

## EL FORMATO DEL ÍTEM Y LA NATURALEZA DEL CONSTRUCTO MEDIDO EN LA SOLUCIÓN DE PROBLEMAS

Elsy Urdaneta\*

### RESUMEN

*En esta investigación se formuló un modelo teórico tentativo que propone variables procedentes del ámbito cognitivo, del metacognitivo y relativas a la tarea como factores que determinan la ejecución de un ítem. El objetivo consiste en determinar cómo influyen estas variables en la respuesta al ítem de un test y si esta influencia es diferente según el formato utilizado al plantear el ítem. Se trabajó con el modelo lineal generalizado con función de enlace logit, utilizando un modelo de regresión logística binaria para los tres tipos de formato de ítem, con información proveniente de la aplicación de dos cuestionarios a una muestra de 476 estudiantes de 4º año de bachillerato de centros educativos del Estado Trujillo y de la información aportada por los profesores acerca del nivel de conocimiento declarativo y procedimental de cada estudiante. Se concluyó que las variables incluidas en el modelo constituyen factores que guardan una relación bastante estrecha con la forma en que un sujeto realiza la tarea planteada. Cada formato demanda del sujeto examinado diferentes conductas o diferentes niveles en el empleo de su cognición o metacognición. Del mismo modo, la*

---

\* Doctora en Metodología de las Ciencias del Comportamiento (Universidad Autónoma de Madrid), Diplomado de Estudios Avanzados en Metodología de las Ciencias del Comportamiento (UAM). Profesora Asociada en el Área de Estadística en el Núcleo "Rafael Rangel" de la Universidad de Los Andes Trujillo-Venezuela. Coordinadora de Investigación y Postgrado en el NURR-ULA. E-mail: elsyurdaneta@ula.ve

Recibido: 05/04/2016

Aprobado: 01/07/2016

*percepción sobre la tarea demandada es diferente según el formato en que se presenta la misma.*

**Palabras clave:** formato del ítem, cognitivo, metacognitivo, motivación, volición.

## THE FORMAT OF ITEM AND THE NATURE OF CONSTRUCTION IN SOLVING PROBLEMS

### ABSTRACT

*In this research, a tentative theoretical model was formulated that proposes variables from the cognitive, metacognitive and task-related as factors that determine the execution of an item. The objective is to determine how these variables influence the response to the item of a test and if this influence is different according to the format used when raising the item. The generalized linear model with logit link function was used, using a binary logistic regression model for the three types of item format, with information from the application of two questionnaires to a sample of 476 students of 4th year of high school in Trujillo state educational institutions and the information provided by teachers about the level of declarative and procedural knowledge of each student. It was concluded that the variables included in the model are factors that have a fairly close relationship with the way a subject performs the task. Each format demands from the examined subjects different behaviors or different levels in the use of their cognition or metacognition. In the same way, the perception about the task demanded is different according to the format in which it is presented.*

**Key words:** *itemformat, cognitive, metacognitive, motivation, volition.*

### Introducción

La psicología cognitiva plantea que el saber es más que la mera acumulación de conocimiento factual y la utilización de procedimientos rutinarios: significa ser capaz de integrar conocimientos, destrezas

y procedimientos de forma que sean útiles para interpretar situaciones y resolver problemas (Pellegrino, Chudowsky y Glaser, 2001). Por ello, se considera la ejecución de tareas mentales a dos niveles: cómo se procesa la información y cómo se organiza el conocimiento específico asociado a dominios particulares del conocimiento.

Los constructos asociados a la solución de problemas (aptitud, inteligencia, rendimiento) no son unidimensionales, sino que reflejan una compleja combinación de procesos, estrategias, estructuras del conocimiento y otras variables no cognitivas. Son muchas las fuentes que determinan la dificultad de una tarea y deben ser de tal manera que puedan ser sistemáticamente manipuladas para investigar los distintos procesos cognitivos implicados. Como se ha puesto claramente de manifiesto en investigaciones sobre el asunto (Martínez, 1999; Snow, 1992; entre otras) las diferencias individuales no están únicamente en las aptitudes, en cada componente del procesamiento de información o en los conocimientos previos, sino también en el uso de estrategias al resolver la tarea en cuestión, en los recursos psicológicos básicos de que disponga cada sujeto, en sus motivaciones, en la manera de percibirse a sí mismos y en la forma como pueden ser influidos por variables contextuales relacionadas con la tarea.

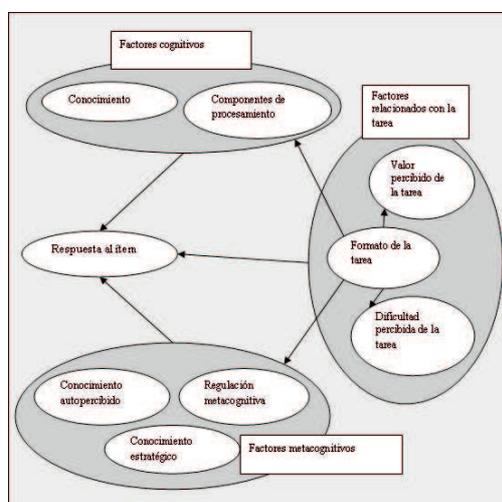
Partiendo de un modelo teórico, cuyo planteamiento es claramente deudor de las propuestas de Snow, Corno y Jackson (1996) y que además incorpora aportes de otros autores que van en la misma línea (e.g. Masters y Mislevy, 1993; Nickerson, 1994; O'Neil y Schacter, 1997; Schoenfeld, 1987; Sternberg, 1985, 1998) al considerar la importancia de variables de diversa índole en la resolución efectiva de problemas, esta investigación procura explorar cómo resuelve el sujeto tareas planteadas en distintos formatos, en función de la gestión que hace de sus habilidades cognitivas y no cognitivas.

