hay K atributos latentes dicotómicos α_k en la matriz Q , entonces habrá un máximo de 2^K perfiles latentes posibles para la clasificación de los examinados; en consecuencia el número

total de perfiles latentes se incrementa de modo exponencial conforme el número de posibles clases medidas por la prueba se incrementa (Templin & Henson, 2006; von Davier, 2005, 2008).

Dada la naturaleza diagnóstica de los CDMs, la matriz Q idealmente debe ser desarrollada y validada por expertos en temas relativos a los constructos medidos por la prueba (Tatsuoka, 1990). Por ejemplo, la Tabla 1 incluye una lista de cotejo de doce signos y síntomas ($J = 12$) ligados al diagnóstico de trastornos psicológicos de depresión o de ansiedad generalizada ($K = 2$); dicha matriz Q debe ser generada y validada por expertos en salud mental con el fin de determinar si estos signos y síntomas en la lista de cotejo son suficientes para poder hacer un diagnóstico clínico correcto. Tras la validación de la matriz Q , la lista de cotejo y el modelo de diagnóstico cognitivo elegido para el análisis ayudarán a clasificar a cada persona evaluada de acuerdo a la identificación o ausencia de uno o ambos trastornos psicológicos.

TABLA 1.
Ejemplo de lista de cotejo en forma de matriz Q .

Signos y Síntomas	Ansiedad Generalizada	Depresión
Estado anímico depresivo	0	1
Falta de interés y disfrute de actividades cotidianas	0	1
Fatiga prevalente	1	1
Autoestima reducido	0	1
Sentimientos de culpa	1	1
Dificultad para dormir	1	1
Disminución en el apetito	0	1
Ideaciones de autoagresión	0	1
Irritabilidad	1	0
Nerviosismo o preocupación	1	0
Sentimientos de inquietud constante	1	0
Dificultad en concentración en actividades	1	1

Origen e investigación actual en modelos de diagnóstico cognitivo

Los CDM tienen su origen y pueden ser expresados como una variación del modelo general de clases latentes (Rupp & Templin, 2008; Templin & Henson, 2006; von Davier, 2009). El modelo general de clases latentes incluye dos componentes: un *componente estructural* que describe el total de posibles perfiles latentes que se pueden obtener a partir de la definición de la probabilidad conjunta de los atributos latentes $P(\alpha)$, y un *componente de medición* entendido como la probabilidad condicional para las variables observadas (por ejemplo, los reactivos de una prueba) dado que los atributos latentes α_k satisfacen el supuesto de independencia condicional (Rupp & Templin, 2008). De esta manera, el modelo general de clases latentes cuando los atributos latentes α_k se definen como dicotómicos puede ser expresado como

$$P(Y_{ij} = y_{ij}) = \sum_{k=1}^{2^K-1} P(\alpha) \prod_{j=1}^J \pi_{jk}^{y_{ij}} (1 - \pi_{jk})^{1-y_{ij}} \quad (1)$$

donde la probabilidad de que el examinado i presente la respuesta y_{ij} (por ejemplo, que elija la respuesta correcta) al reactivo Y_j se expresa del lado izquierdo de la ecuación (A) y depende del componente estructural (B) y el componente de medición (C) del modelo de clases latentes. Es de notar que el componente estructural del modelo incluye el número total menos uno $2^K - 1$ de los posibles perfiles latentes $P(\alpha) = P(\alpha_1, \alpha_2, \dots, \alpha_k)$.

$$(A) \quad P(Y_{ij} = y_{ij})$$

$$(B) \quad \sum_{k=1}^{2^K-1} P(\alpha)$$

$$(C) \quad \prod_{j=1}^J \pi_{jk}^{y_{ij}} (1 - \pi_{jk})^{1-y_{ij}}$$

Los CDM derivan del modelo expresado en la Ecuación (1) a través del componente de medición, mismo que toma forma de una

variable aleatoria con distribución Bernoulli con un parámetro π_{jk} que se define como la probabilidad de respuesta a cada uno de los reactivos (Rupp, Templin, & Henson, 2010). El parámetro π_{jk} es comúnmente definido en psicometría como la función de respuesta al reactivo, y tomará distintas formas dependiendo del CDM específico que se utilice para modelar los datos.

