Rev. de Psicol. Gral. y Aplic., 1989, 42 (4), 475-482

ANÁLISIS DE LOS TESTS DE INTELIGENCIA UTILIZANDO LA TEORÍA DE LA INFORMACIÓN

C. SANTISTEBAN
Universidad Complutense de Madrid

J. M. SALINAS
Universidad de Granada

Resumen

Mediante la información de Shannon se analizan los elementos que configuran los tests de inteligencia. Desde este punto de vista, preguntas y respuestas son consideradas como mensajes que recibe y emite el sujeto que se evalúa. Tras una breve exposición de la teoría de Shannon, se examina la cantidad de información contenida en las preguntas y en las respuestas, tanto bajo los supuestos de la teoría clásica de los tests, como en las condiciones que proporcionan los modelos del rasgo latente. Los elementos teóricos así obtenidos permiten emplear esta perspectiva en el estudio de los datos que, en una muestra de estudiantes, suministra el test de Raven.

Abstract

The configuration of elements in intelligence tests are analyzed by means of the Shannon information function. From this approach questions and answers are taken as messages received and emitted by the subject under evaluation. The theoretical concepts obtained in this way are used in the interpretation of data from Raven test applied to a sample of students.

Introducción

La Teoría de la Información de Fisher es bien conocida en Estadística, así como otras teorías que incluyen todos los problemas que son susceptibles de utilizar la palabra «información» en su sentido más usual. Dentro de la Teoría de la Información podemos considerar que subsisten tres concepciones que se han dado en llamar: Estadística, Métrica-Estructural y Selectiva; cada una se presenta en un contexto propio, pero no hay problema de rivalidad entre ellas puesto que se han adaptado a diferentes usos y se ha comprobado que en muchos casos coinciden. En Estadística Inferencial se hace uso de las funciones de información, en conexión con el concepto de «verosimilitud»; estas funciones, así como la teoría de la Información Estructural y Métrica, desarrollada por Gabor (1946) y Mc Kay (1950), es posible utilizarlas para aplicaciones en Psicología. Sin embargo, la llamada Información Selectiva, particularmente asociada a los nombres de Wiener y Shannon, es la aproximación que se sigue en este artículo, por considerarla la más inmediatamente reveladora de la «cantidad de información» en los tests.

La Teoría de la Información es una rama de la Teoría de la Probabilidad originada a partir de dos artículos de Claude E. Shannon (1948), en los cuales se propone e investiga un nuevo modelo matemático de comunicación. Una de las más importantes innovaciones de este modelo es el considerar los componentes de un sistema de comunicación como entidades probabilísticas. Shannon propone en su artículo una medida cuantitativa de la cantidad de información, avalada por un experimento probabilístico basado en la clásica entropía de Boltzmann (1896) a partir de la Física Estadística. En esta concepción, la cantidad de información está fuertemente conectada con la cantidad de incertidumbre. Un acto de comunicación proporciona información, en la medida en que reduce una condición de ignorancia o incertidumbre acerca del estado del sistema en consideración.

Supongamos que se ha percibido un suceso. Su ocurrencia puede enseñar algo; como mínimo a partir de ese momento se sabe que dicho suceso se ha producido. La teoría de la información se va a limitar a este mínimo buscando medir la información que aporta el hecho de que se haya verificado. Desde esta perspectiva, cuanto más imprevista sea la realización de un acontecimiento, más informativa dará su constatación. Si se conocía con certeza que este resultado debía producirse, su ocurrencia no revela nada; por el contrario, si no hay seguridad acerca de su aparición, cuando éste se presenta, se desvela una incógnita y se aumenta el grado de conocimiento acerca de ese suceso. Es posible medir lo imprevisto de un suceso a partir de la probabilidad de que ocurra, por ello, se dirá que la información aportada por un suceso es tanto mayor cuanto menor es su probabilidad de ocurrencia. En cierta manera, la cantidad de información es determinada por la cantidad de incertidumbre, o más precisamente, por la cantidad en que la incertidumbre ha sido reducida.

Fundamentos teóricos

En 1948 Claude E. Shannon asegura que se puede hacer una aproximación a la medición de estas nociones tan abstractas como información e incertidumbre; para ello parte de las siguientes consideraciones: sean dos sucesos, x e y, con probabilidades de ocurrencia p(x) y p(y) respectivamente. La información que proporcionará la percepción de estos sucesos $I_x = (f(p(x))$ e $I_y = f(p(y))$ será una función f de sus probabilidades de ocurrencia. Se trata de determinar esa función f de acuerdo con el concepto de información que se establece, en donde se exige que la función f cumpla las siguientes condiciones:

a) Si $p(x) = 1$, entonces $f(p(x)) = 0$.

