
¿Quién se comió mi prueba?

Falsos positivos y el caso del 16FP de Cattell

César Andrés Caamaño, Virginia Vallejo & Raisia Evangeline Caamaño
Psykhê Centro de Investigaciones e Intervención Sistémica

Estudio Presentado en el 6to Congreso Dominicano-Puertorriqueño de Psicología,
en Santo Domingo, República Dominicana, Octubre de 2007

El 16FP de Cattell es, probablemente, la herramienta de evaluación psicológica más utilizada en la República Dominicana. Como toda prueba psicológica, debe poseer adecuados niveles de validez y confiabilidad para evitar que la información que se brinda sobre una persona lleve a decisiones equivocadas que le perjudiquen su salud mental, física y social (APA, 2002; Bryant, 2000; Cohen & Swerdlik, 1988/2001; Strube, 2000).

El presente estudio explora estas dos características en tres poblaciones distintas (N = 164, 71 y 75 para el 2003, 2004 y 2005, respectivamente). El análisis factorial arrojó 41 componentes para la población del 2003, 35 para la población del 2004 y 37 para la del 2005. En ninguno de los casos se logró convergencia en la rotación de estos componentes. La confiabilidad alpha de cada factor estuvo entre -.02 y .39 para el 2003 (Mdn = .13), entre -.24 y .63 para el 2004 (Mdn = .20) y entre -.21 y .45 para el 2005 (Mdn = .18).

Los resultados encontrados demuestran que las versiones del 16FP evaluadas están lejos de cumplir con los criterios mínimos de validez y confiabilidad, por lo cual su uso no es sólo contrario a la ética, sino que lleva a descripciones y toma de decisiones erróneas, las cuales perjudican a las personas evaluadas y a la psicología como ciencia. Los conceptos de Cattell, tan ampliamente reconocidos dentro de nuestra

comunidad, deben ser re-estudiados con fines de elaborar nuevas versiones de la prueba que sean válidas y confiables.

Palabras clave: 16FP, Cattell, validez, confiabilidad, ética, psicometría, prueba psicológica, dieciséis factores de personalidad.

En la década de 1940 se entendía que todos los aspectos de la personalidad humana que tienen o han tenido alguna importancia, interés o utilidad han sido ya grabadas en la sustancia del lenguaje (Cattell, 1943). Debido a esto, Raymond Bernard Cattell presentó una respuesta pionera a la pregunta de cuántos términos relacionados con la personalidad quedan luego de un análisis factorial; su respuesta fue 16 (Cattell, 1946; Cohen & Swerdlik, 1998/2001). Las investigaciones de Allport y Odbert (como se cita en Cohen & Swerdlik, 1998/2001; Cattell, 1943; Torres Viñals, 1994) habían sugerido que existían más de 18,000 nombres y términos de rasgos de la personalidad en el idioma inglés, pero, que de éstos, sólo una cuarta parte (más o menos 4,500) eran rasgos de personalidad reales o palabras y términos que designaban tendencias determinantes, generalizadas y personalizadas. Cattell agregó a esta lista algunos nombres y términos de rasgos y luego hizo que jueces estimaran las diferencias entre todas las palabras (Cattell, como se cita en Cohen & Swerdlik, 1998/2001; Cattell, 1943). El resultado fue una reducción en el tamaño de la lista a 171 nombres y términos de rasgos. Posteriormente realizó varios análisis factoriales, reduciendo con esto la lista a 36 rasgos superficiales, y luego a 16 dimensiones básicas o rasgos fuente (Cattell, 1943; 1945; 1946; Cohen & Swerdlik, 1998/2001). Las investigaciones de Cattell culminaron con la publicación de una prueba llamada Cuestionario de Dieciséis Factores de Personalidad (*Sixteen*

Estamos eternamente agradecidos con la memoria del Dr. Puro M. Peralta, quien hizo posible la realización de este estudio.

Correspondencia concerniente a este artículo debe ser dirigida a César Andrés Caamaño, Psykhê Centro de Investigaciones e Intervención Sistémica, Avenida Núñez de Cáceres, Residencial Las Praderas III, Edificio 7, Apartamento 303, Santo Domingo, República Dominicana. E-mail: ccaamano@psykhe.org

Personality Factor Questionnaire; 16FP), del cual se han publicado revisiones en los años 1956, 1962, 1968 y 1993.

El 16FP es un cuestionario para medir dimensiones normales de la personalidad a través del análisis factorial. Sandín y Chorot (como se cita en Gomara Pérez, 1994) explican que “lo que hace el análisis factorial es agrupar los reactivos que tienen relación entre sí en unidades más generales. Al aplicar esta metodología a la investigación de la personalidad los factores han sido interpretados como rasgos” (sección de introducción y breve reseña histórica, ¶ 15); la conducta es una manifestación indirecta de los constructos internos que denominamos rasgos. En este modelo, los rasgos no son medibles directamente, pero las respuestas del sujeto se interpretan como signos. “Se presupone que la personalidad constituye la variable intermediaria, el nexo de unión entre el estímulo y la respuesta observable. La variable intermediaria puede considerarse como un proceso... que causa el fenómeno de la conducta observada” (Tous, como se cita en Gomara Pérez, 1994, sección de introducción y breve reseña histórica, ¶ 16). Cattell define el rasgo como una tendencia a reaccionar, relativamente de forma permanente y amplia (Sandín & Chorot, como se cita en Gomara Pérez, 1994); engloba la capacidad, el temperamento, la dinámica, los estados de ánimo y las situaciones (Kline, como se cita en Gomara Pérez, 1994). Para Cattell, la personalidad es la determinante de la conducta en una situación dada y el componente básico de la personalidad son los rasgos (Gomara Pérez, 1994).

