



Interdisciplinaria

ISSN: 0325-8203

ISSN: 1668-7027

interdisciplinaria@fibercorp.com.ar

Centro Interamericano de Investigaciones Psicológicas y
Ciencias Afines

Argentina

Barragán, Sandra

Modelación con teoría de grafos para la unidimensionalidad de un instrumento de evaluación

Interdisciplinaria, vol. 35, núm. 1, 2018, -Junio, pp. 7-33

Centro Interamericano de Investigaciones Psicológicas y Ciencias Afines
Argentina

Disponible en: <https://www.redalyc.org/articulo.0a?id=18058784002>

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

redalyc.org
UAEM

Sistema de Información Científica Redalyc

Red de Revistas Científicas de América Latina y el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso
abierto

Modelación con teoría de grafos para la unidimensionalidad de un instrumento de evaluación

Modeling with graph theory to unidimensionality for an assessment instrument

Sandra Barragán*

*Magister en Ciencias Matemáticas.

Profesora Asociada II adscrita al Departamento de Ciencias Básicas y Modelado de la Universidad de Bogotá Jorge Tadeo Lozano. Grupo de fundamentos y didáctica de las ciencias.

E-Mail: Sandra.barragan@utadeo.edu.co

Universidad de Bogotá Jorge Tadeo Lozano.

Carrera 4 N°22-61 Módulo 6 Oficina201. Bogotá – Colombia.

Resumen

En el marco de las políticas públicas en Colombia sobre calidad de la educación superior se encuentran los lineamientos para que el Estado reconozca los programas de pregrado como de alta calidad, entre tales lineamientos se consideran los sistemas de evaluación de los estudiantes basados en políticas institucionales claras y transparentes que propicien una identificación integral de las condiciones académicas y actitudinales de los estudiantes por lo que es imprescindible tecnificar y modernizar los procesos de evaluación para fortalecer la emisión de inferencias confiables sobre el logro académico. En este artículo se describe una aplicación de la Teoría de Grafos usando algunas de sus definiciones y teoremas básicos con la intención de potenciar su uso para el estudio del supuesto de unidimensionalidad de la Teoría de Respuesta al Ítem como alternativa para robustecer la evaluación. La unidimensional fue entendida en el sentido de la homogeneidad escalar y de criterio para las relaciones de orden definidas por la dominación tanto para ítems como para evaluados: para los ítems conforme al número de evaluados que los contestaron correctamente y a la dificultad; para los evaluados respecto al número de respuestas correctas y a la habilidad. Se aplicó la combinación de las dos teorías a un caso de estudio que contempla el instrumento de la prueba diagnóstica del Examen de Clasificación de Matemáticas Básicas en la Universidad Jorge Tadeo Lozano, con esto se ofrece un valor agregado por

que se evidencian los resultados de forma algorítmica y visual con ayuda de herramientas informáticas especializadas.

Palabras clave: Evaluación, Homogeneidad, Modelo de Rasch, Teoría de Grafos, Teoría de Respuesta al Ítem, Unidimensionalidad.

Abstract

In Colombia, public policies about quality of higher education involve student assessment systems as a relevant characteristic. These systems should have clear and transparent institutional policies requiring comprehensive identification of attitudinal and academic conditions about examinee performance. Consequently, on the way of continuous improvement of academic services offered by institutions of higher education, an assessment process adequately technified allowing valid inferences about academic achievement is required. In this sense, the Rasch measurement model of Item Response Theory is a modern alternative to strengthen educational measurement estimating the ability of the student and the difficulty of the item on a comparable scale. Unidimensionality, local independence and internal consistency are assumptions made in Rasch measurement model. Unidimensionality assumption has several definitions, one of them is the occurrence of a dominant factor influencing test performance. Another definition is considered in the present study. Here, unidimensionality was

interpreted as scalar and criterial homogeneity for the order relations defined by domination for items and assessed applicants. Graph Theory is an ideal mathematical modelling approach to this assumption inasmuch as represent intangible interactions as required. In order to achieve this, Graph Theory and Item Response Theory were combined to examine the qualifying test for the Basic Mathematics subject at Jorge Tadeo Lozano University as a case study. This test was composed by 45 items divided in three blocks. With several combinations of these blocks, three virtual booklets were obtained each one comprising 30 of them. From the application of June 2011, a test sample of 509 responses chains was obtained. Three data bases were processed one for each booklet, the first spanned 175 responses chains, the second 170 and the third one 164. To begin with the examination on the Rasch measurement model, the parameters reliability, separation, Cronbach Alpha and item residual correlation were estimated to gauge and determine performance test for getting acceptable values in each booklet applied. Tatsuoka was followed to undertake the path to unidimensionality from Graph Theory gathering a real case experience processed with suitable software. Additionally, the sensitivity of the order relation was verified through: 1) ordering by number of correct responses per item (1I order) and items difficulty (2I order); 2) ordering by number of correct responses per applicant (1E order) and estimated ability (2E order). Furthermore, the linear models were obtained collating these orders. In like manner, three Guttman scales and their adjacency matrices were schematized one for each booklet. Subsequently, the respective graphs were processed and represented using Gephi just as a specialized tool that enables running some algorithms like Force Atlas. Afterwards, the second power of each adjacency matrix was found using Matlab 2014b and domination matrices were calculated for both items and applicants in the aforementioned orders for a total of 12 matrices. Consistency index developed by Cliff was computed for the domination matrices. As a result, moderate consistency was observed. Significant domination for the entries of these matrices was analyzed through McNemar test in order to have an asymmetric dominance relation. Moreover, a reachable matrix was calculated for each one of

these significant domination matrices as a limit of a sequence of boolean powers. Finally, dominance hierarchies were illustrated with vertex degrees and compared with student maps by the means of Winsteps 3.73. The combination of Graph Theory and Item Response Theory allowed a deeper comprehension of unidimensionality assumption. Thereupon, universities can optimize their resources offering to applicant differential academic options per individual position in the ability scale. The results can be used to outline advantages for the applicants who can evidence their position in the ability scale and identify the different areas to improve.

Key words: Evaluation, Graph theory, Homogeneity, Item Response Theory, Rasch Model, Unidimensionality.

Introducción

En Colombia, conforme a las políticas de ley y las definidas por el Consejo Nacional de Educación Superior (CESU) el Consejo Nacional de Acreditación (CNA) ha formulado los Lineamientos para la acreditación de programas de pregrado, para el reconocimiento por parte del Estado de la calidad de las instituciones de educación superior y de programas académicos (Ramírez et al., 2013). Tales lineamientos contienen una característica de alta calidad destinada al sistema de evaluación de los estudiantes inmersa en un factor sobre Procesos Académicos. El CNA resalta que dicho sistema de evaluación “debe permitir la identificación de las competencias, especialmente las actitudes, los conocimientos, las capacidades y las habilidades adquiridas de acuerdo con el plan curricular y debe ser aplicado teniendo en cuenta la naturaleza de las características de cada actividad académica” (Ramírez et al., 2013, p. 33). Con lo cual, las políticas públicas llevan a los programas académicos y a las Instituciones de Educación Superior (IES) a la evaluación integral del desempeño de un estudiante.

Por otro lado, si se reconocen al menos dos de los desafíos planteados por Bogoya (2006) referidos al sistema de evaluación educativa

tales como: (1) tener información robusta, confiable y oportuna en cada evaluación aplicada y (2) crear y pulir nuevas estrategias y modelos de evaluación, se presenta una oportunidad para combinar diferentes teorías y técnicas con el fin de estimar el nivel de logro académico en la educación superior. Estos análisis sistemáticos refinan las inferencias, garantizan la precisión de la evaluación y aseguran la certidumbre del diagnóstico, lo que permite planificar intervenciones eficaces (Muñiz, 2010).

La TRI, como teoría psicométrica ha sido aplicada en proyectos internacionales como PISA (OECD, 2009), TIMSS (Olson, Martin & Mullis, 2008) y SERCE (SERCE, 2010) y en proyectos colombianos como SABER PRO (ICFES, 2014).

Consecuentemente, en el trabajo que se informa se aplicó la Teoría de Grafos para el estudio del supuesto de unidimensionalidad de la Teoría de Respuesta al Ítem (TRI) como en Tatsuoka, M. (1986), con el apoyo de herramientas informáticas especializadas que hacen expeditos los cálculos y la visualización permitiendo así profundidad en el análisis y el tratamiento de datos como lo requiere un sistema de evaluación.

Teoría de Respuesta al Ítem

El modelo de Rasch de la TRI estima la probabilidad p con la que un evaluado contesta correctamente un ítem. Para ítems de selección múltiple con una respuesta correcta, p es la función exponencial de la distancia entre el valor de la habilidad del estudiante y la dificultad del ítem (Ecuación 1); β_v es la habilidad del evaluado y δ_i es la dificultad del ítem i ($i=1,2,\dots,m$) (Wright & Stone, 1998).

$$p[X_{vi} = 1 | \beta_v, \delta_i] = \frac{e^{\beta_v - \delta_i}}{1 + e^{\beta_v - \delta_i}} \quad (1)$$

En la TRI, la dificultad de los ítems es una función de la probabilidad de marcar la respuesta correcta y de la habilidad de aprendizaje; definición que no incluye la apreciación sobre la cantidad de esfuerzo particular

requerido para contestarlos (DeMars, 2010; Wright & Stone, 1998).

El modelo de Rasch se caracteriza por (1) las estimaciones de los parámetros de los ítems son independientes del grupo de evaluados que responden una prueba, (2) las estimaciones de las habilidades de los evaluados son independientes de los ítems particulares usados en una prueba y (3) se conocen con precisión las estimaciones de la habilidad y la dificultad (Hambleton, 1990; Pardo, 2001). Lo anterior provee un conocimiento profundo y disgregado de los ítems, el instrumento y los evaluados.

Para aplicar el modelo de Rasch se requiere que los instrumentos de evaluación satisfagan los supuestos de consistencia interna, independencia local y unidimensionalidad.

La consistencia interna mide la coherencia de los ítems de la prueba lo que significa que saber cómo se desempeña una persona en un ítem permite predecir cómo lo harán las demás (Ruiz, 2005). Para estimar indicadores de la consistencia interna existen diferentes técnicas, una de ellas es calcular el coeficiente Alpha de Cronbach con el que se estima el límite inferior del coeficiente de confiabilidad (Linacre, 2008).