Esto indica que si es x el suceso seguro, su probabilidad de ocurrencia es la unidad, y por tanto, tal como se indicaba en la introducción, no proporciona ninguna información.

b) Si $p(x) < p(y)$, entonces $f(p(x)) > f(p(y))$.

Esta segunda condición está en concordancia con lo indicado de que la información aportada por un suceso era tanto mayor cuanto menor es su probabilidad de ocurrencia.

c) f debe ser aditiva, para sucesos independientes.

Parece lógico exigir para f una forma tal que la cantidad de información obtenida de la verificación conjunta de dos sucesos independientes sea la suma de las cantidades de información aportadas por cada uno de ellos; por consiguiente, la tercera condición que debe cumplir la función f es:

Si x e y son independientes, entonces

$$f(p(x)\,p(y)) = f(p(x)) + f(p(y))$$

La función que satisface tales condiciones es una función logarítmica, que para cualquier suceso x, y una constante k, denotamos por:

$$I_x = -k \log p(x) \quad \text{con } k > 0$$

Si consideramos ahora un experimento probabilístico con n posibles resultados: $x_1, x_2, \ldots x_n$, con probabilidades $p_1, p_2, \ldots p_n$, satisfaciendo las condiciones $p_i \geqslant 0$; $i = 1, \ldots, n$; $\Sigma_i\, p_i = 1$.

Se puede representar un espacio finito de probabilidad en el esquema

$$X = \begin{bmatrix} x_1 & \ldots & x_n \\ p_1 & \ldots & p_n \end{bmatrix} = \begin{bmatrix} x_1 & \ldots & x_n \\ p(x_1) & \ldots & p(x_n) \end{bmatrix}$$

Tomando en consideración lo expuesto anteriormente, la ocurrencia del resultado x_1 eliminaría una cierta cantidad de incertidumbre,que vendría medida por la función f, o lo que es lo mismo, la aparición del suceso x_1 nos proporciona una información

$$I_{x_1} = -\log p_1 \text{ para el valor } k = 1$$

y de forma general, para cualquier resultado x_i

$$I_{x_i} = -\log p_i;\ i = 1, 2, \ldots, n$$

Si el interés no está centrado en la cantidad de indeterminación que desvela la constación de un resultado particular, sino en el grado de incertidumbre que presenta globalmente el esquema X, se considera la cantidad media de información que obtendremos al repetir el experimento un gran número de veces.

De esta forma, la incertidumbre global de un esquema X, que designaremos H(X), viene dada por la esperanza matemática de la cantidad de información de sus resultados

$$H(X) = -\Sigma_i\, p_i \log p_i$$

Esta aserción se ejemplariza de una forma inmediata si se consideran, por ejemplo, los dos esquemas siguientes:

$$X = \begin{bmatrix} x_1 & x_2 \\ 0{,}5 & 0{,}5 \end{bmatrix};\quad Y = \begin{bmatrix} x_1 & x_2 \\ 0{,}96 & 0{,}04 \end{bmatrix}$$

Se puede observar que el suceso x_1 proporciona mayor información en el esquema X que en el Y; por el contrario, el suceso x_2 suministra una mayor cantidad de información en el esquema Y, pero globalmente el esquema X tiene mayor incertidumbre, y por consiguiente más capacidad de producir información que el Y.

En efecto, si consideramos la expresión anterior y logaritmos decimales, tenemos que:

$$H(X) = -0{,}5(-0{,}30103) + 0{,}5(-0{,}30103) = 0{,}30103$$

$$H(Y) = -0{,}96(-0{,}01773) + 0{,}04(-1{,}39794) = 0{,}07294$$

con lo que se comprueba que $H(X) > H(Y)$

Shannon introduce el término entropía, para definir estas funciones.

Definición

Sea una distribución finita de probabilidad $p_i \geqslant 0$; $i = 1,2, \ldots, n$ tal que $\Sigma p_i = 1$. La correspondiente ENTROPIA viene dada por la expresión:

$$H_n = H_n(p_1, \ldots, p_n) = -\Sigma_i\, p_i \log p_i$$

El logaritmo puede tomarse en cualquier base mayor que 1, pero habitualmente y por razones que veremos posteriormente se toma en base dos.

