

INVESTIGACIONES

*TEST INFORMATIZADOS Y EL REGISTRO DEL TIEMPO DE RESPUESTA,
UNA VIA PARA LA PRECISION EN LA DETERMINACION
DEL NIVEL DE LOGRO DE UN SABER MATEMATICO*

Computerized quiz and the timing of response, a way for the accuracy in determining the achievement of a mathematical knowledge

*Eduardo Cabrera¹, José González¹, Eduardo Montenegro¹, Alejandro Nettle¹,
Miguel Guevara²*

¹Equipo Laboratorio [EXPERIMENTAL] de Saberes Matemáticos (i.e., Equipo *Lab[e]saM*) inserto en el Departamento de Matemática y Física, Facultad de Ciencias Naturales y Exactas, Universidad de Playa Ancha, Valparaíso, Chile.

²Ingeniero Informático, Académico del Departamento de Ingeniería, Facultad de Ciencias Naturales y Exactas, Universidad de Playa Ancha, Valparaíso, Chile.
E-mail: labesam.upla@gmail.com

Resumen

Preocupados por la educación de calidad, se presenta un proceso de precisión en la determinación del nivel de competencias matemáticas de los alumnos, por medio de una reparametrización del modelo logístico en función del tiempo de respuesta, parámetro estimado a través de la rendición de test virtuales. Dicho trabajo tiene como base teórica el procesamiento de los saberes según *Lab[e]saM* complementado por la incorporación de tecnologías.

Palabras clave: saberes según *Lab[e]saM*, Función Logística, test informatizados, Teoría de Respuesta al Item, Modelo Logístico de tres parámetros.

Abstract

Worried about an education of quality, a process of precision in the determination of the level of mathematical competitions of the students appears, by means of a parameterization of the logistic model based on the response time, parameter estimated through the virtual tests surrender. This work has as theoretic bases the processing of knowledge according to *Lab[e]saM* complemented by the incorporation of technologies.

Key words: knowledge according to *Lab[e]saM*, logistic function, computerized test, Item response model, three parameters logistic model.

1. INTRODUCCION

Actualmente en el ámbito de la evaluación educacional se está interesado en determinar con mayor precisión el nivel de competencias de nuestros alumnos, dando origen a una serie de estrategias y teorías, resaltando entre estas la Teoría de Respuesta al Ítem. Esta teoría plantea la existencia de una relación funcional entre la probabilidad de acierto a un determinado ítem y el nivel de competencia del alumno, siendo esta relación identificada a través del modelo Logístico.

Este modelo y modificaciones de él, como el modelo Logístico de uno, dos o tres parámetros; ha sido el resultado de esta incansable búsqueda de precisión.

Esta última década, tomó un fuerte realce la incorporación de herramientas computacionales en la aplicación de test, siendo la razón fuerza, el hecho de poder obtener resultados y mediciones en forma inmediata, es decir, por la facilidad de cálculo, entre otras importantes utilidades como la reducción de falla humana en la calificación, facilidad de administración de calificaciones o modelo adaptivo del cuestionario según las respuestas que va entregando el estudiante.

La propuesta tiene como sustento la teoría de respuesta al ítem, pero con la diferencia en la incorporación de un nuevo parámetro en el modelo caracterizado como tiempo de respuesta y el sustento teórico definido por *Lab[e]saM*, que se resumirá en el punto 5, de algunos análisis preliminares.

2. EL PROBLEMA DE INVESTIGACION

En la actualidad el uso de la tecnología ha permitido desarrollar una serie de estrategias metodológicas para la medición del logro de una competencia, permitiendo además considerar una serie de factores y cálculos inmediatos, por tal razón hemos querido utilizar esta información, en especial el tiempo de respuesta, como un factor influyente en el nivel de competencia del alumno. En la actualidad ya no importa solamente que el profesional responda de manera correcta una situación problemática, sino que la resuelva en el menor tiempo posible.

Por tal razón, el problema de este trabajo es: *“Definir una variación del Modelo Logístico, en la cual se considere el parámetro tiempo de respuesta, considerada en esta visión como un factor relevante a la hora de la determinación del nivel de competencia de un alumno”*.

3. OBJETIVOS DE LA INVESTIGACION QUE SE PRESENTA

- Promover el uso de recursos tecnológicos para los procesos de evaluación.
- Presentar un modelo de probabilidad adaptado a la consideración del tiempo de respuesta como un parámetro discriminativo en la determinación del nivel de competencia.
- Caracterizar el parámetro temporal en el modelo.
- Simular la metodología de trabajo bajo la propuesta.
- Promover el uso de la Teoría de Respuesta al Ítem.

4. METODOLOGIA DE LA INVESTIGACION

La concepción de los saberes y la Teoría de Respuesta al Item (*i.e.*, IRT)¹ serán el soporte teórico y andamiaje de nuestra propuesta metodológica en el análisis de la información en función del nivel de competencia de los alumnos.

El *parámetro temporal o tiempo de respuesta* conforma el objeto de diseño de nuestra investigación, pues sobre él se desarrolla la cuantificación o estimación del nivel de competencia del alumno a través del método de estimación de máxima verosimilitud.

