

Febrero 2015  
Vol.2, No.1

# revie

Revista de Investigación y Evaluación Educativa

ISSN 2409-1553

23. ¿Existe con la capacitación recibida y las funciones desempeña?  
 A. Si  
 B. No



Instituto Dominicano de Evaluación e  
Investigación de la Calidad Educativa

## revie

Revista de Investigación y Evaluación Educativa

Revista Digital de suscripción gratuita del Instituto Dominicano de Evaluación e Investigación de la Calidad Educativa (IDEICE)

**Periodicidad Semestral**

**Edición**

Febrero 2015, Vol.2, No. 1

**Dirección Ejecutiva**

Dr. Julio Leonardo Valeirón Ureña

**Consejo Editorial**

Mtra. Dinorah de Lima Jiménez

Dr. Julián Álvarez Acosta

Dr. Luis Camilo Matos De León

**Corrección de estilos**

Alicia Delgado y Mestres

**Coordinación General**

Ing. Dilcia Armesto Núñez

**Fotografía**

Yeimy Olivier Salcedo

**Diseño y Diagramación**

Natasha Mercedes Arias

ISSN: 2409-1553

IDEICE

Ave. César Nicolás Penson No. 30, Gazcue  
Santo Domingo, D.N.

Teléfono: +1 (809) 732-7152

[www.ideice.gob.do](http://www.ideice.gob.do)

Santo Domingo, Rep. Dom.



Esta obra está bajo una licencia de Licencia Creative Commons Atribución-Non-Comercial-SinDerivar 4.0 Internacional.

El volumen 2 de *revie*, afianza la vocación del Instituto Dominicano de Evaluación e investigación de la Calidad Educativa, IDEICE, de ir aportando conocimientos e informaciones, producto de evaluaciones en investigaciones en el sector educativo. Se concreta de esta manera su misión: ser el soporte científico para aportar conocimientos pertinentes para la toma de decisiones.

En el sector educativo, como en muchos otros sectores del Estado, parecía tolerable el que las propias ideas fruto de una experiencia particular, en ocasiones carentes de objetividad, fueran los criterios orientativos del norte de las acciones de los líderes educativos. Hoy la investigación y las evaluaciones investigativas permiten abordar la realidad educativa con el objetivo de conocerla más profundamente, estableciendo las relaciones entre los diferentes elementos del sistema, y con ello orientar de manera más efectiva, a la toma de decisiones, tanto en el ámbito local como regional y/o nacional.

El IDEICE, en consecuencia, reafirma su vocación investigativa para con ello no solo conocer la realidad educativa, sino propiciar su transformación. En este orden, los artículos presentados en este número de *revie* siguen trillando la trocha que abre ése horizonte.

El Dr. José Vicente Díaz, en su artículo “El papel del análisis de ítems en la elaboración de las pruebas de conocimientos y aptitud” presenta, de manera particular, a través de un instrumento usado como forma, y cuyo resultado describe, como fondo, la realidad de la enseñanza de la matemática, y consigo el aprendizaje, así como la puesta en perspectiva de los desafíos que en el caso del sistema educativo dominicano

no tiene en esta disciplina, extensible a todas las demás.

El estudio de la Dra. Sandra González “Evaluación del impacto del Coordinador Docente en los centros educativos de República Dominicana” concluye de manera positiva respecto a la decisión de nombrar “coordinadores docentes” en las escuelas, con el propósito de dar seguimiento y orientación al trabajo de aula. Se enfatiza la necesidad de “profesionalizar” esta función docente con el propósito de hacerla más efectiva.

El estudio del Dr. Morales y colaboradores “Una aproximación de la incidencia del docente de matemáticas en el logro de aprendizajes estudiantiles”, desde la perspectiva de otras variables presenta un cuestionamiento al desempeño del docente en el rendimiento matemático de los estudiantes. En ese sentido, pareciera que el sistema educativo está de espaldas al fundamento paradigmático de las revoluciones científicas que han tenido a la matemática como la herramienta de su mejor quehacer.

Finalmente, el Dr. Héctor Valdés Veloz en su artículo “Introducción a la Neurociencia” destaca los aportes de esta nueva ciencia a la noción de “aprendizaje”, mostrando un cambio de perspectiva, desde la recuperación y almacenamiento de la cultura construida por un grupo social, hasta la consideración de que por aprendizaje se ha de entender los diferentes cambios que se producen en el pensamiento y en el comportamiento.

Se hace entrega de un número más de *revie* bajo la consideración de que toda labor humana es mejorable en el tiempo, la historia de la ciencia así lo muestra.

**Dr. Julio Leonardo Valeirón**

Director Ejecutivo IDEICE

4

*El papel del análisis de ítems en la elaboración de las pruebas de conocimiento y aptitud*

*Dr. José V. Díaz Esteve*

*Evaluación del impacto del coordinador docente en los centros educativos de República Dominicana*

*Dra. Sandra González Pons*

33

46

*Una aproximación de la incidencia del docente de Matemáticas en el logro de aprendizajes estudiantiles*

*Dr. Daniel Morales Romero*

*Lic. Claudia Curiel*

*Lic. Carmen Maura Taveras*

*Samuel Bonilla, M.B.*

*Introducción a la Neurodidáctica*

*Dr. Héctor Valdés Veloz*

67





# DR. JOSÉ V. DÍAZ ESTEVE

[jose.v.diaz@uv.es](mailto:jose.v.diaz@uv.es)

*Profesor de la Facultad de Psicología.  
Universidad de Valencia, España.*

# El papel del análisis de ítems en la elaboración de las pruebas de conocimiento y aptitud

## Resumen

El Análisis de ítems (A.I.) es un paso importante en la construcción de pruebas de conocimiento y de aptitud, que son muy utilizadas en Educación y Psicología.

Dado el caso, que en estas disciplinas la medición suele hacerse por definición, se hace necesario determinar con precisión: los dominios de conocimientos o las destrezas y rasgos; ya que éstos deben ser conceptos bien elaborados en las pruebas de conocimiento, o construcciones bien definidas en las pruebas de aptitud.

Este artículo intenta presentar una síntesis de los diversos procesos que hay que tener en cuenta para todo análisis de *ítems*, utilizando (para visualizar mejor) una prueba de conocimiento en Matemáticas, la cual se aplicó a 1,496 alumnos que solicitaron ingresar en el Instituto Superior de Formación Docente Salomé Ureña (ISFODOSU) en junio del 2013.

Esta prueba consta de 40 ítems, extraídos del banco de ítems de las “Pruebas Nacionales”, del Ministerio de Educación de República Dominicana. En tal sentido, urge señalar que el valor de la media teórica de esta prueba sería de 20, suponiendo que sus ítems han sido debidamente calibrados a través de un buen Análisis de ítems.

## Abstract

The Item Analysis (AI) is an important building testing of knowledge and ability, which are widely used in education and psychology.

Due to these disciplines measurement is usually done by definition, and it requires to determine precisely: the fields of knowledge, skills or traits, as these should be well-developed in tests of knowledge, or concepts well defined constructed in the aptitude tests.

This article attempts to present a synthesis of the various processes that every item analysis must take into account. Used to better visualize in a test in Mathematics by 1,496 students, who applied in June 2013 to enter ISFODOSU.

This test consists of 40 items, taken from the item bank of the National Tests of Ministry of Education from Dominican Republic. It is to note that the value of the theory of this test would be 20, assuming that their items have been properly calibrated through a good Item Analysis.

## Key Words:

Analysis of Items, Knowledge tests, Tryouts, National tests.

## Palabras claves:

Análisis de ítems; pruebas de conocimiento; pruebas de aptitud; pruebas nacionales.



## INTRODUCCIÓN

En Psicometría suelen distinguirse varios aspectos en cualquier Análisis de ítems:

### I. Análisis sustantivo de los ítems

Dada la condición antes indicada respecto a que en estas disciplinas la **medición se suele hacer por definición**, la primera tarea del constructor de pruebas sería definir los contenidos y procesos mentales en el que se quiere enmarcar los ítems.

Lo que implica elaborar previamente un **modelo de procesamiento teórico** sobre el cual se va a basar la construcción del instrumento de medida.

### II. Análisis cuantitativo de los ítems (A.I.)

Una vez se ha elaborado el instrumento de medida, éste debe ser aplicado a una muestra suficientemente grande y representativa de la población la que se quiere aplicar para **calibrar sus ítems**. Esta calibración consistirá en calcular los índices de dificultad y de capacidad discriminativa de los ítems, y el índice de fiabilidad del test. Estos cálculos permiten seleccionar o rechazar los *ítems* que han de formar parte del instrumento de medida.

### III. Análisis de las alternativas

Obtener una buena prueba o test es una tarea ardua y larga, que exige redactar ítems capaces de **medir los atributos psicológicos** o los **dominios de conocimientos** señalados. La tarea de mejorar los ítems del instrumento suele hacerse mediante

el llamado “análisis de las alternativas”, que analiza el poder de atracción que tiene cada una de las opciones que se ofrecen en cada ítem, si éstos son de selección múltiple.

