

# METODOLOGÍA

## Validez convergente de dos pruebas de evaluación de la minuciosidad

José Manuel Hernández López, José Hector Lozano Bleda, Pei-Chun Shih y José Santacreu Mas  
Universidad Autónoma de Madrid

La falta de correspondencia entre la medición objetiva del comportamiento y la medición basada en el autoinforme del individuo, unida al uso generalizado del cuestionario como instrumento de evaluación de la personalidad, plantea un problema en relación a la validación de instrumentos de evaluación objetiva en dicho ámbito de estudio. El objetivo del siguiente trabajo es la obtención de una medida de validez convergente para una prueba informatizada de evaluación objetiva de la minuciosidad a partir del desarrollo de otra prueba funcionalmente equivalente y morfológicamente distinta. Se exponen las características de la prueba de evaluación diseñada. La muestra la conformaron 267 titulados universitarios (194 hombres y 73 mujeres). El alto valor de validez convergente obtenido ( $r = .638$ ,  $p < .001$ ) supone un primer paso en el proceso de validación de ambas pruebas.

*Convergent validity of two tests of conscientiousness.* The absence of correspondence between objective measurement of behavior and measurement based on self-report along with generalized use of questionnaires as personality assessment instruments raises a problem with regard to the validity of assessment instruments in this area of study. The goal of this work is to obtain a convergent validity measure for an objective assessment computerized task of conscientiousness based on the development of a functional equally but morphologically different task. The features of the designed assessment task are commented upon. The sample was made up of 267 university graduates (194 males and 73 females). The high convergent validity value obtained ( $r = .638$ ,  $p < .001$ ) is a first step in the validation process of both tasks.

### La evaluación de la personalidad: el comportamiento del individuo, lo que dice y lo que hace

El estudio de la personalidad está ligado a la posibilidad de demostrar que la conducta de los individuos es función de tendencias de comportamiento relativamente consistentes y estables que ocurren en diversas situaciones y en diferentes momentos. Dichas tendencias de comportamiento se manifiestan en situaciones en las que las contingencias ambientales no determinan la conducta concreta a realizar. Para que una tendencia de comportamiento se considere una variable de personalidad es imprescindible que pueda ser categorizada con coherencia, es decir, con una correspondencia clara entre el constructo y las conductas a la que se refiere. En

este sentido, la variable de personalidad puede ser expresada en términos de un patrón comportamental que la describe a priori y sobre el cual habría de verificar el ajuste que tiene el comportamiento específico de un individuo concreto. Tradicionalmente, el estudio de esta correspondencia que, globalmente, hemos llamado validez relacionaba el concepto verbal (por ejemplo, arriesgado, honesto, cobarde o introvertido) con un conjunto de conductas que necesariamente lo describían. Desde nuestra visión teórica, la relación se produce entre la definición de un patrón comportamental y el conjunto de conductas realizadas.

Desde una perspectiva comportamental de estudio de la personalidad cabe diferenciar entre la historia de aprendizaje sintetizada a través de tendencias de comportamiento y la historia de aprendizaje sintetizada a partir de proposiciones verbales (Santacreu, Rubio y Hernández, 2006). Si bien esta última permite ser evaluada mediante el autoinforme del individuo, aquélla habrá de ser necesariamente evaluada a partir de datos obtenidos mediante la observación del comportamiento. Para operativizar dichas tendencias se utiliza el concepto de estilo interactivo entendido como la manera personal e idiosincrásica en que un individuo afronta si-

tuaciones en las que su comportamiento no está pre-determinado por las contingencias ambientales (Ribes y Sánchez, 1992). Así pues, el estilo interactivo constituye la unidad de evaluación para la síntesis en tendencias de comportamiento de la historia de aprendizaje en el estudio funcional de la personalidad (Santacreu, Hernández, Adarraga y Márquez, 2002) de la misma manera que el rasgo constituye la unidad de evaluación para el análisis de la historia de aprendizaje sintetizada en proposiciones verbales en el estudio estructural de la personalidad.

Desde una perspectiva *multiplista* de la evaluación psicológica (Cook, 1985; Fernández-Ballesteros, 1994), y respetando la triangulación como estrategia de recogida de datos (Denzin, 1970; Stieglitz, 2003), se aboga por la necesidad de contar con múltiples fuentes de información, procedimientos e instrumentos de evaluación que, en última instancia, aportan distintos tipos de datos. La evaluación de la personalidad ha considerado la posibilidad de utilización de datos provenientes del autoinforme de los individuos, del registro de sus comportamientos o de sus «huellas conductuales» (Cattell, 1965; Cattell y Kline, 1977; Block, 1993). Ello remite a la conocida distinción entre datos Q (datos de autoinforme recogidos a través de cuestionarios), datos T (de observación directa del comportamiento en situaciones de test) y datos L (observación del comportamiento en situaciones naturales o historias de vida). Dejando al margen el último tipo de datos mencionado, es conocida la existencia de estudios que muestran la falta de correspondencia entre medidas de personalidad obtenidas mediante autoinforme (datos Q) y medidas objetivas basadas en datos observables (datos T) (Cattell y Klein, 1977; Skinner y Howarth, 1975). Ello ha redundado en una merma de la validez convergente de diversos procedimientos de recogida de información que, supuestamente, estaban evaluando la misma dimensión de personalidad. Con independencia de la decisión sobre cuál estaba midiendo bien (lo que nos introduciría en la discusión sobre la validez de constructo), la encrucijada que plantea esta falta de correlación ha derivado en la escasa utilización de los datos objetivos para la evaluación de dimensiones de la personalidad, lo que ha provocado una carencia en el diseño de pruebas objetivas impidiendo obtener indicadores sólidos en términos de validez convergente, al contrario de lo sucedido en la evaluación de las aptitudes y competencias.