Entre los primeros CDM figura la propuesta de Maris (1999), quien desarrolla un modelo conjuntivo de múltiple clasificación de clases latentes (MCLCM, por sus siglas en inglés) utilizando la idea de matrices Q de Tatsuoka (1990). Después, Junker y Sijtsma (2001) proponen el modelo de insumos deterministas y salidas aleatorias (DINA, por sus siglas en inglés) donde el parámetro π_{jk} del componente de medición es expresado como

$$\pi_{jk} = P(Y_{ij} = 1 | s_j, g_j, \alpha_i) = \left[(1 - s_j)^{\prod_{k=1}^K \alpha_{ik}^{q_{jk}}} \right] \left[g_j^{1 - \prod_{k=1}^K \alpha_{ik}^{q_{jk}}} \right] \quad (2)$$

donde el parámetro de adivinación g_j del reactivo se define como la probabilidad de contestar correctamente al reactivo (o bien, de presentar un síntoma en una lista de cotejo) dado el examinado no cuenta con todos los atributos latentes α_{ik} relacionados con ese reactivo de acuerdo con la matriz Q (Huo & de la Torre, 2014). El parámetro de desliz s_j del reactivo se define como la probabilidad de contestar incorrectamente (por ejemplo, $Y_{ij} = 0$) a pesar de que el examinado cuenta con todos los atributos latentes. Como se definió anteriormente, α_{ik} son los atributos dicotómicos latentes, y q_{jk} son los elementos el vector fila dentro de la matriz Q que corresponde al reactivo j . Para que el reactivo pueda ser considerado diagnósticamente informativo de los atributos latentes α_k que intenta medir, los parámetros del reactivo deben cumplir con la regla de desigualdad $(1 - s_j) > g_j$ (Huo & de la Torre, 2014).

Dentro entre los primeros CDM, el modelo DINA es particularmente relevante pues mucha de la investigación inicial y reciente se ha enfocado en extensiones de éste modelo. Por ejemplo, de la Torre y Douglas (2004) propusieron un modelo DINA de orden superior donde el supuesto de independencia entre los atributos latentes queda de lado al incluir una variable latente continua de orden superior similar a la encontrada en los modelos de teoría de respuesta al reactivo; de la Torre (2008) y Huo y de la Torre (2014) desarrollaron una aproximación en la cual el modelo DINA puede contar con múltiples matrices Q para distintos tipos de examinados; nuevamente, de la Torre (2011) propuso una generalización del modelo DINA que resulta más flexible en términos de la especificación de los atributos como efectos principales y con interacciones; DeCarlo (2011) propuso una reparametrización del modelo DINA utilizando los principios de la teoría de detección de señales y demostró su aplicación en la especificación de la matriz Q (DeCarlo, 2012); Huang y Wang (2014) han extendido el modelo para dar lugar a una versión con efectos aleatorios; Park y Lee (2014) propusieron una reparametrización del modelo DINA que incluye covariables; y Culpepper (2015) ha propuesto la estimación bayesiana del modelo DINA.

Además del modelo DINA, Junker y Sijtsma (2001) también desarrollaron un modelo similar con especificación de adivinación y desliz a nivel de los atributos latentes al que denominado de insumos aleatorios y salidas deterministas (NIDA, por sus siglas en inglés). Hartz (2002) propuso la estimación del Modelo Unificado Reducido (RUM, por sus siglas en inglés). Templin y Henson (2006) desarrollaron una alternativa al modelo DINA denominado DINO que difiere en el modo en que interactúan los atributos latentes; recientemente, Köhn & Chiu (2016) demostraron que el modelo DINO y el modelo DINA son en realidad equivalentes,

por lo que las estimaciones de un modelo pueden ser expresadas en términos del otro.

Tras los modelos iniciales surgieron una serie de modelos generales de diagnóstico cognitivo, desarrollados con el fin de contar con funciones de respuesta al reactivo más generales a partir de las cuales se pueden derivar modelos específicos de acuerdo con la teoría cognitiva a probar empíricamente. Destaca el trabajo de von Davier (2005) con el modelo diagnóstico general (GDM, por sus siglas en inglés), Henson, Templin y Willse (2009) con el modelo log-lineal de diagnóstico cognitivo (LCDM, por sus siglas en inglés), y de la Torre (2011) con el modelo DINA generalizado (G-DINA, por sus siglas en inglés). Las diferencias entre estos modelos radican principalmente en la parametrización matemática y del software estadístico utilizado para la estimación de los mismos.