En el modelo a someter a prueba se incorpora un conjunto importante de las variables que distintas investigaciones precedentes señalan como posibles factores influyentes en la ejecución de tareas, con el fin de buscar un modelo parsimonioso, que permita manejar adecuadamente la información disponible, pero que resulte lo más completo e informativo posible.

En la Figura 1 se presenta un esquema del modelo propuesto. Como se puede observar, las variables provenientes de la dimensión cognitiva que se incluyen son el conocimiento (declarativo y de procedimientos) y los componentes de procesamiento de la ejecución cognitiva. Las variables incluidas que pertenecen al ámbito no cognitivo son (1) las referentes a los aspectos metacognitivos (en particular, el conocimiento estratégico, la regulación metacognitiva y la percepción del sujeto sobre sí mismo) y (2) las relativas a la tarea, tanto desde la perspectiva del sujeto (en cuanto al valor o dificultad percibida sobre la misma), como desde lo relacionado directamente con el formato en que se presente la tarea.

Lo que supone este planteamiento que se somete a prueba es que los procesos mentales –cognitivos y no cognitivos- que se activan en el sujeto cuando se enfrenta a una tarea se pueden manifestar diferencialmente en función del formato del ítem.

**Figura 1. Modelo propuesto**



Cuando el sujeto responde al ítem, utiliza sus conocimientos acerca del dominio específico al que se enfrenta y los procedimientos que ayuden a la solución del problema, recurriendo a las estrategias

o heurísticos que considere adecuadas para cada situación, así como poniendo en juego su habilidad para controlar todas las actividades metacognitivas implicadas en la solución del problema en cuestión, planificando su forma de trabajo, controlando la efectividad de su plan y verificando la plausibilidad de su respuesta. Estas acciones se ven influidas por sus creencias acerca de su competencia para resolver la tarea, por la importancia o valor atribuido a ésta y por su percepción acerca de la dificultad de la misma, todos ellos factores decisivos en la motivación del sujeto y, en consecuencia, en su éxito para resolverla; con la posibilidad de que se revelen diferencias individuales dependiendo del formato de la tarea.

### **Factores cognitivos**

Dentro del ámbito cognitivo la selección de las variables a incluir en el modelo se realizó en consonancia con el planteamiento fundamental de la psicología cognitiva, que propone como principales elementos en la actuación la organización del conocimiento y el procesamiento de la información.

El estudio de las estructuras del conocimiento tiene una relevancia fundamental y representa el más importante aporte de las teorías cognitivas a la medición en el ámbito educativo, al contribuir a la descripción de la adquisición, organización y uso del conocimiento en dominios específicos (Pellegrino, Chudowsky y Glaser, 2001). Particularmente el conocimiento declarativo y el conocimiento procedimental son considerados esenciales para una ejecución exitosa, constituyéndose en categorías de la arquitectura del conocimiento ampliamente aceptadas, las cuales siempre son incluidas como las formas fundamentales de organización y almacenamiento del conocimiento (Yilmaz y Yalcin, 2012).

Por su parte, el estudio de los componentes del procesamiento cognitivo es absolutamente clave en las investigaciones cognitivas, abordándose desde diversos enfoques. En particular, parece especialmente apropiado el enfoque propuesto por Sternberg (1985, 1998) en su teoría triárquica de la inteligencia, que considera la existencia de un numeroso conjunto de procesos que se activarán en función de la

tarea que deba realizarse y que se agrupan, de modo resumido, en tres fases básicas: la codificación (el sujeto representa mentalmente la tarea identificando los elementos que permitirán la solución del problema), la combinación (el sujeto realizará las operaciones necesarias para alcanzar y valorar una solución) y respuesta (se expresa el resultado de las fases anteriores).

### **Factores metacognitivos**

La relevancia de la metacognición como un determinante en la ejecución de tareas la posiciona como un factor que debe ser incluido dentro del modelo que se propone, pues estos procesos actúan como guía en el proceso y mejoran la eficiencia de las conductas orientadas hacia la solución de problemas (Davidson, Deuser y Sternberg, 1994).

La planificación y el control ejercido durante la ejecución de la tarea son importantes elementos de regulación cognitiva que contribuyen a un mejor desempeño y que son ampliamente reconocidos en numerosos estudios (e. g. Davidson, Deuser y Sternberg, 1994; Händel, Artelt, y Wienert, 2013; Nickerson, 1994; Scherer y Tiemann, 2012). En su teoría triárquica, Sternberg (1985, 1998) señala que los componentes de ejecución sólo se desencadenan si son mediados por los procesos de control ejecutivo.

De igual manera, el conocimiento estratégico, representado por los heurísticos utilizados para la solución de problemas, resulta clave para obtener la solución de modo más efectivo (Krathwohl, 2002; Nickerson, 1994; Schoenfeld, 1983).

El conocimiento autopercibido o creencias acerca de las propias competencias es una variable que también se considera un determinante de gran importancia en la solución de problemas, tanto desde la perspectiva de la metacognición, como desde la perspectiva de la motivación y la volición, generando acciones cognitivas decisivas en la solución de problemas (Kupermintz, 2002; Lau y Roeser, 2002; Schoenfeld, 1983, 1985; Sugrue, 1995).

## Factores relacionados con la tarea

El modo en que las diferencias conativas entre los estudiantes influyen la actividad cognitiva, tanto en la escuela como fuera de ella, hace de estas variables un importante elemento a explorar para poder interpretar con mayor exactitud el significado de las puntuaciones.