Esta cantidad H_n (incertidumbre promedio) puede interpretarse como una medida de la incertidumbre o como una medida de la información media susceptible de ser obtenida del esquema y verifica las siguientes condiciones:

1. Es siempre positiva:

$$H_n(p_1, \ldots p_n) \geqslant 0$$

2. Si $p_j = 1$ y $p_i = 0$ $(1 \leqslant i \leqslant n \,;\, j \neq i)$

$$\text{entonces, } H_n(p_1, \ldots, p_n) = 0$$

O sea, la entropía es cero si una de las categorías va a ocurrir con probabilidad 1.

3. $H_{n+1}(p_1, \ldots, p_n, 0) = H_n(p_1, \ldots, p_n)$

Es decir, cuando un suceso tiene probabilidad cero de aparición, su aportación a la incertidumbre es nula.

Por otra parte, la función H así definida posee ciertas propiedades que por su importancia deben ser mencionadas, como son: independencia, continuidad y aditividad.

Independencia: La aportación de un suceso x_i del esquema a la incertidumbre viene dada por el término $p_i \log p_i$, lo que nos indica que dicha aportación depende única y exclusivamente de su probabilidad de aparición, sin referencia alguna a su naturaleza, dimensión, o cualquier otra magnitud. Por ello, la entropía se trata de una definición matemática independiente del sustrato físico, o lo que es lo mismo, mide la información capaz de ser suministrada por un esquema, con independencia del soporte material que lo contenga.

Continuidad: La función $p_i \log p_i$ es continua, por ello pequeñas variaciones en la probabilidad de ocurrencia de un suceso entrañan solamente pequeñas variaciones en la incertidumbre del sistema.

Aditividad: La cantidad de información de dos esquemas independientes X e Y es la suma de las entropías de cada uno de ellos.

En efecto, si el esquema X viene dado por una distribución de probabilidad $p_i \geqslant 0$; $i = 1, 2, \ldots, n$ $\Sigma_i p_i = 1$; e Y, por la distribución $p_j \geqslant 0$; $j = 1, 2, \ldots, k$; $\Sigma_j p_j = 1$

y designamos por p_{ij} las probabilidades de las parejas de sucesos, entonces:

$$H(X,Y) = -\Sigma_{ij} p_{ij} \log p_{ij}$$

en virtud de la independencia $p_{ij} = p_i p_j$, y en consecuencia:

$$\log p_{ij} = \log p_i + \log p_j$$

por consiguiente,

$$H(X,Y) = -\Sigma_{ij} p_{ij} \log p_i - \Sigma_{ij} p_{ij} \log p_j$$

o lo que es igual,

$$H(X,Y) = -\Sigma_j p_j \Sigma_i p_i \log p_i - \Sigma_i p_i \Sigma_j p_j \log p_j$$

como quiera que $\Sigma_j p_j = 1$ y $\Sigma_i p_i = 1$ tenemos que

$$H(X,Y) = -\Sigma_i p_i \log p_i - \Sigma_j p_j \log p_j$$

es decir,

$$H(X,Y) = H(X) + H(Y)$$

Se pueden generalizar las funciones y hablar de incertidumbre multivariante y transmisión de la información multidimensional. El caso más simple es el caso bivariante en que la entropía de Shannon se define como:

$$H(x,y) = -\Sigma p(x,y) \log p(x,y)$$

donde p(x,y) denota la probabilidad conjunta.

Si los dos sucesos X e Y son independientes, tenemos la incertidumbre conjunta máxima

$$H_{max}(x,y) = -\Sigma p(x,y) \log p(x,y)$$

Si por el contrario, ambos sucesos X e Y están correlacionados debemos considerar la denominada incertidumbre contingente

$$H(x{:}y) = H_{max}(x,y) - H(x,y)$$

Por otra parte, la incertidumbre condicional de ambos sucesos queda definida en la forma siguiente:

$$H_x(y) = -\Sigma_x p(x) \Sigma_y p_x(y) \log p_x(y)$$

$$H_y(x) = -\Sigma_y p(y) \Sigma_x p_y(x) \log p_y(x)$$

Las formas multivariantes son una generalización de la bivariante.

Cantidad de información de un mensaje. Unidad de información

El mayor interés de la medida de la cantidad de información aparece cuando tratamos de transmitir el conocimiento adquirido acerca de la ocurrencia de un suceso. Para que el resultado de una experiencia pueda ser enviado se necesita un código que permita expresar dicho resultado como una sucesión de signos o símbolos que constituyan el mensaje, de tal forma que el destinatario sea capaz de reconstruir el resultado que se le quiere transmitir con el mensaje recibido.