La independencia local indica que la respuesta de un evaluado a un ítem no requiere que éste recurra a información de otros ítems para contestarlo correctamente (SERCE, 2010). Este supuesto se examina mediante coeficientes de correlación inter-ítem. El parámetro de correlación se emplea en dos perspectivas: entre las respuestas a dos ítems (correlación inter-ítem) y entre las de un ítem y el conjunto (correlación ítem-prueba) (SERCE, 2010).

El cumplimiento del supuesto de unidimensionalidad en el sentido de la homogeneidad de criterio implica que los ítems que conforman un instrumento deben contribuir conjuntamente a medir algo en común, es decir que solo un rasgo latente o constructo se encuentra en la base del conjunto de ítems (Linacre, 2008). Este supuesto garantiza que para cada evaluado el modelo asigna un único valor de habilidad y cualquier otro factor que

afecte la respuesta del ítem debe tratarse como un error aleatorio (DeMars, 2010). Se espera que los ítems estén encauzados a medir el nivel de conocimiento de un evaluado sobre los dominios conceptuales específicos y no otros factores como la motivación que tengan para tomar el examen, la ansiedad, la habilidad para trabajar rápidamente o la tendencia a conjeturar cuando duda de las respuestas (Hambleton, Swaminathan & Rogers, 1991).

La unidimensionalidad en el sentido de la homogeneidad escalar se refiere a un método basado en patrones de respuesta que se aproximan al escalamiento de Guttman en el que el puntaje total corresponde a las respuestas correctas a las n preguntas más fáciles y a nada más (Hattie, 1985). Linacre enfatiza que la unidimensionalidad jamás es perfecta y que la verdadera pregunta está en si la multidimensionalidad es tan grande que requiere separar los ítems en diferentes tests, uno por cada dimensión (Linacre, 2008).

Otra definición estipula que un conjunto de ítems es unidimensional cuando es posible encontrar un vector de valores $\Phi = \{\phi_1\}$ tales que la probabilidad de contestar correctamente un ítem g es $\pi_{ig} = f_g(\phi_1)$; la independencia local se tiene para cada valor de Φ (Hattie, 1985). Para estimar la unidimensionalidad existen muchos métodos; en un estudio hecho por Hattie se comparan 87 de ellos (Hattie, 1984). La Figura 1 resume algunos con sus hipótesis y sus principales autores. El método a través del análisis factorial busca que solo un factor predomine, es decir, que el instrumento mida una sola variable y no una mixtura. Tiene en cuenta la proporción de la variancia total explicada por las medidas y la explicada por el contraste con el primer factor extraído (Linacre, 2008). Un método basado en patrones de respuesta es el índice de consistencia de Cliff (ver ecuación 2). La aplicación de los dos métodos mejoró el entendimiento del supuesto de unidimensionalidad y del efecto del orden que se impone a los ítems y a los evaluados sobre su estimación.

Para profundizar en la unidimensionalidad de un instrumento de evaluación se con-

sideró la Teoría de Grafos como modelación matemática apropiada para estudio de intangibles como la interacción en redes sociales y económicas. En educación ha sido aplicada para jerarquizar conceptos y encontrar los prerequisitos para elaborar planes de estudio y contenidos programáticos.

El análisis psicométrico del tipo tratado en este trabajo se usa en diversos campos. A manera de ilustración se mencionan los trabajos de Abraham y Brenca (2014) que lo aplica en la observación comunicativa en niños con patologías del habla y el Merino y Vilcherrez (2014) que estudia la deseabilidad social para niños a fin de analizar la validez del reporte de las experiencias psicológicas de los niños puesto que algunas veces tratan de mostrar una imagen mejorada de ellos mismos para agradar a los demás.

Teoría de Grafos

A continuación, se presentan algunas definiciones básicas que permiten la aproximación desde la modelación matemática con los grafos. Un grafo $G = (V, E)$ es una estructura matemática que consiste en dos conjuntos finitos V y E . Los elementos de V son llamados vértices (o nodos) y los elementos de E son llamados aristas. Un grafo dirigido es un grafo cuyas aristas son dirigidas (Gross & Yellen, 2006). Una matriz $A = (a_{ij})_{n \times n}$ asociada a un grafo es una matriz de adyacencia, si a_{ij} es el número de aristas que unen i con j (Bondy & Murty, 1976).

Para cada vértice se definen dos guarimbas que posibilitan la clasificación de los vértices. El grado de salida o semigrado externo de un vértice v_i ,

$$d^+(v_i) = \sum_{j=1}^n a_{ij}$$

y es el número de aristas dirigidas que parten de v_i . El grado de entrada o semigrado interno de un vértice v_j ,

$$d^-(v_j) = \sum_{i=1}^n a_{ij}$$

es el número de aristas dirigidas que inciden en v_j (Flament, 1972).

Una cadena entre dos vértices es la ruta entre ellos. Una r -cadena está conformada por r aristas dirigidas que deben recorrerse para acceder de un vértice a otro (Grossman, 1988). Si A es una matriz de adyacencia y A^r es la r -ésima potencia de A , entonces la posición ij de A^r representa el número de formas en que v_i accede a v_j , en r -cadenas (Festinger, 1949).

Con las anteriores definiciones se inició el procesamiento de las cadenas de respuestas de la aplicación de junio de 2011 del Examen de Clasificación de Matemáticas Básicas. Particularmente, se estudiaron los grafos provenientes de las matrices de adyacencia correspondientes a las matrices de dominación tanto para ítems como para evaluados ordenados por números de respuestas correctas, por dificultad y por habilidad obtenidos con el Modelo de Rasch de la TRI, con lo que se revisó el desempeño de los ítems en conjunto y de los evaluados en conjunto.

Caso de estudio

La Universidad de Bogotá Jorge Tadeo Lozano (Utadeo) aplicó el Examen de Clasificación de Matemáticas Básicas entre el primer período académico de 2007 (2007 I) y el primero de 2014 (2014 I) a los aspirantes de los diferentes programas académicos que incluían este curso en su plan de estudios. Dicho examen era una evaluación diagnóstica que implicaba para el estudiante que no lo aprobaba, tomar un curso de 64 horas presenciales. El instrumento de evaluación se calibró mediante la TRI. En el Gráfico 1 se muestran los resultados en los períodos de aplicación.

Método

Participantes

En la aplicación de junio de 2011 del Examen de Clasificación hubo 509 evaluados cuyas cadenas de respuestas se incluyeron en la base del estudio. De los evaluados, el 52.8% eran mujeres y el 46.8% hombres, el .4% faltante corresponde a dos evaluados que no ingresaron a la Utadeo. El 26.5% se matriculó en la Facultad de Ciencias Económicas Administrativas, el 45.6% en la Facultad de Ciencias Humanas, Artes y Diseño, el 26.1% en la Facultad de Ciencias Naturales e Ingeniería, el .6% en la Facultad de Relaciones Internacionales y Ciencias Jurídicas y Políticas y el .8% pertenecía al Convenio Proyecto Enlace Colegio -Universidad. Por otra parte, el 75% de los evaluados tenía entre 16 y 20 años y el 18.1% entre 21 y 25 años. La edad del 6.5% fluctuaba entre los 26 y los 40 años.

Instrumento de evaluación

Para el Examen de Clasificación se elaboró un banco de ítems en torno a los dominios conceptuales correspondientes al Curso de Matemáticas Básicas: números reales, expresiones algebraicas, factorización de expresiones polinómicas y ecuaciones de primero y segundo grado en una incógnita. El diseño de tales ítems se procuró en los dominios cognitivos: reconocimiento, aplicación y razonamiento como en la prueba internacional Trends in International Mathematics and Science Study (TIMSS - Mullis, Martin, Rud dock, Sullivan & Preuschoff, 2009). Cada ítem contaba con un enunciado y cuatro opciones de respuesta con única clave, siguiendo las técnicas de diseño compiladas en Barragán, Bogoya, Contento y Ocaña (2014).

Para la aplicación de junio de 2011, el instrumento de la prueba usó 45 ítems del banco, distribuidos en tres bloques abarcando todos los dominios conceptuales del curso. A partir de ellos se conformaron tres cuadernillos virtuales: el C1 con los bloques 1 y 2, el C2, con

los bloques 1 y 3 y el C3 con los bloques 2 y 3; cada evaluado respondió un cuadernillo de 30 ítems en total. El C1 fue respondido por 175 evaluados, el C2 por 170 y el C3 por 164.

Procedimiento

Las cadenas de respuestas se separaron en tres bases, una por cada cuadernillo de aplicación C1, C2 y C3. Dichas bases se sometieron a un procesamiento inicial mediante la TRI con ayuda del software WINSTEPS 3.73 desarrollado en la Universidad de Chicago por el profesor Jhon Linacre. Con este paquete se obtuvieron las habilidades de los estudiantes, las dificultades de los ítems y los parámetros de los cuadernillos.

Dada la sensibilidad de los cálculos a los órdenes impuestos, se analizaron dos tipos de relaciones de orden para los ítems y dos para los evaluados que se traducen en asignaciones numéricas susceptibles de ser ordenadas de menor a mayor con la relación de orden de los números reales ($a \leq b$). Las relaciones de orden quedaron definidas por cada cuadernillo como sigue:

1.- Para ítems: número de evaluados que contestan correctamente cada ítem (1I) y Dificultad de los ítems (2I) de menor a mayor dificultad.

2.- Para evaluados: número de respuestas correctas por cada evaluado (1E) y Habilidades de los evaluados (2E) de mayor a menor habilidad.

Para el análisis del supuesto de unidimensionalidad se siguió el procedimiento planteado por Tatsuoka, M. (1986) el cual consistió en: (1) considerar las cadenas de respuestas como un grafo bipartito (ya que al conjunto de los vértices se le puede hacer una partición en dos conjuntos disyuntos) que tiene una representación matricial especial que engloba las matrices de dominación de ítems y evaluados. Tales matrices tienen dispuestas sus filas y columnas según el número de respuestas correctas. (2) Calcular el índice de consistencia de Cliff para estas matrices de

dominación. (3) Construir una relación de dominación asimétrica. (4) Extraer conjuntos unidimensionales.

Se reproduce el mencionado procedimiento agregando los órdenes 2I y 2E, lo que se vio potenciado por la información real obtenida con el caso de estudio, así como por capacidad de cálculo y visualización de las herramientas informáticas. Como referencia se incluye el análisis factorial a través de la proporción de la variancia total explicada por las medidas y la explicada por el contraste con el primer factor. Un ejemplo de la implementación del análisis factorial para distinguir las dimensiones latentes se encuentra en Merino y Vilcherrez (2014).