El uso de CDMs en pruebas estandarizadas es relativamente pequeño comparado con los modelos de Teoría Clásica y Teoría de Respuesta al Ítem, a pesar de poseer la capacidad de estimar parámetros psicométricos de los reactivos no considerados por otros modelos psicométricos. En específico, el uso de los CDM se reporta con mayor frecuencia en la literatura psicométrica, principalmente en el contexto de investigación clínica y educativa. Fuentes de datos analizados con distintos CDM incluyen: la Evaluación Nacional de Progreso Educativo (NAEP, por sus siglas en inglés), la prueba de Tendencias en el Estudio Internacional de Matemáticas y Ciencia (TIMSS, por sus siglas en inglés), el Examen para el Certificación de Dominio del Inglés (ECPE, por sus siglas en inglés), inventarios clínicos para el diagnóstico de trastornos psicológicos, y listas de cotejos de síntomas para el diagnóstico clínico (de la Torre, van der Ark, & Rossi, 2015; Lee, Park, & Taylan, 2011; Templin & Henson, 2006; Templin & Hoffman, 2013; Xu & von Davier, 2008).

Sin embargo, más investigación se debe realizar con el fin de hacer a los CDMs una alternativa psicométrica de evaluación más robusta comparada con la Teoría Clásica y la Teoría de Respuesta al Ítem. Existen temas que aún requieren un mayor análisis y discusión dentro del área de investigación en CDM, como son la adecuación de estos modelos para equiparación de pruebas (Xin & Zhang, 2014), la especificación y validación de la matriz Q (Chiu, Douglas, & Li, 2009; de la Torre, 2008; de la Torre & Chiu, 2015; DeCarlo, 2012; Liu, Xu, & Ying, 2012), parametrización alternativa de modelos (de la Torre, 2011; DeCarlo, 2011; Henson, Templin, & Willse, 2009; von Davier, 2014), la relación de los CDM con otros modelos psicométricos (Lee, de la Torre, & Park, 2012; von Davier, 2005, 2008), medidas de ajuste del reactivo y del modelo (de la Torre & Lee, 2013), medidas de clasificación de ítems y examinados (Henson, Roussos, Douglas, & He, 2008), y las bases teóricas de los modelos en ciencias cognitivas (Kuo, Pai, & de la Torre, 2016; Leighton, Gierl, & Hunka, 2004; Rupp, 2007).

Estimación de CDMs y software estadístico

La estimación de los distintos modelos de diagnóstico cognitivo se puede llevar a cabo por métodos de máxima verosimilitud, particularmente con el algoritmo de esperanza-maximización (EM, por sus siglas en inglés; Dempster, Laird, & Rubin, 1977), o por métodos bayesianos con el uso de métodos de cadena de Markov Monte Carlo (MCMC, por sus siglas en inglés) como el algoritmo de muestreo de Gibbs (Geman & Geman, 1984). Aunque no es la intención de este texto entrar en los detalles de la estimación de los modelos, el autor recomienda revisar a de la Torre (2009) en torno la estimación de parámetros para el modelo DINA con el uso del algoritmo EM y Culpepper (2015) para una revisión de la estimación del modelo DINA por métodos bayesianos. De igual manera, McLachlan y Krishnan (1996) es un texto recomendado para entender las bases del algoritmo EM,

Gelman et al. (2013) y Gill (2007) son fuentes introductorias a estadística bayesiana para aquellos no familiarizados con esta aproximación.

Como se mencionó en la sección anterior, la investigación en torno a la estimación de CDM es un tema que requiere más investigación. Durante su revisión de la literatura acerca de los CDM, Rupp y Templin (2008) consideran como temas de centrales de investigación la identificación de los modelos, la convergencia de la estimación, y la parametrización de las variables latentes. En términos de identificación de parámetros, von Davier (2005) ha propuesto criterios para la identificación local de los CDM partiendo de la investigación en modelos de clases latentes.

Respecto a la parametrización de los modelos, la investigación se ha centrado alternativas al modelo saturado en el componente estructural de los CDM, esto ha resultado en modelos que asumen independencia entre los atributos latentes α_k (Maris, 1999), modelo con variables latentes de orden superior que flexibilizan la independencia condicional de los atributos (de la Torre & Douglas, 2004; DeCarlo, 2012), modelos con distribuciones alternativas para los atributos latentes (Hartz, 2002; Templin & Henson, 2006), y modelos con jerarquías entre los atributos (von Davier, 2010). Además, Xu y Zhang (2016) recientemente demostraron los requisitos mínimos que debe cumplir la matriz Q para poder identificar los modelos de diagnóstico cognitivo.