El método de codificación más simple es, evidentemente, utilizar un símbolo diferente para cada uno de los posibles resultados. En este caso, la cantidad media de información, por símbolo del mensaje, es igual a la entropía del esquema del que se esté considerando su resultado. Por ello, Shannon definió la cantidad de información de un mensaje X como:

$$H(x) = -\Sigma_i p_i \log_2 p_i$$

donde p_i es la probabilidad del símbolo x_i.

Esta cantidad H(x) representa la cantidad media de información por símbolo y no la cantidad de información de un mensaje. La información de un mensaje se obtendrá multiplicando H(x) por el número total de símbolos utilizados en ese mensaje. Es decir, la función H caracteriza no a un mensaje en particular, sino al conjunto de símbolos con la misma distribución de probabilidad, que es lo que podíamos llamar, en sentido amplio, un alfabeto.

Consideremos ahora el caso de la codificación binaria, en la que la transmisión de información se realiza utilizando un alfabeto compuesto por dos símbolos, 0 y 1. En este caso, la cantidad media de información por símbolo es:

$$H(x) = -p(0)\log_2 p(0) - p(1)\log_2 p(1)$$

Si $p(0) = p$, la fórmula anterior se transforma en

$$H(x) = -p\log_2 p - (1-p)\log_2(1-p)$$

que es una función que depende únicamente de p. La función alcanza su máximo cuando $p = 1/2$, en cuyo caso:

$$H(x) = -2\frac{1}{2}\log_2\frac{1}{2} = 1$$

ésta es la unidad de información; se denomina bit (sugerido por el estadístico Tukey como contracción de binary digit) y puede ser definida como la máxima cantidad de información que puede obtenerse con un alfabeto binario.

La incertidumbre que se está considerando es univariada y es un concepto importante en el tratamiento de muchos problemas psicológicos (Jensen, 1979, 1982; Lally y Nettelbeck, 1977). Como quiera que los conceptos de información son significativos en la conducta humana, el número de categorías en cualquier situación de elección tendrá importantes consecuencias para la conducta; por consiguiente, debemos conocer cuántas posibilidades de elección o conducta existen, así como sus probabilidades. Concretamente, en el contexto de los texts psicológicos, podemos especificar exactamente la cantidad de información potencial de la respuesta; pero no así cuándo se usan cuestiones abiertas.

La Teoría de la Información en los tests de inteligencia

Nos planteamos aquí algunos problemas relacionados con los tests de inteligencia de acuerdo con la teoría matemática de la información. Hay que señalar que a pesar de la amplia utilización de la Teoría de la Información en gran número de campos de la Psicología, son escasísimos los trabajos que abordan la teoría de los tests desde esta perspectiva. Puede señalarse como excepción, algo alejado ya en el tiempo, el trabajo de Hick (1951) «Information theory and intelligence tests».

Dentro de los tests de inteligencia podemos preguntamos acerca de la información que contiene la pregunta, la información contenida en la respuesta y el proceso (inductivo, deductivo o inductivo-deductivo) por el que el individuo llega a la respuesta correcta.

El fenómeno psicológico asociado con el acto de solución no es objeto de estudio en la Teoría de la Información. La «aptitud de un individuo para percibir relaciones» u otras aptitudes implicadas en el procedimiento nos dan una aproximación, según los objetivos del test, del conocimiento o habilidad en el sujeto que se aplica, sin discernir si se ha llegado a este «conocimiento» a través de la información o de un proceso de maduración.

En el lenguaje de la Teoría de la Información el sujeto recibe una señal que es la pregunta, y a partir de ella se le requiere para que construya el mensaje original. La ocultación de la respuesta, o de parte de la respuesta es de naturaleza probabilística, desde el punto de vista del sujeto. El sujeto del test es lo que se suele llamar un «filtro predictivo», y lo que a menudo se quiere comprobar es hasta qué punto es un buen filtro para cierto tipo de «inputs», refiriéndonos a «predictivo» solamente en el sentido de aquel futuro que pertenece al marco de referencia del sujeto.