Las matrices fueron estudiadas desde la perspectiva de la Teoría de Grafos con Matlab (2014b) de la compañía MathWorks y para el trabajo de representación y de ejecución de algoritmos específicos de los grafos, se usó Gephi como plataforma de visualización interactiva. Para organización de los datos se empleó EXCEL de la suite de Microsoft Office.

Resultados y discusión

Resultados del Modelo de Rasch

La calibración de las 509 cadenas de respuestas y los 45 ítems proporcionó las dificultades de los ítems en el intervalo [-1,8727;1,8025] y que las habilidades de los evaluados en el intervalo [-4,8265;3,8338].

La Tabla 1 muestra los parámetros obtenidos para cada uno de los cuadernillos. En ella observa que la confiabilidad es mayor a .60 en cada cuadernillo, igual que la consistencia interna, estimada con el coeficiente Alpha de Cronbach. Para los tres cuadernillos todas las correlaciones inter-ítem fueron menores que .50. Así los parámetros presentan valores aceptables considerando los tamaños de las bases procesadas. Intencionalmente se suprimen los parámetros que dan cuenta de la unidimensionalidad, pues estos se analizarán más adelante.

Órdenes para los ítems y para los evaluados

En la Tabla 2 se describen los órdenes en los cuadernillos de acuerdo a su definición.

Para evidenciar la relación entre los órdenes se realizaron los diagramas de dispersión para ítems y evaluados en cada cuadernillo, se encontraron los modelos lineales que los representan. El Gráfico 2 muestra los diagramas para C1 y la Tabla 3 recoge las asociaciones lineales entre los órdenes. La importancia de establecer la relación entre cada par de órdenes radica, por ejemplo en que en C1 se encontraron cinco estudiantes con habilidades iguales a -1,0245; sin embargo, de ellos uno tuvo 7 respuestas correctas, tres tuvieron 8 y uno tuvo 9, fundamentado en que los ítems tienen dificultades diferentes.

Conforme a las cadenas de respuestas, las dificultades y las habilidades para cada cuadernillo se estableció un Escalograma de Guttman. En cada cuadernillo se ordenaron los estudiantes con el Orden 2E y los ítems con el Orden 2I, para disponerlos en un arreglo rectangular en el que las filas contienen las cadenas de respuestas de los evaluados y las columnas, las respuestas a cada ítem; las respectivas entradas se marcaron con 1 si la respuesta era correcta y con 0, en otro caso.

Se esperaba que un evaluado respondiera en forma correcta, con una probabilidad mayor que .50, aquellos ítems cuya dificultad fuera menor que su habilidad; el mismo evaluado debía responder en forma no correcta, también con una probabilidad mayor que .50, los ítems cuya dificultad fuera mayor que la habilidad referida (Bogoya, Barragán, Contento & Ocaña, 2014). La Tabla 4 contiene la primera parte del escalograma del C3 en razón del orden 164x30 con el que cuenta.

Matrices de dominación para ítems y para evaluados

Para cada escalograma se configuró una matriz S_k donde $k=1, 2$ y 3 definida como

$$S_k = (s_{ij})_{n \times m} = \begin{cases} 1, & \text{si el estudiante } i \text{ contestó} \\ & \text{correctamente el ítem } j \text{ del} \\ & \text{cuadernillo } k \\ 0, & \text{en otro caso} \end{cases}$$

S_1, S_2 y S_3 quedaron de tamaño 175 x 30; 170 x 30 y 164 x 30 respectivamente. A partir de dichas matrices se encontraron las matrices complemento $\bar{S}_k = (\bar{s}_{ij})_{n \times m}$

$$\text{tal que } \bar{s}_{ij} = 1 - s_{ij}$$

Ahora con las matrices

$$S_k \text{ y } \bar{S}_k^t \text{ en la que } \bar{S}_k^t$$

representa la matriz transpuesta de \bar{S}_k se definieron tres matrices por bloques A_k donde $k = 1, 2$ y 3 ; los símbolos \mathbf{O} son matrices nulas del tamaño adecuado para que las A_k sean cuadradas.

$$A_1 = \begin{pmatrix} \mathbf{0} & \bar{S}_1^t \\ S_1 & \mathbf{0} \end{pmatrix}_{205 \times 205}$$

$$A_2 = \begin{pmatrix} \mathbf{0} & \bar{S}_2^t \\ S_2 & \mathbf{0} \end{pmatrix}_{200 \times 200}$$

$$A_3 = \begin{pmatrix} \mathbf{0} & \bar{S}_3^t \\ S_3 & \mathbf{0} \end{pmatrix}_{194 \times 194}$$

Las matrices A_k son matrices de adyacencia, que en el caso de estudio son entendidas como matrices de dominación pues cada una de las entradas a_{ij} se interpretan como

$$A_k = (a_{ij})_{rxr} = \begin{cases} 1, & \text{si el ítem (evaluado) } i \\ & \text{domina al evaluado (ítem) } j \\ 0, & \text{en otro caso} \end{cases} \quad (3).$$

Un ítem domina a un evaluado si éste no contesta correctamente el ítem; un evaluado domina un ítem si lo contesta correctamente. Así un ítem no puede ser dominado por otro ítem y un evaluado no puede ser dominado por otro evaluado. Esta interpretación pro-

viene de los bloques que conforman A_k (los bloques, están constituidos por las matrices S_k y \bar{S} y las dos matrices nulas).

Con ayuda de Gephi se representaron los grafos correspondientes. Como ejemplo, la Figura 2 muestra el grafo asociado a la matriz A_2 del C2 que cuenta con 200 vértices y 5.100 aristas. Para una visualización mejor se ejecutaron diferentes algoritmos de los grafos. En particular, se empleó el algoritmo Force Atlas que no tiene condición de parada y que analiza cada nodo ubicándolo donde sea oportuno (Jiménez, 2014). Con una clasificación de los vértices por el grado de entrada, se aplicó una distribución de atracción la que envía los vértices *Hub* hacia la periferia y los vértices *Authority* hacia el centro. En atención a lo cual los ítems y los evaluados con menos respuestas correctas se observan en la periferia. Así, el ítem 13 es un *Hub* ya que domina a todos los evaluados en el C2. Aunque el dominio conceptual por el que indagaba era la radicación, el ítem 13 resultó ser el ítem más difícil del C2 con dificultad 1,0624. En estudios previos se había mostrado que dicho ítem presentaba funcionamiento diferencial moderado con respecto a las variables de contexto: sexo, rango de edad, naturaleza jurídica del colegio de procedencia y facultad de inscripción (Bogoya, Ocaña, Barragán & Contento, 2014). Por otra parte, el ítem 12 es un nodo *Authority* porque tiene la mayor cantidad de respuestas correctas y la menor de las dificultades, una explicación posible es que el ítem indaga por una aplicación básica de la regla de tres simple. Además, el evaluado 193 es un *Hub* en vista de que obtuvo 4/30 respuestas correctas y tiene una habilidad de -1,856, con lo que fue de los más dominados por los ítems. En contraste con el evaluado 65 que contestó correctamente 27/30 y que fue levemente dominado por los ítems.

Para emplear el Teorema de Festinger se calculó A_k^2 , el cuadrado de la matriz A_k cuyas entradas se interpretan como el número de formas en que i accede a j , en 2-cadenas.

Se obtuvo

$$A_1^2 = \begin{pmatrix} \bar{S}_1^t S_1 & \mathbf{0} \\ \mathbf{0} & S_1 \bar{S}_1^t \end{pmatrix}_{205 \times 205}$$

$$A_2^2 = \begin{pmatrix} \bar{S}_2^t S_2 & \mathbf{0} \\ \mathbf{0} & S_2 \bar{S}_2^t \end{pmatrix}_{200 \times 200} \text{ y}$$

$$A_3^2 = \begin{pmatrix} \bar{S}_3^t S_3 & \mathbf{0} \\ \mathbf{0} & S_3 \bar{S}_3^t \end{pmatrix}_{194 \times 194}$$

En consecuencia, se pueden apreciar los dos aspectos siguientes:

1.-

$\bar{S}_k^t S_k$ es la matriz de dominación de los ítems: la entrada ij indica el número de evaluados que son dominados por el ítem i y que a su vez dominan el ítem j . Para el presente caso,

$\bar{S}_k^t S_k$ para $k=1, 2$, y 3 son de orden 30×30 . La Tabla 5 muestra la parte superior de la matriz,

$\bar{S}_1^t S_1$ donde se ha encuadrado la entrada 1,9 cuyo valor es 22. El valor de la entrada precisa que 22 evaluados no contestaron correctamente el ítem 12 que está ubicado en la posición 1 con dificultad -1,8727, en el Orden 2I; estos mismos 22 evaluados contestaron correctamente el ítem 2, que se encuentra en la posición 9 pues tiene una dificultad de -0,5785. Esta es una situación importante para examinar en términos de la consistencia pues no resulta natural pensar que estos 22 estudiantes no contestan correctamente el ítem más fácil de este cuadernillo, pero si contestan bien, uno más difícil.

2.-

$S_k \bar{S}_k^t$ es la matriz de dominación de los evaluados: aquí la entrada ij muestra el número de ítems i que son dominados por el evaluado j , mismos ítems que dominan al evaluado j .

$$S_1 \bar{S}_1^t \quad S_2 \bar{S}_2^t \quad S_3 \bar{S}_3^t$$

son matrices cuadradas de órdenes 175, 170 y 164, respectivamente. La Tabla 6 tiene re-

saltada la entrada 1,11, con un valor de 6. Esta entrada revela que el evaluado de mayor habilidad (2,6248) de los 175 que presentaron el C1 contestó correctamente 6 ítems que no fueron bien respondidos por el evaluado en la posición 11 cuya habilidad es 1,3285. Los evaluados fueron organizados con el Orden 2E.

Para continuar el proceso en el análisis se reorganizaron las filas y las columnas de las matrices

$$\bar{S}_k^t S_k \text{ y } S_k \bar{S}_k^t$$

mediante los órdenes 1I, 2I y 1E, 2E respectivamente. En la Tabla 7 se renombran las matrices después de las ordenaciones.