El repertorio de variables de este tipo que pueden influir en la conducta del sujeto ante una tarea es realmente enorme. Sin embargo, en el modelo propuesto se optó por el planteamiento de Eccles y Wigfield (1995, 2002) al incluir el valor percibido de la tarea y la dificultad percibida de la tarea, pues según exponen estos autores, éstas son las variables fundamentales que promueven los procesos motivacionales y volitivos, suscitando conductas de logro o de disuasión según se considere la tarea en relación a estas variables.

El hecho de que una tarea resulte entretenida, atractiva, permita demostrar las capacidades y sea considerada de utilidad, es decir, el valor que se le dé a la tarea, contribuirá a motivar al sujeto a realizarla. De igual modo, si el sujeto percibe que el esfuerzo y la dificultad de la tarea puedan sobrepasar sus capacidades, la voluntad de enfrentar la tarea resulta mermada. En este sentido, se pueden considerar estas variables como atributos motivacionales y volitivos con influencia determinante en la ejecución del ítem.

Asimismo, el formato del ítem puede representar un factor decisivo de la actuación del sujeto frente a una tarea particular, al constituirse en elemento que puede estar asociado a diferencias individuales en las variables tanto cognitivas como no cognitivas que se activan en el momento de la ejecución (Martínez, 1999; Snow, 1989).

En tal sentido, el estudio del formato refiere al asunto de la validez por cuanto existe la preocupación de examinar si distintos formatos son capaces de medir un constructo al mismo nivel de amplitud y profundidad. Tal como afirma Messick (1993), la equivalencia del rasgo medido con distintos formatos es un tema relacionado con la validez de constructo, de tal modo que las investigaciones sobre el formato del

ítem y su influencia en la ejecución de la tarea se perfilan como un importante aporte en el estudio del significado de las puntuaciones.

Así pues, se plantea como objetivo de investigación identificar, a través de la comprobación empírica del modelo teórico planteado, qué factores influyen en la respuesta al ítem de un test y determinar si esta influencia es diferente según el formato utilizado al plantear el ítem, lo cual permite aproximar una explicación acerca del significado psicológico de las puntuaciones obtenidas cuando el constructo es medido con ítems de distinto formato

### **Método**

Para lograr el objetivo propuesto se realizó una investigación de campo con un diseño de medidas repetidas a fin de reducir la variabilidad atribuible a las diferencias individuales entre sujetos, garantizando de esta manera la precisión y potencia de la prueba.

### **Sujetos**

Se aplicaron las pruebas a una muestra no probabilística formada por 476 participantes, estudiantes de 4° año de bachillerato de siete centros educativos ubicados en el Estado Trujillo en Venezuela. La selección de los estudiantes de 4° año obedeció al hecho de que la mayoría de los estudiantes de este curso tienen 15 años, que es la edad para la cual están diseñadas las preguntas utilizadas en la prueba de competencias del estudio PISA, programa desarrollado por la Organización para la Cooperación y el Desarrollo Económico (OCDE), cuyo objetivo es evaluar la formación de los alumnos cuando llegan al final de la etapa de enseñanza obligatoria.

La muestra tiene una distribución equitativa en relación al sexo de los examinados (50.2 % masculino y 49.8% femenino). La mayoría de los examinados tiene una edad de 15 años (73.7 %) y el resto con edades muy próximas. En cuanto a la titularidad de los centros educativos se tienen porcentajes relativamente similares: centros privados (42.6 %) y centros públicos (57.4 %).

## **Instrumentos**

Para hacer posible la recolección de la información sobre las variables a estudiar se utilizaron tres instrumentos: una prueba de competencias para observar la ejecución de los estudiantes en cada tipo de formato de ítem y obtener la puntuación en la variable solución de problemas, un cuestionario general *ad hoc* para obtener información sobre un número importante de las variables relacionadas al modelo y una hoja de registro de los profesores, para estimar el nivel de conocimiento declarativo y procedimental de los estudiantes.

## **Prueba de competencias**

La prueba de competencias permitió obtener la respuesta del sujeto a cada uno de los ítems formulados en tres distintos tipos de formato. Se optó por incluir un único ítem por cada tipo de formato considerado en el estudio con el objeto de optimizar el tiempo de aplicación. Los formatos utilizados en esta prueba son: elección múltiple (EM), respuesta corta (RC) y respuesta abierta (RA).

Las preguntas seleccionadas fueron extraídas del grupo utilizado en la edición 2003 del estudio PISA (INECSE, 2005) del dominio solución de problemas, al no estar asociadas directamente a ningún área curricular concreta y no demandar recursos especiales que pudieran encarecer los costos de aplicación, ni tampoco requerir tiempos excepcionales para realizar la tarea. Además, todas estas preguntas ya habían sido probadas en relación a su validez y confiabilidad para medir los constructos deseados en los análisis desarrollados dentro del estudio PISA, mostrando una altísima eficiencia en los usos proyectados.

En particular, se seleccionaron la pregunta 8 del problema denominado *Campamento* y la pregunta 15 del problema denominado *Vacaciones*. El problema *Vacaciones* suministra un mapa de carreteras y una tabla de distancias y se pregunta sobre la ruta más corta para viajar de un destino a otro. En el ítem de EM se les da la opción correcta y tres distractores y en el ítem RC se les pide que digan el menor kilometraje a recorrer. En ambos problemas deben calcular los recorridos por las distintas rutas que aparecen en el mapa y seleccionar la menor.

