

El Modelo Alternativo de los Cinco Grandes: estudios de fiabilidad y validez del Zuckerman-Kuhlman Personality Questionnaire (ZKPQ) en población general y clínica

Sergi Valero i Ventura

Servei de Psiquiatria
Hospital Universitari Vall d'Hebron, Barcelona

Tesis doctoral dirigida por la Dra. Montserrat Gomà i Freixanet

Departament de Psiquiatria i Medicina Legal
Facultat de Medicina
Universitat Autònoma de Barcelona
2009





Índice

	Pág.
1. Sobre la estructura de este documento	4
2. Referencias de las publicaciones	5
3. Introducción	6
3.1. El Modelo de los Cinco Grandes	7
3.2. El Modelo Alternativo de los Cinco Grandes	8
4. Estudios de fiabilidad y validez del ZKPQ	11
4.1. Sobre la fiabilidad	11
4.1.1. Fiabilidad como consistencia interna	14
4.1.2. Fiabilidad como estabilidad de la medida	15
4.2. Sobre la validez	15
4.2.1. Validez de contenido	16
4.2.2. Validez de constructo	17
4.2.2.1. Análisis factorial	18
4.2.2.2. Matriz multirasgo-multimétodo	19
4.2.3. Validez discriminante	20
4.3. Baremos	22
5. Discusión	24
6. Consideraciones finales y estudios futuros	35
7. Referencias	36
8. Apéndices	42
I. <i>Art-1</i> : Gomà-i-Freixanet, M., Valero, S., Puntí, J. & Zuckerman, M. (2004). Psychometric properties of the Zuckerman-Kuhlman Personality Questionnaire in a Spanish sample. <i>European Journal of Psychological Assessment, 20</i> (2), 134-146.	43
II. <i>Art-2</i> : Gomà-i-Freixanet, M., Wismeijer, A. A. J. & Valero, S. (2005). Consensual validity parameters of the Zuckerman-Kuhlman Personality Questionnaire: Evidence from self-reports and spouse reports. <i>Journal of Personality Assessment, 84</i> (3), 279-286.	57
III. <i>Art-3</i> : Gomà-i-Freixanet M, Soler, J., Valero, S., Pascual, J. C. & Pérez, V. (2008). Discriminant validity of the ZKPQ in a sample meeting BPD diagnosis vs. normal-range controls. <i>Journal of Personality Disorders, 22</i> (2), 178-190.	65
IV. <i>Art-4</i> : Gomà-i-Freixanet, M. & Valero, S. (2008). Spanish normative data of the Zuckerman-Kuhlman Personality Questionnaire in a general population sample. <i>Psicothema, 20</i> (2), 324-330.	79
V. <i>Art-adjunto</i> : Gomà-i-Freixanet, M., Valero, S., Muro, A. & Albiol, S. (2008). Zuckerman-Kuhlman Personality Questionnaire: psychometric properties in a general population sample. <i>Psychological Reports, 103</i> , 845-856.	87

1.
*Sobre la
estructura de este
documento*

El presente documento se ha estructurado entorno a dos grandes bloques de información: una introducción y una discusión. La presencia de cinco artículos, cuatro constitutivos propiamente del compendio de publicaciones que justifica esta tesis doctoral, más un quinto artículo, identificado como artículo adjunto, y que se ha incluido con el objetivo de proporcionar mayor consistencia a la satisfacción de los diversos objetivos psicométricos contemplados, ha supuesto la decisión de elaborar un documento con apartados más genéricos que los correspondientes a los estándares de un informe científico. La razón fundamental ha sido la de no redundar en aspectos que quedan clara y extensamente expuestos en las respectivas publicaciones. Es por ello que se ha prescindido de la exposición de las diversas muestras, materiales o procedimientos, o de una exhaustiva descripción de todos los resultados obtenidos. Además, los objetivos de estudio aparecen dispuestos a lo largo de la introducción, aprovechando la exposición de cada uno de los conceptos psicométricos que vertebran este documento. Se ha considerado que los contenidos presentados en este trabajo debían suponer básicamente una elaboración conceptual y general que proporcionara una perspectiva lo más armónica posible del conjunto de elementos incluidos en las publicaciones agrupadas en esta tesis doctoral.