#### La validación de las pruebas de evaluación de la personalidad

Partiendo de estas asunciones se plantea un problema relativo a la validación (en términos de validez convergente) de instrumentos de evaluación basados en la observación de tendencias de comportamiento en la investigación en psicología de la personalidad. Los investigadores que obtienen datos a partir del autoinforme de los individuos, con motivo del uso generalizado del cuestionario como instrumento de evaluación en dicho ámbito de estudio (Hernández, Santacreu y Rubio, 1999), han obtenido reiteradamente altos valores correlacionales entre cuestionarios que evalúan el mismo rasgo de personalidad. Pero la ausencia de correspondencia entre datos obtenidos a partir de la observación del comportamiento y datos obtenidos a partir del autoinforme del individuo imposibilita la validación de pruebas objetivas mediante la correlación con índices obtenidos de pruebas de evaluación basadas en el autoinforme. Se hace, por tanto, necesario recurrir a procedimientos de validación que busquen la correlación de datos procedentes de diversas pruebas de evaluación objetiva de la personalidad. El pro-

pósito de este artículo es aportar datos que avalen la validez convergente de pruebas que evalúen la historia de aprendizaje manifestada de forma sintética en tendencias de comportamiento consistentes y estables.

#### La evaluación del comportamiento minucioso, ordenado y concienzudo

El objetivo concreto del presente trabajo es la presentación de una prueba objetiva diseñada para la evaluación del estilo interactivo *ejecución de una tarea de manera ordenada, organizada, siguiendo un patrón sistemático* (Hernández, 2000) que permita obtener una medida de la validez convergente mediante la correlación significativa con una prueba anterior (Hernández, Sánchez-Balmisa, Madrid y Santacreu, 2003). Dicho estilo interactivo podría considerarse equiparable, en términos teóricos, a la dimensión de Minuciosidad del Modelo de Cinco Factores de la Personalidad (Digman, 1990; Goldberg, 1992), más concretamente a alguna de las facetas que componen dicho factor en el cuestionario NEO-PI-R (Costa y McCrae, 1992; 1995). Ahora bien, en términos empíricos, el trabajo de Sánchez-Balmisa, Hernández, Madrid, Peña y Santacreu (2003) constató la inexistencia de correlación significativa de esta medida con la obtenida a partir de la aplicación de la escala de responsabilidad del cuestionario de personalidad BFQ (Caprara, Barbaranelli y Borgnoni, 1993). Dicho resultado va en la línea de la distinción establecida entre la síntesis de tendencias de comportamiento y la síntesis de proposiciones verbales.

La prueba objetiva informatizada originalmente diseñada para evaluar la minuciosidad, conocida como *Test de Minuciosidad (TM-A)*, presentaba unas propiedades psicométricas caracterizadas por una amplia variabilidad en la distribución de las puntuaciones, ausencia de correlación de los índices de competencia y motivación con el estilo interactivo y valores de consistencia interna equiparables a los obtenidos por pruebas consolidadas en la evaluación tradicional de la personalidad (Hernández et al., 2003). La prueba cumplía, por tanto, con los requerimientos exigidos para la evaluación de estilos interactivos, los cuales, básicamente, se concretan en los siguientes. En primer lugar, que la prueba o tarea presente apertura de contingencias, es decir, que no existan determinantes contextuales del comportamiento. Concretamente en este punto se recomienda, como en otras pruebas objetivas, que la tarea sea desconocida para la persona, al menos, en la variable de personalidad que se fuera a medir; que no informe de las consecuencias por ejecutar una u otra respuesta y que en las instrucciones de la prueba no haya establecido un criterio de eficacia para la ejecución de la tarea en un tiempo determinado (Santacreu, Hernández y Rubio, 2004). En segundo término, que la resolución eficaz de la situación no dependa de la puesta en marcha de una competencia. Así, hay que asegurar que todos los individuos resuelven la tarea para así evaluar la forma en que materializan dicha eficacia. En resumen, nos interesa el «cómo» actúan para identificar su estilo de comportamiento. Se trata de que la propia tarea de evaluación no influya o contamine el objeto de evaluación, que no es otro que la síntesis de la historia de aprendizaje a través de las tendencias de comportamiento del individuo.

Por otro lado, en un trabajo anterior (Hernández, Rubio, Revuelta y Santacreu, 2006) pusimos a prueba la validez interna de la prueba *TM-A* atendiendo a su ajuste a un modelo politómico basado en el índice propuesto por Rudas, Clogg y Lindsay (1994) para el análisis de tablas de contingencia. Los resultados indicaron

un adecuado ajuste al modelo ( $G^2= 2.730,77$ ;  $g.l.= 2.158$ ;  $G^2/g.l.= 1,26$ ).