Cálculo del índice de consistencia de Cliff

El índice de consistencia de Cliff (1977) permite estimar la unidimensionalidad como la homogeneidad escalar. Dicho índice se caracteriza porque ser una función del orden de los ítems (Tatsuoka, K. 2009) y por pertenecer al intervalo $[-1, 1]$ donde -1 representa la inconsistencia perfecta y 1 la consistencia perfecta (Tatsuoka, M. 1986).

$$c_{ti} = \frac{\Sigma_{i < j} (\bar{S}_k^t S_k)_{ij} - \Sigma_{i > j} (\bar{S}_k^t S_k)_{ij}}{\Sigma_{i \neq j} (\bar{S}_k^t S_k)_{ij}} \quad (2)$$

En la Tabla 8 se resumen los cálculos c_{ti} de los índices de consistencia de Cliff en las diferentes combinaciones de cuadernillos $k = 1, 2$ y 3 y de órdenes $pI = II$ y $2I$

De la misma forma se calculó el índice de consistencia de Cliff para las matrices de dominación de los evaluados en las combinaciones de cuadernillos $k = 1, 2$ y 3 y de los órdenes $pE = 1E$ y $2E$ (ver Tabla 9).

En la Tabla 8 se observa que los ítems de los tres cuadernillos con cualquiera de los dos órdenes son moderadamente inconsistentes y en la Tabla 9 se lee que los evaluados con los tres cuadernillos con cualquiera de los dos órdenes, son moderadamente consistentes.

Por otro lado, en la Tabla 10 se muestran las relaciones entre las variancias explicadas por el contraste con el primer componente y por las medidas para los tres cuadernillos. Es importante tener en cuenta que a mayor número de cadenas de respuestas, el valor del error de la estimación disminuye. El indicador estadístico para la relación entre variancias tiene un valor de aprobación de 4,00 (Linnacre, 2008), por eso se observa que para C1 y C3 se confirma moderadamente la unidimensionalidad en el sentido unifactorial, pero para el C2 no se puede afirmar lo mismo.

Subconjuntos unidimensionales de ítems y de evaluados

Para extraer subconjuntos de ítems que sean más fuertes desde el punto de vista de la unidimensionalidad considerada como un orden fundamentado en la dominación, se realizó el estudio de estructuras jerárquicas con base en matrices de dominación significativa como en Tatsuoka (Tatsuoka, M., 1986).

Las matrices de dominación $N_{Ck,PI}$ y $N_{Ck,PE}$ (ver Tabla 7) no corresponden a la definición de unidimensionalidad porque representan una relación de dominación en ambos direcciones ya que los términos ij y ji en general son distintos de cero.

Para iniciar la jerarquización de la relación de dominación para ítems y para evaluados es necesario que las relaciones de dominación sean asimétricas; es así que se define que el ítem i domina al ítem j solo si la entrada $(N_{Ck,PI})_{ij}$ es significativamente más grande que la entrada $(N_{Ck,PI})_{ji}$ con la prueba de McNemar para la diferencia entre proporciones correlacionadas (Tatsuoka, M., 1986). De forma análoga se define para los evaluados.

La prueba de McNemar es un test de significancia enunciando que si

$$z_{ij} = \frac{(n_{ij} - n_{ji})}{\sqrt{n_{ij} + n_{ji}}}$$

y si el cuadrado de z_{ij} excede el valor del estadístico χ^2 con un grado de libertad, existe diferencia

significativa entre las entradas ij y ji (Agresti, 1990). Esto significa aplicar la ecuación (3) para los ítems y la (4) para evaluados, para cada par ij en donde $i < j$.

$$z_{ij} = \frac{(N_{Ck,pI})_{ij} - (N_{Ck,pI})_{ji}}{\sqrt{(N_{Ck,pI})_{ij} + (N_{Ck,pI})_{ji}}} \quad (3)$$

$$z_{ij} = \frac{(N_{Ck,pE})_{ij} - (N_{Ck,pE})_{ji}}{\sqrt{(N_{Ck,pE})_{ij} + (N_{Ck,pE})_{ji}}} \quad (4).$$

De esta forma se calcularon las matrices Z correspondientes a las matrices de dominación $N_{Ck,pI}$ para ítems y $N_{Ck,pE}$ para evaluados; a partir de ellas se reconoció y marcó la dominación entre ítems y entre evaluados con un nivel de significancia de 0,05 para lo cual se definieron las matrices $N_{Ck,pI}^*$ y $N_{Ck,pE}^*$ en las que $(N_{Ck,pI}^*)_{ij} = 1$ si el respectivo z_{ij} era mayor a 3,8414 y 0 en otro caso. Análogamente se hizo para las entradas de $N_{Ck,pE}^*$.

A manera de ilustración, la Figura 3 muestra partes de dos matrices triangulares superiores: en la izquierda, la matriz Z asociada a $N_{C3,II}$ que es la matriz de dominación de los ítems luego de la reordenación de los mismos de acuerdo al orden 1E; en la derecha aparece la matriz $N_{C3,II}^*$ que es la matriz de dominación significativa de los ítems. Dos entradas aparecen encuadradas y en negrilla para mostrar la asignación.

Como la unidimensionalidad es entendida como un orden basado en la relación de dominación significativa se hace necesario determinar la estructura jerárquica de los ítems con lo que aunado al índice de Cliff quedan establecidos los subconjuntos unidimensionales. Para la jerarquización de esta relación, a $N_{Ck,pI}^*$ y $N_{Ck,pE}^*$ se les adicionó la matriz identidad I_n del tamaño compatible, de modo que se calcularon las matrices $R_{Ck,pI}$ y $R_{Ck,pE}$ como en (5) y (6)

$$R_{Ck,pI} = N_{Ck,pI}^* + I_{30} \quad (5)$$

$$R_{C1,pE} = N_{C1,pE}^* + I_{175}; R_{C2,pE} = N_{C2,pE} + I_{170}$$

$$\text{y } R_{C3,pE} = N_{C3,pE}^* + I_{164} \quad (6)$$

Haciendo uso de Matlab, a cada una de estas matrices se les calcularon las sucesivas potencias Booleanas, formando una sucesión de matrices. La matriz para la cual esta sucesión converge es llamada matriz de accesibilidad. La sucesión genérica se muestra en la ecuación (7)

$$R_{Ck,pI}^1 b, R_{Ck,pI}^2 b, R_{Ck,pI}^3 b, R_{Ck,pI}^4 b, \dots \rightarrow P_{Ck,pI} \quad (7).$$

Los elementos no nulos de esta matriz límite P indican que los ítems correspondientes son accesibles uno por el otro. Esto fue interpretado como una dominación entre ítems que puede ser directa si es en una sola etapa (si la sucesión se estabiliza en la primera potencia booleana), o indirecta si involucra otros ítems (tantos como diga la potencia booleana en la que la sucesión se estabiliza).

A saber, la esquina superior izquierda de la matriz de dominación significativa de los ítems del C1 ordenados por dificultad $N_{c1,2l}$ está en la Figura 4 y la esquina superior izquierda del límite de la sucesión de las potencias Booleanas de $R_{c1,2l} = N_{c1,2l}^* + I_{30}$ aparece en la Figura 5. Esta última es entendida como la matriz de accesibilidad de los ítems del C1. La sucesión se estabilizó en la primera potencia booleana, lo que permitió afirmar que la entrada encuadrada en la Figura 7, muestra que el ítem 24 domina significativamente al ítem 8 de forma directa. Esto se ratifica con la entrada encuadrada en la Figura 6.

Para jerarquizar la relación de dominación significativa se ejecutaron dos procedimientos: en el primero se revisaron las entradas de las matrices de accesibilidad para identificar los ítems que dominan y son dominados; en el segundo se calcularon, mediante Gephi, los grados de entrada y salida para cada uno de los vértices de correspondientes a las matrices de accesibilidad. Para el primer procedimiento se revisaron las en-

tradas de las matrices de accesibilidad $P_{ck,pl}$ teniendo en cuenta que son matrices triangulares superiores y que si la entrada x_{ij} es igual a 1 indica que el ítem x_i domina significativamente al ítem x_j . Es por esto que para la matriz de accesibilidad $P_{c1, 2l}$ para $N_{c1, 2l}$ se encontró que la estructura jerárquica de la dominación significativa de los ítems que figuraban en el C1 es la que aparece en la Tabla 11. Aquí se puede identificar que los ítems 29 y 24 encabezan la jerarquía porque dominan significativamente a los demás y porque entre ellos no se dominan (la entrada 1,2 de $P_{c1, 2l}$ es 0). Asimismo, los ítems que son dominados significativamente por todos los otros ítems son el 26 y el 12.

En el segundo procedimiento a la matriz de accesibilidad para $N_{c1, 2l}$ de la Figura 5, se le realizó la jerarquización mediante el grado de los vértices como aparece en la Figura 6, el tamaño del nodo y la intensidad del color negro la representan. Para efectos de visualización y de mejor distribución se aplicó el algoritmo Force Atlas con una fuerza de repulsión de 10.000 y ajustado por tamaños de acuerdo al grado.

Al comparar con el mapa de ítems obtenido con WINSTEPS 3.73, se observó una estructura similar como lo muestran las Figuras 7 y 8. La posición en el mapa de ítems se interpretó como en el siguiente ejemplo: en el C1, dado que un evaluado contesta correctamente el ítem 17, la probabilidad de contestar correctamente el ítem 22 es menor. En la base completa para el C1, se encontró que 32/175 evaluados respondieron bien el ítem 22 pero no el ítem 17 en contraste con los 53/175 evaluados que contestaron correctamente el ítem 17 y no el ítem 22.

Conclusiones

El uso combinado de la Teoría de Grafos y la TRI representa una alternativa metodológica y técnica para evaluar de forma integral a los estudiantes de Educación Superior lo que posibilita que el mismo evaluado y la IES identifiquen el nivel de desempeño académico individual. Con este conocimiento el evaluado puede reconocer el grado de cumplimiento de los objetivos de formación y proponerse un plan para lo que le haga falta. A su vez, la IES puede optimizar sus recursos y ofrecer opciones académicas diferenciadas con base en la escala de habilidad y en la posición individual.

Si bien la implementación de la Teoría de Grafos y la TRI requiere conocimientos específicos y manejo de software especializado, su aplicación en análisis de instrumentos de evaluación ofrece aspectos detallados de la prueba, de los ítems y de los evaluados apoyándose en la Psicometría. Todo esto fortalece los sistemas de evaluación de los estudiantes, sistemas que están previstos como una característica para los programas de pregrado en los lineamientos para acreditación de alta calidad que son una de las formas vigentes en que el Estado Colombiano materializa las políticas públicas sobre calidad de la educación.

Herramientas informáticas como WINSTEPS 3.73, Matlab (2014b) y Gephi apoyan en el procesamiento de los datos garantizando eficiencia y eficacia en la obtención de conclusiones e inferencias.