Fases de la Investigación:

1. Análisis preliminar.
2. Concepción de la investigación.
3. Simulaciones.
4. Análisis *a posteriori* y evaluación.

5. ALGUNOS ANALISIS PRELIMINARES

La concepción teórica *Lab[e]saM* está soportada por el marco de la Teoría de las Situaciones Didácticas (Brousseau, 2004), y particularmente en la aproximación antropológica propuesta por Chevallard (1991), caracterizando así tres facetas:

- *Saber Sabio* generado en las comunidades científicas. Para esta investigación consideraremos que el saber Sabio posee infinitas *dimensiones*² no manejables en el sentido estadístico.
- *Saber a Enseñar* se orienta a la labor del enseñante (*i.e.*, el que enseña) que debe producir la recontextualización y elección responsable de los conocimientos que pretenden transformarse en conocimiento para el alumno, es decir, la labor del enseñante es elegir y construir un modelo (y como tal, reduce la complejidad) adecuado al contexto y al *saber sabio*, para transformarlo en un saber adecuado para enseñar.
- *Saber Aprendido*, representa el modelo construido por el alumno como producto del trabajo intelectual en su interacción con el modelo enseñado.

Como postulado de la investigación se presenta que la calidad del modelo definido por el alumno depende en esencia de la calidad de la *transposición didáctica*; por un lado, el *saber aprendido* se medirá bajo la forma de preguntas y respuestas del tipo Bernoulli (*i.e.*, dicotómicas), orientadas a las *dimensiones* independientes que constituyen un saber, por otro se asume que los alumnos poseen las competencias básicas para una correcta decodificación.

¹ La Teoría de Respuesta al Item (IRT) se atribuye al trabajo del matemático danés Georg Rasch (1960), aunque se encuentran también importantes semejanzas en el trabajo de Louis Leon Thurstone, 1912 y Alfred Binet, 1905

² En este trabajo el sustantivo *dimensión* alude a cada una de las características esenciales del saber.

Pensamos que, a través de este trabajo, podemos iniciar un camino hacia la caracterización de las variables que tienen mayor incidencia en el modelo que el alumno construye en una situación de enseñanza, y en la calidad de éste.

Formalización:

Sean Θ el universo de los *saberes sabios*, donde cada $\theta \in \Theta$ tiene asociado un conjunto X_θ , constituido por todas las *dimensiones* de θ , que suponemos es de tamaño no finito, y

$$\mathfrak{S} = \bigcup_{\theta \in \Theta} 2^{X_\theta}$$

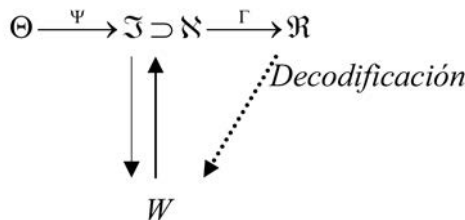
donde 2^{X_θ} es la familia de las partes de X_θ . Además; sea $\Psi : \Theta \rightarrow \mathfrak{S}$ una aplicación que denominaremos *Operador de Elección* donde la imagen de $\theta \in \Theta$, a través de Ψ , elige un subconjunto finito y no vacío de X_θ que se espera caracterice a θ .

Observemos que \mathfrak{S} es una σ -álgebra sobre $\{X_\theta : \theta \in \Theta\}$. Por otro lado, se define un *Operador de Transposición* Γ , cuyo dominio es \mathfrak{K} , una familia de conjuntos finitos de \mathfrak{S} , esto es, $\mathfrak{K} \subset \mathfrak{S}$, cuyas proyecciones definen una familia de *saberes enseñados*, que a través de un proceso de decodificación el alumno lo transforma en *saber aprendido*.

Esta familia de *saberes aprendidos*, W , contiene algunas *dimensiones* y *seudodimensiones*³ del saber, cuya $\#W \leq \#\mathfrak{K}$, según lo anterior interesa determinar cuáles fueron aquellas *dimensiones* que constituyen a W que son parte de \mathfrak{K} , además determinar qué tan distante se encuentran estas *seudodimensiones* de lo que debieran haber sido y por último cuál es la *distancia* existente entre el *saber aprendido* y *saber a enseñar*, es decir, cuál es la *distancia* entre el modelo definido por el alumno a través de la transposición (*saber aprendido*) y el *saber a enseñar*.

Adviértase que si $\#W \geq \#\Psi(\Omega)$, la *distancia* entre el *saber aprendido* y el *saber a enseñar* se considerará nula.

Visualización:



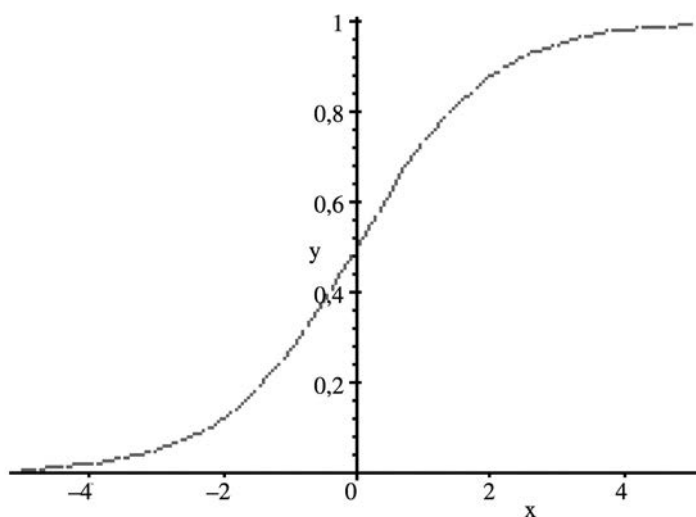
³ En este caso utilizamos la palabra *seudodimensión* para hacer alusión a la dimensión de saber que posee el alumno, y que no coincide con las dimensiones del *saber enseñado* que recibió.