Finalmente, hay que señalar que existen otras posibles clases de análisis, dentro de la psicometría cognitiva, como son:

### IV. Análisis integrativo

Llamado así porque en se intenta integrar los datos obtenidos en los pasos anteriores, para inferir los procesos mentales o componentes cognitivos que subyacen detrás de las respuestas.

### V. Análisis de errores

Busca ver los tipos de errores que los sujetos comenten consistentemente, así como las causas de los mismos.

### VI. Análisis de los patrones de respuestas

Busca hallar el tipo de patrones que pueden encontrarse en el conjunto de datos que se han recogido, y en los datos que nos ofrece el instrumento de medida elaborado (ver Poner nombre completo y obra Díaz 2002).

En los siguientes párrafos se pretende aplicar las diversas técnicas de un análisis de ítems a la prueba de conocimiento en Matemáticas, la forma A que se aplicó a 1,469 solicitantes para ingresar al ISFODOSU en el verano del año 2013.

## I. ANÁLISIS SUSTANTIVO DE LOS ÍTEMS

Una parte importante del análisis de ítems es el llamado análisis sustantivo, que busca identificar la categoría a la que los ítems pertenecen, de acuerdo al modelo de procesamiento propuesto, que permite agrupar los ítems de acuerdo al **contenido curricular** que se intenta evaluar; o, dentro de la psicometría cognitiva, a la actividad mental que debe tener el sujeto para responder bien al ítem.

Esta actuación debe ser hecha básicamente por el equipo constructor de la prueba y por el psicómetra que coopera con ellos. Su trabajo consistirá, pues, en construir primero un **modelo de**

**procesamiento** que defina con claridad el dominio de conocimientos que se pretende evaluar, o los procesos mentales que tiene el alumno para solucionar bien cada *ítem* de la prueba.

Dado el caso que se analiza una prueba de conocimiento en Matemáticas, elaborada sobre un conjunto de *ítems* que provienen del banco de ítems de las Pruebas Nacionales, se propone el siguiente ejemplo de modelo de contenidos en el que se presenta primero una demarcación y categorización global de esos contenidos y luego una especificación de los mismos dentro de cada categoría.

### I.1 Elaboración del modelo de contenidos para la prueba de conocimiento en Matemáticas.

#### Demarcación y categorización global de los contenidos

1° MC: teoría de conjuntos.

2° MN: sistemas numerales y simbologías matemáticas.

5° MO: operaciones, cálculos y estaciones con clases de números.

4° MM: sistemas de medida.

5° MA: álgebra.

6° MG: geometría.

7° MT: trigonometría

8° MCC: el espacio bidimensional definido por coordenadas cartesianas.

9° ME: contenidos sobre conceptos elementales de Estadística.



## I.2 Especificación de los contenidos globales en las sub- categorías

<b>1º Teoría de conjuntos (1º MC).</b>	<b>2º Sistemas numerales MN y simbologías.</b>
<b>MC.1:</b> conceptos básicos: pertenencia y no pertenencia.	<b>MN.1:</b> conceptos: base del sistema, secuencias, reglas de formación.
<b>MC.2:</b> forma de expresarlos: extensión y comprensión.	<b>MN.2:</b> secuencias.
<b>MC.3:</b> operaciones con conjuntos: unión, intersección, diferencias, complemento.	<b>MN.3:</b> combinaciones y permutaciones.
<b>MC.4:</b> uso de diagramas para expresar relaciones.	<b>MN.4:</b> simbologías.
<b>MC.5:</b> aplicaciones a situaciones de la vida real.	<b>MN.5:</b> resolver problemas de pensamiento lógico-matemático.
<b>3º Operaciones; cálculos con números (MO).</b>	<b>4º Sistemas de Medidas (MM).</b>
<b>MO.1:</b> principales operaciones: MO1.1: suma-resta, MO1.2: multiplicación-división, MO1.3 potencia-raíz.	<b>MM.1:</b> concepto del sistema de medida. Los más utilizados: europeo, inglés.
<b>MO.2:</b> Operar con números: MO2.1: enteros, MO2.2: decimales, MO2.3: fracciones, MO2.4: complejos, MO2.5 matrices, MO2.6: logaritmos, MO2.7: irracionales.	<b>MM.2:</b> medidas de longitud, superficies, volúmenes.



<b>3º Operaciones; cálculos con números (MO).</b>	<b>4º Sistemas de Medidas (MM).</b>
<b>MO.3:</b> cálculo y estimaciones: MO3.1 cálculo mental, MO3.2 regla de tres. MO3.3 probabilidades, MO3.4 integral y diferencial.	<b>MM.3:</b> otras medidas de: líquidos, temperaturas, presión atmosférica.
<b>MO.4:</b> Combinaciones y permutaciones.	<b>MM.4:</b> medidas de ángulos.
	<b>MM.5:</b> medidas del tiempo.
<b>5º Álgebra (MA)</b>	<b>6º Geometría (MG)</b>
<b>MA.1:</b> la expresión algebraica, elementos y clases.	<b>MG.1:</b> concepto de elementos geométricos: punto, segmento, rectas, ángulos planos, perímetros.
<b>MA.2:</b> simplificar expresiones algebraicas.	<b>MG.2:</b> solución de problemas con líneas.
<b>MA.3:</b> racionalizar expresiones algebraicas.	<b>MG.3:</b> solución de problemas con ángulos.
<b>MA.4:</b> resolver ecuaciones de primer grado.	<b>MG.4:</b> solución de problemas con triángulos.
<b>MA.5:</b> resolver ecuaciones de segundo grado.	<b>MG.5:</b> solución de problemas con cuadriláteros.
<b>MA.6:</b> resolver inecuaciones algebraicas con $>$ ó $<$	<b>MG.6:</b> solución de problemas con polígonos.
<b>MA.7:</b> resolver sistemas de ecuaciones.	<b>MG.7:</b> solución de problemas de áreas.
	<b>MG.6:</b> solución de problemas de volúmenes.



7º Trigonometría (MT)	8º Coordenadas cartesianas (MCC):
MT.1: funciones, igualdades trigonométricas. Funciones inversas.	MCC.1: conceptos básicos: ejes, valoración de puntos, líneas, figuras.
MT.2: razones trigonométricas: seno, coseno, tangente, cotangente, secante, cosecante, definiciones, fórmulas.	MCC.2: traslación de figuras.
MT.3: aplicación al cálculo de elementos del triángulo rectángulo.	MCC.3: rotación de figuras.
MT.4: solución de problemas trigonométricos.	MCC.4: simetría.
	MCC.5: proyección.
	MCC.6: estimación de valores: ángulos.

9º Estadística (ME)
ME.1: distribución de frecuencias, gráficas.
ME.2: medidas de tendencia central.
ME.3: medidas de dispersión.
ME.4: índices de correlación.

La importancia de elaborar un modelo de procesamiento, reside en que éste permite obtener una visión global de los *dominios de conocimientos* que se quieren abarcar con la prueba; ponderar su peso específico dentro de la mismas, y sobre todo dar un código a cada ítem, que ayuda para la creación de un banco de ítems que facilita su almacenamiento y su manejo informático.

### 1.3 Elaboración del modelo de procesamiento para la prueba de conocimiento en Matemáticas

Entre los objetivos que se tienen, al construir las pruebas de conocimiento, se pueden señalar estos:

- 1º evaluar **el dominio de conocimientos** que tiene el alumno,
- 2º constatar el **nivel de habilidad o destrezas** que ha desarrollado el mismo,
- 3º determinar *las actitudes o sentimientos* que tiene el sujeto ante esta disciplina académica.

De ahí que los objetivos operacionales específicos de cualquier prueba de conocimiento se deben organizar bajo tres grandes aspectos:

- 1º **aspectos referentes** al conocimiento, en el que se propone estimar el dominio de conocimientos que tiene el alumno en este caso, sobre: conceptos básicos en matemáticas, signos, símbolos, estructura de los números, procedimientos para resolver problemas, etc.
- 2º **aspectos referentes a las habilidades o destrezas**, entre los que se propone: evaluar destrezas mentales, tales como: tener dominio de las operaciones clásicas (suma-resta, multiplicación-división, potencia-radición) con las diversas clases de números; así como las capacidades de clasificar, descubrir las reglas de for-

mación de una secuencia, manejar estructuras espaciales, el dominio del razonamiento lógico, la flexibilidad mental, la habilidad de solucionar de problemas.

3º **aspectos referentes al desarrollo de actitudes y sentimientos**, que busca: conocer de los alumnos la confianza en sí mismo, la curiosidad intelectual, el sentimiento de interés por lo novedoso y complejo, el sentido de relevancia de los datos, el deseo de profundización en la comprensión de ciertos conocimientos, y la perseverancia en la búsqueda de la solución de situaciones problemáticas a través de aplicaciones matemáticas.

En este apartado solamente se han esbozado los citados objetivos operacionales que pueden guiar hacia la construcción del modelo de procesamiento citado. Este trabajo debe ser hecho conjuntamente por el equipo de constructores de *items* en armonía con las sugerencias de un psicólogo especializado en procesos mentales, que permita encuadrar los mismos, si es factible, dentro de la nueva "Psicometría cognitiva" (Nombre Anderson, 1983; Nombre Embretson, 1993).