Tabla 1
Parámetros para cada cuadernillo

Cuadernillo	Confiabilidad	Separación	Coeficiente <i>Alfa de Cronbach</i>	Intervalo de variación de la correlación entre dos ítems
C1	0,72	1,52	0,74	[-0,241, 0,206]
C2	0,79	1,96	0,78	[-0,286, 0,283]
C3	0,82	2,16	0,83	[-0,276, 0,405]

Tabla 2
Órdenes e intervalos para cada cuadernillo

Cuadernillo	Intervalos de oscilación		Orden 1E	Orden 2E
	Orden 1I	Orden 2I		
	Número de evaluados que contestan correctamente cada ítem	Dificultad de los ítems	Número de respuestas correctas por cada evaluado	Habilidades de los evaluados
C1	[27,141]	[-1,8727, 1,8025]	[1,27]	[-4,8265, 2,6248]
C2	[0,147]	[-1,8727, 1,5858]	[3,28]	[-2,2076, 3,8338]
C3	[30,122]	[-1,2480, 1,8025]	[1,28]	[-3,4537, 3,0621]

Tabla 3
Comparativo entre los órdenes para ítems y para evaluados

Modelos lineales para los comparativos					
Orden 2I vs Orden 1I			Orden 2E vs Orden 1E		
<i>y</i> =		<i>r</i> =	<i>y</i> =	<i>r</i> =	
C1	$-34,241x + 86,376$	0,983	$5,0417x + 14,154$	0,969	
C2	$-35,543x + 83,218$	0,936	$5,1868x + 13,214$	0,983	Positivas
C3	$-31,352x + 81,903$	0,971	$5,5695x + 14,009$	0,991	y fuertes

Tabla 4
Parte inicial del escalograma de gutman para el c3

		Dificultad												
		-1,248												
		-0,915												
		-0,884												
		-0,716												
		-0,713												
		-0,701												
		-0,548												
		-0,460												
		-0,344												
		-0,33												
		-0,115												
		CLAVE	A	B	D	B	C	A	A	B	B	C	D	
Estu.	Hab.		It26	It18	It27	It30	It32	It20	It40	It44	It43	It38	It17	
376	3,062	1	1	1	1	1	1	1	1	1	1	1	1	1
302	2,596	1	1	1	1	1	1	1	1	1	0	1	1	1
308	2,596	1	1	1	1	1	1	1	1	1	1	1	1	0
234	2,248	1	1	1	1	1	1	1	1	1	1	1	1	1
248	2,248	1	1	1	1	1	1	1	1	1	1	1	1	1
296	2,248	1	1	1	1	1	1	1	1	1	1	1	1	1
286	1,963	1	1	1	1	1	0	1	1	1	1	1	1	1
333	1,963	1	1	1	1	1	0	1	1	1	1	1	1	1
221	1,719	1	1	1	1	1	1	1	1	1	1	1	1	1
254	1,719	1	1	1	1	1	1	1	1	1	0	0	0	1
343	1,719	1	1	1	0	1	0	1	0	1	1	1	0	0
240	1,501	1	1	1	1	1	1	1	1	1	1	1	1	1
241	1,501	1	1	1	1	1	1	1	1	1	1	1	1	1
275	1,501	1	1	1	1	1	1	1	1	1	1	1	1	1

Estu.=Estudiante, Hab.= Habilidad

Tabla 5
Esquina superior izquierda de la matriz $\bar{S}_1^t S_1$ de dominación de los ítems del C_1

0	18	22	12	16	16	17	16	22	14	9	14	16	15	7	10	12	8
30	0	24	22	20	20	16	21	28	19	17	14	16	15	15	21	17	13
44	34	0	36	31	33	33	37	24	25	16	22	24	14	17	21	25	22
43	41	45	0	47	41	33	36	35	29	24	29	29	29	20	31	28	15
38	30	31	38	0	31	26	33	33	24	21	16	24	18	20	18	20	21
41	33	36	35	34	0	29	35	38	26	20	22	20	20	16	18	22	18
47	34	41	32	34	34	0	31	37	28	26	18	19	18	18	22	27	18

Tabla 6

Esquina superior izquierda de la matriz $S_1 \bar{S}_1^t$ de dominación de los ítems del C_1

0	4	3	5	4	5	5	6	4	7	6	5	7	7	7	7	9	8	8	9	9	
2	0	3	4	3	5	4	5	5	6	6	4	6	7	5	7	6	6	7	8	9	9
1	3	0	6	3	6	4	7	4	6	5	5	6	8	6	8	7	8	6	8	10	9
2	3	5	0	3	3	4	3	4	6	5	4	5	5	5	4	5	6	7	6	6	9
1	2	2	3	0	4	3	5	3	6	4	4	5	6	5	6	5	7	6	7	9	10
1	3	4	2	3	0	3	2	4	6	2	4	4	3	5	4	3	6	7	6	7	8
2	3	3	4	3	4	0	5	5	7	4	5	6	5	5	6	6	7	5	9	9	9

Tabla 7

Nuevos nombres de las matrices de dominación de los ítems y de los evaluados después de la reordenación

Matri z	Nombre de la matriz luego de la reordenación de los ítems		Matri z	Nombre de la matriz luego de la reordenación de los evaluados	
	Orden 1I	Orden 2I		Orden 1E	Orden 2E
$\bar{S}_1^t S_1$	$N_{C1,1I}$	$N_{C1,2I}$	$S_1 \bar{S}_1^t$	$N_{C1,1E}$	$N_{C1,2E}$
$\bar{S}_2^t S_2$	$N_{C2,1I}$	$N_{C2,2I}$	$S_2 \bar{S}_2^t$	$N_{C2,1E}$	$N_{C2,2E}$
$\bar{S}_3^t S_3$	$N_{C3,1I}$	$N_{C3,2I}$	$S_3 \bar{S}_1^t$	$N_{C3,1E}$	$N_{C3,2E}$

Tabla 8

Índice de consistencia de Cliff para las matrices de dominación de los ítems

Cuadernillo	Intervalos de oscilación			
	Orden 1I	Orden 2I	Orden 1E	Orden 2E
	Número de evaluados que contestan correctamente cada ítem	Dificultad de los ítems	Número de respuestas correctas por cada evaluado	Habilidades de los evaluados
C1	[27,141]	[-1,8727, 1,8025]	[1,27]	[-4,8265, 2,6248]
C2	[0,147]	[-1,8727, 1,5858]	[3,28]	[-2,2076, 3,8338]
C3	[30,122]	[-1,2480, 1,8025]	[1,28]	[-3,4537, 3,0621]

Tabla 9
Índice de consistencia de Cliff para las matrices de dominación de los evaluados

	$\sum_{i < j} (N_{Ck, pE})_{ij}$	$\sum_{i > j} (N_{Ck, pE})_{ij}$	$\sum_{i \neq j} (N_{Ck, pE})_{ij}$	c_{t1}
	Orden 1E	Orden 2E	Orden 1E	Orden 2E
C ₁	141.389	140.998	59.309	59.700
C ₂	134.905	134.765	53.557	53.697
C ₃	131.209	131.209	44.524	44.524
			Suma total	Orden 1E
			200.698	0,409
				Orden 2E
				0,405
				0,431
				0,430
				0,493
				0,493

Tabla 10
Relación entre las varianzas para los cuadernillos

	Porcentaje de varianza explicada		Relación entre varianzas
	Por las Medidas	Contraste con el primer componente	
C1	23,3	5,5	4,236
C2	22,4	6,5	3,446
C3	25,1	5,4	4,641

Tabla 11
Jerarquía por dominación significativa para la matriz de Accesibilidad $P_{C1, 2l}$ para $N_{C1, 2l}^*$

Niveles	Ítems
1	I29, I24
2	I6
3	I11
4	I13, I8, I10, I4
5	I1, I28, I22, I25, I28, I22
6	I9
7	I19, I17
8	I16
9	I3
10	I2
11	I20, I30, I15, I7, I27, I18
12	I26, I12

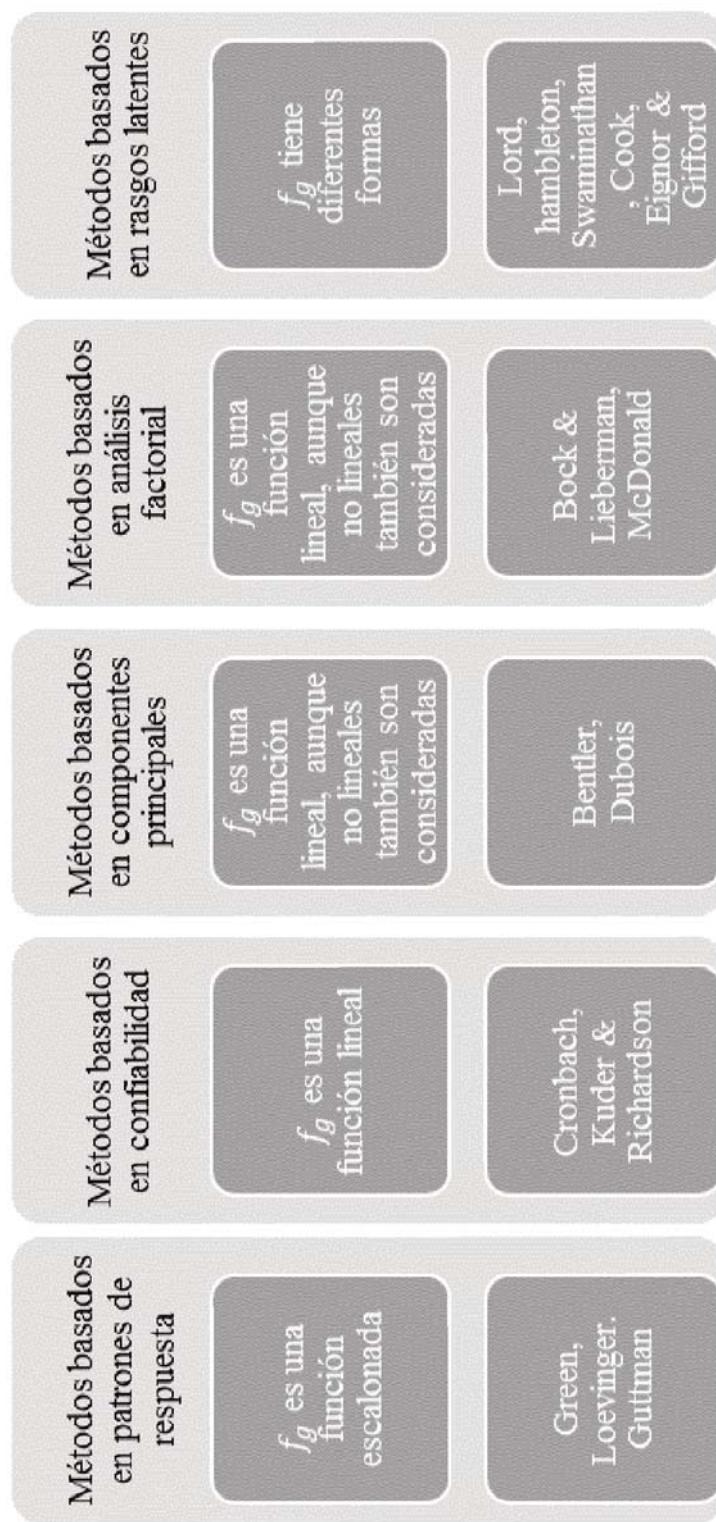


Figura 1. Descripción de los métodos empleados para estudiar la unidimensionalidad