Para representar el ítem de respuesta abierta se utilizó la pregunta 8 del problema denominado *Campamento*, el cual tenía varias respuestas correctas. El problema consistía en distribuir, respetando tres condiciones, a un grupo de 46 jóvenes de ambos sexos que deberían estar acompañados al menos por un adulto de los ocho disponibles en siete habitaciones, por lo cual existe más de una respuesta correcta.

Para asignar puntuaciones a cada ítem se procedió de modo similar a como se hace en el estudio PISA. Para el ítem de elección múltiple se considera la opción *b* como la respuesta correcta y se le adjudicaba un punto, el resto de opciones se consideran respuestas incorrectas y se les adjudicaba cero puntos. Para el ítem de respuesta corta se considera *900 Km.* como la respuesta correcta y se le adjudicaba un punto, otras respuestas se consideraban incorrectas y se les adjudicaba cero puntos. Para el ítem de respuesta abierta se adjudicaba un punto para la respuesta correcta, que ha de satisfacer todas las condiciones planteadas en el enunciado. Se adjudicaba cero puntos para otras respuestas.

La confiabilidad de las puntuaciones obtenidas con estos ítems se determinó tanto desde la perspectiva de la Teoría Clásica de los Tests mediante el coeficiente alfa y desde la perspectiva de la Teoría de la Respuesta al Ítem utilizando un modelo de dos parámetros que nos revela los valores de discriminación y dificultad en las curvas características y de información correspondientes. Como argumento de validez, además de lo relativo al contenido de la prueba, se realizó un análisis factorial exploratorio para demostrar la unidimensionalidad del constructo.

### **Cuestionario general**

El cuestionario general se elaboró con el objeto de reunir la mayor cantidad de información sobre variables relacionadas con: (1) la percepción del sujeto sobre la tarea a realizar, (2) aspectos metacognitivos y (3) los componentes de ejecución o procesamiento cognitivo que se activan para responder a las tres preguntas de la prueba de competencias. Para la elaboración de este cuestionario se partió del estudio de las teorías y modelos que sobre estas variables se han desarrollado.

Se especificaron seis aspectos a medir con este cuestionario general: los relacionados con la percepción del sujeto sobre la tarea (*valor percibido de la tarea y dificultad percibida de la tarea*), los concernientes al ámbito metacognitivo (*conocimiento autopercebido, regulación metacognitiva y conocimiento estratégico*) y dentro del ámbito cognitivo los *componentes de procesamiento cognitivo*.

El cuestionario general queda conformado por 44 preguntas que debían ser respondidas por el estudiante en relación a cada una de las tareas que había realizado en la prueba de competencias, a fin de tener información sobre todas las variables para cada uno de los formatos estudiados.

Para estimar la confiabilidad de las puntuaciones y argumentar la validez de las inferencias se calculó el coeficiente alfa y se realizaron análisis factoriales confirmatorios.

### **Hoja de registro de los profesores**

Para obtener información acerca del nivel de conocimiento declarativo y procedimental de los estudiantes se solicitó una valoración cuantitativa a sus profesores de Matemáticas, ya que la información procedente de la prueba de competencias, formada solo por tres ítems, no proporcionaba una estimación adecuada de los mismos.

A cada profesor se le suministró una hoja de registro constituida por un listado con los nombres de los estudiantes examinados correspondientes a su grupo y dos columnas, una para registrar el nivel de conocimiento declarativo y otra para el conocimiento procedimental.

### **Plan de análisis**

En primer lugar se procedió a la definición y construcción de las variables hipotetizadas como predictoras de la respuesta al ítem. La variable respuesta está representada por las respuestas de los sujetos a los ítems de la prueba de competencia (tres preguntas cada una con un formato diferente) y las variables independientes serían cada una de

las correspondientes a la medida de los factores que se plantean en el modelo como posibles determinantes de la respuesta al ítem, a saber:

- Dentro del ámbito cognitivo, el conocimiento declarativo, el conocimiento procedimental y los componentes de procesamiento cognitivo (codificación: COD, combinación: COM y respuesta: RES).
- Dentro del ámbito metacognitivo, el conocimiento autopercebido (CAP), el conocimiento estratégico (CES) y la regulación metacognitiva (planificación: PLANy control: COM).
- Dentro del ámbito relativo a la tarea, el valor y la dificultad percibidas en la misma.

Las medidas de conocimiento declarativo (CD) y procedimental (CP) que se utilizaron son las valoraciones que hicieron de estas variables los profesores de matemáticas de los estudiantes examinados.

El resto de las variables del modelo hipotetizándose construyeron sumando las puntuaciones en los ítems indicadores de cada componente o factor.

Una vez realizadas todas las tareas previas de revisión y preparación de los datos para la prueba empírica del modelo se ejecutó el modelado de las variables a objeto de poder dar respuesta a las preguntas que se plantea esta investigación.

Para someter a comprobación empírica el modelo teórico formulado, se utilizó un modelo lineal generalizado con el logit como función de enlace. Dado que para los ítems de competencias solo existen dos categorías de respuesta (correcta o incorrecta), se trabajó con la regresión logística binaria. Para estimar el modelo de regresión logística para cada uno de los formatos siguió el procedimiento descrito por Ato y cols. (2005, p. 57).