## II. ANÁLISIS CUANTITATIVO DE LOS ÍTEMS EN LA PRUEBA DE CONOCIMIENTO EN MATEMÁTICAS FORMA A

En psicometría se señalan dos métodos para efectuar el Análisis cuantitativo de los ítems (A.I):

- El primero, de acuerdo a la **Teoría Clásica de los Test (TCT)**, pretende hallar los valores paramétricos de los ítems, utilizando la Matemática tradicional para estimar: el índice de dificultad (ID: que expresa el % de aciertos al ítem), y el coeficiente de consistencia interna (ICI) u homogeneidad (IHI), que indica el valor de la correlación punto biserial ( $r_{pb}$ ) entre el valor de la respuesta dada a cada ítem y la suma de respuestas buenas en todos los ítems de la prueba total para cada sujeto.

- y, el segundo, que de acuerdo a la **Teoría de la Respuesta al ítem (TRI)** intenta encontrar los valores paramétricos, utilizando la matemática bayesiana (cálculo de probabilidades) para estimar: el parámetro  $a_i$  que indica la capacidad discriminativa del ítem, el parámetro  $b_i$  que indica el nivel de aptitud que se necesita para responder bien al ítem, y finalmente el parámetro  $c_i$  que expresa la probabilidad de responder bien al azar.

### II.1 Análisis cuantitativo de los ítems utilizando el método de la TCT en la forma de Matemática A

El análisis de ítems (A.I.) se utiliza para estudiar los valores paramétricos de los ítems (índices de dificultad y consistencia interna), y de este modo seleccionar los que son aceptables e útiles a los propósitos para formar parte de la prueba. En este apartado se aplica el método tradicional basado en la TCT.

Los criterios de selección para los dos índices se expresan a continuación:

1º **Índice de dificultad:** se obtiene calculando el por ciento (%) de aciertos que tiene cada ítem.

Los valores de los **índices de dificultad** se suelen situar en las dos zonas de la curva normal (inferior y superior), para establecer los siguientes criterios de selección y rechazo:

	inferior	superior
-zona de rechazo total:	0.00 a 0.05	0.96-1.00
-zona de aceptación crítica:	0.06 a 0.25	0.76-0.95
-zona de aceptación moderada	0.26 a 0.45	0.56-0.75
-zona de aceptación óptima:	0.46 a 0.50	0.51-0.55

## 2° Coeficiente de consistencia interna (ICI) o índice de homogeneidad.

El ICI se obtiene tomando como criterio de referencia la correlación entre las puntuaciones en el ítem y el total de prueba.

Para que se puedan dar las siguientes pautas para realizar la selección de los ítems, tomaremos en cuenta:

■ -Zona de rechazo total: **ítems a rechazar**, aquellos cuyas correlaciones sean negativas o nulas (en este caso para un  $gl > 900$ ,  $r_{pb} < 0.06$ , al  $ns: 0.05$ ).

■ -Zona de aceptación crítica: ítems que deben

ser utilizados con precaución, y son los que presentan los valores  $r_{pb}$  entre 0.06 y 0.19.

■ -Zona de aceptación moderada: ítems aceptables, cuyos valores  $r_{pb}$  están entre 0.20 y 0.40.

■ -Zona aceptación Máxima: ítems buenos y muy buenos, cuando los valores  $r_{pb}$  están desde 0.40 hacia arriba.

### Resultados de los índices de la TCT en la prueba de Matemáticas, forma A, mediante el programa Excel aplicada en ISFODOSU en el verano del 2012.

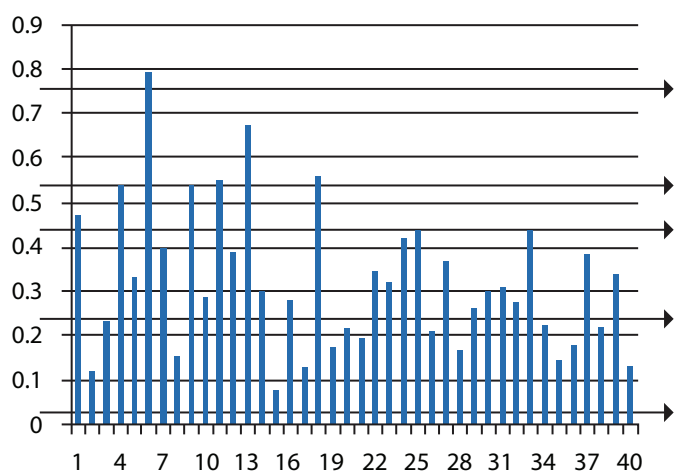
Contenido de las filas: 1 n° del ítems, 2.ID, 3.ICI

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0.47	0.12	0.23	0.54	0.33	0.79	0.40	0.15	0.54	0.29	0.55	0.39	0.67	0.30	0.08	0.28	0.13	0.56	0.17	0.22
0.24	0.10	0.12	0.32	0.27	0.28	0.32	0.13	0.44	0.19	0.41	0.18	0.38	0.11	0.03	0.11	-0.05	0.38	0.09	0.10
21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	18	19	30
0.19	0.35	0.32	0.42	0.44	0.21	0.37	0.17	0.26	0.30	0.31	0.28	0.43	0.23	0.15	0.18	0.38	0.22	0.34	0.13
0.18	0.18	0.30	0.23	0.19	0.04	0.25	0.02	0.13	0.24	0.27	0.19	0.37	0.17	-0.05	0.08	0.27	0.21	0.15	0.12

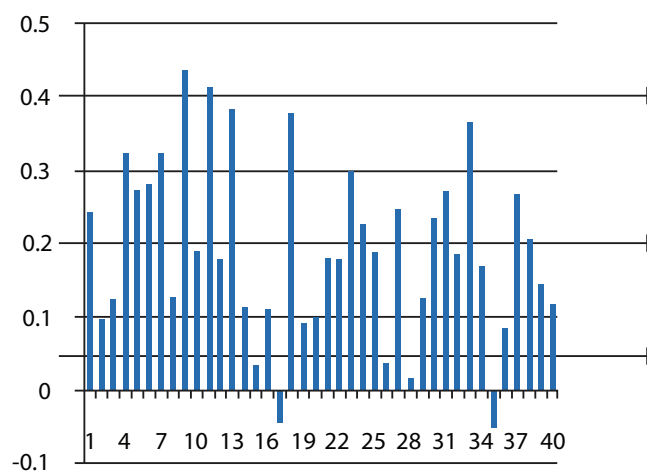


A continuación se presentan los histogramas de los valores de los índices de dificultad y de consistencia interna en MATEM- forma A:

**ID. Promedio: 0.322.**



**ICI. Promedio: 0.192.**



Vistos estos datos, dado que en la TCT para rechazar algún ítem debe fallar, al menos, uno de los dos índices, se puede ver que en la prueba MATEM-forma A los ítems pueden distribuirse de este modo:

Ítems que están en la	ID	nº	ICI	nº
zona de rechazo total:	< .5 y >.95	00	<.06	04
zona de aceptación crítica:	.06 a .25 y .76-.95.	16	.07-.20	18
zona de aceptación moderada:	.26 a .45 y .56-.75.	20	.21- .40	16
zona de aceptación óptima:	.46 a .50 y .51-.55	<u>04</u>	.41-.99	<u>02</u>
		<b>40</b>		<b>40</b>

\* El Índice de Dificultad (ID) promedio está en 0.322, lo que indica que es una prueba que presenta una dificultad alta, ya que está muy por debajo del promedio teórico que sería de 0.50.

\* El Índice de Consistencia Interna (ICI) u Homogeneidad (IHI) presenta un promedio de 0.192, lo que indica que es una prueba cuyos valores de consistencia interna están una gran parte en la zona crítica.

\* Y, un índice de fiabilidad de 0.398, que se considera muy bajo, lo que afecta la predictibilidad de los resultados.

## Conclusiones

Dado que de acuerdo a la TCT para rechazar un ítem éste debe estar en la zona de rechazo en alguno de los dos parámetros, se rechazarían solo cuatro ítems en la prueba.

Pero sería recomendable reformar algunos de los 18 ítems que están en la zona de aceptación crítica. Esto puede hacerse mejor utilizando el método de la TRI, ya que ofrece criterios de rechazo más objetivos y precisos que el simple conteo como se ha hecho antes.

## II.2 Análisis cuantitativo de los ítems utilizando el método de la TRI en la prueba de conocimiento en Matemáticas, forma A

En esta sección se busca efectuar el análisis de ítems utilizando el método de la Teoría de la Respuesta al Ítem (TRI), a través del cual se intenta estimar con más precisión los valores de los tres parámetros que este método ofrece.

Pero antes de proceder al mismo resulta oportuno ofrecer unas nociones elementales sobre la TRI (para ampliar información ver Hambleton & Swaminathan, 1987, NOMBRE Muñiz, 1990, NOMBRE Díaz, 1997).

La TRI, que tiene un carácter eminentemente tecnológico, permite efectuar una mejor cons-

trucción de tests y realizar un análisis de los resultados mucho más potente que el utilizado con el método la Teoría Clásica de los Tests (TCT).