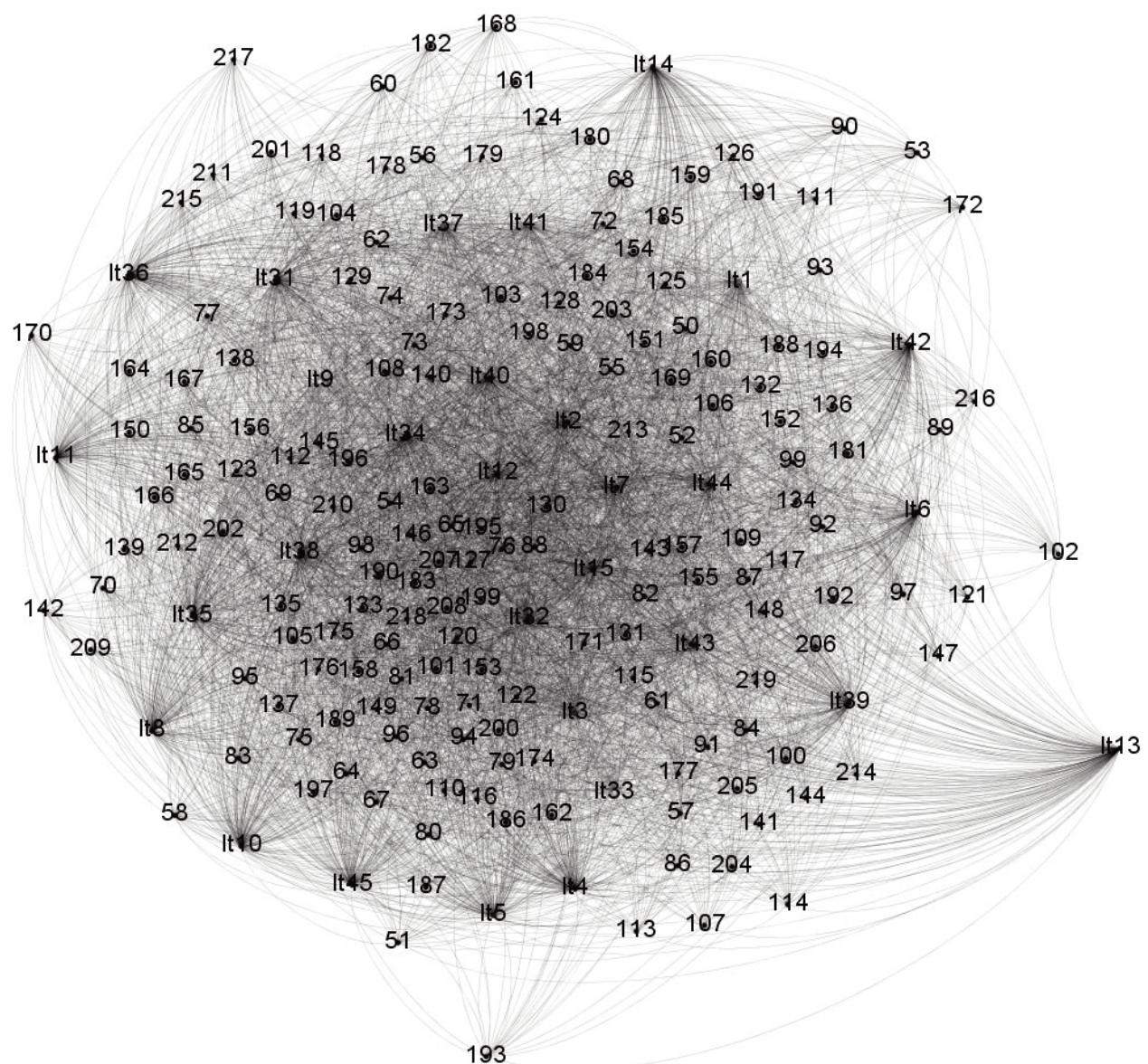


Figura 2. Grafo asociado a la matriz de dominación A2

0,00	-0,32	-1,73	-1,78	-2,56	-3,00	-3,56	-3,94	-4,26	-4,53	-4,48	-4,57	-5,25	-5,17	-7,55	-5,66	-6,06	-6,67	-6,65	-6,56	-7,07	-7,09	
0,00	0,00	-1,41	-1,58	-2,38	-2,43	-3,36	-3,81	-3,98	-3,81	-4,23	-4,19	-4,70	-4,59	-7,42	-5,68	-5,74	-6,04	-6,69	-6,04	-6,74	-6,54	
0,00	0,00	-0,25	-0,96	-1,17	-2,00	-2,22	-2,25	-2,32	-2,73	-2,83	-3,35	-3,40	-7,35	-4,06	-4,15	-4,72	-5,28	-5,10	-5,24	-5,53		
0,00	0,00	0,00	0,62	-0,83	-1,63	-2,03	-2,11	-2,18	-2,46	-2,59	-3,08	-3,21	-7,07	-3,94	-3,93	-4,78	-4,86	-5,06	-4,90	-5,58		
0,00	0,00	0,00	0,00	-0,25	-1,15	-1,22	-1,50	-1,52	-1,86	-2,08	-2,79	-2,52	-6,86	-3,12	-3,58	-4,20	-4,28	-4,59	-4,61	-4,69		
0,00	0,00	0,00	0,00	0,00	-0,93	-1,12	-1,21	-1,34	-1,81	-1,81	-2,50	-2,46	-7,07	-3,01	-3,30	-4,13	-4,33	-4,42	-4,50	-4,64		
0,00	0,00	0,00	0,00	0,00	0,00	-0,25	-0,38	-0,52	-0,82	-1,02	-1,72	-1,61	-6,93	-2,12	-2,49	-3,39	-3,69	-3,44	-3,67	-4,04		
0,00	0,00	0,00	0,00	0,00	0,00	0,00	-0,13	-0,28	-0,70	-0,70	-1,31	-1,41	-6,56	-2,18	-2,54	-3,30	-3,45	-3,22	-3,87	-4,03		
0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	-0,14	-0,55	-0,63	-1,25	-1,34	-6,32	-2,06	-2,61	-3,20	-3,58	-3,35	-3,49	-4,35		
0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	-0,38	-0,48	-1,12	-1,23	-6,71	-1,81	-2,15	-3,01	-3,30	-3,20	-3,61	-3,73		
0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	-0,13	-0,74	-0,82	-6,56	-1,46	-1,86	-2,56	-2,81	-2,83	-3,15	-3,25		
0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	-0,61	-0,69	-6,40	-1,13	-1,57	-2,39	-2,52	-2,56	-2,72	-2,93		
0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	-0,11	-6,78	-0,61	-1,05	-1,89	-2,05	-2,18	-2,43	-2,49	
0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	-0,08	-0,48	-0,92	-1,54	-1,94	-1,91	-2,15	-2,46		

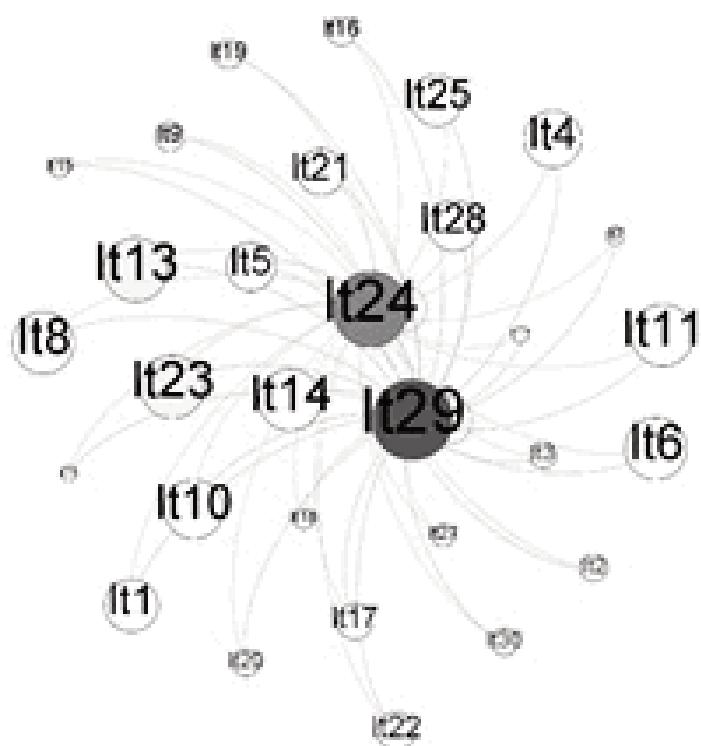
Figura 3A. Matriz $N_{C3,II}$ de la matriz Z asociada a $N_{C3,II}$

Figura 3. Matriz $N_{C3,11}$ de la matriz Z asociada a $N_{C3,11}$

	It29	It24	It14	It23	It6	It11	It13	It8	It10	It4	It25	It1	It28
It29	0	0	1	1	1	1	1	1	1	1	1	1	1
It24	0	0	1	0	0	1	1	1	1	1	1	1	1
It14	0	0	0	1	0	0	0	0	0	0	1	1	1
It23	0	0	0	0	0	0	1	1	1	1	1	1	1
It6	0	0	0	0	0	0	1	1	1	1	1	1	1
It11	0	0	0	0	0	0	0	0	0	0	1	1	1
It13	0	0	0	0	0	0	0	0	0	0	1	0	0
It8	0	0	0	0	0	0	0	0	0	0	1	0	0