En términos de estadísticas de ajuste de CDMs, Rupp, Templin y Henson (2010) proponen el uso del criterio de información de Akaike (AIC), el criterio de información Bayesiano (BIC), y el valor de residuales cuadrados normalizados para los examinados. Además, de la Torre y Lee (2013) han propuesto medidas de ajuste para los reactivos bajo el modelo DINA generalizado.

La literatura en torno a los CDM reporta el uso de distintas paqueterías de software estadístico para la estimación de estos modelos. La Tabla 2 hace un resumen de las paqueterías de software estadístico comúnmente reportadas en la literatura, los modelos estimados en cada tipo de software, y la aproximación estadística utilizada para la estimación de los modelos. Como se aprecia en la Tabla 2, la mayor parte del software estadístico utiliza el Algoritmo EM permitiendo la estimación de

los parámetros de una manera rápida comparada con los métodos bayesianos. Dependiendo del costo de la paquetería estadística, hay software libre de licencia como R (R Core Team, 2012) y Open BUGS (Thomas, O'Hara, Ligges, & Sturtz, 2006), software que se puede conseguir vía el contacto de su autor (Von Davier, 2005), y software con licencia como Mplus (Muthén & Muthén, 2012) y Latent GOLD (Vermunt & Magidson, 2005).

TABLA 2.

Paquetería estadística para la estimación de modelos de diagnóstico cognitivo.

Paquetería estadística	Utilizando en	Modelo(s) estimado(s)	Estimación
MDLTM	Von Davier (2005, 2014)	GDM, RUM, MCLCM, DINA	Algoritmo EM
Mplus	Templin & Hoffman (2013), Chiu, Köhn, & Wu (en prensa)	LCDM, RUM reducido.	Algoritmo EM
Latent GOLD	DeCarlo (2011), Park & Lee (2014)	DINA reparametrizado	Probabilidad máxima a posteriori
Open BUGS	DeCarlo (2012), Culpepper (2015)	DINA, DINA reparametrizado	Bayesiana por muestreo de Gibbs
Paquete 'CDM' en R	Robitzsch, Kiefer, George, & Ünlü, (2014)	GDM, G-DINA, DINA, DINO	Algoritmo EM

Tutorial didáctico

En esta sección se presenta de modo didáctico el uso del modelo DINA a través del paquete 'CDM' (Robitzsch, Kiefer, George, & Ünlü, 2014) en R (R Core Team, 2012) para el análisis de datos. La selección del modelo DINA sobre otros CDM se basa en la parsimonia del modelo, larga literatura y extensiones en torno al mismo. El apéndice A incluye un ejemplo de código en R para estimar el modelo DINA en R.

Simulación de datos

En este ejemplo, 500 bases de datos cada una con tamaño de muestra (es decir, el número de examinados simulados) $N = 1000$ fueron simuladas utilizando en combinación las funciones de probabilidad presentadas en las ecuaciones (1) y (2). Con respecto al componente estructural del

modelo expresado en la ecuación (1), tres atributos categóricos latentes α_k fueron generados de modo independiente con una función logística; el tamaño de cada clase latente fue $P(\alpha_1) = 0.100$, $P(\alpha_2) = 0.622$, y $P(\alpha_3) = 0.450$. Dado que la función $P(\alpha_{ik})$ produce probabilidades de presencia de cada atributo para cada examinado, éstas fueron comparadas con un número aleatorio producido por una distribución Uniforme $U[0, 1]$ de modo tal que si $P(\alpha_{ik}) \geq U_{ik}$ entonces el atributo $\alpha_{ik} = 1$, de otro modo $\alpha_{ik} = 0$.

En los términos requeridos para un investigador aplicado, estas especificaciones matemáticas de la simulación indican la frecuencia con la que esperamos estimar cada atributo latente medido por la prueba en la muestra de examinados y, después, cómo convertir de

modo aleatorio las probabilidades de presencia de cada atributo en cada examinado en indicadores dicotómicos de los atributos latentes α_k . Por ejemplo, si un investigador tiene evidencia de que un constructo psicológico latente sólo ocurre en un 2% de la muestra o de la población medida (por ejemplo, un trastorno anímico poco común), puede generar datos en los que $P(\alpha_k) = 0.02$ con el fin de probar alguna hipótesis bajo un CDM.