A) *Información contenida en la pregunta*

Puesto que la información es una función de la probabilidad del mensaje o símbolo, pero no una propiedad del mensaje o símbolo en sí mismo, no se puede derivar una medida absoluta a priori de la información. Hay instrucciones puramente deductivas como «2 es a 4 como 4 es a x», si las relaciones se entienden en el sentido general de la proporcionalidad. Pero cualquier otro tipo de cadena «A es a B como C es a X» no implica convencionalmente proporcionalidad y hay que saber si es una relación única y en qué consiste. Si se da esta unicidad podremos saber la cantidad de información, porque es evidente que existe una relación entre la información dada y la requerida. Es más fácil completar 2, 4, 6, X que completar la secuencia 2, 4, X, Y. Por tanto, no se puede conocer la información dada a alguna cuestión si no hay alguna suposición referente a la noción que el sujeto puede tener y de las respectivas probabilidades relativas implicadas, lo que conduce a la necesidad de recurrir a la experimentación.

B) *Información contenida en la respuesta*

1) Respuestas dicotomizadas

En el contexto de la teoría de los tests, si se hace una codificación de la respuesta en acierto (A) y fallo (F) asignando los valores 1 y 0 respectivamente, cualquier secuencia de unos y ceros, por ejemplo 1, 0, 1, 0 se puede considerar como una codificación del resultado A F A F. A la secuencia observada se le puede llamar «señal», y el número binario que se desea reconstruir es el mensaje. Si x representa el conjunto de posibles mensajes, e y representa las señales, la información recibida acerca de la capacidad del sujeto de acuerdo con la teoría de Shannon es:

$$R = H(x) - H_y(x) = H(y) - H_x(y) = H(x) + H(y) - H(x,y)$$

donde

$H_y(x)$ es la entropía de x cuando y es conocida
$H_x(y)$ es la entropía de y cuando x es conocida

H(x,y) es la entropía conjunta basada en todos los posibles pares (x,y)

El resultado observado es el resultado de un proceso no markoviano, ya que si el primer dígito representa que el sujeto se ha equivocado, el segundo puede que dé respuesta correcta, tiende a corregir la señal, etc. La probabilidad en cualquier estado depende de todos los estados previos. En este sentido, un test de inteligencia podría construirse como un proceso ramificado.

En el caso de que las probabilidades de acierto y fallo en los ítems del test, p_i y q_i, puedan derivarse de alguna modelización teórica del proceso psicológico por el que los sujetos emiten sus respuestas, es posible utilizar las funciones de entropía para valorar a priori la información que proporcionará el test respecto del sujeto o grupo de sujetos a que va dirigido. La posibilidad de cuantificar, en relación con las variables consideradas en la teoría, la cantidad de información que aporta cada ítem, nos da la opción de construir pruebas con características determinadas previamente y ajustadas a los fines que se persigan. En este sentido, los modelos de rasgo latente permiten controlar la información que suministra el test en función de los parámetros aptitud del sujeto, dificultad del ítem, poder discriminante, etcétera.

2) Cuestiones de elección múltiple

Consideremos ahora preguntas en las cuales el número de alternativas que se ofrecen como posible respuesta es t; de forma general, se puede decir que cuanto mayor es t, mayor es la capacidad potencial que tiene dicha cuestión de producir información, por lo que t no debe ser muy pequeño.

La máxima entropía resulta si todas las t son igualmente posibles, pero la máxima información útil se obtiene igualando las probabilidades de las categorías que en realidad se emplean en la puntuación; esta situación puede representarse en el siguiente esquema:

$$\begin{bmatrix} C & F & \dots & F \\ \frac{1}{2} & \frac{1}{2(t-1)} & \dots & \frac{1}{2(t-1)} \end{bmatrix}$$

Entonces la entropía total por cuestión es:

$$H(x) = -\frac{1}{2}\log\frac{1}{2} - \Sigma\frac{1}{2(t-1)}\log\frac{1}{2(t-1)} =$$

$$= \frac{1}{2}\log 2 + \frac{1}{2}\,\Sigma\,\frac{1}{t-1}(\log 2 - \log(t-1)) =$$

$$= \frac{1}{2}\log 2 + \frac{1}{2}(\log 2 + \log(t-1)) =$$

$$= \log 2 + \frac{1}{2}\log(t-1)$$

Donde $\frac{1}{2}\log(t-1)$ es la información no relevante debida a la clase de respuestas incorrectas, supuestas equiprobables. Pero si una subdivisión es significativa en términos de grados o clases de equivocación, se puede encontrar una mayor cantidad de información útil.