### II.2.1 Supuestos fundamentales de la TRI

1º Existe en los sujetos, que van a ser evaluados, un **rasgo latente** o atributo, rasgo que suele ser simbolizado en la TRI por  $\theta$  (theta).

2º Luego se elige un conjunto de comportamientos observables capaces de medir este rasgo (ítems). Al conjunto total de estos ítems se le llama **dominio del atributo**. Si se extrae del mismo una muestra representativa de  $n$  ítems se obtiene un test.

3º Aplicada esta muestra de ítems a los sujetos, respuestas diversas a los mismos, condicionadas al nivel de desarrollo en la aptitud que se quiere medir de los sujetos.

### II.2.2 Procedimientos

La TRI especifica el 3º supuesto de esta forma:

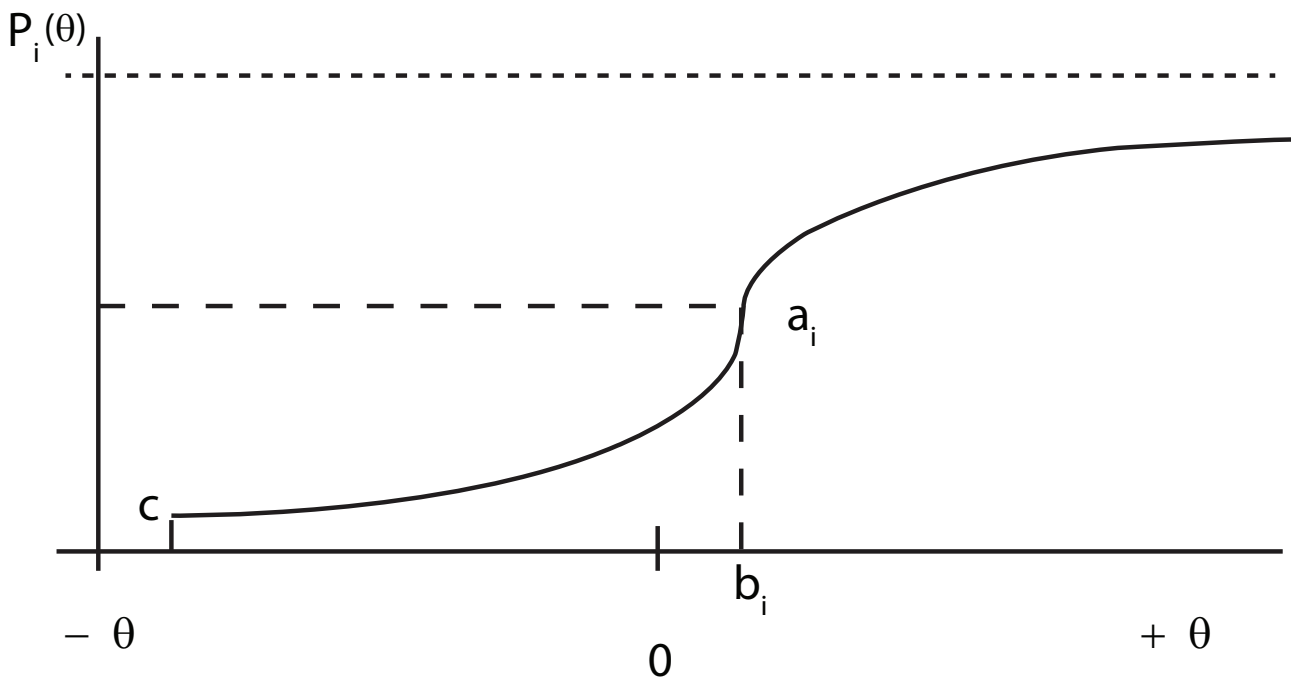
*La relación existente entre las respuestas a los ítems de la prueba y la aptitud  $\theta$  que tiene el sujeto para responder al ítem, se expresa a través de una función creciente que indica la probabilidad de acertar el ítem, condicionada al nivel de aptitud del sujeto.*

Relación que suele representarse gráficamente como una línea curva (cóncavo-convexa) denominada Curva Característica del Ítem (CCI), como puede apreciarse en la figura siguiente:



**Figura 1.**

CCI de un ítem de un test de aptitudes



En esta CCI se aprecian los siguientes elementos:

- Los **dos ejes** de las coordenadas cartesianas, como son:

- la abscisa ( $\theta$ ), que es el eje horizontal, que viene a ser a una línea sobre la que se pueden escalar los valores atribuidos a la aptitud cuya unidad es 1 (theta),

- la ordenada ( $P_i(\theta)$ ), que es el eje vertical, que representa los valores de la probabilidad de acertar el ítem, condicionada al nivel de aptitud del sujeto  $P_i(\theta)$ .

- Y una línea curva que se asemeja a una **S** achatada, que une los conceptos que representan en las coordenadas  $\theta$  y  $P_i(\theta)$ . Esta línea cóncava-convexa suele definirse por tres parámetros:

- a) El *parámetro*  $b_i$ , que equivale al valor en  $\theta$  del punto de inflexión de la CCI; es decir, la proyección de dicho punto sobre la abscisa.

Este parámetro indica el grado de *dificultad* del ítem; esto es, el nivel de aptitud que necesita para responder bien al ítem. Los valores del *parámetro*  $b_i$  teóricamente pueden estar entre:



$-\infty \leq b_i \leq \infty$  aunque en la práctica se toman los valores que están entre -3.50 y +3.50.

b) El *parámetro*  $a_i$  es el valor proporcional a la pendiente de la recta tangente a la CCI, en el punto de inflexión de la curva.

El valor del *parámetro*  $a_i$  viene a ser la derivada primera de la ecuación, que representa la CCI en el punto de inflexión.

Este parámetro, también llamado parámetro discriminativo del ítem, indica el grado en que la respuesta al ítem varía con el nivel de aptitud del sujeto. Toma siempre valores positivos, ya que si tomara valores negativos la curva sería decreciente, por lo que sus valores deben estar entre:  $0 \leq a_i \leq \infty$ .

c) El *parámetro*  $c_i$ , denominado de asintoticidad, representa la probabilidad de acertar el ítem, cuando la aptitud es nula; es decir, cuando el sujeto contesta al azar.

Este parámetro, también llamado índice de conjetura o pseudo-azar, equivale a la altura del punto donde la CCI corta a la ordenada (Y).

Generalmente su valor se toma aproximadamente, como  $c=1/k$  siendo k el número de alternativas del ítem. Sus valores estarían, pues, entre:  $0 \leq c_i \leq 1.0$ .

La CCI expresa, pues, la relación entre las respuestas dadas a los ítems y el nivel de aptitud que debe tener el sujeto para responder bien al

ítem por medio de una ecuación exponencial que indica la probabilidad de acertar un ítem o alcanzar una puntuación determinada en función de la aptitud del sujeto:  $P(\bar{X}, \theta)$ ,

Donde:

$\bar{X}$ : representa el sistema de puntuaciones que obtienen los sujetos en el ítem o test,

$\theta$ : los valores dados al rasgo que se pretende medir (escala métrica de  $\theta$ ),

P: la relación de probabilidad existente entre ambos elementos.

Cuando la probabilidad de acertar está condicionada al nivel de aptitud de los sujetos, la expresión anterior toma esta forma:  $P(\bar{X}|\theta)$ .

### II.2.3 Condiciones esenciales en la TRI

Aceptados los supuestos señalados, es necesario averiguar si en el modelo psicométrico adoptado cumple las siguientes condiciones:

- la independencia local,
- la unidimensionalidad del rasgo o atributo, y,
- la ausencia de presión temporal.

El concepto de **independencia local** tiene que ver con que no haya ninguna relación de dependencia entre los elementos que están colocados en un mismo campo.

Concepto que aplicado a los tests, se puede hablar tanto de la independencia local de los ítems, como de la independencia local de suje-



tos; es decir, que la respuesta que se dé a cada ítem no está ligada a la dada en otro ítem, y que las respuestas de un sujeto no estén contaminadas por las respuestas dadas por otros sujetos de la muestra.

La segunda condición la **unidimensionalidad**, es indispensable en la mayoría de los modelos, ya que la probabilidad de que los sujetos acierten a los ítems debe depender principalmente del nivel de la aptitud de los sujetos, y no de otros factores.

De modo que el concepto “unidimensionalidad” está incrustado en la misma concepción de la CCI. Dada la dificultad de encontrar tests plenamente unidimensionales se ha optado por considerar como tales a aquellos que tengan un primer factor dominante.

Se han propuesto muchos métodos para demostrar la unidimensionalidad de los tests. Hattie (1985) cita ochenta y siete, casi todos ellos basados en el análisis factorial, pero éstos no resultan totalmente satisfactorios, sobre todo cuando se trata de ítems dicotómicos (ver: Martínez Arias, 1995).

De ahí que en la actualidad se esté trabajando intensamente en la línea de modelos multidimensionales para estimar los parámetros de los ítems (ver: Reckase, 1979, 1983, 1985, 1997; Maydeu, 1996).