Partiendo de la base de la consistencia y la validez interna de la prueba *TM-A*, el presente trabajo muestra el desarrollo de otra tarea informatizada funcionalmente equivalente (en el sentido de que evalúa el mismo estilo interactivo) y morfológicamente distinta (en el sentido de sus diferencias formales) a la anterior, que, conservando las garantías psicométricas en términos de consistencia interna, permita disponer de datos relativos a la validez convergente entre ambas.

### Método

#### Participantes

La muestra está formada por 267 sujetos, titulados universitarios, aspirantes en un proceso de selección para un puesto de alta cualificación técnica: 194 varones, de edades comprendidas entre los 19 y 43 años ( $\bar{X}= 29.11$ ;  $DT= 3.852$ ) y 73 mujeres de edades comprendidas entre los 22 y los 37 años ( $\bar{X}= 28.03$ ;  $DT= 3.431$ ).

#### Instrumento

La prueba objetiva que se ha diseñado en esta ocasión con el fin de evaluar el estilo interactivo *ejecución de una tarea de manera ordenada, organizada, siguiendo un patrón sistemático* (a la que se ha denominado como *Test de Minuciosidad-Fichas (TM-F)*) consiste en una tarea en la que se presenta un panel que contiene varios tipos de fichas, las cuales otorgan puntos al ser pulsadas. El objetivo explícito de la tarea consiste en obtener la mayor cantidad de puntos pulsando sobre el menor número de fichas, maximizando así la eficiencia del comportamiento. Dado que no todos los botones otorgan el mismo número de puntos, el comportamiento máximamente eficiente consistirá en pulsar exclusivamente sobre los botones que aportan la máxima puntuación. Se pretende, por tanto, que el participante aprenda la relación de contingencia establecida entre el tipo de botón y los puntos obtenidos. Para ello, habrá de pulsar tentativamente con el ratón sobre los distintos tipos de fichas durante los primeros ensayos hasta aprender qué tipo de ficha otorga la mayor puntuación. Una vez establecida dicha relación, el candidato habrá de identificar y pulsar sobre aquellas fichas que proporcionan la mayor cantidad de puntos.

La prueba original de evaluación del estilo interactivo (*TM-A*) consistía en identificar y pulsar con el ratón, de entre una variedad de imágenes distintas, aquellas que fuesen iguales a una que era presentada como modelo (Hernández et al., 2003).

Las figuras 1 y 2 muestran dos ejemplos de matrices correspondientes a ambas pruebas de evaluación.

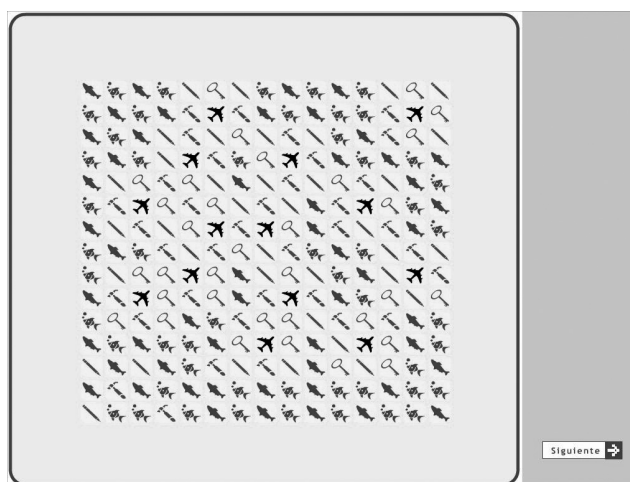
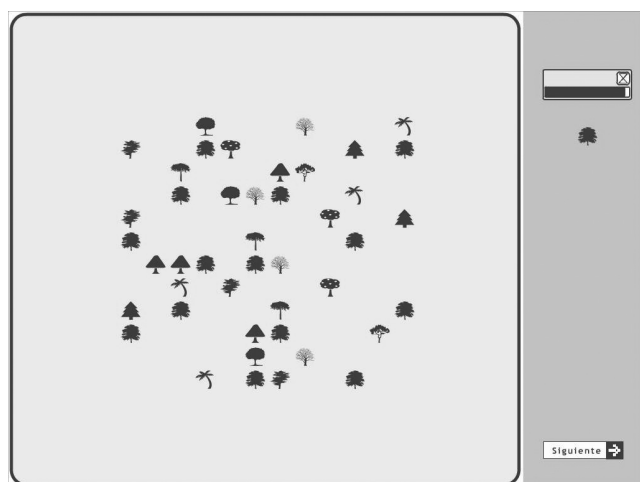
Como puede apreciarse, si bien la prueba original se presentaba como una tarea de evaluación de la capacidad de discriminación perceptiva, la prueba diseñada en esta ocasión se presenta como una tarea de evaluación de la capacidad de aprendizaje. De este modo, se pretende enmascarar el verdadero objetivo de evaluación de la prueba, es decir, la medición del estilo interactivo, cumpliendo así con uno de los requisitos clásicos establecidos por Cattell (1965) para el diseño de pruebas objetivas de evaluación de la personalidad.

Atendiendo a las especificaciones establecidas para la evaluación de estilos interactivos se han tenido en cuenta los siguientes aspectos en el diseño de la nueva prueba de evaluación:

Instrucciones: se indica a los participantes que han de obtener el mayor número de puntos posible pulsando sobre el menor número de fichas en cada ensayo. De acuerdo a lo planteado, la tarea es de contingencias abiertas respecto a la ejecución del comportamiento a evaluar (seguir o no una estrategia minuciosa) y cabe esperar, por tanto, que los candidatos muestren variabilidad en el mismo.