Por otro lado, desde un punto de vista estocástico, consideremos que el poseer la *dimensión i-ésima* del saber es una variable aleatoria del tipo Bernoulli, donde la probabilidad de éxito está modelada por la siguiente función de distribución,

$$P(\theta_{ij}) = \frac{e^{(\theta_{ij} - b_i)}}{1 + e^{(\theta_{ij} - b_i)}}$$

que es función de θ_{ij} , que representa el nivel de competencia de la *j-ésima* persona en la *i-ésima dimensión* del saber θ , además b_i representa el nivel de importancia de la *i-ésima dimensión* del saber en estudio, gráficamente se tiene:

Gráfico 1



Permitiendo visualizar que a mayor nivel de competencia en la *dimensión* del saber en estudio (en el gráfico el eje de absisas), mayor es la probabilidad de éxito (eje de ordenadas). Así, la probabilidad de que la *j-ésima* persona construya la *i-ésima* dimensión está dada por:

$$P(X_{ij} = x) = (P(\theta_{ij}))^x (1 - P(\theta_{ij}))^{1-x}$$

donde $x = 0,1$ denota, respectivamente, fracaso o éxito.

5.1. ALGUNOS ASPECTOS HISTORICOS

En muchos sistemas educativos se da la necesidad de disponer de un mecanismo para evaluar la adquisición de conocimiento de los alumnos, algo fundamental para identificar el éxito o fracaso en el proceso de aprendizaje. Los tests son probablemente hoy en día el modo más habitual de efectuar dicha tarea, pudiendo estar ligado a diferentes contextos educativos (Cortada de Kohan, N., 2003).

Históricamente, la primera teoría fue la que hoy se conoce como Teoría Clásica de los Test (T.C.T), que surgió a comienzos del siglo XX. Si bien no recibió su forma axiomática hasta mediados de los 60, la TCT se fundamenta en el denominado modelo clásico que establece una relación lineal entre el nivel del rasgo del examinado y la puntuación obtenida en el test realizado (López, J.; T. Pérez & A. Armendáriz, 2004).

$$X = V + \varepsilon$$

Este modelo se expresa en términos de la puntuación empírica del test, el elemento central sobre el que gira toda la teoría. Concretamente, se considera que la puntuación empírica del sujeto, esto es, el valor observado en el test, es igual a la suma de dos componentes hipotéticos y desconocidos *a priori*: la puntuación verdadera o nivel del rasgo real del evaluando y un determinado error de medida (Muñiz, J., 1997).

La principal limitación de la TCT es que en su contexto las características del test y las del examinado son dependientes, es decir, las mediciones obtenidas dependen por lo general de la naturaleza del test utilizado; y a la inversa, las propiedades de los test dependen de los sujetos a los cuales se aplica. Así, para la TCT, el nivel del rasgo del alumno se mide mediante el número de respuestas acertadas en el test realizado. Por lo tanto, los resultados estarán siempre relacionados con el test administrado: si las preguntas son difíciles, el nivel del rasgo de los examinados resultará ser bajo porque habrá pocos que la respondan correctamente; y viceversa. Así, la dificultad estimada de un test según la TCT dependerá de quién lo realice: si los examinados son muy listos, responderán correctamente a las preguntas y, por tanto, el test será considerado fácil. Estas dependencias entre test y examinados supone que las medidas obtenidas en la TCT no se pueden utilizar en otros contextos de manera generalizada, y por tanto comparar individuos que hayan hecho exámenes distintos resultará muy difícil (Hambleton, R., 1990).

La Teoría de Respuesta al Item supera esta limitación, pues propone modelos matemáticos orientados a las preguntas (que en este contexto reciben el nombre de ítem), en contraposición a la TCT, para el cual la evaluación de conocimientos gira en torno al test como unidad. Se trata de una teoría relativamente joven, data de comienzos de los años 60, habiendo sido implementada por primera vez durante los 80 (López, J.; T. Pérez, & A. Armendáriz, 2004).

Dicho de paso, la Psicometría y Edumetría se ocupan de los problemas de medición en Psicología y Educación, utilizando la estadística como pilar básico para la elaboración de teorías y para el desarrollo de métodos y técnicas específicas de medición (Olea, J., V. Ponsoda, F. Abad, J. Revuelta, B. Gil & C. García, 2004).

5.2. VISION GENERAL DE AMBAS TEORIAS

En este apartado se analizarán dos de las principales teorías que se utilizan en el ámbito de la medición educacional: La Teoría Clásica (TC) y la Teoría de Respuesta al Ítem (TRI).