Figura 4. Matriz $N_{C1,2I}$

	It29	It24	It14	It23	It6	It11	It13	It8	It10	It4	It25	It1	It28
It29	1	0	1	1	1	1	1	1	1	1	1	1	1
It24	0	1	1	1	0	1	1	1	1	1	1	1	1
It14	0	0	1	1	0	0	1	1	1	1	1	1	1
It23	0	0	0	1	0	0	1	1	1	1	1	1	1
It6	0	0	0	0	1	0	1	1	1	1	1	1	1
It11	0	0	0	0	0	1	0	0	0	0	1	1	1
It13	0	0	0	0	0	0	1	0	0	0	1	0	0
It8	0	0	0	0	0	0	0	1	0	0	1	0	0

Figura 5. Matriz de accesibilidad para $N_{C1,21}$ Figura 6. Jerarquía por grado para la matriz de accesibilidad para $N_{C1,21}$

Niveles	Ítems
1	It29
2	It24
3	It13, It23
4	It8, It6, It11, It14
5	It10
6	It4
7	It1
8	It28, It25
9	It5
10	It21
11	It22
12	It17
13	It16, It19, It9
14	It30, It20, It2, It3
15	It27
16	It18, It15
17	It7
18	It12, It26

Figura 7. Jerarquía encontrada

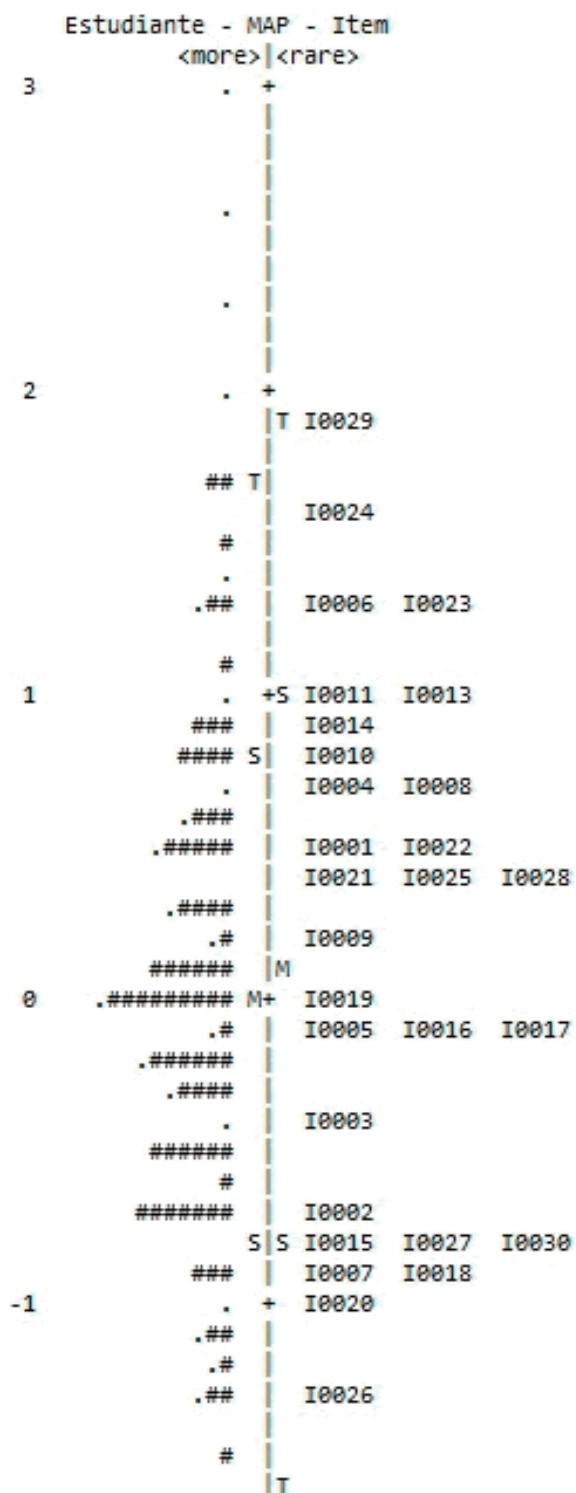


Figura 8. Mapa de ítems de WINSTEPS 3.73

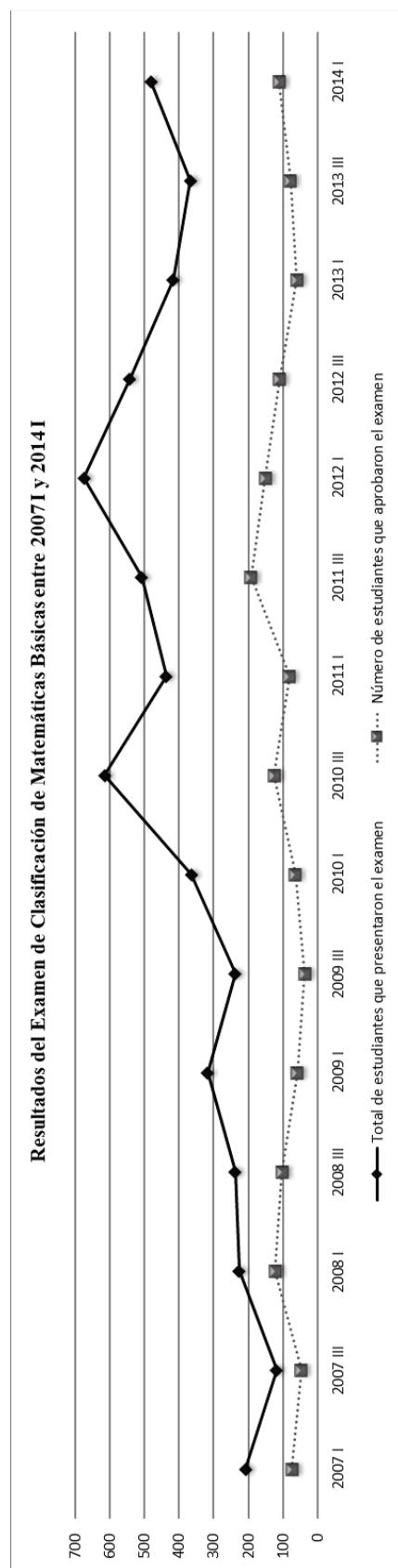


Gráfico 1. Resultados del Examen de Clasificación de Matemáticas Básicas

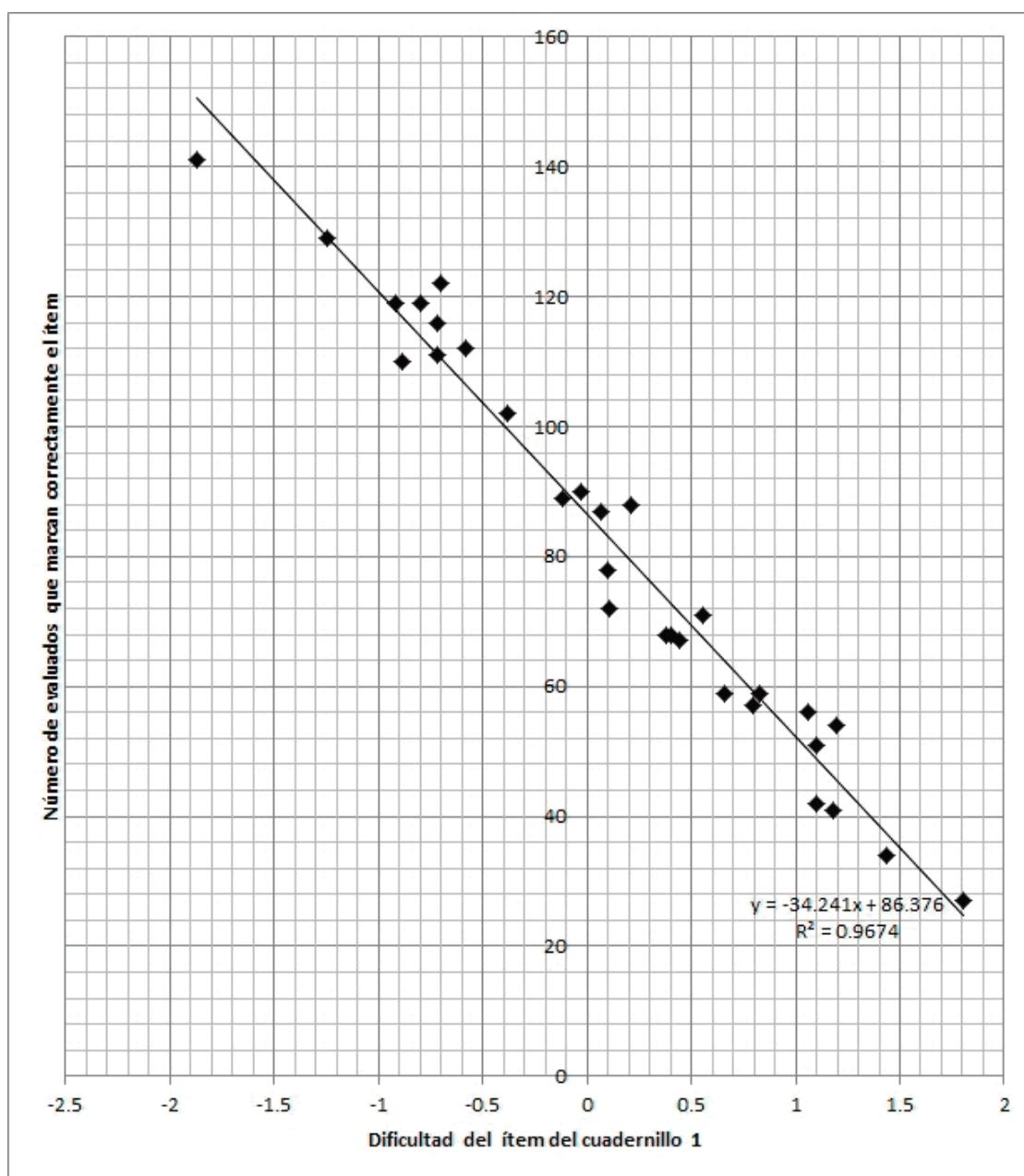


Gráfico 24. Diagramas de dispersión para las relaciones en los órdenes para ítems y evaluados para C_I