Formato: la prueba consta de 16 ensayos representados por 16 matrices cuadradas de 15 columnas por 15 filas (225 celdas). Las matrices incluyen tres categorías de fichas que representan diferentes figuras (aviones, peces y objetos), las cuales, al ser pulsadas, otorgan diferente cantidad de puntos. Los aviones otorgan siempre 6 puntos; los peces otorgan 2 o 0 puntos, con una relación de contingencia del 50%; y los objetos otorgan 1 o 0 puntos, de nuevo con una relación de contingencia del 50%.

Como puede apreciarse en las figuras 3 y 4, los 14 botones que otorgan la máxima puntuación (*targets*) se distribuyen en una submatriz de 7 filas por 7 columnas, conteniendo cada fila y cada columna 2 *targets*, respectivamente. Dicha distribución reproduce con exactitud la presentada por los *targets* de la prueba original (los que son iguales al modelo). De forma idéntica en ambas pruebas, la matriz va rotándose en cada ensayo, adoptando 4 posiciones diferentes que se van sucediendo a lo largo de los 16 ensayos.



Figuras 1 y 2. Ejemplos de matrices de las pruebas Test de Minuciosidad-Árboles y Test de Minuciosidad-Fichas

Así, por ejemplo, la matriz del primer ensayo es equivalente a la matriz de los ensayos 5, 9 y 13, y así sucesivamente.

La duración de los ensayos en ambas pruebas es de 25 segundos. Dicho tiempo se considera más que suficiente para la realización de la tarea con independencia de la estrategia utilizada durante la ejecución de acuerdo con los resultados obtenidos en el estudio anterior (Hernández et al., 2003).

Durante la tarea, los participantes pueden pasar al siguiente ensayo pulsando el botón correspondiente. De no hacerlo, transcurridos los 25 segundos de duración establecidos para cada ensayo, la prueba automáticamente da paso al siguiente ensayo.

Con la intención de asegurar el aprendizaje de la relación de contingencia, se ha pretendido facilitar la discriminación perceptiva de las fichas que contienen las máximas puntuaciones. De este modo, se han presentado los aviones en color negro y orientados a 45°, en contraste con el color gris y la orientación a 135° de las figuras pertenecientes a las restantes fichas (véase la figura 2).

Retroalimentación: tras pulsar sobre cada ficha, la figura representada en ella se sustituye por el número de puntos obtenidos al pulsarla. Durante el transcurso de la tarea no existen contadores que reflejen el tiempo transcurrido, el total de fichas pulsadas o el total de puntos obtenidos. Sólo al finalizar cada uno de los 6 primeros ensayos se presenta información respecto al número de puntos obtenidos y al total de fichas pulsadas en el ensayo. Con ello se pretende facilitar el aprendizaje al objeto de alcanzar el criterio de eficiencia relativo a los puntos obtenidos. Como puede apreciarse, no se presenta información alguna relativa al comportamiento objeto de estudio, es decir, al seguimiento o no de una estrategia minuciosa.

#### Procedimiento

Ambas pruebas de evaluación fueron aplicadas durante el proceso de selección mencionado. La distancia temporal entre la aplicación de una y otra prueba fue aproximadamente de 1 hora y 40 minutos, período de tiempo en el cual los candidatos realizaron otras tareas de evaluación.

Los candidatos ejecutaron las pruebas de manera individual en una terminal informática. Para su realización no eran necesarios conocimientos informáticos, siendo únicamente preciso el manejo del ratón. Los datos correspondientes a la ejecución de cada candidato se registraron automáticamente en una base de datos para su posterior tratamiento.

#### Variables

##### *La puntuación en el estilo interactivo*

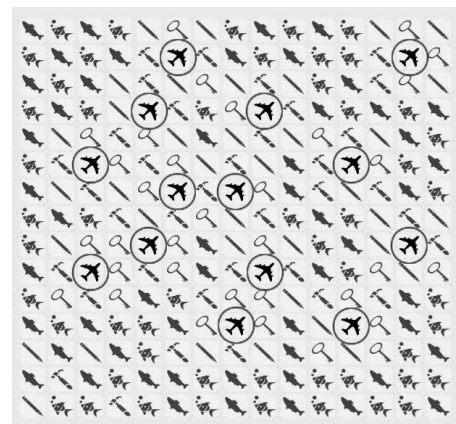
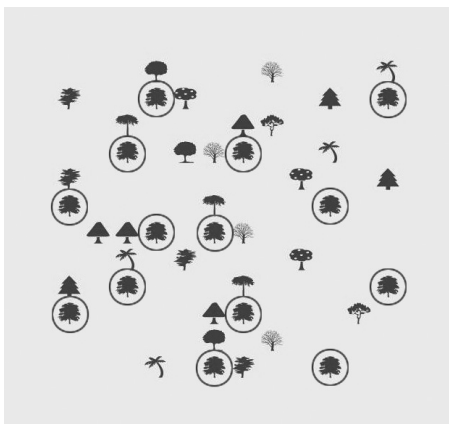
El índice de minuciosidad tiene por objeto cuantificar en qué medida los participantes presentan un estilo interactivo caracterizado por una ejecución ordenada, organizada y siguiendo un patrón sistemático. La clave en la operativización de dicho índice es la trayectoria seguida en la sucesión de pulsaciones sobre los *targets*. Partiendo de la existencia de 14 *targets* en la matriz, tras pulsar sobre uno de ellos sólo se otorgará un punto si la siguiente pulsación es sobre el otro elemento *target* correspondiente, o bien a la misma fila, o bien a la misma columna que el pulsado previamente. En caso de pulsarse sobre cualquiera de los 11 *target* restantes no se obtendría puntuación. Todo ello hace que la puntuación máxima del estilo interactivo en cada ensayo sea de 7 y la mínima de 0. Dado que el índice de minuciosidad lo constituye el sumatorio de las puntuaciones de todos los ensayos, el rango de puntuaciones para el conjunto de la prueba es de 0-112, correspondiente a la máxima puntuación en los 16 ensayos ( $7 \times 16 = 112$ ).