En la fase de especificación se formuló el componente sistemático del modelo, definido por el conjunto de variables predictoras del modelo. En la fase de selección se procedió a la estimación del modelo de regresión logística trabajando paso a paso, utilizando siempre y de

forma sistemática dos métodos distintos de estimación, a fin de verificar que se obtenían los mismos resultados: el método de Wald hacia adelante y el método de razón de verosimilitudes hacia atrás. Para la fase de evaluación se valoraron los distintos modelos obtenidos examinando la contribución de cada variable al modelo mediante el estadístico de Wald, el ajuste global del modelo mediante la razón de verosimilitud G y el coeficiente pseudo  $R^2$  de Nagelkerke y la eficacia predictiva del modelo mediante el porcentaje de casos correctamente pronosticados (% CP).

## **Resultados**

El volumen de información que suministraron los datos recolectados es muy rico, pero por cuestiones relativas a espacio editorial solo se reportan en este artículo los hallazgos más relevantes de cara al objetivo planteado en la investigación.

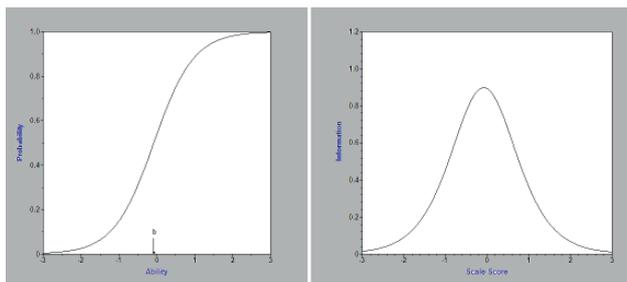
A continuación se presentan los resultados obtenidos al examinar las propiedades psicométricas de las puntuaciones en las dos pruebas utilizadas en el estudio. En el caso de la prueba de competencias, se ha examinado también la calidad métrica de cada una de las 3 preguntas de la prueba, dado el papel clave que desempeñan en la investigación.

### **Prueba de competencias**

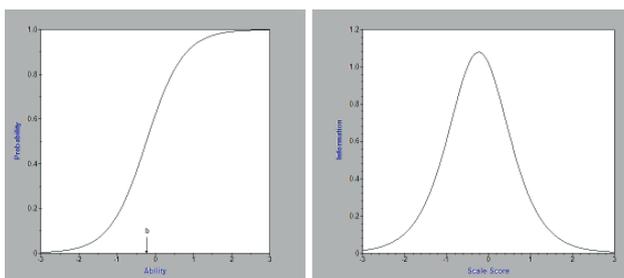
En la Figura 2 se presentan la curva característica y la curva de información de los 3 ítems de la prueba de competencias. Los tres ítems presentan una dificultad semejante (0.530 para EM, 0.578 para RC y 0.505 para RA) y una buena capacidad discriminativa (0.380, 0.393 y 0.305 respectivamente) ligeramente inferior en el ítem de respuesta abierta, con una precisión máxima en la zona central de la escala.

**Figura 2. Curva característica y función de información de cada ítem**

**Curva característica y función de información del ítem EM**



**Curva característica y función de información del ítem RC**



El valor obtenido para el coeficiente alfa en las puntuaciones de la prueba cognitiva es de 0.519. El análisis factorial exploratorio revela la presencia de un único factor, en el que cargan los tres ítems con pesos factoriales altos y bastante similares: 0.751 para EM, 0.792 para RC y 0.691 para RA.

En resumen, las evidencias recogidas tanto acerca de la precisión como de la validez de las puntuaciones obtenidas con los ítems de esta prueba apoyan su utilización para el logro del objetivo de investigación planteada en este estudio. Por otra parte, es necesario señalar que estos ítems han sido ampliamente probados ya que provienen de una prueba internacional de altísima calidad, como lo son las pruebas aplicadas en el programa PISA de la OCDE.

## **Cuestionario general**

Los valores del coeficiente alfa para las puntuaciones en cada grupo de ítems según variable a medir oscilan entre 0.603 y 0.929 y permiten afirmar que las puntuaciones obtenidas con los ítems de la prueba general son confiables y que están en perfecta consonancia con los reportados para las escalas revisadas que usualmente se utilizan para la medición de estas variables.

Los resultados de los análisis factoriales confirmatorios realizados permiten corroborar la estructura interna hipotetizada para cada variable medida. Los índices RMSEA, que indican error, tienen valores que en ningún caso superan a 0.142, los índices GFI y AGFI, que indican ajuste del modelo, en ningún caso tienen valores inferiores a 0.98, con valores de chi cuadrado bajos y con  $p < 0.05$  para la mayoría de los casos.

En resumen, los resultados obtenidos para el coeficiente alfa y para los análisis confirmatorios representan evidencias favorables acerca de la confiabilidad de las puntuaciones obtenidas con la prueba general y de la validez de las inferencias a que pueda dar lugar.

## **Comparación de los modelos obtenidos en los distintos formatos**

En la Tabla 1 se presenta un resumen de los resultados de los modelos de regresión que mejor ajustaron para cada formato estudiado. Como se advierte del examen de esta tabla, la variable combinación se muestra como un predictor común para cualquiera de los tres formatos, informando de la importancia que tienen para el sujeto la facilidad a la hora de organizar la información, de realizar los cálculos pertinentes y de verificar la adecuación de la solución.

**Tabla 1. Resumen de los resultados para EM, RC y RA**

	FORMATOS DE ÍTEM		
	EM	RC	RA
Variables en el modelo	CD, COM y CON	COD y COM	CP, COM, PLAN, CON y VTF
H-L (p)	0.615 (0.961)	0.058 (1.000)	3.408 (0.906)
G (p)	421.262 (0.000)	560.948 (0.000)	298.998 (0.000)
R <sup>2</sup> de Nalgelkerke	0.810	0.989	0.888
% CP	90.7	99.5	94.7

En el estudio de las variables pertenecientes al bloque relativo al conocimiento se observa que la respuesta al ítem de EM es influenciada por el conocimiento declarativo, mientras que la respuesta al ítem de RA lo es por el conocimiento procedimental. En el modelo para el formato de RC ninguna de las variables relativas al conocimiento muestra un aporte que signifique una mejoría importante del modelo.