Primero que todo, se hace necesario definir algunos términos. Distintas pruebas miden diferentes características de los examinados, por ejemplo: conocimiento matemático, razonamiento científico, capacidad de memorización, vocabulario, etc. En medición educacional, por “*Rasgo*” se entenderá a la característica del examinado que va a ser medida a través de la prueba y, por consiguiente, ésta se utiliza para representar de manera genérica cualquiera de esas características. Por otro lado, en lo que respecta a los ítems de una prueba, hay parámetros (o descriptores) que permiten describir sus atributos particulares. Los parámetros de un ítem que suelen ser más frecuentemente utilizados y, por consiguiente estimados son:

- a) Nivel de Dificultad.
- b) Nivel de Discriminación de un Ítem.
- c) Efectos del Azar.

Teoría Clásica

Una forma de medición que se utiliza con mucha frecuencia en pruebas de gran escala, como las de admisión a la universidad (por ejemplo, en el SAT de Estados Unidos y en la PSU en Chile), es la llamada Teoría Clásica (TC). En teoría clásica, el indicador del rasgo de un estudiante corresponde al puntaje que se obtuvo en la prueba, construido a partir del número de respuestas correctas (o número de respuesta correctas netas) que obtuvo. Como indicador de la dificultad de una pregunta, en este sistema se utiliza la razón entre el número de personas que acertaron correctamente el ítem y el total de alumnos que dieron una respuesta a éste. Por otra parte, el índice de discriminación de un ítem se calcula como la correlación entre el puntaje de la respuesta a éste y el puntaje en la prueba total. Además, en los test formados por ítem de opción múltiple, en las que sólo una alternativa es la correcta, podemos sobrestimar la puntuación directa de una persona dado que alguno de sus aciertos ha podido producirse por azar. El problema entonces consiste en establecer un procedimiento para descontar del número total de aciertos que se han producido por azar. Si se asume que no se conoce la respuesta correcta a un ítem, todas las alternativas de respuesta son equiprobables y la probabilidad de acertar al azar ese ítem se puede establecer como: $\frac{1}{n}$, donde n es el número de alternativas del ítem.

Estos descriptores de los ítems pueden calcularse ya sea en el contexto de una prueba piloto o experimental, o en la prueba definitiva u operacional.

Como se puede apreciar, en la TC el grado del rasgo de una persona depende del grupo de ítems (vale decir, de sus diferentes parámetros) que contiene la prueba. Por ejemplo, si la prueba es fácil, un mismo alumno tendrá un puntaje mayor que si la prueba es difícil. Con esto resulta difícil hacer comparaciones entre estudiantes que han rendido pruebas diferentes. A su vez, la estimación de los parámetros depende del grupo de personas a las cuales se aplicó el test. Así, un mismo ítem puede ser catalogado como

fácil si el grupo que rindió la prueba es excepcionalmente hábil, pero como difícil si el grupo que rindió la prueba es desaventajado. Con respecto a la discriminación, un ítem puede aparecer muy discriminativo en el contexto de un grupo con nivel heterogéneo del rasgo, pero poco discriminativo si el grupo que rindió la prueba es homogéneo (es decir, si todos los estudiantes tienen un nivel del rasgo similar). Esta debilidad, que se genera en la estimación del rasgo de los examinados y los parámetros de los ítems, ha llevado a buscar un método que permitiese una medida del rasgo que sea independiente de los ítems a los que estos se han enfrentado, una caracterización de los ítems independiente de la población a la que se aplica, y al mismo tiempo una medida más fiel de la precisión con que se está midiendo el rasgo.

La satisfacción de estos requerimientos se ha intentado encontrar en la Teoría de Respuesta al Ítem o TRI.

Teoría de Respuesta al Ítem

Los modelos de la TRI se centran en los ítems e intentan establecer, para cada uno de ellos, la probabilidad de ser respondidos correctamente. Esta probabilidad depende del rasgo del examinado y de ciertas características de los ítems, entre las que se encuentra el grado de dificultad, discriminación y la probabilidad de ser respondido correctamente por azar por una persona de bajo nivel del rasgo. Hay varios modelos en TRI de distinta complejidad. El modelo más simple es aquel que sólo diferencia los ítems según su grado de dificultad. Sin embargo, otros modelos permiten además incluir otros parámetros de los ítems, como su grado de discriminación y la probabilidad de responderlo correctamente al azar.

Aparte de lo anterior, la TRI hace posible conocer el nivel de precisión que un ítem aporta a la estimación para cada nivel del rasgo. En términos técnicos, esto es, lo que se conoce como Información de Fischer, que contextualizado se llama Información del Ítem. Mientras mayor es la información que aporta el ítem a un determinado nivel del rasgo, mayor es la precisión en la estimación de ese nivel.

La principal ventaja teórica de la TRI es que mediante su utilización se lograría que un estudiante obtuviese siempre la misma estimación de su nivel del rasgo, independientemente de las preguntas que le tocó responder. También, con la TRI un ítem tendría siempre los mismos parámetros que lo describen (dificultad, discriminación, azar), independientemente del grupo que rindió la prueba. Esta notable propiedad se llama invarianza, siendo ésta la principal ventaja que distingue a la TRI de la Teoría Clásica. Sin embargo, es preciso destacar que la invarianza se cumple siempre y cuando se satisfagan ciertos supuestos y requisitos que se enunciarán a continuación.