Con respecto al componente de medición del modelo, la matriz Q, los parámetros de desliz y de adivinación del

reactivo como los mostrados en la Tabla 3. La matriz incluye un total de catorce reactivos, los seis primeros miden únicamente un atributo latente, mientras que el resto de los reactivos que miden dos o más atributos (reactivos 7 a 14). Los valores de los parámetros de adivinación y de desliz del reactivo para cada reactivo fueron establecidos en el rango entre 0.00 y 0.20 con el fin de tener reactivos que logren discriminar entre examinados que poseen y no poseen los atributos latentes (Lee, de la Torre, & Park, 2012).

TABLA 3.
Matriz Q y parámetros de adivinación y desliz.

Reactivo	Matriz Q			Parámetros del reactivo	
	α_1	α_2	α_3	Adivinación g_j	Desliz s_j
1	1	0	0	0.00	0.20
2	1	0	0	0.10	0.05
3	0	1	0	0.05	0.15
4	0	1	0	0.15	0.10
5	0	0	1	0.10	0.20
6	0	0	1	0.05	0.05
7	1	1	0	0.20	0.00
8	1	1	0	0.05	0.10
9	1	0	1	0.15	0.20
10	1	0	1	0.10	0.10
11	0	1	1	0.05	0.20
12	0	1	1	0.15	0.15
13	1	1	1	0.10	0.15
14	1	1	1	0.15	0.05

Para los propósitos de este ejemplo, se buscó determinar la capacidad de estimación del tamaño de las tres clases latentes $P(\alpha_k)$, y de los parámetros de adivinación g_j y desliz s_j del reactivo del modelo DINA utilizando la librería 'CDM' en R. Con este objetivo, medidas de valor promedio del estimador, sesgo absoluto, y error cuadrático medio (Rizzo, 2008) fueron calculados para cada parámetro y son reportadas en la siguiente sección.

RESULTADOS

Los resultados en la Tabla 4 detallan la estimación de los parámetros del reactivo. Como se muestra, en promedio todos los parámetros de adivinación y desliz son correctamente estimados, inclusive en los casos en los que se consideró que el reactivo no contaba con ningún grado de adivinación o de desliz. En general, la estimación de los parámetros de adivinación se logra con menor sesgo absoluto y error cuadrático medio comparada con la estimación de los parámetros de desliz.

TABLA 4.
Recuperación de parámetros de adivinación y desliz del reactivo.

Reactivo	Adivinación				Desliz			
	Par	Estimador promedio	Sesgo absoluto	ECM	Par	Estimador Promedio	Sesgo absoluto	ECM
1	0	0.0001	0.0001	0.0000	0.2	0.2005	0.0371	0.0022
2	0.1	0.1004	0.0084	0.0001	0.05	0.0496	0.0500	0.0006
3	0.05	0.0517	0.0157	0.0004	0.15	0.1493	0.0153	0.0004
4	0.15	0.1493	0.0213	0.0007	0.1	0.0973	0.0116	0.0002
5	0.1	0.0981	0.0115	0.0002	0.2	0.1995	0.0170	0.0005
6	0.05	0.0502	0.0116	0.0002	0.05	0.0506	0.0115	0.0002
7	0.2	0.2006	0.0108	0.0002	0	0.0010	0.0010	0.0000
8	0.05	0.0499	0.0059	0.0001	0.1	0.0994	0.0332	0.0017
9	0.15	0.1504	0.0096	0.0001	0.2	0.1979	0.0506	0.0042
10	0.1	0.1002	0.0078	0.0001	0.1	0.1000	0.0387	0.0023
11	0.05	0.0496	0.0070	0.0001	0.2	0.2003	0.0209	0.0007
12	0.15	0.1502	0.0114	0.0002	0.15	0.1497	0.0195	0.0006
13	0.1	0.1002	0.0077	0.0001	0.15	0.1461	0.0569	0.0050
14	0.15	0.1495	0.0092	0.0001	0.05	0.0504	0.0343	0.0019

Nota: 'Par' se refiere al valor del parámetro utilizado para la simulación de datos, 'ECM' se refiere al error cuadrático medio.

Resultados similares para el tamaño de las clases latentes del modelo son presentados en la Tabla 5. Nuevamente, los valores promedio estimados para cada clase latente son muy cercanos al valor del parámetro utilizado para la generación de los datos simulados. Además, el sesgo absoluto y error cuadrático medio de los estimadores es muy pequeño, confirmando una buena recuperación de los parámetros relativos a los atributos latentes del modelo.

TABLA 5.
Recuperación del tamaño de las clases latentes.

Attribute	Parámetro	Estimador promedio	Sesgo absoluto	ECM
α_1	0.100	0.1001	0.0079	0.0001
α_2	0.622	0.6204	0.0158	0.0004
α_3	0.450	0.4502	0.0147	0.0003

Nota: 'ECM' se refiere al error cuadrático medio.