Existe una tercera condición señalada por

Hambleton & Swaminathan (1987) que exige la ausencia de presión temporal en la ejecución del test (Speededness), ya que si un sujeto no responde a un ítem nunca se sabrá con certeza si el fracaso es debido a la falta de capacidad del sujeto o a otros factores, como sería la falta de tiempo para resolverlo.

Esta condición debe ser considerada como una consecuencia de la anterior. Por lo que es preferible dar tiempo suficiente y controlar las respuestas de los sujetos que no alcanzan el final del test. Más tarde, Hambleton, Swaminathan y Rogers (1991), ante la dificultad de obtener tests donde todos los ítems sean contestados por todos los sujetos, recomiendan utilizar aquellos protocolos que al menos un 75 % de sujetos completen el test, y aquellos ítems que sean completados al menos por el 80% de los examinados.

#### **II.2.4 Estimación de los parámetros de las CCI de los ítems en la forma de matemática A**

Para estimar los valores paramétricos de las CCI, de los ítems del POMA.3, se ha utilizado el programa XCALIBRE, que está incluido en el paquete estadístico que ofrece el “ASC Item and Test Analysis Package of Assessment Systems Corporation.”

A continuación se presentan algunos datos entresacados del “output” del citado programa, aplicado la matriz de respuestas dadas a la

prueba Matemáticas forma A para una muestra de 1,466 estudiantes, quienes tomaron esta prueba para ingresar en ISFODOSU en el curso académico 2013-2014.

En este artículo se presentan en un solo cuadro todos los datos que el XCALIBRE ofrece en el análisis cuantitativo de los ítems, efectuado con la prueba de conocimiento de Matemáticas forma A, que se ha aplicó en junio 2013 a los postulantes para ingresar en los Centros de Educación Superior de ISFODOSU.

En este cuadro la información presenta los datos en filas y columnas. En las filas y las columnas se sitúan las siguientes informaciones de los ítems:

**1ª Una identificación numeral**, que indica la posición del ítem en la prueba, y una **literal** que especifica el contenido, que en este caso es M, es decir que es un prueba de Matemáticas.

**2ª** Un conjunto de filas en las que se ofrecen estas informaciones:

**a.** El **código** asignado al ítem, de acuerdo al modelo de procesamiento adoptado.

**b.** La **clave**, que es la respuesta correcta asignada por el editor de la prueba.

**c.** Los **indicadores** propios del XCALIBRE, expresados por estas letras:

**\*R:** que señala el ajuste de la CCI real a la teórica (fit), se rechaza si  $\text{fit} > 2.00$

**\*P:** que aparece cuando algún parámetro del ítem en concreto no se ajusta a alguno de estos criterios:

-**a<sup>i</sup>**  $> 0.30$ ,

-**b<sup>i</sup>**, el valor debe estar dentro de estos límites: -2.95 a +2.95

-**c<sup>i</sup>** el valor debe ser  $< 0.40$

**\*K:** cuando existe un posible error en la clave, ya que alguno de los distractores tiene un PC de respuestas y una correlación ( $r_{i\theta}$ ) más altos que la clave.

**3ª Análisis de las alternativas** que contiene esta información:

**\*PC:** porciento de sujetos que elige cada alternativa: A, B, C, D y Oth.

Los mejores ítems son aquellos que presentan:

- el mayor PC, que se debe hallar en la clave.

- los PC en los distractores deben ser más bajos y con diversos valores entre sí.

- la dificultad del ítem se consigue presentando alternativas cercanas conceptualmente a la clave, y controlando el error que cometería el sujeto al responder en cada alternativa.

- alternativas con iguales o similares valores en PC deben ser evitadas, pues tienen un poder de atracción similar.



**4ª Correlación** del valor en el ítem con el valor de la aptitud ( $r_{i\theta}$ ), que el sujeto tiene en la prueba total, en el caso que la muestra sea grande ( $N=1,496$ ), se señala que el valor de la correlación debe ser mayor de 0.0

Los mejores ítems son aquellos que presentan correlaciones positivas significativas en la clave, y negativas en distinto nivel en los distractores.

Para facilitar la percepción de los datos se construido un cuadro síntesis, que se ha dividido en cuatro subcuadros que contiene cada uno 10 ítems, de este modo:

**Ítems del 1 al 10**

Ítem	1-M	2-M	3-M	4-M	5-M	6-M	7-M	8-M	9-M	10-M
<b>A. Contenido</b>	G.1	A.1	O.2.4	A.2	A.5	N.4	N.3	G.3	C.4	C.3
<b>B. Clave</b>	C	A	D	B	D	B	A	C	A	B
<b>C. Indicador</b>	---	PK	PK	---	---	---	---	P	---	---
$a_i$ (P.discr.)	.49	1.02	0.90	0.53	0.74	0.53	0.73	0.96	0.89	0.77
$b_i$ (P.dificul)	1.57	3.00	3.00	0.71	2.03	-1.36	1.56	3.00	0.42	2.71
$c_i$ (P.azar.)	.27	0.13	0.22	0.24	0.22	0.22	0.23	0.15	0.20	0.23
Ajuste (fit)	.99	1.52	0.85	0.65	0.26	0.67	0.46	1.10	0.35	0.28
<b>D. Alterna</b>										
PC A	.20	.12~	.16	.31	.21	.11	.40~	.51	.54~	.18
PC B	.13	.06	.22	.54~	.33	.79~	.29	.08	.19	.29~
PC C	.47~	.48	.37	.37	.11	.03	.20	.15~	.13	.26
PC D	.19	.34	.23~	.10	.33~	.07	.11	.26	.14	.26
PC Oth	.01	.00	.01	.01	.02	.01	.01	.01	.01	.01
<b>E. Correlac</b>										
$r_{i\theta}$ A	-.12	.02~	-.17	-.23	-.08	-.25	.31~	.05	.59~	-.09
$r_{i\theta}$ B	-.10	-.13	-.07	-.36~	-.09	.43~	-.26	-.09	-.30	.12~
$r_{i\theta}$ C	.22~	.05~	.19	-.11	-.11	-.16	-.04	.06~	-.27	-.01
$r_{i\theta}$ D	-.05	.00	.01~	-.16	.25~	-.26	-.04	-.04	-.23	-.05
$r_{i\theta}$ Oth	-.05	-.07	-.03	-.02	-.03	-.10	-.06	-.05	-.08	-.02

Ítems del 11 al 20

Ítem	11.M	12.M	13.M	14.M	15.M	16.M	17.M	18.M	19.M	20.M
<b>A. Contenido</b>	N.2	O.2.5	A.7	O.2.6	T-2	A.2	O.2.5	O.2.5	CC.5	T.3
<b>B. Clave</b>	A	D	B	A	B	B	D	B	D	A
<b>C. Indicador</b>	---		---	P	RPK	PK	RPK	---	P	PK
$a_i$ (P.discr.)	0.87	0.68	0.67	0.80	1.15	0.86	1.18	0.73	0.94	0.98
$b_i$ (P.dificul)	0.37	2,57	-.33	3.00	3.00	3.00	3.00	0.33	3.00	3.00
$c_i$ (P. azar.)	0.21	0.30	0.22	0.27	0.11	0,25	0.15	0.21	0.17	0.21
Ajuste (fit)-	0.38	0.91	0.44	0.81	2,23	1,11	2.51	0.54	1.04	1.54
<b>D. Alterna</b>										
PC A	0.55~	0.21	0.15	0.30~	0.14	0.48	0.47	0.14	0.25	0.22~
PC B	0.25	0.20	0.67~	0.28	0.08~	0.12	0.16	0.56~	0.26	0.37
PC C	0.11	0.20	0.07	0.19	0.11	0.28~	0.23	0.19	0.29	0.25
PC D	0.08	0.39~	0.11	0.21	0.66	0.11	0.13~	0.10	0.17~	0.15
PC Oth	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.00	0.02	0.02
<b>E. Correlac.</b>										
$r_{i\Theta}$ A	0.58~	0.03	-0.34	0.03~	-0.08	0.15	0.26	-0.30	0.02	-0.05~
$r_{i\Theta}$ B	-0.33	-0.12	0.54~	0.02	-0.03~	-0.13	-0.08	0.53~	-0.05	-0.01
$r_{i\Theta}$ C	-0.23	-0.03	-0.22	-0.01	0.05	0.01~	-0.11	-0.22	-0.03	0.03
$r_{i\Theta}$ D	-0.24	0.10~	-0.24	-0.03	0.05	-0.10	-0.15~	-0.22	-0.03~	0.02
$r_{i\Theta}$ Oth	-0.05	-0.01	-0.07	-0.02	-0.01	-0.06	-0.02	-0.07	-0.05	0.02

Ítems del 21 al 30

Ítem	21-M	22-M	23-M	24-M	25-M	26-M	27-M	28-M	29-M	30-M
<b>A. Contenido</b>	G.1	T.2	O.2.5	N.2	C.4	CC.5	CC.2	G.3	T.1	G.1
<b>B. Clave</b>	C	B	D	B	C	A	A	B	B	C
<b>C. Indicador</b>	PK	---	---	---	---	PK	---	RPK	P	---
$a_i$ (P.discr.)	0.86	0.75	0.85	0.62	0.51	1.07	0.63	1.07	0.81	0.67