Supuestos de la TRI

1. Unidimensionalidad: este supuesto consiste en que en una prueba todos los ítems están midiendo uno y sólo un rasgo en los examinados.
2. Independencia Local: postula que, dado un nivel del rasgo, las respuestas a los ítems no pueden estar correlacionadas entre sí.
3. Igualdad de Educación: que todos los alumnos que rindan la prueba hayan tenido experiencias educacionales similares.

4. Que la prueba no haya sido apurada, es decir, que los tiempos de estimados en el desarrollo de la prueba se respeten.
5. Efectos de Contextos no controlados: se refiere a que algunas preguntas se comportan de modo diferente según la posición que tengan en la prueba.

Ventajas y Desventajas de la TRI

La Teoría de Respuesta al Ítem presenta una serie de potenciales ventajas sobre la teoría clásica. La principal de ellas es la invarianza de los puntajes de la prueba y de las características de las preguntas. También surge, gracias a las curvas de información, herramienta exclusiva de la TRI, la posibilidad de optimizar el proceso de selección de preguntas según el objetivo que se busca.

Todas las ventajas anteriores se pierden cuando los supuestos y requisitos de la TRI no se cumplen. Para que haya invarianza, principal ventaja de la TRI sobre la TC, es fundamental que exista, como se señalaba anteriormente, unidimensionalidad e independencia local. Sin embargo, en muchas ocasiones la naturaleza de las disciplinas mismas les impide someterse a tales restricciones. Es por ello que en la práctica los supuestos y requisitos de la TRI se transgreden (Dussaillant, F. 2003) a menudo. Este incumplimiento de los supuestos lleva no sólo a perder la invarianza sino que afecta directamente a la estimación del rasgo e introduce errores en aplicaciones secundarias de la teoría.

6. CONCEPCION DE LA INVESTIGACION

Este nuevo parámetro tiene como postulados de la propuesta las siguientes relaciones:

- Si en un ítem se requiere de mayor tiempo para su respuesta implicará mayor dificultad para el alumno.
- Si en un ítem se requiere de mayor tiempo para su respuesta implicará mayor discriminación.

6.1. PROPUESTA

La propuesta metodológica que se presenta tiene su sustento en la Teoría de Respuesta al Ítem.

La propuesta se basa en el hecho de poder cuantificar el tiempo de emisión de respuesta, desde el momento que la persona recibe el ítem hasta que emite una respuesta, este tiempo será caracterizado por t_{ij} , que representa el tiempo que demora la persona j -ésima en emitir una respuesta al ítem i del test.

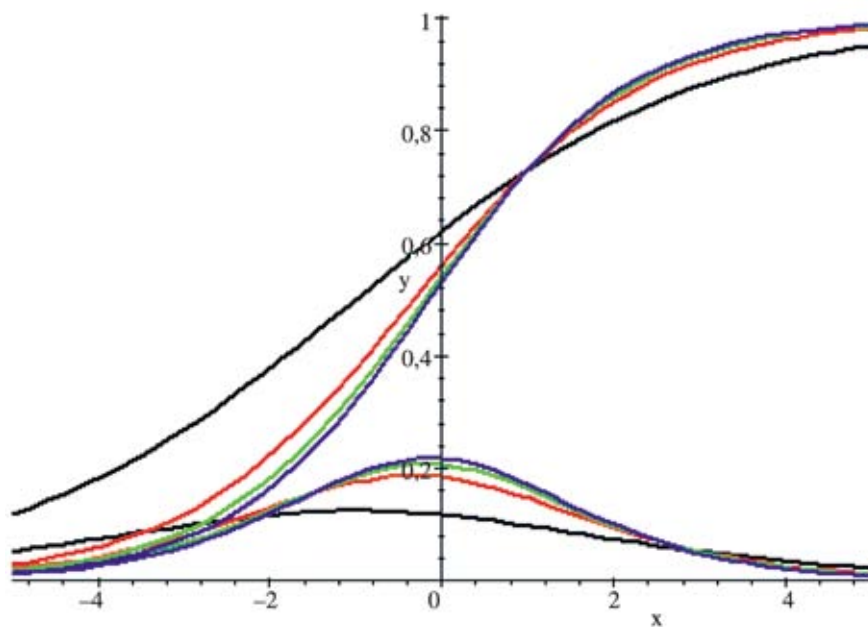
La incorporación de este nuevo parámetro permitirá desarrollar un enfoque de la teoría de los test hacia las pruebas computarizadas, cuyas consecuencias serán el poder diferenciar a dos personas que tradicionalmente se encuentran en igual condición.