DISCUSIÓN Y CONCLUSIÓN

En este texto se resumieron elementos teóricos y aplicados de los modelos de diagnóstico cognitivo, poniendo énfasis en las bases matemáticas del modelo, su

aplicación, y áreas de investigación actual. La aplicación de los CDMs podría ayudar al investigador aplicado cuando tiene una teoría cognitiva multidimensional sobre su fenómeno de estudio, especialmente cuando se utilizan modelos simples como el DINA o modelos generales y flexibles como el GDM.

A pesar de la recepción positiva que estos modelos recibieron en la primera década del siglo XXI (Rupp & Templin, 2008), se debe reconocer que aún presentan limitaciones frente a los modelos de teoría clásica y de teoría de respuesta al reactivo que deberán de superar tanto en el plano teórico como en su aplicación. Por ejemplo, aún no hay consenso respecto a cómo definir cada celda de la matriz Q – elemento central a todos los CDM – cuando no existe consenso entre los validadores de la prueba en torno a los atributos medidos por cada reactivo (de la Torre & Chiu, 2015; DeCarlo, 2012; Liu, Xu, & Ying, 2012). Además, poca investigación se ha enfocado en la relación entre la definición teórica de los atributos categóricos latentes en los

CDM y los constructos psicológicos que se buscan medir, por lo que un abordaje teórico que retome el enfoque de la validez de constructo (Messick, 1989) y el escalamiento categórico de los CDM ayudaría a dar mayor soporte a esta perspectiva psicométrica.

De modo similar, la teoría de respuesta al reactivo se encuentra bien consolidada en procesos aplicados como el establecimiento de bancos de reactivos y el desarrollo de pruebas computarizadas adaptativas por su capacidad de escalamiento de la variable latente continua ϑ (van der Linden & Glas 2000), dejando a los CDM con poco espacio para innovación en estos procesos.

REFERENCIAS

- Chiu, C. Y., & Douglas, J. (2013). A nonparametric approach to cognitive diagnosis by proximity to ideal response patterns. *Journal of Classification*, 30(2), 225-250.
- Chiu, C., Douglas J., & Li, X. (2009). Cluster analysis for cognitive diagnosis: Theory and applications. *Psychometrika*, 74(4), 633-665.
- Chiu, C.-Y., Köhn, H.-F., & Wu, H.-M. (en prensa). Fitting the Reduced RUM with Mplus: A Tutorial. *International Journal of Testing*.
- Culpepper, S. A. (2015). Bayesian Estimation of the DINA Model with Gibbs Sampling. *Journal of Educational and Behavioral Statistics*, 40(5), 454-476.
- de La Torre, J. (2008). An Empirically Based Method of Q-Matrix Validation for the DINA Model: Development and Applications. *Journal of Educational Measurement*, 45(4), 343-362.
- de la Torre, J. (2009). DINA model and parameter estimation: A didactic. *Journal of Educational and Behavioral Statistics*, 34(1), 115-130.
- de la Torre, J. (2011). The generalized DINA model framework. *Psychometrika*, 76(2), 179-199.
- de la Torre, J., & Chiu, C.Y. (2015). A General Method of Empirical Q-Matrix Validation. *Psychometrika*, 81(2), 253-273.
- de la Torre, J., & Douglas, J. A. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika*, 69(3), 333-353.
- de la Torre, J., & Lee, Y. S. (2013). Evaluating the Wald Test for Item-Level Comparison of Saturated and Reduced Models in Cognitive Diagnosis. *Journal of Educational Measurement*, 50(4), 355-373.
- de la Torre, J., van der Ark, L. A., & Rossi, G. (2015). Analysis of clinical data from cognitive diagnosis modeling framework. *Measurement and Evaluation in Counseling and Development*. Publicación en línea. DOI:10.1177/0748175615569110
- DeCarlo, L. T. (2011). On the analysis of fraction subtraction data: The DINA model, classification, latent class sizes, and the Q-matrix. *Applied Psychological Measurement*, 35(1), 8-26.
- DeCarlo, L. T. (2012). Recognizing Uncertainty in the Q-Matrix via a Bayesian Extension of the DINA Model. *Applied Psychological Measurement*, 36(6), 447-468.
- Dempster, A.P., Laird, N.M., & Rubin, D.B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society*, 39(1), 1-38.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. Boca Raton, FL: CRC press.
- Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721-741.