C) *Modelos de rasgo latente*

Se ha considerado anteriormente la información contenida en la respuesta desde el punto de vista de la teoría clásica de los tests. Así, la entropía de Shannon se expresa como función de las probabilidades de respuesta correcta p y de respuesta errónea q acerca de las que no se realizaba ninguna hipótesis.

Si se aborda ahora el problema de analizar la información contenida en la respuesta, desde la perspectiva de los modelos de rasgo latente tendremos expresiones explícitas que nos proporcionarán, para cada sujeto y cada ítem, las probabilidades de obtener una respuesta correcta o equivocada.

Por ejemplo, si se considera el modelo de Rash tendremos que las probabilidades de éxito y de fracaso al contestar un ítem son:

$$\text{Prob(éxito)} = \exp(b-d)/1 + \exp(b-d)$$
$$\text{Prob(fracaso)} = 1/1 + \exp(b-d)$$

donde b denota la habilidad o aptitud del sujeto y d la dificultad del ítem.

Los sucesos éxito/fracaso aparecen en la corrección de la prueba con las probabilidades anteriormente indicadas y de acuerdo con la Teoría de la Información, la cantidad de información que se podría obtener viene dada por la siguiente expresión, en la cual se han tomado logaritmos neperianos por razones de simplicidad

$$H(x) = -\frac{\exp(b-d)}{1+\exp(b-d)}\ \ln\ \frac{\exp(b-d)}{1+\exp(b-d)} -$$

$$-\frac{1}{1+\exp(b-d)}\ \ln\ \frac{1}{1+\exp(b-d)} =$$

$$= \frac{\exp(b-d)}{1+\exp(b-d)}\{\ln(1+\exp(b-d)) - (b-d)\} -$$

$$-\frac{1}{1+\exp(b-d)}\ln(1+\exp(b-d)) =$$

$$= \ln(1+\exp(b-d)) - (b-d)\frac{\exp(b-d)}{1+\exp(b-d)}$$

La representación de la función H(x) muestra que la incertidumbre que puede desvelar la respuesta a un ítem es función de la aptitud del sujeto y de la dificultad de la pregunta.

Aplicación: Cálculo de la cantidad de información en la respuesta usando el test de Raven

El test de Raven es el test de inteligencia que más se adecua a los propósitos investigadores relativos a la inteligencia. Por eso lo hemos adoptado para calcular la información contenida en las respuestas, considerando distintos puntos de vista: *a*) el caso de reducción de las alternativas por el empleo de respuestas dicotomizadas; *b*) el que considera las significaciones asociadas a las nueve alternativas posibles, y *c*) la información obtenida a partir de una primera aproximación con un modelo logístico de dos parámetros.

Método

La muestra utilizada fue de 90 sujetos pertenecientes a los cursos 1.º, 2.º y 4.º de las licenciaturas de Biológicas, Matemáticas y Psicología de la Universidad de Granada. A todos ellos se les pasó el Test de Matrices Progresivas de Raven, de Nivel Superior series I y II, en la forma y con las limitaciones prescritas en el manual del mismo.

La tabulación de los resultados, así como el cálculo de la cantidad de información de los ítems, a partir de las distintas opciones consideradas, fue ejecutada en un ordenador IBM Personal Computer AT. En el mismo equipo, se obtuvieron las estimaciones máximo-verosímiles de los parámetros aptitud y dificultad de los ítems del modelo logístico considerado. Los programas para el cálculo de la cantidad de información contenida en los ítems así como para la estimación de los parámetros han sido elaborados por los autores.

Resultados

El cuadro 1 presenta, para el total de sujetos, la tabulación de los resultados obtenidos, tanto para los 12 ítems de la serie I como para los 36 de la serie II. La opción 0 recoge el número de sujetos que no responden al ítem; las otras opciones indican las frecuencias con que eran seleccionadas en cada ítem las 8 alternativas ofrecidas. Las frecuencias en negrita corresponden a las respuestas correctas.

El cuadro 2 proporciona para los 36 ítems de la serie II la cantidad de información contenida en las respuestas, de acuerdo con los siguientes criterios:

1.º Se emplean para su cálculo las nueve alternativas posibles.

2.º Las nueve opciones se dicotomizan considerando sólo dos alternativas para la respuesta, las de acierto y error.