Ítem	21-M	22-M	23-M	24-M	25-M	26-M	27-M	28-M	29-M	30-M
$b_i$ (P.dificul)	3.00	2.61	1.78	2.18	2.32	3.00	1.90	3.00	3.00	2.46
$c_i$ (P. azar.)	0.18	0.28	0.20	0.30	0.31	0.20	0.23	0.17	0.24	0.22
Ajuste (fit)-	0.64	0.74	0.29	0.91	1.39	1.71	0.70	2.13	0.58	0.72
<b>D. Alterna</b>										
PC A	0.38	0.19~	0.18	0.42~	0.08	0.21~	0.37~	0.22	0.18	0.32
PC B	0.14	0.35~	0.24	0.19	0.07	0.43	0.16	0.17~	0.26~	0.20
PC C	0.19~	0.23	0.25	0.13	0.44~	0.23	0.18	0.27	0.24	0.30~
PC D	0.27	0.22	0.32~	0.24	0.40	0.13	0.29	0.33	0.31	0.17
PC Oth	0.01	0.01	0.02	0.03	0.01	0.01	0.01	0.01	0.01	0.01
<b>E. Correlac.</b>										
$rt\Theta$ A	-0.12	-0.07	-0.18	0.14~	-0.16	-0.07~	0.27~	-0.07	0.04	-0.11
$rt\Theta$ B	-0.13	0.10~	-0.05	-0.03	-0.14	0.09	-0.11	-0.11~	0.05~	-0.02
$rt\Theta$ C	0.09~	-0.05	-0.12	0.03	0.13~	0.02	-0.16	0.04	-0.03	0.20~
$rt\Theta$ D	0.17	-0.00	0.32~	-0.16	0.04	-0.07	-0.04	0.12	-0.03	-0.07
$rt\Theta$ Oth	-0.06	-0.00	-0.04	-0.00	-0.04	-0.01	-0.09	-0.05	-0.05	-0.05

### Ítems del 31 al 40

Ítem	31-M	32-M	33-M	34-M	35-M	36-M	37-M	38-M	39-M	40-M
<b>A. Contenido</b>	CC.1	G.3	CC.1	G.1	O.2.5	T.3	G.1	G.1	O.2.2	N.3
<b>B. Clave</b>	D	A	C	A	A	D	C	A	B	B
<b>C. Indicador</b>	---	---	---	P	RPK	PK	---	---	P	PK
$a_i$ (P.discr.)	0.70	0.78	0.76	0.80	1.19	0.96	0.60	0.83	0.76	1.02
$b_i$ (P.dificul)	1.99	2,85	1.10	3.00	3.00	3.00	1,91	2,93	3.00	3.00
$c_i$ (P. azar.)	0.19	0.23	0.21	0.20	0.16	0.18	0.24	0.19	0.29	0.14
Ajuste (fit)-	0.81	0.28	0.32	0.57	2.73	1.17	0.68	0.47	0.80	1.57
<b>D. Alterna</b>										
PC A	0.30	0.28~	0.21~	0.22~	0.15~	0.34	0.34	0.22~	0.35	0.20

Ítem	31-M	32-M	33-M	34-M	35-M	36-M	37-M	38-M	39-M	40-M
PC B	0.23	0.16	0.23	0.22	0.57	0.34	0.18	0.22	0.34~	0.13~
PC C	0.16	0.12	0.44	0.25	0.11	0.13	0.38~	0.36	0.20	0.09
PC D	0.31~	0.44	0.12~	0.29	0.16	0.18~	0.08	0.19	0.08	0.55
PC Oth	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.02	0.02	0.02
<b>E. Correlac.</b>										
rt⊖ A	-0.13	0.10~	-0.18	0.07~	-0.17~	-0.01	-0.18	0.13~	0.00	0.02
rt⊖ B	-0.06	-0.02	-0.23	0.01	0.29	0.03	-0.04	-0.05	0.04~	0.02~
rt⊖ C	-0.12	-0.02	0.43	-0.03	-0.08	-0.01	0.25~	-0.05	-0.05	-0.07
rt⊖ D	0.30~	-0.05~	-0.12	-0.04	-0.15	0.00~	-0.06	-0.01	0.02	0.02
rt⊖ Oth	-0.11	-0.06	-0.08	-0.04	-0.07	-0.04	-0.04	-0.03	-0.06	-0.03

**Observaciones sobre los datos del cuadro anterior**

**A. De contenido:** al código que aparece de cada ítem de esta fila hay que colocar delante M, como MG.1 para el 1º, MA.1 para el 2º, MO.2.4 para el 3º, de acuerdo al cuadro de contenidos antes ofrecido **con estos datos**.

~ **identificación** de la clave (A;B;C;D, Oth, que corresponde a otro tipo de respuesta.

**Color rojo**, indica que el ítem que no cumple con el criterio del parámetro, como : bi=3.00, es un valor que está fuera del intervalo 2.95 y +2.95

**Color violeta**, ítem que no cumplen con criterio ajuste CCI, ya que fit >2.00.

**Color verde**, ítem que no cumplen con criterios PC y correlación rt⊖ ya que estos deben ser menores en los distractores, que en la clave, dato que corresponde al indicador **K** (keit-clave).



### Comentarios al cuadro síntesis de datos anterior:

#### 1º Análisis integrativo de los ítems por códigos:

Se ha dado el nombre de análisis integrativo de los ítems a esta parte del A.I. que intenta integrar los datos obtenidos en los pasos anteriores, para inferir los procesos mentales o componentes cognitivos que subyacen detrás de las respuestas.

El cuadro siguiente intenta reunir en un solo cuadro la distribución de los ítems que hay en la prueba, en las nueve categorías señaladas en el modelo de procesamiento adoptado, así como el número de ítems rechazados en el análisis cuantitativo, realizado por el método de la TRI.

#### Cuadro de frecuencias de ítems totales y rechazados en cada categoría

Ítems	1ºMC	2º MN	3º MO	4ºMM	5º MA	6º MG	7º MT	8ºMCC	9ºME
Nº ítems	3	5	8	0	5	9	5	5	0
Recha.	...0	...1	... 5	... 0	... 2	... 3	... 4	... 2	...0

#### Comentarios:

De modo que en toda la prueba el 42.5 % (17/40) son rechazados o deben ser modificados, los cuales se distribuyen de este modo entre las nueve categorías de contenidos establecidas:

1º -MC (Teoría de conjuntos): Ítems rechazados por el sistema (0/3)

ítems	9	10	25
Código	C.4	C.3	C,4
Indic.			

**Recomendaciones:** inicialmente ninguno de estos ítems debe ser rechazado o modificado.

2º -MN (Sistema numeral) Ítems rechazados por el sistema (1/5).

ítems	6	7	11	24	40
Código	N.4	N.3	N.2	N.2	N.3
Indic.					PK

**Recomendaciones:** en esta categoría el único ítem que presenta problemas es el ítem 40, que presenta estos indicadores:

P (valor  $b_i$  muy alto=3.00>2.95) y K, porque hay otra alternativa con mayor poder de atracción que la clave B y es la C, que tiene una correlación muy baja (0.02), por lo que se recomienda revisar bien el texto o cambiar el ítem, ya que no aporta nada a la fiabilidad del test.



3º -MO (destreza en operación de matemáticas): Ítems rechazados por el sistema (5/8).

ítems	3	12	14	17	18	23	35	39
Código	0.2.4	0.2.5	0.2.6	0.2.5	0.2.5	0.2.5	0.2.5	0.2.2
Indic	PK		P	RPK			RPK	P

**Recomendaciones:** en esta categoría, cinco de los ocho ítems son rechazados de este modo:

- el ítem **3** (referido a operaciones con números decimales) es rechazado por estos indicadores: **P**: valor  $b_i$  muy alto= $3.00 > 2.95$  y **K**, que indica que hay otra alternativa con mayor poder de atracción que la clave D. Esta es la C, que tiene una correlación  $r_{i\Theta}$  satisfactoria, por lo que se recomienda cambiar la clave D por la alternativa C.

- el ítem **14** (referido a operaciones con matrices) es rechazado por este indicador: **P**: valor  $b_i$  muy alto= $3.00 > 2.95$ , es decir que es un ítem muy difícil, dificultad debida quizás a que ese contenido no se ha estudiado debidamente en el bachillerato.

- el ítem **17** (referido a operaciones con números irracionales) es rechazado por estos indicadores: **R** ( $=2.5$ ), lo que indica que la CCI real no se ajusta a la teórica, **P** el valor  $b_i$  es muy alto= $3.00 > 2.95$ , y **K**.

Hay otra alternativa con mayor poder de atracción que la D y es la A, que tiene una correlación  $r_{i\Theta}$  satisfactoria, por lo que se recomienda cambiar la clave a la alternativa A.

- el ítem **35** (referido a operaciones con números irracionales) presenta estos indicadores: **R** ( $=2.73$ ), lo que indica que la CCI real no se ajusta a la teórica, **P** el valor  $b_i$  muy alto= $3.00 > 2.95$ , y **K**. Hay otra alternativa con mayor poder de atracción que la B, que tiene una correlación  $r_{i\Theta}$  satisfactoria, por lo que se recomienda cambiar la clave a la alternativa B.