##### *Variables relativas a la dificultad de la prueba*

Para que la nueva prueba diseñada TM-F permita evaluar la variable minuciosidad y orden en la ejecución es imprescindible que todos los participantes aprendan a pulsar, en primer lugar, las fichas correspondientes a los aviones. Como indicador de la dificultad de la prueba, o, si se prefiere, del grado de competencia requerido a los candidatos, se calcula un índice de aprendizaje. Dicho índice permite evaluar la eficacia con la que los candidatos resuelven la tarea, de acuerdo con las instrucciones: obtener la mayor puntuación posible pulsando en el menor número de fichas.

El índice de aprendizaje se calcula mediante el sumatorio del número de *targets* (aviones) pulsados en las 14 primeras pulsaciones partido por el número total de *targets* en cada ensayo (14 *targets* en todos los ensayos). Dado que el rango de valores posible es de 0 a 1 en cada ensayo, el rango del índice para los 16 ensayos es de 0 a 16.

De acuerdo con todo lo anterior, sólo si se cumple la condición de que el participante pulsa, en primer lugar, los 14 botones que representan aviones, se podrá establecer un paralelismo entre los índices de minuciosidad de las dos pruebas. Es decir, únicamente



Figuras 3 y 4. Distribución de los targets en TM-A y TM-F

habiendo llevado a cabo la tarea que la prueba requiere será posible la evaluación del estilo interactivo, que consiste, precisamente, en la manera en que el participante realiza dicha tarea.

El que todos los participantes muestren un alto nivel de aprendizaje permitirá concluir que el nivel de competencia de la muestra ha quedado controlado, no interfiriendo así la dificultad de la tarea en la medición del estilo interactivo.

Resultados

Análisis de dificultad de la prueba

En primer lugar, procedimos a analizar el grado en el cual los participantes han aprendido a pulsar en primer lugar los aviones, condición *sine qua non*, como se mencionó anteriormente, para poder comparar los índices del estilo interactivo de ambas pruebas. El índice de aprendizaje de la muestra tiene un valor medio de 14.8 (D.T.= 1.31).

En la figura 5 se muestra la evolución a lo largo de los ensayos del índice de aprendizaje. Se observa cómo la muestra alcanza el criterio de aprendizaje en torno al cuarto o quinto ensayo, manteniéndose a continuación el comportamiento estable durante el resto de la prueba. Esta evolución justifica que, a efectos de cómputo del índice del estilo interactivo, hayamos utilizado los 12 últimos ensayos de la prueba, lo que nos aporta un indicador cuyo rango abarca una puntuación de 0 a 84 (7x12 ensayos). Para mantener la equivalencia entre las dos pruebas se han seleccionado también sólo los últimos 12 ensayos en la prueba *TM-A*.

Grado de discriminación de la prueba

Uno de los indicadores de que la prueba cumple con el requisito de apertura de contingencias es el hecho de que las puntuacio-

nes en el estilo interactivo presenten variabilidad en su distribución. La tabla 1 muestra los estadísticos descriptivos de los índices de minuciosidad correspondientes a las dos pruebas obtenidos para los 12 últimos ensayos de ambas. La variabilidad de las puntuaciones del índice de minuciosidad de la prueba de fichas queda manifiesta en el alto valor de la desviación típica, 17.85, y en los valores máximo y mínimo, 4 y 84, respectivamente. Como se observa, las medidas de tendencia central, así como las medidas de dispersión de ambos índices, resultan muy parejas.

Las figuras 6 y 7 muestran los histogramas de la distribución de las puntuaciones correspondientes a los índices de evaluación del estilo interactivo de ambas pruebas.

Como se observa en los gráficos, ambos índices dan lugar a una distribución de puntuaciones muy similar. La forma de la distribución se asemeja a la que acostumbra a presentar el índice de minuciosidad de la prueba *TM-A*, es decir, una curva normal ligeramente desplazada a la derecha (Hernández et al., 2003). En esta ocasión, se observa una pequeña acumulación de casos en la cola derecha de la distribución, aproximándose a la forma de una distribución bimodal. Los índices de asimetría y curtosis tipificados adoptan valores de 8.9 y 3.4, respectivamente, indicando alejamiento de la normalidad.