- Gill, J. (2007). *Bayesian methods: A social and behavioral sciences approach*. Boca Raton, FL: CRC press.
- Junker, B. W., & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement, 25*(3), 258–272.
- Köhn, H.-F., & Chiu, C.-Y. (2016). A proof of the duality of the DINA model and the DINO model. *Journal of Classification, 33*, 171-184.
- Kuo, B. C., Pai, H. S., & de la Torre, J. (2016). Modified Cognitive Diagnostic Index and Modified Attribute-Level Discrimination Index for Test Construction. *Applied Psychological Measurement, 40*(5), 315-330.
- Hartz, S. M. (2002). *A Bayesian framework for the unified model for assessing cognitive abilities: Blending theory with practicality*. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign, Urbana-Champaign, IL.
- Henson, R. A., Templin, J. L., & Willse, J. T. (2009). Defining a family of cognitive diagnosis models using log-linear models with latent variables. *Psychometrika, 74*(2), 191-210.
- Henson, R., Roussos, L., Douglas, J., & He, X. (2008). Cognitive diagnostic attribute-level discrimination indices. *Applied Psychological Measurement, 32*(4), 275-288.
- Huang, H.Y. & Wang, W.C. (2014). The Random-Effect DINA Model. *Journal of Educational Measurement, 51*(1), 75–97.
- Huo, Y., & de la Torre, J. (2014). Estimating a Cognitive Diagnostic Model for Multiple Strategies via the EM Algorithm. *Applied Psychological Measurement, 38*(6), 464-485.
- Lee, Y. S., de la Torre, J., & Park, Y. S. (2012). Relationships between cognitive diagnosis, CTT, and IRT indices: an empirical investigation. *Asia Pacific Education Review, 13*(2), 333-345.
- Lee, Y.-S., Park, Y. S., & Taylan, D. (2011). A cognitive diagnostic modeling of attribute mastery in Massachusetts, Minnesota, and the U.S. national sample using the TIMSS 2007. *International Journal of Testing, 11*, 144-177.
- Leighton, J. P., Gierl, M. J., & Hunka, S. M. (2004). The attribute hierarchy method for cognitive assessment: A variation on Tatsuoka's rule-space approach. *Journal of Educational Measurement, 41*(3), 205-237.
- Liu, J., Xu, G., & Ying, Z. (2012). Data-driven learning of Q-matrix. *Applied Psychological Measurement, 36*, 548–564.
- Maris, E. (1999). Estimating multiple classification latent class models. *Psychometrika, 64*(2), 187–212.
- McLachlan, G. J., & Krishnan, T. (1996). *The EM algorithm and extensions*. New York, NY: Wiley.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (pp. 13-103). Old Tappan, NJ: Macmillan.
- Muthén, L. K., & Muthén, B. O. (2012). *Mplus statistical modeling software: Release 7.0*. Los Angeles, CA: Muthén & Muthén.
- Park, Y. S., & Lee, Y.-S. (2014). An Extension of the DINA model using covariates: Examining factors affecting response probability and latent classification. *Applied Psychological Measurement, 38*(5), 376-390.
- R Core Team (2012). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>
- Rizzo, M. L. (2008). *Statistical Computing with R*. Boca Raton, FL: Chapman & Hall/CRC.
- Robitzsch, A., Kiefer, T., George, A.C., & Ünlü, A. (2014). *CDM: Cognitive Diagnosis Modeling*. R package version