Cuestionario 16FP

El 16FP es una herramienta que desde su introducción, hace más de 40 años, ha sido utilizado para una amplia variedad de aplicaciones, entre las que se incluyen:

- Proveer información para la guía vocacional general que ayuda a determinar para cuáles ocupaciones un individuo está mejor inclinado.
- Asistir en la selección de personal y en el desarrollo de carrera a través de la medición de cinco dimensiones primarias frecuentemente identificadas con predecir el potencial y el estilo de gerencia.
- Asistir al diagnóstico clínico, a la prognosis y a la planificación de la terapia. El 16FP

ayuda a proveer a los clínicos de una medida de rango normal de la ansiedad, el ajuste y los problemas conductuales.

- Ayudar a identificar factores de la personalidad que pueden predecir la compatibilidad marital y la satisfacción. Los resultados destacan las áreas donde existen problemas o donde los mismos son potenciales.
- Ayudar a identificar estudiantes con potenciales problemas académicos, emocionales y sociales. (Krug, 2000; Pearson Education, 2007a).

Esta herramienta requiere entre 35 y 50 minutos para que se completen 185 ítems de respuestas múltiples (26 ítems adicionales para generar el reporte de consejería de parejas). Se entiende que porque la relación entre los ítems y el rasgo medido no es obvia, se le dificulta al examinado modificar deliberadamente sus respuestas para alcanzar un desempeño deseado. La herramienta está diseñada para ser aplicada a personas de 16 o más años de edad, con un nivel de lectura de al menos el quinto grado de educación básica. Las escalas resultantes consisten en 16 dimensiones bipolares de personalidad, cinco factores globales y tres escalas de validez. Las normas utilizadas en Estados Unidos corresponden a un muestreo aleatorio estratificado, de 10,261 adultos, que refleja el Censo de ese país para el año 2000. (Pearson Education, 2007a)

El 16FP está clasificado como una herramienta de evaluación tipo 3. Esta tipología requiere que el usuario de la prueba (y el que la adquiere) tenga una licencia para practicar la psicología de manera independiente, o que el usuario haya completado un grado doctoral (en algunos casos es aceptable el grado de maestría; ver Colegio Dominicano de Psicólogos [CODOPSI], 2002) en un programa correspondiente a alguno de los campos de estudio indicados para la prueba que incluya entrenamiento (a través de docencia de salón y de experiencia práctica) en la administración e interpretación de instrumentos clínicos. En caso de que ninguna de estas calificaciones sea cumplida, el usuario debe proveer pruebas de que obtiene el derecho de administrar pruebas de este nivel dentro de su jurisdicción. (Pearson Education, 2007b)

En República Dominicana existen varias formas del 16FP, las cuales están diseñadas

considerando el nivel de formación académica de la persona que es evaluada (Evaluaciones Psicológicas Sistémicas [EPSI], 1995; Instituto de Psicología Clínica y Comportamiento Organizacional [IPCCO], 1997; Instituto de Servicios Psicosociales y Educativos [ISPE], s.f.). A pesar de las variantes, la forma más ampliamente utilizada es la denominada *forma C*, por lo cual, en este estudio, nos enfocaremos en este formato.

La forma *C* del 16FP es distribuida, en la República Dominicana, por tres instituciones distintas, las cuales, sin embargo, comercializan exactamente idénticas versiones de la prueba (EPSI, 1995; IPCCO, 1997; ISPE, s.f.). Sorprendentemente, los manuales de ISPE e IPCCO son textualmente los mismos (IPCCO, 1997; ISPE, s.f.). La versión dominicana difiere grandemente de la versión oficial, más reciente y

actualizada, del 16FP que se utiliza en los Estados Unidos, a pesar de que los manuales indican que el modelo original para su adaptación fue el 16FP desarrollado por Cattell (EPSI, 1995; IPCCO, 1997; ISPE, s.f.). La forma *C* consta de 105 ítems de respuestas múltiples (80 menos que la versión oficial estadounidense) que resultan en las 16 dimensiones bipolares de personalidad descritas por Cattell. No incluye información sobre los cinco factores globales, ni sobre las tres escalas de validez ni sobre el reporte de consejería de parejas. De esta forma evidenciamos que la diferencia observada no es sólo en cuanto a la cantidad de ítems, sino también en cuanto al conjunto de información general que se obtiene de la aplicación del instrumento. En la Tabla 1 presentamos la comparación de la terminología utilizada por ambas versiones del instrumento en los factores comunes.