La introducción se ha constituido, en primer lugar, entorno a una breve exposición relativa al modelo de los Cinco Grandes de Personalidad, ya que se trata del modelo de personalidad con una mayor penetrancia bibliográfica en la actualidad, para pasar posteriormente a una exposición de cuáles son los fundamentos teóricos y empíricos del Modelo Alternativo de los Cinco Grandes, modelo entorno al cual se justifica el trabajo desarrollado en las cinco publicaciones presentadas. Un segundo bloque, claramente más extenso, recopilará todos los conceptos psicométricos que se consideran más relevantes y que han sido objeto principal de cada una de las publicaciones compiladas. Será en este espacio en el que, además de exponer algunas de las consideraciones que se han percibido como imprescindibles para el análisis de cada una de estas propiedades psicométricas, se presentarán los objetivos de estudio que fundamentan cada publicación.

La discusión se aborda exponiendo las observaciones y juicios que se han percibido como más importantes y que, por supuesto, se derivan de cada una de las diversas publicaciones, sin obviar la exposición de las críticas y limitaciones que también se han determinado como necesarias. Esta discusión tiene un cuerpo único, es decir, no está dividida en apartados con el objetivo de facilitar la conexión entre sí de los distintos conceptos psicométricos así como de las consideraciones prácticas que se derivan. Este bloque narrativo concluye con un conjunto de reflexiones críticas generales sobre el trabajo desarrollado hasta el momento y con una breve exposición de las ideas fundamentales que guiarán nuestros estudios en el futuro. Seguidamente, y dando fin al documento, se proporciona una copia de los artículos que dan origen y sentido a la presente tesis doctoral.

2. Las cinco publicaciones en las que se fundamenta este documento son las que se referencian a continuación:
Referencias de las publicaciones compendiadas

Art-1:

Gomà-i-Freixanet, M., Valero, S., Puntí, J. & Zuckerman, M. (2004). Psychometric properties of the Zuckerman-Kuhlman Personality Questionnaire in a Spanish sample. *European Journal of Psychological Assessment*, 20(2), 134-146.

Art-2:

Gomà-i-Freixanet, M., Wismeijer, A. A. J. & Valero, S. (2005). Consensual validity parameters of the Zuckerman-Kuhlman Personality Questionnaire: Evidence from self-reports and spouse reports. *Journal of Personality Assessment*, 84(3), 279-286.

Art-3:

Gomà-i-Freixanet M, Soler, J., Valero, S., Pascual, J. C. & Pérez, V. (2008). Discriminant validity of the ZKPQ in a sample meeting BPD diagnosis vs. normal-range controls. *Journal of Personality Disorders*, 22(2), 178-190.

Art-4:

Gomà-i-Freixanet, M. & Valero, S. (2008). Spanish normative data of the Zuckerman-Kuhlman Personality Questionnaire in a general population sample. *Psicothema*, 20(2), 324-330.

Art-adjunto:

Gomà-i-Freixanet, M., Valero, S., Muro, A. & Albiol, S. (2008). Zuckerman-Kuhlman Personality Questionnaire: psychometric properties in a general population sample. *Psychological Reports*, 103, 845-856.

A lo largo del texto, y con el objetivo de identificar con facilidad el artículo al se está haciendo referencia en cada instante, se dispondrá el identificador del artículo correspondiente entre paréntesis (*Art-1, Art-2, Art-3, Art-4* y *Art-adjunto*)¹.