Considerando $b_i = 0$, $a_i = 1$ y $c_i = 0$ parámetros de dificultad, discriminación y azar respectivamente en el modelo logístico de tres parámetros, se propone el siguiente modelo, denominado Modelo Logístico Reacondicionado, que incorpora el parámetro temporal t_{ij} , como:

$$P(\theta_{ij}) = \frac{e^{\frac{1}{t_{ij}} + \left(1 - \frac{1}{t_{ij}}\right)\theta_{ij}}}{1 + e^{\frac{1}{t_{ij}} + \left(1 - \frac{1}{t_{ij}}\right)\theta_{ij}}}; t_{ij} > 0; t_{ij} \in ZI; \forall ij$$

Gráficamente, es posible verificar los efectos t_{ij} sobre los parámetros clásicos de dificultad y discriminación en el modelo logístico de uno y dos parámetros, esto es:

Gráfico 2



- Curva de color negro representa al modelo y la densidad, condicionados a $t_{ij} = 2$
- Curva de color rojo representa al modelo y la densidad, condicionados a $t_{ij} = 4$
- Curva de color verde representa al modelo y la densidad, condicionados a $t_{ij} = 6$
- Curva de color azul representa al modelo y la densidad, condicionados a $t_{ij} = 8$

Es posible observar que a medida que aumenta el parámetro temporal, es decir, el tiempo que transcurre desde que la persona recibe el ítem hasta que emite una respuesta, se producen perturbaciones respecto de los parámetros de dificultad y discriminación, estableciéndose una relación proporcional en forma directa.

Análíticamente, se define al parámetro temporal t_{ij} como un parámetro de localización y de escala, a partir del cual se definen una familia de densidades que indizan a t_{ij} , esto es:

$$f_f(\theta) = \left\{ f\left(\frac{1}{t} + \left(1 - \frac{1}{t}\right)\theta\right) = f_{t_{ij}}(\theta) = \frac{\left(1 - \frac{1}{t}\right)e^{\frac{1}{t_{ij}} + \left(1 - \frac{1}{t_{ij}}\right)\theta}}{\left(1 + e^{\frac{1}{t_{ij}} + \left(1 - \frac{1}{t_{ij}}\right)\theta}\right)^2}; t_{ij} \in \mathbb{Z}^+, \forall ij \right\}$$

Notemos que

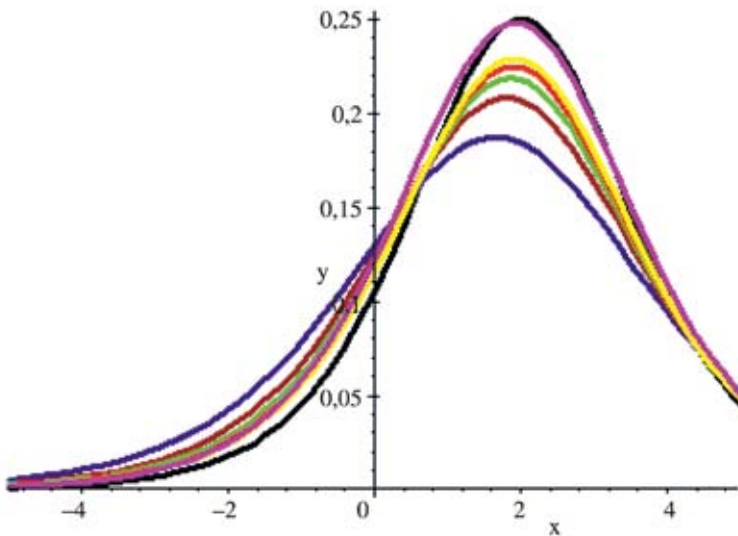
$$P_t(\theta_{ij}) = \frac{e^{\frac{1}{t_{ij}} + \left(1 - \frac{1}{t_{ij}}\right)\theta_{ij}}}{1 + e^{\frac{1}{t_{ij}} + \left(1 - \frac{1}{t_{ij}}\right)\theta_{ij}}}; t_{ij} > 0; t_{ij} \in \mathbb{Z}; \forall ij,$$

satisface lo siguiente:

$$\lim_{t_{ij} \rightarrow \infty} \frac{e^{\frac{1}{t_{ij}} + \left(1 - \frac{1}{t_{ij}}\right)\theta_{ij}}}{1 + e^{\frac{1}{t_{ij}} + \left(1 - \frac{1}{t_{ij}}\right)\theta_{ij}}} = \lim_{t_{ij} \rightarrow \infty} P_t(\theta_{ij}) = \frac{e^\theta}{1 + e^\theta} = P(\theta_{ij})$$

Es decir, la libertad temporal nos permite recuperar el modelo original de tres parámetros, gráficamente se tiene:

Gráfico 3



El gráfico anterior muestra la convergencia de las densidades, ante variaciones del parámetro temporal, a la densidad independiente de éste (Curva de color negro).

Modelo Logístico Temporal

Este modelo representa la probabilidad de que la persona j acierte al ítem i , con un nivel de competencia θ , condicionado a un tiempo de respuesta t_{ij} .

$$P(\theta_{ij}) = \frac{e^{\frac{1}{t_{ij}} + \left(1 - \frac{1}{t_{ij}}\right)\theta_{ij}}}{1 + e^{\frac{1}{t_{ij}} + \left(1 - \frac{1}{t_{ij}}\right)\theta_{ij}}}; t_{ij} > 0; t_{ij} \in ZI; \forall ij$$

Debemos notar, que en este modelo no se presentan los parámetros de dificultad, discriminación y azar, con el objetivo de resaltar la influencia del tiempo de respuesta en el modelo.

Una vez caracterizado el parámetro temporal, el siguiente paso es aplicar las pruebas, registrar el tiempo de respuesta y estimar el nivel de competencia de cada estudiante a través de la estimación máximo verosímil.