Tabla 1. Comparación de la terminología utilizada en las 16 dimensiones bipolares de personalidad del 16FP, entre la versión estadounidense y la versión dominicana

Factor	Dimensiones	Polaridades de la versión estadounidense^a	Polaridades de la versión dominicana^b
A	Calidez	<i>Reserved vs. warm</i>	Reservada vs. abierta
B	Razonamiento	<i>Concrete vs. abstract</i>	Inteligencia baja vs. Inteligencia alta
C	Estabilidad emocional	<i>Reactive vs. emotionally stable</i>	Afectada por los sentimientos vs. Emocionalmente estable
E	Dominancia	<i>Deferential vs. dominant</i>	Sumisa vs. dominante
F	Vivacidad	<i>Serious vs. lively</i>	Sobria vs. Descuidada
G	Consciente de las reglas	<i>Expedient vs. rule-conscious</i>	Despreocupada vs. escrupulosa
H	Audacia social	<i>Shy vs. socially bold</i>	Cohibida vs. emprendedora
I	Sensitividad	<i>Utilitarian vs. sensitive</i>	Sensibilidad dura vs. Sensibilidad blanda
L	Vigilancia	<i>Trusting vs. vigilant</i>	Confiable vs. suspicaz
M	Abstracción	<i>Grounded vs. abstracted</i>	Práctica vs. imaginativa
N	Privado	<i>Forthright vs. private</i>	Franca vs. astuta
O	Aprensión	<i>Self-assured vs. apprehensive</i>	Apacible vs. aprensible
Q1	Apertura al cambio	<i>Traditional vs. open to change</i>	Conservadora vs. analítico-crítica
Q2	Auto dependencia	<i>Group-oriented vs. self-reliant</i>	Dependiente vs. autosuficiente
Q3	Perfeccionismo	<i>Tolerates disorder vs. perfectionistic</i>	Autoconflictiva vs. controlada
Q4	Tensión	<i>Relaxed vs. Tense</i>	Relajada vs. tensa

^a Pearson Education, 2007a. ^b EPSI, 1995; IPCCO, 1997; ISPE, s.f.

Calidad de una prueba psicológica

Existen varios criterios que definen que una prueba psicológica tenga buena calidad. Cohen y Swerdlik (1998/2001) indican que entre los criterios lógicos más comunes se encuentran: a) instrucciones claras para su aplicación, calificación e interpretación; b) economía en el tiempo que se lleva aplicarla, calificarla e interpretarla; c) que mida lo que pretenda medir; y d) que mejore la calidad de vida de quien responde la prueba.

Pero más allá que las características lógicas, una prueba psicológica debe cumplir con criterios técnicos, los cuales tienen que ver con la psicometría, o “ciencia de la medición psicológica” (Cohen & Swerdlik, 1998/2001, p. 38). Una prueba es sólida cuando cumple con dos aspectos clave: la confiabilidad y la validez.

La confiabilidad es el atributo de consistencia de una evaluación (Cohen & Swerdlik, 1998/2001). Este atributo es vital para el trabajo científico y de investigación pues cuando se utilizan instrumentos no confiables se oscurece la habilidad para encontrar las relaciones entre los eventos y los conceptos (Strube, 2000). La confiabilidad implica una libertad de error aleatorio (Anastasi, como se cita en Alexander Servat, 2003). Cuando una prueba es confiable, uno puede asegurar que funcionará de la misma forma en el futuro; que será consistente en sus características (Strube, 2000). Un coeficiente de confiabilidad es un índice de confianza; es una proporción que indica la razón entre la varianza de la puntuación verdadera en una prueba y la varianza total (Cohen & Swerdlik, 1998/2001).

Por otro lado, para que una prueba sea válida, sus preguntas o cuestionario deben evaluar lo que se quiere medir y sólo lo que se quiere medir (Magnusson, como se cita en Alexander Servat, 2003). La validez consiste en el juicio que se hace en base a la evidencia sobre la significación o la veracidad de algo; se refiere a un juicio concerniente a lo bien que una prueba mide lo que pretende medir (Cohen & Swerdlik, 1998/2001). Dentro de nuestro contexto, validez concierne a que una conclusión o inferencia particular que uno desea hacer es precisa, razonable y correcta (Bryant, 2000). Esto es de vital importancia a la hora de establecer relaciones entre eventos y conceptos pues hasta que un conjunto amplio de evidencia soporten

una conclusión particular, uno debe asumir que dicha conclusión no es válida (Bryant, 2000).

La comunidad científica define que si no se logra la confiabilidad en un instrumento, no se puede lograr validez (Alexander Servat, 2003), por ende, éste es el primer análisis que debe pasar cualquier instrumento antes de que se señale como robusto. Si el instrumento no logra alcanzar un coeficiente adecuado de confiabilidad no debe ser utilizado (American Psychological Association [APA], 2002).

Cuando un psicólogo utiliza pruebas de las cuales desconoce sus niveles de validez y de confiabilidad está violando todos los preceptos éticos y profesionales de la ciencia (APA, 2002; Colegio Dominicano de Psicólogos [CODOPSI], 2002). El propio *Código de Ética y Disciplina* del CODOPSI (2002) establece, respecto a las pruebas psicológicas, que “se deben incluir todos los datos relativos a su tipificación, validez, confiabilidad y formas de uso” (p. 10, artículo 34). La violación a este artículo constituye una falta grave de acuerdo al Código de Ética y Disciplina que rige el ejercicio profesional de la psicología en nuestro país (CODOPSI, 2002). Veámoslo de este modo, el profesional de la psicología tiene varias obligaciones éticas, que, a manera de principios, se pueden englobar en las siguientes:

- a) Todo psicólogo está obligado a proveer un beneficio para quienes trabaja y debe cuidar de no hacerles daño.
- b) Todo psicólogo está consciente de sus responsabilidades profesionales y científicas para con la sociedad y para con las comunidades específicas en las cuales trabaja.
- c) Todo psicólogo busca promover exactitud, honestidad y sinceridad en la ciencia, la enseñanza y la práctica de la psicología.
- d) Todo psicólogo reconoce que la imparcialidad y la justicia corresponden a todas las personas para acceder y beneficiarse de las contribuciones de la psicología y de la misma calidad en los procesos, procedimientos y servicios que se ofrecen por los psicólogos.
- e) Todo psicólogo respeta la dignidad y el valor de todas las personas, así como el derecho de los individuos a la privacidad, confidencialidad y auto-determinación. (APA, 2002)

Más específicamente, existen estándares éticos relativos a la administración, adaptación, corrección e interpretación de las evaluaciones psicológicas. El profesional de la psicología sólo debe utilizar instrumentos para los cuales su validez y confiabilidad han sido establecidas para ser utilizados con miembros de la población evaluada (ver APA, 2002, acápite 9.02). Similarmente, el profesional de la psicología no debe basar sus evaluaciones, decisiones de intervención o recomendaciones, en datos o resultados de pruebas y medidas que están obsoletos para el propósito que se persigue. Tampoco debe basar estas decisiones o recomendaciones en pruebas o medidas que están obsoletas y no adecuadas para el propósito que se persigue (ver acápite 9.08).

El Código de Ética y Disciplina del CODOPSI es un documento ligero y general, que no abunda en las profundidades de los principios que deben regir un comportamiento ético. El mismo es una adaptación del Código que estableció la Asociación Dominicana de Psicología en el año 1980 y no comprende muchas de las situaciones conflictivas en las que se puede ver involucrado el psicólogo de hoy día. Es por eso que el código de ética que establece la APA representa una mejor herramienta para ayudarnos a definir y modelar la conducta ética del psicólogo dominicano. Seguro hay quienes alegan que estos preceptos éticos corresponden solamente a los profesionales que son miembros de la APA, sin embargo disentimos de esto, pues por su profundidad, amplitud y exhaustividad en comprender múltiples situaciones, procurando siempre el bienestar de las personas que utilizan los servicios de los psicólogos y a los profesionales de la psicología por igual, brindan un marco realista que permite discriminar qué es ético y qué no. Tenemos que comprender que nuestro mundo es globalizado y que la práctica profesional en nuestro país debe alcanzar el mismo nivel ético y productivo de nuestros pares internacionales.

Confiabilidad y validez del 16FP forma C

De los tres manuales de la forma C del 16FP, el de ISPE (s.f.), el de IPCCO (1997) y el de EPSI (1995), sólo el último presenta indicaciones sobre la validez y confiabilidad de la herramienta (ver APA, 2002; CODOPSI, 2002). En el caso de la confiabilidad, el manual de EPSI indica que, de manera general, “las formas del 16FP... indican siempre valores superiores a 0.60 [sic] en todos

los casos” (p. 10). Sin embargo, no presenta indicación directa de la confiabilidad de cada uno de los factores en las distintas formas evaluadas, de modo que no podemos establecer de manera específica cuál es la confiabilidad de cada factor.

Lo mismo pasa en el caso de la validez, en la que indica que la validez concurrente de los factores oscila entre .65 y .85 y que “en algunos casos se obtuvieron coeficientes de .45, pero [que] estos son los menos” (EPSI, 1995, p. 10). No se hace evidente si, como plantea Cattell (1943; 1944; 1945; 1946), en el proceso de adaptación y validación se obtuvieron los 16 factores a través del análisis factorial y cuál fue la validez convergente y discriminante de cada factor.

Independientemente de los indicadores de confiabilidad y validez antes mencionados, nos llama la atención otro tipo de consideraciones. El propósito del 16FP es comparar el desempeño de quien responde con el de otros que lo han respondido, de modo que para que sea una buena prueba sus datos normativos deben ser adecuados; este tipo de pruebas “requieren de una descripción completa de la norma de grupo o muestra normativa” (Cohen & Swerdlik, 1998/2001, p. 40). Las normas provistas por EPSI (1995), ISPE (s.f.) e IPCCO (1997), extrañamente idénticas considerando que son manuales de casas distintas, indican que para varones se utilizó una muestra de 2,162 personas y que para hembras una muestra de 127 (con una edad media de 25 años, en ambos casos). No indican más información relativa respecto a de dónde provienen los sujetos, si toda la población dominicana está correctamente representada, si hay diferencias por clase social, nivel socio-económico y nivel educativo, entre otros. Más aun, estos documentos datan de los años 1995 y 1997, de modo que podemos considerar estas normas obsoletas para los tiempos que nos corren.

Método

Propósito del estudio

El propósito de este estudio era determinar si el 16FP cuenta con suficientes niveles de validez y de confiabilidad para ser utilizado en la población dominicana. Para determinar la validez de constructo se decidió realizar un análisis factorial con el cual se esperaba identificar los 16 factores descritos por Cattell, utilizando el método de

extracción de análisis de componentes principales. Para determinar la confiabilidad se realizó la estimación del coeficiente de confiabilidad alfa de Cronbach.