Fiabilidad

Otro de los objetivos de este estudio es comprobar la consistencia con la que los diferentes ensayos de la prueba miden el estilo interactivo, independientemente de la posición que ocupan los targets en la matriz. El coeficiente *alpha* de Cronbach presenta un valor de .945, y los análisis de fiabilidad *dos mitades* y *par-impar*, arrojan resultados de .860 y .904, respectivamente. Estos valores de fiabilidad, junto a los correspondientes a la prueba *TM-A* (tabla

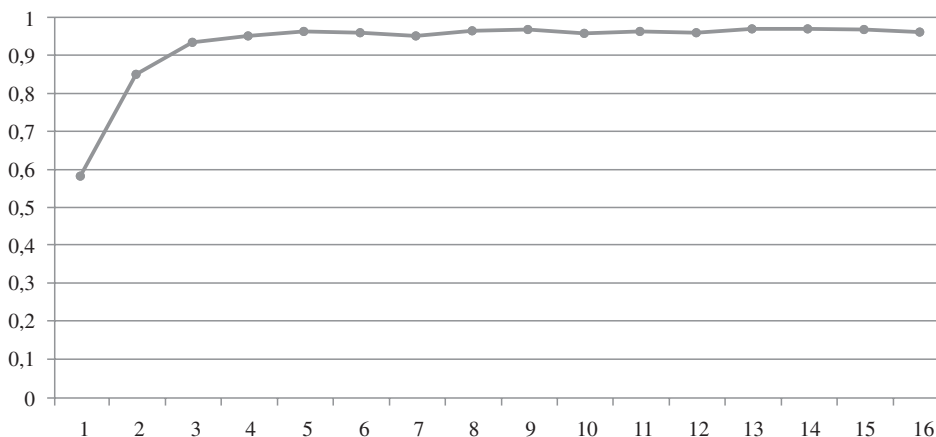
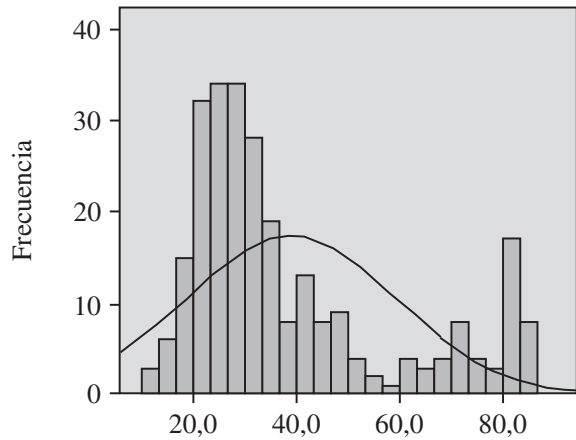
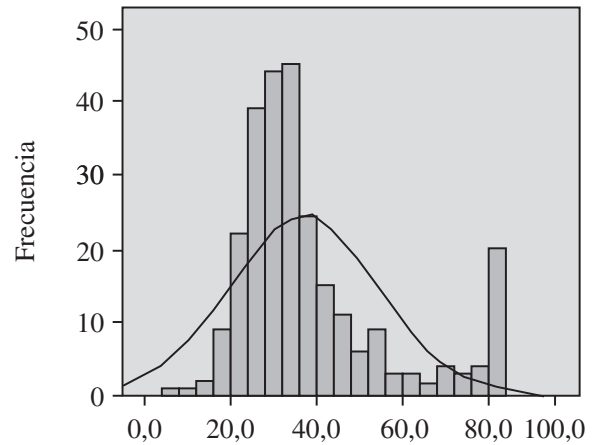


Figura 5. Evolución por ensayos de la media del índice de aprendizaje

	N	Media	Mediana	Moda	D.T.	Mínimo	Máximo	Asimetría		Curtosis	
								Estadístico	Error típico	Estadístico	Error típico
TM-A	267	38,742	31,000	28,0	20,4198	12,0	84,0	1,097	,149	-,080	,297
TM-F	267	38,397	33,000	31,0	17,8498	4,0	84,0	1,327	,149	1,016	,297



TM-A



TM-F

Figuras 6 y 7. Distribución de puntuaciones del estilo interactivo en las pruebas TM-A y TM-F

Tabla 2 Medidas de fiabilidad			
	Alpha de Cronbach	Dos mitades	Par-impar
TM-A	,956	,884	,907
TM-F	,945	,860	,904

2), pueden equipararse a los obtenidos por pruebas consolidadas en la evaluación tradicional de la personalidad.

Análisis de los ítems

En la figura 8 y en la tabla 3 se presenta la evolución de la media de los 12 ensayos en ambas pruebas, así como las desviaciones