CUADRO 1

Serie I

Ítem	Opciones								
	0	1	2	3	4	5	6	7	6
1					1				**89**
2		1			**88**	1			
3		1				**87**		1	1
4		**88**	2						
5	1	3	**80**			2		2	2
6	2	1	10			**72**	1	2	2
7	5			1			**84**		
8	7	2	3	**71**			2	1	4
9	14		1	6		3	4	**55**	7
10	26			2	2			2	**58**
11	42	6		2	5	1	8	**23**	3
12	64	1	2	1		1	**18**		3

Serie II

Ítem	Opciones								
	0	1	2	3	4	5	6	7	8
1				1		**89**			
2		**88**	1					1	
3					1		1	**85**	3
4		2		1	**80**	4		2	1
5		3	1	**82**		1		3	
6		**87**	1				1	1	
7			10				**80**		
8	1	**67**	6	7	2	2	2	1	2
9	1	1				1			**87**
10	2		2	1	**76**	2		3	4
11		2		2		**82**	2	2	
12		1		1	1	1	**80**	3	3
13			**75**	2	1	6	3	2	1
14	1	**75**			4	5	2	3	
15		1	**73**				2	1	13
16	2	4		4	**66**	7	1	3	3
17	1			29	2		**57**		1
18	3	12	1	3	3	3	1	**63**	1
19	1	1		**72**	4	6		3	3
20			16	1	5	2	2	2	**63**
21	3	7		2	10	2			**66**
22	8	1	1	3	2	5	3	**45**	22
23	6	5	1	9		5	**56**	2	6
24	6	32	6	**37**		2		4	3
25	1	4	18		5			**56**	6
26	10		**48**	2		11	16		3
27	10	3	2	6	20	5	7	**37**	
28	23	1	2	11	12	**30**	4	3	4
29	20	3	12	9	2	3	**18**	19	4
30	35	3	6	1	2	**36**	2	1	4
31	35	**4**	1	2	**28**	4	5	5	6
32	36	7	2	2	11	7	1	7	**17**
33	47	2	4	6	2	**22**	3	4	
34	48	**18**	4	7	2	4	6		1
35	47	5	2	**24**	5		1	4	2
36	63	7	**2**		3	4	7	2	2

3.º Se utilizan para su cómputo las probabilidades marginales de éxito o fracaso que proporciona para cada ítem el modelo logístico de dos parámetros.

CUADRO 2

Ítem	I	II	III
1	0,0881	0,0881	0,0887
2	0,1760	0,1537	0,1548
3	0,3857	0,3095	0,3114
4	0,7390	0,5033	0,5058
5	0,5938	0,4237	0,4351
6	0,2637	0,2108	0,2123
7	0,5033	0,5033	0,5058
8	1,4964	0,8200	0,8223
9	0,2637	0,2108	0,2123
10	1,0074	0,6236	0,6262
11	0,6105	0,4327	0,4351
12	0,7667	0,5033	0,5058
13	1,0316	0,6500	0,6527
14	1,0082	0,6500	0,6527
15	0,9145	0,6991	0,7018
16	1,5353	0,8366	0,8389
17	1,2101	0,9481	0,9495
18	1,6184	0,8813	0,8833
19	1,1890	0,7219	0,7245
20	1,4232	0,8813	0,8833
21	1,3746	0,8316	0,8389
22	2,1323	1,0000	1,0000
23	1,9365	0,9565	0,9578
24	2,0638	0,9771	0,9761
25	1,6542	0,9565	0,9578
26	1,9351	0,9968	0,9972
27	2,4259	0,9771	0,9761
28	2,5466	0,9183	0,9165
29	2,7889	0,7219	0,7193
30	2,0707	0,9710	0,9699
31	2,3712	0,8945	0,8925
32	2,5295	0,6991	0,6965
33	2,0537	0,8024	0,7999
34	2,0885	0,7219	0,7193
35	1,9771	0,8366	0,8344
36	1,6627	0,1537	0,1527

En el cuadro 3 se refleja la cantidad de información que se consigue en cada ítem a partir de las respuestas de aquellos sujetos que han obtenido una puntuación total de 26 en la serie II. En primer lugar, considerando las frecuencias de aciertos y errores obtenidos, y en segundo lugar, a partir de las probabilidades de éxito o fracaso que proporciona el modelo logístico para una aptitud de 0,6293, que es la que corresponde a los sujetos con 26 aciertos.