Población

Las pruebas de confiabilidad y validez se realizaron para tres poblaciones distintas. La primera población, constituida por 164 personas, estuvo conformada por los aspirantes a cadetes, para el año 2003, en la Academia Militar Batalla de las Carreras (AMBC); su edad media fue de 18.8 años ($DE = 1.11$ años) y el 93.3% era de sexo masculino. La segunda población estuvo también constituida por aspirantes a la AMBC ($N = 71$), pero correspondían al año 2004; su edad media fue de 18.6 años ($DE = 1.12$ años) y el 87.3% era de sexo masculino. La tercera población estuvo constituida por los aspirantes a encuestadores para la Oficina Nacional de Estadística (ONE; $N = 75$), quienes trabajarían en el año 2005 con la Encuesta Nacional de Hogares de Propósitos Múltiples (ENHOGAR-2005); su edad media fue de 35.2 años ($DE = 9.47$ años) y el 46.7% era de sexo masculino. Para los tres grupos se utilizó la forma C del 16FP como herramienta de selección de los candidatos.

Resultados

Al analizar los datos en busca de evidencia que sustente la validez de constructo de la forma C del 16FP, encontramos que en ninguna de las poblaciones se replicó una composición factorial que correspondiera a los 16 factores descritos por Cattell (EPSI, 1995; IPCCO, 1997; ISPE, s.f.). Para los aspirantes a cadetes del año 2003 el análisis factorial arrojó 41 componentes que explican el 74.9% de la varianza. Los primeros 16 factores para esta población explicaban sólo el 41.1% de la varianza. Para los aspirantes a cadetes del año 2004 el análisis factorial arrojó 35 factores, los cuales explicaban el 85.2% de la varianza; los primeros 16 factores explicaban sólo el 56.0% de la varianza. Para los aspirantes a encuestadores de la ONE el análisis factorial arrojó 37 factores, los cuales explicaban el 85.6% de la varianza; los primeros 16 factores constituían sólo el 54.4 % de la varianza explicada. En ninguno de los casos se logró convergencia en la rotación de estos componentes.

La confiabilidad alpha de Cronbach de cada factor se presenta en la Tabla 2. Para el año

2003 el coeficiente estuvo entre $-.02$ y $.39$ ($Mdn = .13$), para el año 2004 entre $-.24$ y $.63$ ($Mdn = .20$) y para el año 2005 entre $-.21$ y $.45$ ($Mdn = .18$). En ninguno de los casos alcanzó un nivel aceptable (i.e., un índice cercano de 1.00; ver Cronbach, 1964).

Tabla 2. Coeficientes de confiabilidad alpha de Cronbach por factor para cada población con la forma C del 16FP

Factor	Población		
	2003 ($N = 164$)	2004 ($N = 71$)	2005 ($N = 75$)
DM	.39	.20	.40
A	.02	.18	.09
B	.27	.06	.24
C	.37	.63	.41
E	.08	.14	.01
F	.09	-.10	.10
G	.19	-.19	.08
H	.13	.26	.45
I	.12	.24	.31
L	.21	.17	-.10
M	.04	.04	-.21
N	.08	.24	-.18
O	.37	.35	.17
Q1	.00	-.24	.18
Q2	.25	.33	.18
Q3	.21	.37	.43
Q4	-.02	.21	.22

DM = medida de distorsión intencional de la prueba

Discusión

Concordando con la impresión generalizada de algunos profesionales de la psicología, así como de otras áreas profesionales, la forma C del 16FP no presenta evidencia de validez y de confiabilidad. Los resultados alcanzados en las estimaciones del coeficiente de confiabilidad son inaceptables, comprometiendo así toda la utilidad del instrumento.

Al obtenerse niveles tan bajos en la confiabilidad puede esperarse que si se aplica el instrumento dos veces a una misma persona, éste

presente dos perfiles de personalidad totalmente distintos, sólo por el hecho de la inestabilidad interna de los rasgos. Esto representa consecuencias altamente negativas a la hora de evaluar a una persona, pues los resultados no serán consistentes con la conducta que se evidencia de ella ni con evaluaciones futuras, y por lo tanto, no se puede confiar en dichos resultados.

Cronbach (1964) ya había criticado estos bajos índices de confiabilidad de la primera versión del 16FP, señalando que las escalas tenían niveles extremadamente bajos de confiabilidad (.45 – .55) y que no recomendaba el instrumento para la evaluación de individuos. Más recientemente, Domínguez, Nin y Bello (2006) encontraron que es inaceptable la confiabilidad del Cuestionario de Actitudes Personales FPACD, un instrumento basado en los ítems del 16FP y utilizado por la Secretaría de Estado de Interior y Policía de la República Dominicana para determinar quién puede obtener un permiso de porte de arma de fuego.

Pensamos que dos alternativas pueden explicar estos bajos niveles de confiabilidad: a) que los ítems estén mal agrupados dentro de sus factores y que deben re-agruparse en otra configuración para funcionar adecuadamente; o b) que los ítems no reflejan los rasgos de personalidad que se han planteado en la estructura del 16FP y que deben cambiarse por completo. La primera alternativa, sobre la configuración de los ítems, es rechazada a partir de los resultados obtenidos por el análisis factorial. Cuando realizamos este análisis colocamos todos los ítems juntos sin identificar a cuál factor creemos nosotros que cargan, dejando que el análisis nos indique la configuración óptima de todos los ítems. Los resultados nos mostraron que las configuraciones posibles no se corresponden a la descrita por Cattell, por lo que la explicación de 16 factores subyacentes no se logra a partir de los ítems utilizados en la actualidad por la prueba. De hecho, los resultados indicaron que el conjunto de ítems bien podrían estar midiendo mucho más de 16 aspectos, lo cual introduce niveles de error al describir la personalidad.