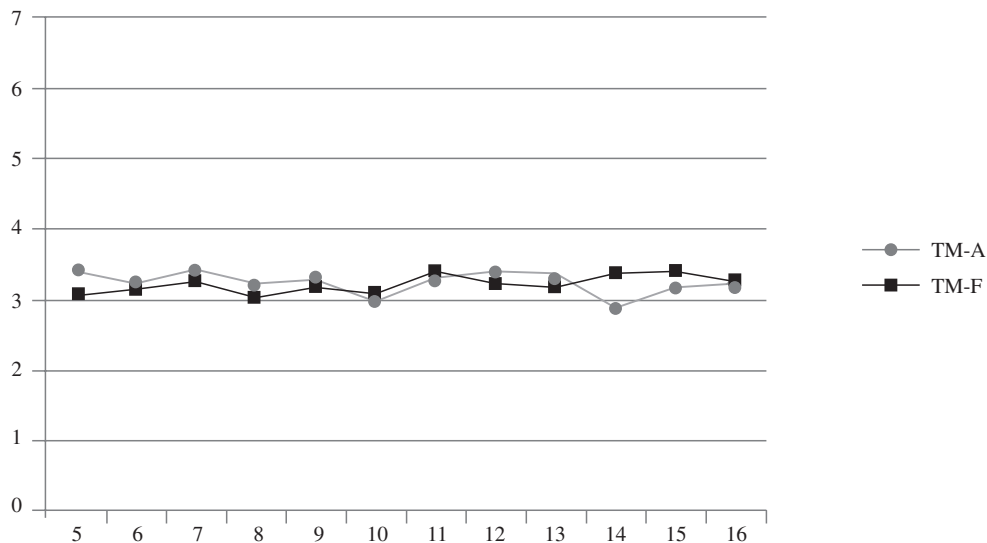


Figura 8. Evolución de los índices de minuciosidad por ensayos de ambas pruebas

Tabla 3 Puntuaciones medias y desviaciones típicas por ensayos de los índices de minuciosidad de ambas pruebas													
		5	6	7	8	9	10	11	12	13	14	15	16
TM-A	Media	3,38	3,21	3,42	3,21	3,27	2,97	3,29	3,39	3,36	2,88	3,16	3,20
	D.T.	1,91	2,16	2,04	2,00	1,94	2,24	2,07	1,97	1,99	2,28	2,11	2,15
TM-F	Media	3,03	3,13	3,25	3,00	3,15	3,08	3,40	3,21	3,15	3,36	3,39	3,24
	D.T.	1,75	1,84	1,86	1,92	1,75	1,95	1,94	1,97	1,79	1,96	1,93	1,96

típicas de cada ensayo. Como puede apreciarse, la puntuación se muestra bastante estable a lo largo de los 12 ensayos. Esta evolución justifica, por un lado, el alto valor de consistencia interna obtenido en ambas pruebas y, por otro, su evidente paralelismo.

#### *Validez convergente y divergente*

Respecto al análisis de la validez convergente entre las dos pruebas diseñadas para evaluar el estilo interactivo, el coeficiente de correlación de *Pearson* entre los índices de evaluación adopta un valor de .638,  $p < .001$ .

En términos de validez divergente, la puntuación de la prueba de minuciosidad, es decir, el comportamiento minucioso, ordenado y concienzudo no muestra relación ( $r = .054$ ,  $p = .382$ ) con una prueba de aprendizaje de contingencias en una tarea de discriminación visual diseñada para la evaluación de la persistencia motivacional (descrita en Hernández, García-Leal, Rubio y Santacreu, 2004).

#### Discusión y conclusiones

Los resultados de este trabajo representan un primer paso en el proceso de validación de pruebas objetivas de evaluación de la personalidad. Las dos pruebas presentadas demuestran ser discriminativas, consistentes y fiables. Además, se relacionan estadísticamente entre sí ( $r = .638$ ,  $p < .001$ ). El alto valor de validez convergente obtenido nos permite afirmar que ambas están midiendo un mismo estilo interactivo. Dicho con otras palabras, las estrategias de actuación que ponen en marcha los individuos enfrentados a dos tareas morfológicamente distintas aunque funcionalmente equivalentes son las mismas.

Nuestros resultados, no obstante, no permiten asegurar que estas pruebas estén evaluando la misma dimensión de minuciosidad

que las tradicionales pruebas de evaluación basadas en el autoinforme de los individuos. Hemos comprobado en anteriores trabajos que los valores correlacionales son lo suficientemente bajos como para afirmar que se refieren a aspectos distintos. Ello supondría un refrendo a los resultados obtenidos desde la década de los sesenta por Cattell y va en la línea de la existencia de dos formas distintas de sintetizar la historia de aprendizaje: en lenguaje y en comportamiento manifiesto.

Una de las grandes ventajas de las pruebas objetivas de evaluación de la personalidad alude al enmascaramiento de la variable evaluada, lo que impide los efectos de la simulación de la respuesta por parte de los candidatos, que pudieran sesgar los resultados obtenidos. Por otro lado, al mantener las contingencias abiertas, ninguna de las pruebas pre-determinan, por las características del contexto, cuáles son las estrategias más eficaces para su resolución. Ello se comprueba con el bajo valor del coeficiente de correlación entre el rendimiento en la tarea y la puntuación en el indicador del estilo interactivo.

A efectos de completar el proceso de validación de estas pruebas quedaría pendiente la comprobación de su validez predictiva. La dimensión de minuciosidad, según el modelo de cinco factores, ha mostrado su capacidad predictiva sobre el rendimiento laboral (Barrick y Mount, 1991; Salgado, 1997) y sobre el rendimiento académico (Cucina y Vasilopoulos, 2005; Dollinger y Orf, 1991; Musgrave-Marquart, Bromley y Dalley, 1997; Wolfe y Johnson, 1995). Como hemos indicado anteriormente, los resultados obtenidos son prometedores y encauzan el trabajo de investigación posterior encaminado a la culminación del proceso de validación externa de nuestras pruebas en la línea apuntada por Elosúa (2003). En este sentido, la constatación de la validez predictiva de nuestras pruebas sobre el rendimiento académico supondría un refrendo en su proceso de validación.