- el ítem **39** (referido a operaciones con números decimales) es rechazado por este indicador: **P** el valor  $b_i$  muy alto= $3.00 > 2.95$ . Hay otra alternativa con un poder de atracción ligeramente mayor a la clave B y es la A, pero que tiene una correlación  $r_{i\Theta}$  nula, por lo que debe revisarse bien el texto del ítem o sustituirlo por otro.

4º -MM (sistema de medidas:0) no aparece ninguna a rechazar.

5º -MA (álgebra:5) Ítems rechazados por el sistema (2/5).

ítems	2	4	5	14	16
Código	A.1	A.2	A5	A.7	A.2
Indic	PK			P	



- el ítem **2** (referido a conceptos básicos de algebra) es rechazado por estos indicadores: **P** el valor  $b_i$  es muy alto= 3.00>2.95, la **K** indica que hay otra alternativa, la **C** que tiene mayor poder de atracción que la clave **A** y una correlación  $r_{i\Theta}$  aceptable, por lo se recomienda sólo cambiar la clave.

- **14** (referido a operaciones con matrices) presenta estos indicadores: **P** el valor  $b_i$  muy alto=3.00>2.95, que tiene una correlación  $r_{i\Theta}$  muy baja (0.03), por lo que se recomienda revisar bien el texto del ítem o sustituirlo por otro.

**6º -MG (Geometría: 9): Ítems rechazados por el sistema (3/9)**

<b>ítems</b>	<b>8</b>	<b>21</b>	<b>28</b>	<b>30</b>	<b>32</b>	<b>34</b>	<b>37</b>	<b>38</b>
Código	G-3	G.1	G.3	G.1	G.3	G.1	G.1	G.1
Indic	P	PK	RPK					

- el ítem **8** (referido a la solución de problemas con ángulos) se rechaza por este indicador: **P** el valor  $b_i$  es muy alto=3.00>2.95, que tiene otra alternativa con un poder de atracción mayor que la clave que es la **A**, además su correlación  $r_{i\Theta}$  es baja (0.05), por lo que se recomienda revisarse bien el texto del ítem o sustituirlo por otro si es necesario.

- el ítem **21** (referido a la solución de problemas con líneas) se rechaza por este indicador: **P** el valor  $b_i$  es muy alto=3.00>2.95, y **K** que tiene otra alternativa con un poder de atracción mayor que la **A**, pero que su correlación  $r_{i\Theta}$  es negativa, por lo que se recomienda revisar bien el texto del ítem o sustituirlo por otro ítem.

- el ítem **28** (referido a la solución de problemas con ángulo) se rechaza por estos indicadores: **R** (=2.13 la CCI real no se ajusta a la CCI teórica), **P** el valor  $b_i$  WA es muy alto=3.00>2.95, y por

**K**. Hay otras alternativas con mayor poder de atracción que la clave **B**, que son **A**, **C** y **D**, pero que tienen unas correlaciones  $r_{i\Theta}$  muy bajas o negativas (0.04, -0.03, -0.03), por lo que se recomienda revisar bien el texto del ítem o sustituirlo por otro ítem.

- **7º MT (trigon.:5) Ítems rechazados por el sistema (4/5).**

<b>ítems</b>	<b>15</b>	<b>20</b>	<b>22</b>	<b>29</b>	<b>36</b>
Código	T.2	T.3	T.2	T.1	T.3
Indic	RPK	RPK		P	PK

- el ítem **15** (referido a conceptos básicos sobre razones trigonométricas) se rechaza por estos indicadores: **R** (=2.23) la CCI real no se ajusta a la teórica, **P** el valor  $b_i$  es muy alto=3.00>2.95, y **K**. Hay otra alternativa con mayor poder de atracción que la clave **B**, y es la **D**, pero que tiene una correlación  $r_{i\Theta}$  muy baja (0.05), por lo

que se recomienda revisar bien el texto del ítem o sustituirlo por otro.

- el ítem **20** (referido a cálculos en el triángulo rectángulo) se rechaza por estos indicadores: **P** el valor  $b_i$  muy alto=  $3.00 > 2.95$  y **K**. Hay otra alternativa con mayor poder de atracción que la clave A, y es la B, pero que tiene una correlación  $r_{i\Theta}$  negativa (-0.05), por lo que se recomienda revisar bien el texto del ítem o sustituirlo por otro.

- el ítem **29** (referido a estimaciones trigonométricas en el triángulo rectángulo) se rechaza por estos indicadores: **P** el valor  $b_i$  es muy alto= $3.00 > 2.95$ . No hay otra alternativa con mayor poder de atracción que la B, pero ésta tiene una correlación  $r_{i\Theta}$  muy baja (0.05), por lo que se recomienda revisar bien el texto del ítem o sustituirlo por otro.

- **36** (referido a estimaciones en el triángulo rectángulo), se rechaza por estos indicadores: **P** el valor  $b_i$  es muy alto=  $3.00 > 2.95$ , y **K**. Hay dos alternativas con mayor poder de atracción: la A y la B, que la clave D, pero que éstas tienen una correlación  $r_{i\Theta}$  negativa (-0.01) y muy baja (0.03) por lo que se recomienda revisar bien el texto del ítem o sustituirlo por otro.

8°- **MCC** (coordenadas cartesianas:5) Ítems rechazados por el sistema (2/5)

	<b>19</b>	<b>26</b>	<b>27</b>	<b>31</b>	<b>33</b>
Código	CC.5	CC.5	CC.2	CC.1	CC.1
Indic	PK	PK			

- el ítem **19** (referido a proyección en espacios definidos por coordenadas cartesianas) se rechaza por estos indicadores: **P** el valor  $b_i$  es muy alto=  $3.00 > 2.95$ , la mayoría de las alternativas tienen mayor poder de atracción A, B y C, que la clave D, pero tienen una correlación  $r_{i\Theta}$  baja (0.02) o negativa (-0.05 y -0.03) por lo que se recomienda revisar bien el texto del ítem o sustituirlo por otro.

- el ítem **26** (referido a proyección en espacios definidos por coordenadas cartesianas) se rechaza por estos indicadores: **P** el valor  $b_i$  es muy alto=  $3.00 > 2.95$ , y **K**. Hay una alternativa con mayor poder de atracción, la B, que la clave A, pero que tiene una correlación  $r_{i\Theta}$  aceptable (0.09), por lo que se recomienda cambiar la clave a la alternativa B.

9° -**ME** (estad.:0) No aparece ningún ítem a rechazar o cambiar.

Sintetizando, de los 40 ítems que tiene la prueba MAT-forma A se aceptan 23 ítems (57.5%), y se rechazan 17 ítems (42.5%). Este rechazo se acentúa en estos contenidos:



MO (5/8): que indica un escaso dominio de realizar operaciones matemáticas con: matrices (3 en O.2.5), números complejos (1 en O.2.4) y logaritmos (1 en O.2.6).

MT (4/5): que indica un escaso conocimiento de los conceptos iniciales de trigonometría (1 en MT.1), un concepto poco claro de las funciones trigonométricas (1 en T.2), y no saber aplicar los conceptos trigonométricos al cálculo de elementos del triángulo rectángulo (2 en MT.3).

MG (3/9): que indica un escaso conocimiento en geometría en solución de problemas con ángulos (2 en G.3), y en los conceptos elementales de geometría (1 en G.1)

MA (2/5): que indica un escaso conocimiento de los conceptos iniciales de álgebra (1 en M.1), y sobre todo en la capacidad de resolver sistemas de ecuaciones (2 en MA.7).

MCC (2/5): que indica una escasa habilidad para hacer proyecciones bidimensionales (2 en MCC).

MN (sistemas numéricos y simbologías), que indica un escaso conocimiento para realizar combinaciones y permutaciones (1 en MN.3).

MM (sistemas de medida) y ME (estadística), que no aparece ningún ítem a rechazar o modificar.

• **Análisis de tendencias a selección inconsistentemente:** una alternativa determinada. Resulta interesante también ver si la distribución de las letras entre las opciones está sesgada o no

en la asignación de las letras a las cuatro alternativas:

**Distribución de frecuencia de claves en el cuadro (N=40):**

A	B	C	D
12	13	7	8

Aunque se ve la colocación de la clave entre las letras, no se encuentra evidencia de sesgo. Sí parece que esta distribución tiende a darle mayor importancia a las dos primeras posiciones.

**Presencia de los indicadores propios del XCALIBRE referentes a los 17 ítems rechazados por:**

Ajuste	Parámetros			Clave
R (fit)	ai	bi	ci	K
5	0	17	0	12

De los 17 ítems que rechaza el programa, cinco hacen referencia a la falta de ajuste (R-fit) de la CCI; doce a problemas de falta de conocimiento del tema tratado, o a un error en la determinación de la clave; y todos esos como es lógico resultan muy difíciles, bien por desconocimiento del tema, bien por una mala redacción del ítem, o bien por un error a la hora de señalar la clave.