Comentarios y conclusiones

Los resultados del cuadro 2 muestran cómo aparece una ganancia sustancial en la cantidad de información, cuando se consideran las nueve alternativas posibles de la respuesta en lugar de la clasificación habitual de las contestaciones en verdaderas o falsas. La única excepción ocurre en el ítem 1 como consecuencia de que en él sólo existen dos categorías de respuestas, la 3 y la 5. El aumento en la información proviene del incremento en el número de mensajes distintos que podemos considerar al disponer de nueve opciones diferentes. Recuérdese que la máxima cantidad de información que se puede obtener con dos dígitos es precisamente la unidad de información, o sea, el bit; por el contrario, con un alfabeto de nueve caracteres puede alcanzarse una cantidad de información de 3,1699 bit en el caso de que todos ellos sean equiprobables. Debe de señalarse, no obstante, que la utilización de toda esa información requiere que cada una de las respuestas posibles esté dotada de una significación propia, en orden a la medida de la inteligencia. Si las ocho alternativas que no son la respuesta correcta carecen de una valoración distinta y sólo se puede apreciar que son erróneas, resultará que estos ocho mensajes se muestran como equivalentes y quedamos constreñidos a utilizar un esquema que contiene únicamente dos sucesos.

Al comparar los valores que se obtienen para la cantidad de información con las frecuencias de aciertos y fallos, con los calculados a partir de las

CUADRO 3

Ítem	I	II
1	0,4690	0,0372
2	0,0000	0,0675
3	0,0000	0,1490
4	0,0000	0,2736
5	0,0000	0,2247
6	0,0000	0,0958
7	0,0000	0,2736
8	0,7219	0,5727
9	0,0000	0,0958
10	0,0000	0,3690
11	0,0000	0,2247
12	0,0000	0,2736
13	0,7219	0,3924
14	0,7219	0,3924
15	0,8813	0,4387
16	0,8813	0,5942
17	0,9710	0,7724
18	0,7219	0,6569
19	0,4690	0,4616
20	0,8813	0,6569
21	0,4690	0,5942
22	0,4690	0,9450
23	0,4690	0,7901
24	1,0000	0,9971
25	0,0000	0,7901
26	1,0000	0,9110
27	0,4690	0,9971
28	0,9710	0,9886
29	0,9710	0,8148
30	1,0000	0,9991
31	1,0000	0,9724
32	0,8813	0,7903
33	0,9710	0,8962
34	0,8813	0,8148
35	1,0000	0,9274
36	0,0000	0,1671

probabilidades que proporciona el modelo logístico de dos parámetros, se observa que son prácticamente iguales. Este resultado es lógico teniendo en cuenta que ambas situaciones consideran las mismas respuestas, acierto o error, y que las estimaciones de los parámetros dificultad y aptitud que determinan los valores de las probabilidades han sido obtenidas de forma que atribuyesen la máxima probabilidad a la muestra utilizada.

Por el contrario, cuando se comparan las cantidades de información determinadas por ambos procedimientos para un nivel de aptitud determinado, como ocurre en el cuadro 3, las diferencias son de cierta entidad, observándose unos valores más uniformes para las obtenidas mediante la utilización del modelo logístico. De forma global se obtiene un ligero aumento de información con este modelo como consecuencia de la que procede de otros grupos con aptitudes diferentes que son utilizados para la estimación de las dificultades. Así, podemos concluir que la utilización de modelos de rasgo latente no aumenta la información contenida en las respuestas pero mejora sus posibilidades de utilización para cada nivel de aptitud.

Referencias

Barnard, G. A. (1952): The theory of information, *J. Roy. Stat. Soc.*

Fisher, R. A. (1950): *Contributions to Mathematical Statistics*, N. Y., J. Wiley.

Gabor, D. (1946): Theory of Communication, *Journal Institute of Technical and Electronics Enginers, 93*, III, 429.

Hick, W. E. (1951): Information Theory and intelligence tests, *British Journal of Psychology, 4*, 3, 157-164.

Jensen, A. R. (1979): g: Outmoded Theory or unconquered frontier, *Creative Science & Technology, 2*, 16-29.

Jensen, A. R. (1982): Reaction time and psychometric g. En Eysenck: *A model for intelligence*, Berlin, Springer-Verlag.

Lally, M., y Nettelbeck, T. (1977): Intelligence, reaction time, and inspection time, *Am. J. of Mental Deficiency, 82*, 273-281.

Mc Kay, D. M. (1950): Quantal aspects of Scientific Information, *Philosophical Magazine, 41*, 289-311.

Proceeding of the Symposium on Information Theory (1950), London.

Shannon, C. E. (1948): A Mathematical theory of comunication, *Bell System Technical Journal, 27*, 379-423.

Shannon, C., y Weaver, W. (1949): *The Mathematical Theory of Comunication*, Urbana, University of Illinois Press.