En cuanto a las variables del bloque metacognitivo, la actividad de regulación cognitiva control se revela como importante para dar respuesta al ítem cuando es formulado en los formatos de EM y RA. No ocurre lo mismo para el formato de RC, pues aun cuando la variable control había sido seleccionada en el modelado parcial, debido a la fuerza predictiva de las variables codificación y combinación, el efecto de cualquiera de las otras variables queda diluido u oculto.

En relación a las variables del bloque relacionado con la percepción sobre la tarea, solamente el modelo para el formato de RA tiene una de ellas, el valor percibido de la tarea considerado en relación al formato en que es presentada. En el modelado parcial, cuando se trataron independientemente del resto de variables, todas las variables de este bloque resultaron seleccionadas como posibles predictores de la respuesta al ítem, cualesquiera fuera el formato considerado.

Otro detalle interesante es el hecho de que el modelo que más predictores retiene es el correspondiente al formato RA, que corresponde a una pregunta de respuesta abierta, mientras que el que menos predictores retiene es RC, que sólo admite una respuesta corta. Comparando los tres modelos también se observa que el modelo con índices de ajuste y de predicción más altos es el hallado para RC, con un porcentaje de casos correctamente pronosticados muy elevado.

### **Discusión y conclusiones**

El estudio del formato de los ítems ha sido un tema de creciente interés a lo largo del tiempo, tal y como lo confirman un número importante de publicaciones (véase Rodríguez, 2002, 2003 o Bennett y Ward, 1993). Como ya se ha mencionado, algunos de estos estudios muestran entre sus hallazgos la existencia de diferencias en las propiedades métricas de las puntuaciones obtenidas con ítems de distinto formato, sin llegar a explicar donde residen éstas. En esta investigación se intentó buscar la naturaleza de esas diferencias, para lo cual, partiendo de una exhaustiva revisión bibliográfica, se planteó un modelo que incluye, con una cobertura suficientemente amplia, aquellos constructos que las distintas teorías señalan como responsables de las conductas de respuesta al ítem. El modelo teórico planteado considera un conjunto de variables de distinta naturaleza (cognitiva, metacognitiva y relacionadas con la percepción del examinado sobre la tarea), que podrían ser activadas diferencialmente según el formato en que se plantea la tarea.

El señalamiento de autores como Wilkinson (1999) o Cudeck y Henley (1991), en relación a que un modelo no es capaz de explicar a cabalidad la complejidad de la conducta humana, resulta perfectamente plausible. No obstante, un modelo puede considerarse una aproximación útil para la explicación y síntesis, la cual contribuye a la formalización de esos múltiples procesos y en consecuencia permite un acercamiento al significado de la naturaleza del constructo cuando es medido con diferentes formatos de ítem. En este sentido, el análisis de los resultados obtenidos al someter a contrastación empírica el modelo siempre se ha realizado atendiendo fundamentalmente al significado del modelo, sin desmedro de la atención necesaria a los índices de ajuste y diagnóstico.

De la comprobación empírica del modelo y con los resultados obtenidos es posible hacer el análisis a dos niveles: uno general, fundado en los resultados del análisis de regresión logística, y uno más particular, basado en los hallazgos parciales encontrados en el camino de la construcción del modelo que dan lugar a otras observaciones que podrían resultar de interés.

En una primera mirada a los modelos de regresión logística obtenidos, se observa que existen coincidencias fundamentales en el comportamiento del sujeto al responder el ítem, independientemente de su formato. En principio, la fase de combinación de los componentes de procesamiento cognitivo implicados en la tarea es una variable que está presente en los modelos obtenidos con los tres formatos, lo que permite sea considerada como el predictor más importante. Este hallazgo es perfectamente coherente con lo esperado. La fase de combinación constituye el conjunto de procesos cognitivos orientados a reunir y organizar la información, hacer los cálculos necesarios y comparar los resultados obtenidos. En la medida en que este conjunto de procesos resulta más fácil para el examinado es más probable que su respuesta sea la correcta. En este sentido, los resultados apuntan a la necesidad de considerar modelos cognitivos para analizar la ejecución de tareas, identificando los componentes del procesamiento cognitivo involucrados en la resolución de la tarea y estableciendo la forma de darle puntuación a cada componente del modelo.

Continuando el examen de los modelos, los resultados obtenidos también permiten afirmar que las variables de naturaleza metacognitiva constituyen constructos importantes que pueden ayudar a explicar la ejecución, al quedar variables de este bloque en el modelo tanto para el formato de EM como para el de RA. El hecho de que no sean retenidas en el modelo resultante para el formato RC se debe a la fuerza que tienen los componentes de codificación y combinación en ese formato. Estos resultados son también coherentes con los obtenidos en otras investigaciones (Dowson y McInerney, 2004; Sternberg, 1998; Pintrich, 2002), que señalan la influencia de los factores metacognitivos y que el formato de respuesta abierta induce a un mayor uso de estrategias de control cognitivo, tal como se ha hallado en los datos de este estudio.

En cuanto a las variables relacionadas con la percepción sobre la tarea, su influencia se oscurece completamente debido al marcado efecto de las variables relacionadas a los aspectos metacognitivos y cognitivos. Solamente el valor percibido de la tarea en relación al formato queda en el modelo para el ítem de evaluación de la actuación.