7. SIMULACIONES

Para este proceso simularemos las respuestas, el tiempo de respuesta y los parámetros de cada ítem, en 3 estudiantes para ejemplificar.

Ahora supongamos que las pruebas fueron suministradas y los resultados son los siguientes:

- Alumno 1. Obtuvo lo siguiente:

Tabla 1

Resultados obtenidos por la persona 1

Item	1	2	3	4	5	6	7	8
Tiempo	5	7	3	8	10	3	6	2
Respuesta	1	1	0	0	0	1	1	0

- Alumno 2. Obtuvo lo siguiente:

Tabla 2

Resultados obtenidos por la persona 2

Item	1	2	3	4	5	6	7	8
Tiempo	3	7	8	10	12	5	7	9
Respuesta	1	1	0	0	0	1	0	0

- Alumno 3. Obtuvo lo siguiente:

Tabla 3

Resultados obtenidos por la persona 3

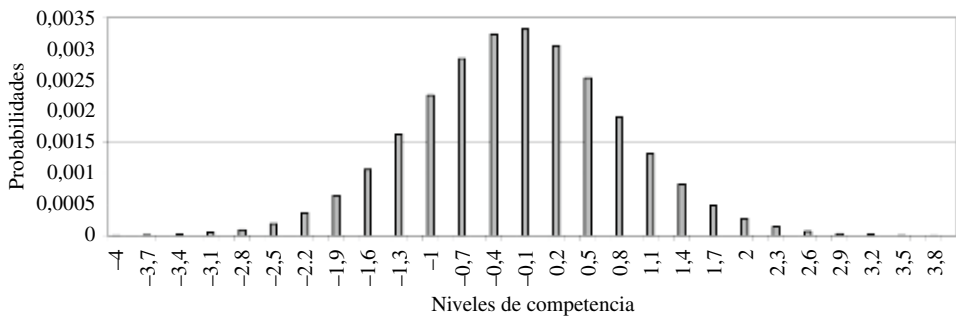
Item	1	2	3	4	5	6	7	8
Tiempo	2	3	3	4	3	5	7	4
Respuesta	1	1	1	0	0	0	1	1

Basado en lo anterior se procede a estimar el nivel de competencias de cada una de las tres personas. En el proceso de estimación se considerarán $b_i = 0$, $a_i = 1$ y $c_i = 0$.

Estimación Persona 1:

Gráfico 4

Nivel de Competencia estimado para la persona 1

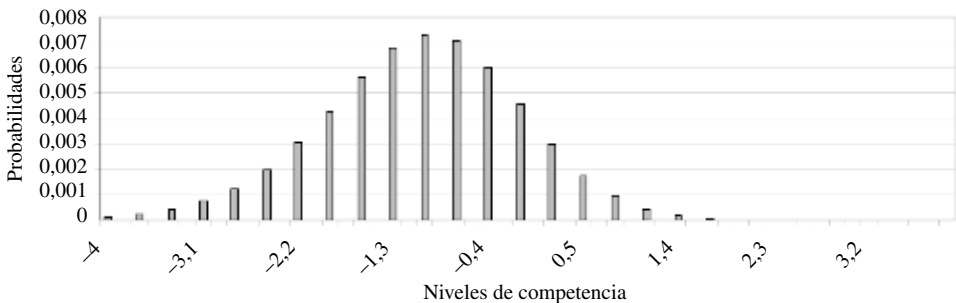


Observemos que el nivel de Competencia que hace más probable el patrón de respuestas de la persona 1 es $\theta_1 = -0.1$.

Estimación Persona 2:

Gráfico 5

Nivel de Competencia para la persona 2

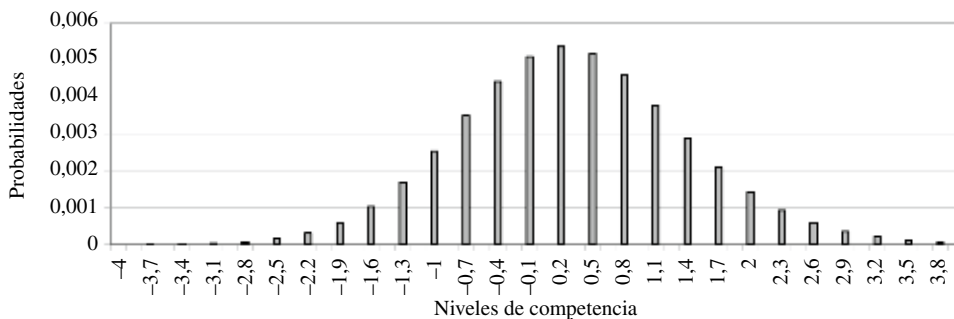


Observemos que el nivel de Competencia que hace más probable el patrón de respuestas de la persona 2 es $\theta_2 = -1.0$. Además es un nivel notoriamente menor a la persona 1.