- 2.7-7. <http://CRAN.R-project.org/package=CDM>
- Rupp, A. A. (2007). The answer is in the question: A guide for describing and investigating the conceptual foundations and statistical properties of cognitive psychometric models. *International Journal of Testing*, 7, 95–125.
- Rupp, A. A., & Templin, J. L. (2008). Unique characteristics of diagnostic classification models: A comprehensive review of the current state-of-the-art. *Measurement*, 6(4), 219-262.
- Rupp, A. A., Templin, J., & Henson, R. A. (2010). *Diagnostic assessment: Theory, methods, and applications*. New York, NY: Guilford.
- Tatsuoka, K. K. (1990). Toward an integration of item-response theory and cognitive error diagnosis. In N. Frederiksen, R. Glaser, A. Lesgold, & M. Shafto (Eds.), *Diagnostic monitoring of skill and knowledge acquisition* (pp. 453-488). Hillsdale, NJ: Erlbaum.
- Templin, J. L., & Henson, R. A. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychological methods*, 11(3), 287-305.
- Templin, J., & Hoffman, L. (2013). Obtaining diagnostic classification model estimates using Mplus. *Educational Measurement: Issues and Practice*, 32(2), 37-50.
- Thomas, A., O'Hara, B., Ligges, U., & Sturtz, S. (2006). Making BUGS open. *R news*, 6(1), 12-17.
- van der Linden, W.J., & Glas, C.A.W. (2000). *Computerized adaptive testing: theory and practice*. Dordrecht: Kluwer Academic Publishers.
- Vermunt, J. K., & Magidson, J. (2005). *Technical guide for Latent Gold 4.0: Basic and advanced*. Belmont, MA: Statistical Innovations Inc.
- von Davier, M. (2005). *A general diagnostic model applied to language testing data* (ETS Research Report RR-05- 16). Princeton, NJ: Educational Testing Service.
- von Davier, M. (2008). The mixture general diagnostic model. In G. R. Hancock & K. M. Samuelsen (Eds.), *Advances in latent variable mixture models* (pp. 255-274). Charlotte, NC: Information Age Publishing.
- von Davier, M. (2009). Some notes on the reinvention of latent structure models as diagnostic classification models. *Measurement*, 7(1), 67-74.
- von Davier, M. (2010). Hierarchical mixtures of diagnostic models. *Psychological Test and Assessment Modeling*, 52(1), 8-28.
- von Davier, M. (2014). The DINA model as a constrained general diagnostic model: Two variants of a model equivalency. *British Journal of Mathematical and Statistical Psychology*, 67(1), 49-71.
- Xin, T., & Zhang, J. (2015). Local Equating of Cognitively Diagnostic Modeled Observed Scores. *Applied Psychological Measurement*, 39(1), 44-61.
- Xu, X., & Davier, M. (2008). *Fitting the structured general diagnostic model to NAEP data*. (ETS Research Rep. No. RR-08-27). Princeton, NJ: ETS.
- Xu, G., & Zhang, S. (2016). Identifiability of Diagnostic Classification Models. *Psychometrika*, 81(3), 625–649.

ANEXO 1.

Estimación del modelo DINA utilizando el paquete 'CDM' en R.

EJEMPLO DEL MODELO DINA

```
# Llama a la librería 'CDM' para poder utilizar el modelo DINA
install.packages("CDM") # Puedes omitir esta línea de código si el paquete ya
# está instalado en tu computadora
library(CDM)
```

```
# Crea tu matriz Q.
q1 <- c(1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 1, 1)
q2 <- c(0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 1, 1)
q3 <- c(0, 0, 0, 0, 1, 1, 0, 0, 1, 1, 1, 1, 1, 1)
QMat <- cbind(q1, q2, q3)
```

```
# Define el directorio donde están almacenados los datos de tu prueba o
# inventario. Es recomendable guardar estos datos en formato de
# comma separated values (*.csv).
```

```
# Datos <- read.csv("C:/Users/Desktop/DINA1.csv")
Datos <- read.csv("DINA1.csv")
```

```
# Utiliza la función din() de la librería 'CDM' para estimar el modelo DINA
# Al usar esta librería, se debe especificar el objeto en R que denota tus
# datos y el objeto que corresponde a tu matriz Q. Además, en el comando
# "rule" se debe especificar que se utilizará el modelo DINA. Si tu base de
# datos contiene información adicional a las respuestas a los reactivos,
# se deben indicar únicamente las columnas de datos que corresponden a
# los reactivos a analizar.
mod_DINA <- din(data = Datos[,1:14], q.matrix = QMat, rule = "DINA")
```

```
# Al utilizar la función summary() en el objeto correspondiente
# al modelo DINA estimado, obtendremos información sobre el ajuste
# del modelo (Devianza, Log-verosimilitud, AIC, BIC), parámetros
# de adivinación y desliz de los reactivos que componen la prueba,
# tamaños de cada atributos latente, correlaciones tetracóricas
# entre los atributos latentes dicotómicos, y distribución de cada
# perfil de atributos latentes.
summary(mod_DINA)
```

```
# Al utilizar la función names() podemos ver toda la información producida
# por din() al estimar el modelo DINA con los datos. Esto puede servir para
# obtener información adicional a la reportada por summary()
names(mod_DINA)
```

```
# Así, la información adicional puede ser solicitada al utilizar el signo de
# pesos '$' después del nombre del objeto en R que corresponde al objeto
# producido por la función din()
mod_DINA$attribute.patt
```