Ahora bien, no se debe perder de vista que la formulación de este modelo no fue hecha con fines predictivos, sino que su propósito fundamental es identificar las variables que pueden explicar el significado de las puntuaciones obtenidas cuando el ítem es medido con distintos formatos, razón por la cual es importante mirar con suficiente detalle los resultados parciales que se desprenden del modelado, así como también todo lo revelado en los análisis preliminares.

Tal como se explicó, el modelado se hizo formato a formato (EM, RC y RA) y por grupos de variables según la dimensión o el ámbito al que pertenecen (cognitivas, metacognitivas y en relación a la tarea). Se comprueba de esta forma que la mayoría de las variables de los distintos bloques considerados en el modelo ejercen una influencia sobre la respuesta al ítem cualquiera sea el formato, pero al ser consideradas de forma conjunta, el efecto de algunas se muestra más importante, diluyendo en algunos casos el efecto del resto de los predictores. También se verifica que el efecto de cualquiera de los factores componentes del modelo es más marcado para los ítems de construcción, en algunos casos para el formato de RC y en otros para el formato de RA, y menos relevante en el ítem con formato de EM.

Se puede observar que el tipo de conocimiento -declarativo o procedimental- que emplean es diferente según el formato afrontado: declarativo para el formato de elección y procedimental para resolver los ítems planteados en formato de respuesta construida. En cuanto a los procesos de ejecución cognitiva, como ya se señalara, el componente que se revela como más importante de todos en cualquiera de los casos es el correspondiente a la fase de combinación, cuya dificultad se valora de modo diferencial en los tres formatos, siendo más difícil para el ítem de respuesta corta y más fácil para el ítem en formato EM. Esto sucede en los 3 componentes de procesamiento: la dificultad percibida varía con el formato del ítem.

Los modelos por bloques revelan que para los formatos de construcción todas las variables del ámbito metacognitivo son predictores importantes, mientras que para el formato de elección sólo el conocimiento autopercebido y el control lo son. En cuanto al uso de estrategias metacognitivas para la solución de problemas, es importante en este punto observar con detalle lo hallado en los análisis de los datos. En primer lugar, se observa que el número de estrategias utilizadas tiene un valor medio superior para la tarea en formato de elección múltiple, pero no existe una relación entre número de estrategias utilizadas y la respuesta al ítem de este formato, relación que sí se observa en los dos formatos de construcción, aun cuando la variable no represente un predictor significativo en el modelo conjunto.

En lo concerniente a las variables relacionadas con la percepción que tiene el sujeto acerca de la tarea se observa que, a pesar de que su efecto desaparece o queda ensombrecido al incorporar variables cognitivas y metacognitivas, resultan predictores significativos tanto en los modelos univariados como en los modelos por bloques. Por otra parte, sin olvidar que el tamaño del efecto en la estimación de las diferencias entre formato para estas variables es muy pequeño, vale destacar que el formato mejor valorado es el de elección múltiple y que los ítems se perciben con mayor dificultad cuando están planteados en un formato de respuesta construida.

Se advierte de lo hallado que las variables estudiadas tienen un comportamiento más similar en los formatos de construcción. Asimismo, se evidencia que la respuesta al ítem, cualquiera sea su formato, está principalmente determinada por las variables de tipo cognitivo, básicamente por los componentes de ejecución en concordancia con lo planteado por Sternberg (1985, 1998). Aun así, se puede afirmar que todos los constructos estudiados constituyen predictores importantes en la ejecución del ítem y que la intensidad del efecto de los distintos factores varía de un formato a otro.

A la vista de los resultados obtenidos se puede concluir que cada formato parece demandar del sujeto examinado diferentes conductas, o diferentes niveles en el empleo de su cognición o metacognición. Del mismo modo, la percepción sobre la tarea demandada es diferente

según el formato en que se presenta la misma, lo cual, como señalan Pellegrino, Chudowsky y Glaser (2001), puede influir en el significado e interpretación de las puntuaciones. Tal como afirma Snow (1993), existe un tejido de elementos que producen diferencias individuales en la respuesta al ítem que tiene relación con el formato.

En consecuencia, para un proceso de medición más integral es recomendable usar pruebas de distintos formatos, pues cada uno de ellos es capaz de revelar diferencias individuales, no solamente en cuanto a conocimientos y procesos cognitivos, sino también en motivación, esfuerzo y mecanismos de control metacognitivo. Suprimir los ítems de construcción en nombre de la economía y de la precisión puede significar un sacrificio en la capacidad de medición del test y podría dar lugar a inferencias con una validez cuestionable.

Dado que cada tipo de pregunta es capaz de producir distintas conductas, el formato de respuesta al ítem debería corresponder con la conducta diana en el objetivo de la medición y garantizar que el significado de las puntuaciones obtenidas se corresponda con el constructo que deseamos medir.

## REFERENCIAS BIBLIOGRÁFICAS

Ato, M., Losilla, J. M., Navarro, J. B., Palmer, A. y Rodrigo, M. F. (2005). *Análisis de datos. Modelo lineal generalizado*. Girona: Edicions a Petició, SL.

Bennett, R. E. y Ward, W. (Eds.) (1993). *Construction versus Choice in Cognitive Measurement: issues in constructed response, performance testing and portfolio assessment*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Benton, S. L. y Kiewra, K. A. (1987) The assessment of cognitive factors in academia abilities. En R. R. Ronning, J. A. Glover, J. C. Conoley y J. C. Witt (Eds.), *The influence of cognitive psychology on testing*. (pp. 145-189). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Bruning, R. H., Schraw, G. J., Norby, M. M. y Ronning, R. R. (2005). *Psicología cognitiva y de la instrucción* (4ª ed.). Madrid: Pearson, Prentice Hall.