En este momento nos cuestionamos si los resultados obtenidos, que señalan que las propiedades psicométricas de la forma *C* del 16FP son inaceptables, pudieran ser producto de haber utilizado el instrumento con una población

distinta a la cual se supone que debe utilizarse. No creemos que esto explique los resultados debido a que el resultado factorial fue similar en tres poblaciones distintas y en tres tiempos distintos, además de que el propósito fue congruente con las justificaciones explicadas en los manuales de aplicación: dos de las poblaciones estaban compuestas por candidatos a puestos académicos, mientras que la tercera estaba constituida por candidatos a puestos laborales.

La revisión de la literatura soporta también nuestros hallazgos, específicamente si nos basamos en seis aspectos importantes que cuestionan la calidad de la prueba: a) el nivel académico que se exige para su utilización; b) la obsolescencia de la edición de la forma *C* del 16FP; c) la discrepancia entre la versión vigente en inglés y la versión dominicana; d) la escasez, confusión y contradicciones en las informaciones para su aplicación, calificación e interpretación; e) la carencia de información sobre su validez y confiabilidad; y f) la carencia de información que describa, completamente, la norma de grupo o muestra normativa.

Primero, ha sido sorprendente la discrepancia que hemos encontrado, en cuanto al nivel académico exigido para utilizar el 16FP, entre la realidad estadounidense y la práctica dominicana. La revisión de la literatura y las experiencias de los profesionales indican que en la República Dominicana se le da un uso liberal a los instrumentos de evaluación psicológica, mientras que en los Estados Unidos se exige un nivel de doctorado y entrenamiento específico (Pearson Education, 2007b); en nuestro país se permite el uso de este instrumento a psicólogos (CODOPSI, 2002), sin que quede definido, claramente, qué nivel define a un profesional como psicólogo (e.g., licenciatura, maestría o doctorado). En la práctica, sabemos de casos donde quienes aplican e interpretan esta herramienta son estudiantes de psicología, estudiantes de psicometría y personal secretarial de las áreas de recursos humanos, y desconocemos la existencia de algún tipo de registro que lleven las casas comercializadoras sobre si quienes adquieren estas herramientas cuentan con las calificaciones mínimas necesarias. Tampoco conocemos de esfuerzos específicos y sistemáticos que realicen estas casas comercializadoras por impartir entrenamientos profundos en la utilización de esta herramienta.

Esta falta de control es grave. Si no puede garantizarse que quien utiliza la prueba tiene

todos los requerimientos para su correcto uso, no puede controlarse el error que se le atribuye al entrevistador o a quien la aplica. Esto significa que no puede garantizarse la calidad de los resultados obtenidos.

Por otro lado, es indudable que, al comparar la herramienta que se utiliza en los Estados Unidos y la que se utiliza en República Dominicana, estamos frente a dos instrumentos distintos, aunque utilicen el mismo nombre. La versión dominicana tiene una menor capacidad de descripción, ofreciendo sólo parte de los reportes que genera el 16FP en su versión más reciente y centrándose sólo en las dimensiones bipolares de la personalidad. Los factores globales, las escalas de validez y el reporte de consejería de parejas ofrecen un espectro mayor que, indudablemente, permiten al psicólogo una mayor intervención.

La versión utilizada en República Dominicana está compuesta por 80 ítems menos. La justificación evidente para esta versión reducida es la brevedad en la aplicación (EPSI, 1995), sin embargo, no compartimos la idea de afectar la confiabilidad y la validez en pos de alcanzar una mayor brevedad en la aplicación de un instrumento de esta relevancia. La economía en el tiempo de aplicación que se logra con 80 ítems es, aproximadamente, de 20 minutos o menos, pero para un instrumento que impacta la vida profesional, laboral y marital de las personas, creemos que es un ahorro que no tiene sentido.

Entendemos que la forma *C* del 16FP se adaptó a partir de los estudios realizados por Cattell, pero no entendemos porqué no se ha seguido adaptando a partir de las versiones más recientes y que han demostrado una mayor utilidad para la ciencia y la práctica de la psicología. Una explicación tentativa podría ser la inseguridad jurídica que caracteriza nuestro país, especialmente en el tema de derecho de autor, la cual impide que las casas que desarrollan instrumentos psicológicos alcancen un retorno de su inversión. Aun así, entendemos que es responsabilidad de estas casas el encontrar medios y tecnologías que les aseguren que sus instrumentos no sean plagiados y que mantengan una inversión constante en mantener actualizados sus instrumentos.

Cohen y Swerdlik (1988/2001) establecieron los criterios lógicos para que una prueba psicológica sea considerada como buena. La claridad de las instrucciones para su aplicación, calificación e interpretación es el

primero de ellos. La versión de la forma *C* del 16FP que se utiliza en nuestro país no cumple este criterio. Las informaciones que proveen los manuales existentes son confusas, contradictorias y con detalles muy escasos sobre la interpretación de los resultados. Todas listan unos breves párrafos que describen las dimensiones bipolares de personalidad, pero que distan mucho de ser una descripción profunda de un rasgo de personalidad. Todas citan otros documentos para que el aplicador de las pruebas pueda profundizar sus conocimientos o aclarar dudas, pero las referencias de los mismos son incompletas y, en dos de los manuales, inexistentes. Hemos visto, en demasiadas ocasiones, que los informes se convierten en transcripciones textuales de los párrafos que describen cada rasgo, pero que poco profundizan en explicar las implicaciones de la estructura personal de dichos rasgos. Entendemos que el criterio del psicólogo es fundamental a la hora de elegir herramientas de calidad y que no debe conformarse con una breve explicación de un rasgo profesional, y cuando esto ocurre, debe abstenerse de utilizar dichos instrumentos.