**4º Diagnostico final de la prueba de conocimiento Matemáticas A**

Tal como aparece esta prueba no parece apta para evaluar el conocimiento de los alumnos en matemáticas, por estas razones:

- casi la mitad de los ítems (17) no son respondidos por los sujetos, ya que al parecer se han escogido ítems muy difíciles  $b_i > 2,95$ , y que resultan bien porque sus contenidos no se han desarrollado adecuadamente en los programas de la mayoría de los centros de secundaria, tales como: realizar operaciones matemáticas con matrices, números complejos, valores angulares, logaritmos, trabajar con funciones trigonométricas, o bien por una ignorancia supina de estos temas en la mayoría de los alumnos.

Sería bien interesante ver si este dato se da también en alumnos que provienen de centros privados, a los que se les atribuye haber recibido una mayor excelencia académica.

- En general se puede decir que la prueba tiene unos índices de dificultad (ID), de consistencia interna (ICI), y de fiabilidad (IF) globales bajos, en los dos métodos de análisis de ítems (AI) propuestos, como puede verse en el cuadro siguiente :

TCT		TRI	
Índice	valor	Par.	valor
ID	0.322	bi	2,22
ICI	0.192	ai	0.81
		ci	0.21
IF	0.398	IF	0.398

Estos valores permiten señalar que la prueba presenta:

- una **dificultad global** muy alta, calculada por los dos métodos:

ID en TCT =0,322, valor muy debajo del ID=0.50 ideal,

$b_i$  en TRI =2.22, que equivale a más de dos desviaciones típicas de la media teórica ( $b_i=0.00$ ).

- un **coeficiente de discriminación** o índice de consistencia interna aceptable, con reservas en los dos métodos:

ICI en TCT =0.192, situado en la zona de aceptación (0.07- y 0.25

$a_i$  en TRI =0.81, que es mayor que 0.30, que es la zona de rechazo.

- un **coeficiente de fiabilidad** o de precisión predictiva muy bajo en ambos métodos: IF en TCT =0,398 y en TRI =0.398, que resulta igual en ambos métodos, ya que se ha calculado a través del coeficiente de Cronbach, que es independiente de la dificultad de los ítems. Dato que indica que la probabilidad de estimar con precisión es del 40%, que es menor que si lo hiciéramos al azar, es decir, al 50%.

Se han propuesto dos métodos para efectuar el análisis de los ítems. Los dos ofrecen datos que permiten realizar la selección y corrección de los ítems con buen fundamento matemático, pero recomendamos utilizar preferentemente el método basado en la Teoría de la Respuesta al Ítem (TRI), que además de toda información



que ofrece el método basado en la Teoría Clásica de los Test (TCT), aporta nueva información, como la estimación del error en el cálculo de los valores paramétricos:  $a_i$ ,  $b_i$  y  $c_i$ , el ajuste al modelo de la CCI en cada ítem, el análisis de las alternativas con índices como los coeficientes bi-seriales mejores, ya que se utiliza directamente el valor  $r_{i\Theta}$  (ítem- aptitud- $\Theta$ ) en lugar de la suma de las respuestas buenas, etc. Pero sobre todo porque el programa XCALIBRE ofrece automáticamente criterios de selección, como son los indicadores literales que saca.

Además, los outputs que saca este programa facilita avanzar en las nuevas técnicas psicométricas, ya sus resultados pueden aplicarse directamente a las múltiples aplicaciones que se están desarrollando en las nuevas perspectivas de la Psicometría.

### Recomendaciones finales

Teniendo en cuenta los datos y comentarios de la prueba de conocimientos en Matemáticas, se recomienda: nombrar una comisión, que considerando las deficiencias encontradas en la misma TCT, y sobre todo en la TRI, reconstruya la prueba, eliminando la mayoría de los ítems que el programa XCALIBRE señala como inaceptables por su dificultad, e introduciendo otros que estén más de acuerdo con los conocimientos matemáticos que deben tener los alumnos, o mas de acuerdo con lo que se les ha enseñado.

El problema de la enseñanza de las Matemáticas en República Dominicana parece ser crónico, pues en todas evaluaciones realizadas (ver referencias bibliográficas, “Análisis de las aplicaciones de Prueba POMA desde el 2002 al 2014”, Díaz, 2007; Díaz 2013, y Díaz, 2014) y en otras se señala siempre la misma dificultad de los alumnos.

Por lo que nos preguntamos si el problema reside:

- en los instrumentos de evaluación utilizados, o, en el bajo nivel de conocimientos de los alumnos.
- ¿En el sistema educativo que no logra enseñar adecuadamente esta disciplina?
- ¿En la tradición cultural del país, que ha discurrido desde sus inicios por senderos literarios y filosóficos, y casi ninguno en las ciencias experimentales?
- ¿En los profesores que no logran enseñar adecuadamente las Matemáticas, como un instrumento necesario para el desarrollo intelectual y científico de los pueblos?

## REFERENCIAS BIBLIOGRÁFICAS

- Anderson, J.R. (1983): *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Díaz, J.V. (1997): *La Teoría de las respuestas a los ítems aplicada a la construcción de tests de aptitudes*. Cristóbal Serrano, Valencia.
- Díaz, J.V. (1998): *Métodos, Diseños y Técnicas de medición y evaluación psicológicas*. Cristóbal Serrano, Valencia.
- Díaz, J.V. (2002): *Procedimientos básicos para la construcción de tests de aptitud y rendimiento*. Cristóbal Serrano, Valencia.
- Díaz, J.V. (2004): *Psicometría: Construcción de instrumentos de medida: cuestionarios, escalas y tests*. Unidad IX. Cristóbal Serrano, Valencia.
- Díaz, J.V. (2005): *Aspectos del análisis de ítems en las diversas fases de la construcción de tests de inteligencia académica*. *Rev. Metodología de las Ciencias del Comportamiento*. Vol. especial. AEMCCO.
- Díaz, J.V., y otros (2007): *Hacia la evaluación de la inteligencia académica del estudiante dominicano: síntesis de estudios realizados con el test IAUD desde el 2002 al 2006*. SEESCYT. Programa Nacional de Publicaciones, Colección Humanidades, Serie: Investigaciones, Vol. 3
- Díaz, J.V. (2013): *El POMA.3, Análisis de sus aplicaciones en el MESCYT desde el año 2007 al 2011*. Libro en proceso de impresión en el Dpto. del MESCYT.
- Díaz, J.V. (2014): *El POMA.4, Análisis de sus aplicaciones en el MESCYT desde el año 2012 al 2014*. Informe entregado al Departamento de Pruebas Diagnósticas del MESCYT a principios del año 2015.
- Embretson, S.E. (1993): *Psychometrics models for learning and cognitive process*. In N.E. Frederiksen, R.J. Mislerly and I.I. Bejar (Eds). *Test theory for a new generation of tests*. Hillsdale, N. J. Lawrence Erlbaum.
- Hambleton, R. K. y Swaminathan, H. (1987). *Item response theory: Principles and Application*. Boston: Kluwer-Nijhoff Publishing.
- Hambleton, R.K. Swaminathan, H. y Rogers, H.J. (1991). *Fundamentals of Item response theory* Newbury: SAGE Publications.



- Hattie, J.A. (1985). Methodological review: assessing unidimensionality of tests and items. *Applied Psychological Measurement*, 9, 139-164.
- Martínez Arias, R.(1995): *Psicometría: Teoría de los tests psicológicos y educativos*. Madrid. Síntesis Psicología.
- Muñiz, J. (1990): *Teoría de respuesta al ítem*. Madrid. Pirámide.
- ITEMANW (1996): *User`s Manual for the ASC Item and test analysis Pack*,. Assessment Systems Corporation, St Paul Minnesota, 55114.
- Reckase, M. D. (1979): *Unifactor latent trait models applied to multifactor tests: Results and implications*. *Journal of educational statistics* 4(3), 207-230.
- Reckase, M.D. (1983). *The definition of difficulty and discrimination for multidimensional item response theory models*. Paper presented at the meeting of American Educational research Association, Montreal, Canada.
- Reckase, M.D. (1985). *The difficulty of test items that measure more than one ability* . *Applied Psychological Measurement*, 9, 401-412.
- Reckase, M.D. (1997). *A Linear Logistic Multidimensional Model for Dichotomus Item Response data*. En W.J. van der Linden y R.K. Hambleton (Eds.). *Handbook of modern items response theory* (pp. 271-286). New York: Springer-Verlag.
- Maydeu, A. (1996). *Modelos multidimensionales de la Teoría de respuesta a los ítems*. En J. Muñiz (ED.). *Psicometría*. (pp. 811-868). Madrid. Universidad.
- MESCYT (Ministerio de Educación Superior, Ciencia y Tecnología), Santo Domingo, República Dominicana.
- SEESCYT (Secretaría de Estado de Educación Superior, Ciencia y Tecnología), 2002: Ley 139-01 de *Educación Superior, Ciencia y Tecnología*, Santo Domingo, República Dominicana.
- SPSS: *Statistical Product and Service Solutions*, v.17 (2008), un programa estadístico informático muy usado en las ciencias sociales. Versión autorizada por la Universidad de Valencia para el uso de sus profesores.