Cudeck, R. y Henley, S. J. (1991). Model selection in covariance structures and the “problem” of simple size: A clarification. *Psychological Bulletin*, 109, 512-519.

Davidson, J. E., Deuser, R. y Sternberg, R. J. (1994). The role of metacognition in problem solving. En J. Metcalfe y A. P. Shimamura (Eds.), *Metacognition: knowing about knowing* (pp. 207- 226). Massachusetts: Massachusetts Institute of Technology Press.

Dowson, M. y McNerney, D. (2004). The development and validation of the goal orientation and learning strategies survey (GOALS-S). *Educational and Psychological Measurement*, 64, 290 -310.

Eccles, J. y Wigfield, A. (1995). In the mind of the actor: the structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology*, 21, 215-225.

Eccles, J. y Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual review of psychology*, 53, 109-132.

Händel, M., Artelt, C. y Wienert, S. (2013). Assessing metacognitive knowledge: Development and evaluation of a test instrument. *Journal for Educational Research Online*, 5(2), 162 – 188. Disponible en: <http://www.j-e-r-o.com/index.php/jero/article/viewFile/365/172>

Instituto Nacional de Evaluación y Calidad del Sistema Educativo (INECSE). (2005). *Prueba de matemáticas y de solución de problemas*. Madrid: Autor. Disponible en: <http://www.mecd.gob.es/dctm/evaluacion/internacional/pisa2003liberados.pdf?documentId=0901e72b801106c6>

Krathwohl, D.R. (2002). A revision of Bloom's taxonomy: an overview. *Theory into Practice*, 41(4), 212-218.

Kupermintz, H. (2002). Affective and conative factors as aptitude resources in high school science achievement. *Educational Assessment*, 8, 123-137.

Lau, S. y Roeser, R. (2002). Cognitive abilities and motivational processes in high school students' situational engagement and achievement science. *Educational Assessment*, 8, 139-162.

Masters, G. N. y Mislevy, R.J. (1993). New views of student learning implications for educational measurement. En N. Frederiksen, R., Mislevy y I. Bejar (Eds.), *Test theory for a new generation of tests* (pp. 219-242). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Martínez, M. (1999). Cognition and the question of test item format. *Educational Psychologist*, 34, 207 - 218.

Messick, S. (1993). Construct validity and constructed-response tests. En R. E. Bennett y W. Ward. (Eds.), *Construction versus Choice in Cognitive Measurement: issues in constructed response, performance testing and portfolio assessment* (pp. 61-73). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Nickerson, R. S. (1994). The teaching of thinking and problem solving. En R. J. Sternberg (Ed.), *Thinking and problem solving* (pp. 409-449). San Diego, California: Academic Press, Inc.

Pellegrino, J. W, Chudowsky, N. y Glaser, R. (Eds.) (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.

Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory into Practice*, 41, 219-225.

Rodriguez, M. (2002). Choosing an item format. En G. Tindal y T. Haladyna (Eds.), *Large-scale assessment programs for all students: validity, technical adequacy, and implementation*. (pp. 213-231). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Rodriguez, M. (2003). Construct equivalence of multiple-choice and constructed-response items: a random effects synthesis of correlations. *Journal of Educational Measurement*, 40, 163-184.

Scherer, R., y Tiemann, R. (2012). Factors of problem-solving competency in a virtual chemistry environment: The role of metacognitive knowledge about strategies. *Computers & Education*, 59, 1199-1214.

Schoenfeld, A. H. (1983). Beyond the purely cognitive: belief systems, social cognition, and metacognition as driving forces in intellectual performance. *Cognitive Science*, 7, 329 -363.

Schoenfeld, A.H. (1987). *Cognitive science and mathematics education*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Schraw, G. y Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19, 460 - 475.

Snow, R. E. (1989). Toward assessment of cognitive and conative structures in learning. *Educational Researcher*, 18(9), 8-14.

Snow, R. E. (1992). Aptitude theory: Yesterday, today, and tomorrow. *Educational Psychologist*, 27, 5-32.

Snow, R. (1993). Construct validity and constructed-response tests. En R. E. Bennett y W. Ward (Eds.), *Construction versus Choice in Cognitive Measurement: issues in constructed response, performance testing and portfolio assessment* (pp. 45-60). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Snow, R. E., Corno, L. y Jackson, D., III. (1996). Individual differences in affective and conative functions. En D. C. Berliner y R. C. Calfee (Eds.), *Handbook of Educational Psychology* (243-310). New York: Macmillan.

Sternberg, R. J. (1985). *Beyond IQ. A triarchic theory of human intelligence*. Nueva York: Cambridge University Press.

Sternberg, R. J. (1998). Principles of teaching for successful intelligence. *Educational Psychologist*, 32, 65 - 72.

Sugrue, B. (1995). A theory-based framework for assessing domain-specific problem-solving ability. *Educational Measurement: Issues and Practice*, 14(3), 29-36.

Wilkinson, L. (1999). *The grammar of graphics*. Springer: Nueva York.  
Yilmaz, I. y Yalcim, N. (2012). The Relationship of Procedural and Declarative Knowledge of Science Teacher Candidates in Newton's Laws of Motion to Understanding. *American International Journal of Contemporary Research*, 2(3), 50-56. Disponible en: [http://www.aijcrnet.com/journals/Vol\\_2\\_No\\_3\\_March\\_2012/5.pdf](http://www.aijcrnet.com/journals/Vol_2_No_3_March_2012/5.pdf)