#### Referencias

- Barrick, M.R., y Mount, M.K. (1991). The Big-Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1-26.
- Block, J. (1993). Studying personality the long way. En D.C. Funder, R.D. Parke, C. Tomlinson-Keasey y K. Widaman (Eds.): *Studying lives through time*. Washington DC: American Psychological Association.
- Caprara, G.V., Barbaranelli, C., y Borgogni, L. (1993). *BFQ: Big Five Questionnaire Manuale*. Firenze: Organizzazioni Speciali.
- Cattell, R.B. (1965). *The scientific analysis of personality*. Harmondsworth: Penguin Books.
- Cattell, R.B., y Kline, P. (1977). *The scientific analysis of personality and motivation*. London: Academic Press.
- Cattell, R.B., y Warburton, E.W. (1967). *Objective personality and motivation tests*. Urbana, Champaign, III: University of Illinois Press.
- Cook, T.D. (1985). Postpositivism critical multiplism. En L. Shotland y M.V. Mark (Eds.): *Social Science and Social Policy*. Palo Alto, Ca.: Sage.
- Costa, P.T., y McCrae, R.R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL, Psychological Assessment Resources.
- Costa, P.T., y McCrae, R.R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of Personality Assessment*, 64, 21-50.
- Cucina, J.M., y Vasilopoulos, N.L. (2005). Nonlinear Personality-Performance relationships and the spurious moderating effects of traitedness. *Journal of Personality*, 73, 227-259.
- Denzin, N.K. (1978). *The research act: A theoretical introduction to sociological methods 2ª ed.* New York: McGraw Hill.
- Digman, J.M. (1990). Personality structure: Emergence of the five factor model. *Annual Review of Psychology*, 41, 417-440.
- Dollinger, S.J., y Orf, L.A. (1991). Personality and performance in «personality». Conscientiousness and openness. *Journal of Research in Personality*, 25, 267-284.
- Elosúa, P. (2003). Sobre la validez de los tests. *Psicothema*, 15, 315-321.
- Fernández-Ballesteros, R. (1994). Características básicas de la evaluación conductual. En R. Fernández-Ballesteros (Ed.): *Evaluación conductual hoy*. Madrid: Pirámide.
- Goldberg, L.R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4, 26-42.
- Hernández, J.M. (2000). *La personalidad: elementos para su estudio*. Madrid: Biblioteca Nueva.
- Hernández, J.M., García-Leal, O., Rubio, V., y Santacreu, J. (2004). La persistencia en el estudio conductual de la personalidad. *Psicothema*, 16, 39-44.
- Hernández, J.M., Rubio, V.J., Revuelta, J., y Santacreu, J. (2006). A procedure for estimating intrasubject behavior consistency. *Educational and Psychological Measurement*, 66, 417-434.
- Hernández, J.M., Sánchez-Balmisa, C., Madrid, B., y Santacreu, J. (2003). La evaluación objetiva de la minuciosidad: diseño de una prueba conductual. *Análisis y Modificación de Conducta*, 29, 455-477.
- Hernández, J.M., Santacreu, J., y Rubio, V. (1999). Evaluación de la personalidad: una alternativa teórico-metodológica. *Escritos de Psicología*, 3, 30-28.
- Musgrave-Marquart, D., Bromley, S.P., y Dalley, M.B. (1997). Personality, academic attribution and substance use as predictors of academic

- achievement in college students. *Journal of Social Behavior and Personality*, 12, 501-511.
- Rudas, T., Clogg, C.C., y Lindsay, B.G. (1994). A new index of fit based on mixture methods for the analysis of contingency tables. *Journal of the Royal Statistical Society, Series B*, 56, 623-639.
- Salgado, J. (1997). The five factor model of personality and job performance in the european community. *Journal of Applied Psychology*, 82, 30-43.
- Sánchez-Balmisa, C., Hernández, J.M., Madrid, B., Peña, D., y Santacreu, J. (2003). *Evaluación Multimétodo de la Minuciosidad*. Actas del II Congreso Nacional de Psicología de la S.E.I.D.I. Barcelona.
- Santacreu, J., Hernández, J.M., Adarraga, P., y Márquez, M.O. (2002). *La personalidad en el marco de una teoría del comportamiento humano*. Madrid: Ediciones Pirámide.
- Santacreu, J., Rubio, V.J., y Hernández, J.M. (2004). Evaluación objetiva de la personalidad: una alternativa a los cuestionarios. *Análisis y Modificación de Conducta*, 30, 803-825.
- Santacreu, J., Rubio, V.J., y Hernández, J.M. (2006). The objective assessment of personality: Cattell's T-data revisited and more. *Psychology Science*, 48, 53-68.
- Skinner, N.S.F., y Howarth, E. (1975). Cross-media independence of questionnaire and objective test personality factors. *Multivariate Behavioral Research*, 8, 23-40.
- Stieglitz, R-D. (2003). Multimodal assessment (including triangulation). En R. Fernández-Basllesteros (Ed.): *Encyclopedia of Psychological Assessment*. London: Sage.
- Wolfe, R.N., y Johnson, S.D. (1995). Personality as predictor of college performance. *Educational and Psychological Measurement*, 55, 177-185.