Además, los manuales existentes no ofrecen información específica sobre las propiedades de validez y confiabilidad con que cuenta la forma *C* del 16FP. La utilización de este instrumento, como sucede con muchos otros en similares circunstancias, constituye una violación a los preceptos éticos, los cuales deben ser defendidos no sólo por las autoridades competentes, sino también por la sanción moral y social que ejercemos cada uno de los profesionales de esta ciencia.

Esta situación absurda, en el sentido de que existe un acuerdo unánime de la necesidad de contar con estos indicadores de calidad, no debería ni siquiera aparecer en nuestro análisis. Sin embargo, aquí estamos, observando que los psicólogos no estamos jugando el papel como evaluadores éticos de las herramientas que se nos ofrecen. Estamos comprando pruebas, haciendo que nuestros estudiantes las compren, vendiéndole la idea a los empresarios de que podemos ayudarles en sus decisiones, ofreciéndole una guía a las parejas que buscan nuestra ayuda profesional... sin embargo, estamos, conscientemente, utilizando instrumentos de los cuales no sabemos nada. Debemos detenernos en este instante y observar lo que hemos estado haciendo. Utilizar pruebas sin saber si sirven o no es similar a ofrecer un

medicamento sin saber cuáles son sus efectos en el ser humano.

Otro aspecto problemático, que impide establecer que la forma *C* del 16FP tiene las condiciones lógicas mínimas para ser utilizada en la población dominicana, es que data del año 1995. A pesar de que ha pasado más de una década, el instrumento que se comercializa sigue siendo el mismo. Es obvio que una población cambia mucho en 12 años. Precisamente, con lo rápido que los avances tecnológicos nos llevan, la forma como veíamos y explicábamos el mundo cambia radicalmente en mucho menos tiempo. Pretender que los reactivos que funcionaron hace más de una década funcionen adecuadamente para describir la personalidad del individuo de hoy es una utopía. Simplemente debemos observar que la versión en inglés va ya por la 5ta edición, lo cual supone una revisión de sus reactivos; nosotros, por el contrario, seguimos utilizando la primera versión de la adaptación, que aunque hubiera sido perfecta en la década del 90, simplemente el tiempo la convierte en obsoleta.

Más aún, al tratarse de instrumentos que se desarrollaron hace 12 años, es de esperarse que los datos de referencia no sean aplicables a la sociedad actual. No podemos evaluar a la generación actual con los patrones de generaciones anteriores. El cambio social que vivimos hace que también, los datos normativos, sean obsoletos.

Los manuales disponibles presentan información insuficiente para comprender las características de la población que se utilizó para crear las normas. Muestran, solamente, la cantidad de personas, las proporciones por sexo y la edad media, mientras que no sabemos nada acerca de otras variables que nos ayudan a poner los resultados en contexto. No podemos describir los rasgos de un individuo si no sabemos con quién lo estamos comparando. Un sujeto que puede parecer extrovertido en una comunidad, se nos presenta introvertido al compararlo con otra; es evidente, si no contamos con información suficiente no podemos realizar este tipo de inferencias.

Insistimos, la forma *C* del 16FP no es válida ni confiable, por lo que debe discontinuarse su utilización hasta que un conjunto amplio de evidencia soporte que el mismo cuenta con adecuados criterios de validez y confiabilidad. Las repercusiones que tiene la utilización de

herramientas de mala calidad en la vida de las personas es incalculablemente negativa, debido a que se toman decisiones que inciden en el futuro profesional, laboral, familiar y social de los individuos. No es posible mejorar la calidad de vida de las personas evaluadas (otro criterio de la calidad lógica de las pruebas) si el informe que resulta de su aplicación está errado.

Los psicólogos y psicólogas nos hemos comprometido a no hacer daño, a promover la exactitud, la honestidad y la sinceridad en cada uno de nuestros actos, en nuestra ciencia, en la enseñanza y en la práctica de la psicología, de modo que también es nuestra responsabilidad el asegurar que, en nombre de nuestra profesión, no ocurran tales actos. No podemos ser menos insistentes en que no se pueden utilizar instrumentos para los cuales no existe la debida evidencia sobre su validez y sobre su confiabilidad. No podemos insistir menos en que no se pueden utilizar datos, pruebas o medidas que están obsoletas. Cuando hacemos uso de estos instrumentos hacemos un daño irreparable en la vida de las personas que depositan su confianza en nosotros.

Aquí se requiere de una posición drástica respecto a la comercialización de las pruebas psicométricas en nuestro país. La responsabilidad no debe dejarse solamente sobre los profesionales de la psicología, quienes, indiscutiblemente, tenemos el deber de vigilar la correcta utilización de los conocimientos psicológicos. La responsabilidad de supervisar y monitorear la conducta ética recae en el CODOPSI, que debe establecer mecanismos de regulación para la comercialización de estos instrumentos y, obviamente, debe establecer niveles de sanción que le permitan cumplir con esta función. Ya desde 1964 Cronbach indicó que es responsabilidad del autor de cualquier prueba psicológica el demostrar que la misma tiene suficiente precisión para el uso que ese autor recomienda. De manera específica, entendemos pertinente puntualizar las siguientes recomendaciones para toda la clase profesional:

1. Actualizar periódicamente los instrumentos, confirmando sus propiedades psicométricas.
2. Asegurar que los instrumentos cuenten con instrucciones claras y precisas.
3. Asegurar que las informaciones sobre las muestras normativas se mantengan

actualizadas y con suficiente descripción que permita asegurar que la persona evaluada corresponde a esa muestra normativa.