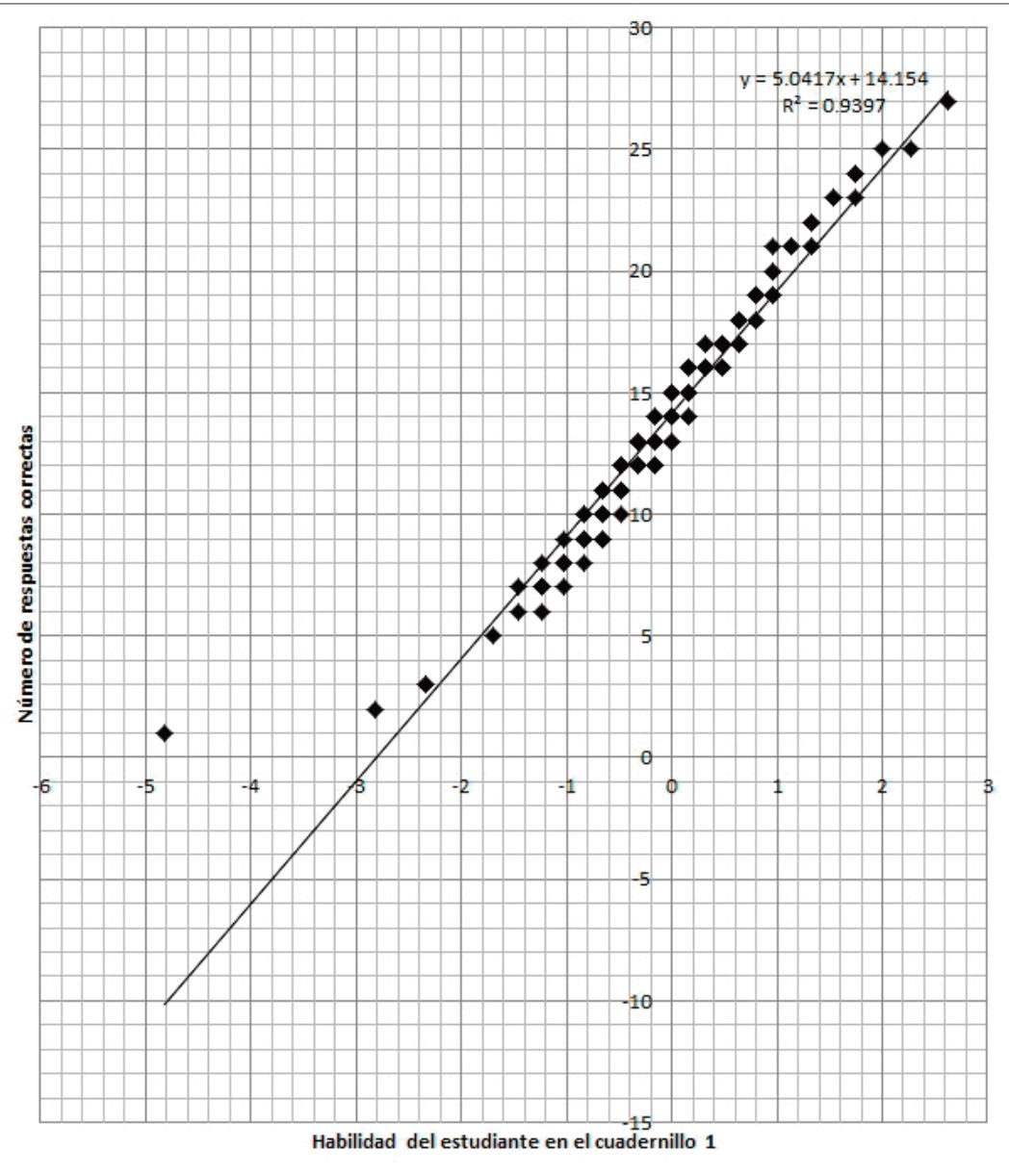


Gráfico 2B. Diagramas de dispersión para las relaciones en los órdenes para ítems y evaluados para C_1

Referencias bibliográficas

Abraham, M.V. & Brenca, R.M. (2014). Análisis psicométrico de la evaluación del aspecto pragmático del lenguaje infantil: Batería ICRA-A [Psychometric analysis of child language pragmatic aspect evaluation: ICRA-A Battery]. *Interdisciplinaria*, 31(1), 139-161. <https://doi.org/10.16888/interd.2014.31.1.9>

Agresti, A. (1990). *Categorical data analysis*. New York: John Wiley & sons.

Barragán, S., Bogoya, D., Contento, M. & Ocaña, A. (2014). *Una aproximación a la construcción de ítems para pruebas en matemáticas* [An approach to the construction of items for math testing]. Recuperado el 8 de febrero de 2017 de http://avalon.utadeo.edu.co/servicios/ebooks/un_a_aproximacion_a_la_construccion_de_items/

Bogoya, D. (2006). Evaluación educativa en Colombia [Educational evaluation in Colombia]. *Memorias del Seminario Internacional de Evaluación* (pp. N1-N27). Cartagena: ICFES.

Bogoya, D., Barragán, S., Contento, M. & Ocaña, A. (2014). Calibración de instrumentos de evaluación - clasificación de matemáticas en la Universidad Jorge Tadeo Lozano [Calibration of assessment instruments - classification in mathematics at the Universidad Jorge Tadeo Lozano]. *Revista Complutense de Educación*, 25(2), 501-519. http://dx.doi.org/10.5209/rev_RCED.2014.v25.n2.41931

Bogoya, D., Ocaña, A., Barragán, S. & Contento, M. (2014). Funcionamiento Diferencial de Ítemes: Examen de matemática - Universidad Jorge Tadeo Lozano [Differential item functioning: math test Universidad Jorge Tadeo Lozano]. *Interdisciplinaria*, 31(1), 121-138. <https://doi.org/10.16888/interd.2014.31.1.8>

Bondy, J. & Murty, U. (1976). *Graph theory with applications*. USA: Elsevier Science Publishing.

Cliff, N. (1977). A theory of consistency of ordering generalizable to tailored testing. *Psychometrika*, 42(3), 375-399. <https://doi.org/10.1007/BF02293657>

DeMars, C. (2010). *Item Response Theory* (Understanding Statistics: Measurement). New York: Oxford University Press.

Festinger, L. (1949). The analysis of sociograms using matrix algebra. *Human Relations*, 153-158. <https://doi.org/10.1177/001872674900200205>

Flament, C. (1972). *Teoría de grafos y estructuras de grupo* [Graphs theory and group structures]. Madrid: Tecnos.

Gross, J. & Yellen, J. (2006). *Graph theory and its applications*. Boca Raton: Chapman & Hall / CRC Taylor & Francis Group.

Grossman, S. (1988). *Aplicaciones de álgebra lineal* [Applications of linear algebra]. México: Grupo Editorial Iberoamérica.

Hattie, J. (1984). An empirical study of various indices for determining unidimensionality. *Multivariate Behavioral Research*, 19, 49-78.

Hattie, J. (1985). Methodology Review: Assessing unidimensionality of test and items. *Applied Psychological Measurement*, 9(2), 139-164.

Hambleton, R. (1990). Item response theory: Introduction and bibliography. *Psicothema*, 2(1), 97-107.

Hambleton, R., Swaminathan, H. & Rogers, J. (1991). *Fundamentals of item response theory*. Newbury Park: SAGE. http://dx.doi.org/10.1207/s15327906mbr1901_3

ICFES. (2014). *Instituto Colombiano para la Evaluación de la Educación* [Colombian Institute for Educational Evaluation]. Recuperado el 11 de marzo de 2016, de <http://www.icfes.gov.co/index.php/institucion/es-educativas/acerca-de-las-evaluaciones/como-se-califican-las-pruebas>

Jiménez, M. (2014). *Ánalisis de comunidades científicas basadas en fuentes de datos online* [Analysis of scientific communities based on online data sources]. (Trabajo fin de grado no publicado). Madrid: Universidad Autónoma de Madrid.

Linacre, J.M. (2008). *A user's guide to Winsteps*. Chicago: John M. Linacre.

Merino, C. & Vilcherrez, L. (2014). Evaluación psicométrica de un instrumento de deseabilidad social para niños [Psychometric assessment of a measure for social desirability for children]. *Interdisciplinaria*, 31(1), 107-120.
<https://doi.org/10.16888/interd.2014.31.1.7>

Mullis, I., Martin, M., Ruddock, G., Sullivan, C. & Preuschoff, C. (2009). *Timss 2011. Assessment Frameworks*. Boston: Lynch School of Education, Boston College.

Muñiz, J. (2010). Las teorías de los test: Teoría Clásica y Teoría de Respuesta a los Items [[The theories of tests: Classic Theory and Response Item Theory]. *Papeles del Psicólogo*, 57-66.

OECD. (2009). *PISA 2006 Technical Report*. París: OECD.

Olson, J., Martin, M. & Mullis, I. (2008). *TIMSS 2007 Technical Report*. Boston: TIMSS & PIRLS International Study Center, Boston College.

Pardo, C. (2001). El modelo de Rasch: Una alternativa para la evaluación educativa en Colombia [The Rasch model: An alternative for educational evaluation in Colombia]. *Acta Colombiana de Psicología*, 1(5), 9-21.

Ramírez, D.M., Gartner, M.L., Bernal, J.E., Zapata, Á., Vallejo, F. A., Prieto, P. A. & Langebaek, C.H. (2013). Consejo Nacional de Acreditación - CNA [National Accreditation Board]. Recuperado el 14 de octubre de 2013 de *Lineamientos para la acreditación de programas de pregrado*. <http://www.cna.gov.co/1741/articles-186359>

Ruiz, C. (2005). Confiabilidad [Reliability]. Recuperado el 3 de junio de 2015 de *Programa institucional de doctorado en educación*: <http://es.calameo.com/read/000261962d0b25b8cdc7b>

SERCE. (2010). Manual de procesamiento de datos y análisis de ítems [Manual data processing and item analysis]. En C. Pardo, *Compendio de los manuales del SERCE* [SERCE manuals compendium]. Santiago de Chile: Interson.

Tatsuoka, K. (2009). *Cognitive assessment: An introduction to the rule space method*. New York: Taylor & Francis Group.

Tatsuoka, M. (1986). Graph theory and its applications in educational research: A review and integration. *Review of Educational Research*, 56(3), 291-329.

Wright, B. & Stone, M. (1998). *Diseño de mejores pruebas utilizando la técnica de Rasch* [Best Test Design Rasch measurement]. México: Ceneval.

Fecha de recepción: 18 de marzo de 2016
Fecha de aceptación: 29 de marzo de 2017