Estimación Persona 3:

Gráfico 6

Nivel de Competencia para la persona 3

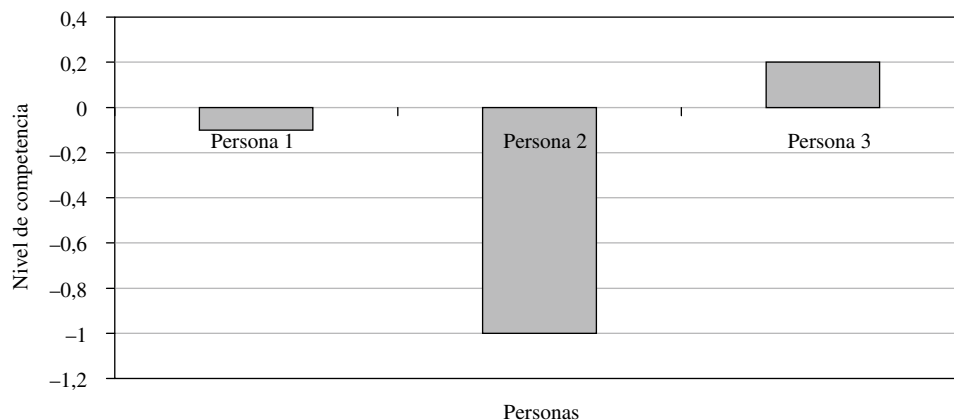


Observemos que el nivel de Competencia que hace más probable el patrón de respuestas de la persona 3 es $\theta_3 = 0.2$. Además es un nivel notoriamente superior a las dos personas anteriores.

En el siguiente gráfico se presentan las variaciones de los niveles de competencia de las tres personas en estudio, claramente se observa la superioridad en la competencia en estudio de la persona 3 respecto de las otras dos.

Gráfico 7

Nivel de Competencia de Cada Persona



8. ANALISIS A *POSTERIORI* Y EVALUACION

Es posible visualizar en las tablas de las simulaciones [Tabla 1, Tabla 2 y Tabla 3], que los tiempos de respuesta son menores en la persona número 3, que gráficamente en el proceso de estimación del nivel de competencia reflejan su efecto en la forma de la distribución, hacia una asimetría de tipo negativa, es decir, el máximo de la función de verosimilitud tiende a presentarse en torno a valores altos en comparación con sus pares, dicho efecto se potencia ante la utilización del parámetro temporal. Si utilizamos el modelo en evaluación de procesos en el tiempo, nos permitirá capturar pequeños incrementos y cambios en tiempos distintos y evaluar procesos de intervención pedagógica.

9. CONCLUSIONES

Preocupados por la educación, la calidad y la precisión en los procesos de medición, el equipo *Lab[e]saM* propone esta metodología de trabajo y la variación del Modelo Logístico, que en este caso sólo se está considerando el tiempo de respuesta como un factor diferenciador entre las puntuaciones de los alumnos.

Vivimos en el siglo de las tecnologías, y no nos debemos de extrañar cuando comience la masificación de pruebas rendidas por medio de computadores y, bajo esa perspectiva, debemos rescatar de ello toda la información posible, de tal manera de precisar y diferenciar el nivel de competencia de cada uno de nuestros alumnos.

Por otro lado, hemos considerado el tiempo de respuesta en el proceso de diferenciación, pues se está transformando en un factor importante a la hora de tomar decisiones respecto de la resolución de problemas, además de ser un reflejo de la calidad de la construcción de un saber para reaccionar ante un determinado estímulo.

10. AGRADECIMIENTOS

Paula Sofía Ponce Grenet
Manuel Freire del Pozo

11. BIBLIOGRAFIA

- Brousseau, G. (2004). *Théorie des Situations Didactiques (Didactiques des Mathématiques 1970-1990)*. (2ª edición), Grenoble, Francia: La Pensée Sauvage, Editions.
- Chevallard, Y. (1991). *La transposition didactique. Du savoir savant au savoir enseigné*. (2ª edición en colaboración con Marie-Alberte Joshua), Grenoble, Francia: La Pensée Sauvage, Editions.
- Cortada de Kohan, N. (2003). Posibilidad de integración de las teorías cognitivas y la psicometría moderna. (pp. 08-23). Argentina: *Revista Argentina de Neuropsicología*.
- Dussaillant, F. (2003). *Técnicas de medición en pruebas de admisión a las universidades*. Santiago, Chile. *C.E.P.*
- Hambleton, R. (1990). *Fundamentals of Item Response Theory*. (pp. 15-31). Educational Measurement.

- López, J., T. Pérez y A. Armendáriz (2004). Evaluación Mediante Test, Departamento de Lenguas y Sistemas Informáticos. Universidad del País Vasco, España. *Revista Iberoamericana de Educación*.
- López, J., T. Pérez y A. Armendáriz (2004). Evaluación Mediante Test, Departamento de Lenguas y Sistemas Informáticos. Universidad del País Vasco, España. *Revista Iberoamericana de Educación*.
- Mabille, A. & V. Albert (1994). Profils de profs: Portarait et styles d'enseignants en sciences. (2ª edición). Louvain-la-Neuve, Bélgica: De Boeck Université.
- Muñiz, J. (1997). Teoría Clásica de los Test. (pp. 26). Universidad de Oviedo, España. Editorial Pirámide.
- Olea, J., V. Ponsoda, F. Abad, J. Revuelta, B. Gil y C. García (2004). Introducción a la Psicometría, Teoría Clásica de los Test y Teoría de la Respuesta al Item. Madrid, España: Edición interna de la Universidad Autónoma de Madrid.