4. Establecer mecanismos de supervisión, control y sanción para que el CODOPSI pueda vigilar el cumplimiento de los estándares éticos, especialmente en lo que se refiere a la evaluación psicológica.
5. Establecer mecanismos para controlar la distribución y uso del material psicométrico.

El CODOPSI tiene un importante rol que jugar en asegurar que las acciones y decisiones tomadas por psicólogos, y a partir de la psicología, no causen daño, mal información, confusión ni perjuicio a las personas. Como hemos dicho antes, es un deber ético utilizar, únicamente, instrumentos para los cuales su validez, confiabilidad y tipificación han sido establecidas para ser utilizados con miembros de la población dominicana (APA, 2002; CODOPSI, 2002) y el no cumplimiento de esto constituye una falta grave, tanto de parte de los que comercializan estos instrumentos como de quienes los utilizan.

Referencias

- Alexander Servat, A. G. (2003). *Implantación estratégica del ISO 9000 versión 2000*. Lima, Perú: Pontificia Universidad Católica del Perú.
- American Psychological Association. (2002). Ethical principles of psychologist and code of conduct. *American Psychologist*, 57(12), 1060-1073.
- Bryant, F. B. (2000). Assessing the validity of measurement. En L. G. Grimm & P. R. Yarnold. *Reading and understanding more multivariate statistics*. Washington, DC, USA: American Psychological Association.
- Cattell, R. B. (1943). The description of personality: basic traits resolved into clusters [versión electrónica]. *The Journal of Abnormal and Social Psychology*, 38(4), 476-506. Recuperado el 26 de septiembre de 2007, de PsycINFO.
- Cattell, R. B. (1944). Interpretation of the twelve primary personality factors [versión electrónica]. *Character & Personality: A Quarterly for Psychodiagnostic & Allied Studies*, 13, 55-91. Recuperado el 26 de septiembre de 2007, de PsycINFO.
- Cattell, R. B. (1945). The description of personality: principles and findings in a factor analysis [versión electrónica]. *American Journal of Psychology*, 58, 69-90. Recuperado el 26 de septiembre de 2007, de PsycINFO.
- Cattell, R. B. (1946). Personality structure and measurement. I. The operational determination of trait unities [versión electrónica]. *British Journal of Psychology*, 36, 88-103. Recuperado el 26 de septiembre de 2007, de PsycINFO.
- Cohen, R. J. & Swerdlik, M. E. (2001). *Pruebas y evaluaciones psicológicas: introducción a las pruebas y a la medición* (J. A. Velásquez, Trad.). México: McGraw-Hill. (Trabajo original publicado en 1988).
- Colegio Dominicano de Psicólogos. (2002). *Código de ética y disciplina* [versión electrónica]. Santo Domingo, República Dominicana: Autor. Recuperado el 20 de septiembre de 2007, de http://idpp.org/idpp_pubs/codopsi/codetica.pdf
- Cronbach, L. J. (1964). A word on 16 PF and reliability: Comment [versión electrónica]. *American Psychologist*, 19(6), 417.
- Domínguez, P., Nin, N., & Bello, I. E. (2006). *Confiabilidad y validez de constructo del Cuestionario de Actitudes Personales FPACD*. Tesis de licenciatura no publicada, Universidad Nacional Pedro Henríquez Ureña, Santo Domingo, República Dominicana.
- Evaluaciones Psicológicas Sistémicas. (1995). *Manual del test 16FP: 16 factores de personalidad*. Santo Domingo, República Dominicana: Autor.
- Gomara Pérez, J. (1994). R. B. Cattell y el 16 PF: Aproximación histórica, metodológica y conceptual. Recuperado el 26 de junio de 2007, de: <http://www.terra.es/personal5/itaka2002/documents/cattell16pf.htm>
- Instituto de Psicología Clínica y Comportamiento Organizacional. (1997). *Test 16FP: manual de aplicación y corrección*. Santo Domingo, República Dominicana: Autor.
- Instituto de Servicios Psicosociales y Educativos. (s.f.). *Manual de aplicación y corrección del test 16FP (C)*. Santo Domingo, República Dominicana: Autor.
- Krug, S. E. (2000). Sixteen Personality Factor Questionnaire. En Kazdin, Alan E. (Ed), *Encyclopedia of psychology, Vol. 7*. (pp. 292-294). Washington, DC, US: American

Psychological Association; Oxford University Press.

Pearson Education. (2007a). 16PF Fifth Edition. Recuperado el 26 de septiembre de 2007, de http://www.pearsonassessments.com/tests/sixtpf_5.htm

Pearson Education. (2007b). Effective January 1, 2007 - Qualification Levels and Requirements. Recuperado el 26 de septiembre de 2007, de

<http://www.pearsonassessments.com/forms/levels.htm>

Strube, M. J. (2000). Reliability and generalizability theory. En L. G. Grimm & P. R. Yarnold. *Reading and understanding more multivariate statistics*. Washington, DC, USA: American Psychological Association.

Torres Viñals, M. (1994). *Instrumentos usuales en la evaluación clínica de adultos*. Barcelona: autora.