¹ Es posible obtener éstos y otros artículos en formato PDF en la página web www.zkpq.com. Esta web, operativa desde el año 2007, proporciona información actualizada sobre los trabajos publicados por nuestro equipo hasta el momento, así como información diversa sobre el modelo y el instrumento (características técnicas, cuestionario, baremos, etc).

Introducción

3. Durante muchos años el análisis taxonómico de la personalidad ha estado básicamente dominado por dos modelos. El modelo de 3 dimensiones de H.J. Eysenck (1947, 1967) y el modelo de R.B. Cattell (1957) constituido por 16 dimensiones. Los dos modelos tienen en común un origen estimulado por la aparición de circunstancias de gran relevancia, tanto desde un punto de vista teórico como metodológico. En primer lugar la aparición, durante la primera mitad de siglo XX, de un conjunto de nuevas estrategias metodológicas de medida en psicología, mediante las cuales los antiguos tipos básicos de temperamento formulados ya en la antigua Grecia (flemático, sanguíneo, melancólico y colérico) se transformaban en conceptos psicológicos nuevos constituidos ahora en dimensiones delimitadas por polos antagonicos a la vez que en entidades taxonómicas de orden superior. En segundo lugar, el desarrollo de técnicas estadísticas multivariadas, en concreto el análisis factorial, técnica que ofrecía una aproximación razonablemente objetiva a la descripción de estas nuevas entidades taxonómicas, determinando y acotando dichas variables de orden superior de entre la cantidad ingente de rasgos evaluados hasta el momento por las numerosas escalas y cuestionarios existentes.

El modelo de Cattell se constituyó en un marco teórico en el que se aúnan rasgos temperamentales, rasgos motivacionales, estados de ánimo, rasgos dinámicos, necesidades e impulsos actitudinales. El mismo autor, años después, se refirió a estos diferentes elementos, de forma sintética, en términos de rasgos temperamentales y rasgos dinámicos. Desgraciadamente este modelo presenta algunas limitaciones tanto conceptuales como empíricas relevantes. Algunas de ellas tienen que ver con el problema de no haberse podido replicar debidamente sus 16 dimensiones entre géneros, edades, métodos o bien porque muchos investigadores, simplemente, no han podido hallar tales factores (Digman, 1990; Eysenck & Eysenck, 1969), lo que sin duda se convierte en un importante problema para el modelo en términos de capacidad de generalización y por tanto de validación del mismo.

El modelo de Eysenck, contrariamente, se ha constituido como uno de los modelos de personalidad científicamente más consistentes y heurísticos de los últimos 60 años siendo, además, el modelo más representativo de los conocidos como modelos factoriales-biológicos. El modelo se articula entorno a las dimensiones Extroversión, Neuroticismo y Psicoticismo (EPQ), lo que le ha otorgado el nombre de *El Modelo de los Tres Grandes*. Dicho modelo se estructura de forma jerárquica, de tal manera que cada una de las tres dimensiones básicas se subdivide en rasgos específicos de personalidad y éstos, a su vez, en hábitos constituidos por reacciones o conjuntos de respuestas específicas. El modelo se constituye, además, como una explicación integradora en términos de fisiología del sistema nervioso central y del aprendizaje humano.

3.1.
***El modelo de los
Cinco Grandes***

Actualmente, el modelo de personalidad normal más comúnmente utilizado es el de los Cinco Grandes (Costa & McCrae, 1985). Según este modelo, los rasgos de personalidad pueden ser descritos en términos de 5 dimensiones básicas denominadas Neuroticismo, Extroversión, Apertura a la Experiencia, Amabilidad y Conciencia. Estas cinco dimensiones básicas se articulan entorno a 30 facetas, 6 para cada dimensión (Costa & McCrae, 1992).

El Modelo de los Cinco Grandes fue creado bajo una concepción léxica según la cual se considera que, puesto que la personalidad es un elemento central en la interacción entre humanos, cualquier rasgo puede ser codificado en el lenguaje natural. Se asume así que estas dimensiones pueden ser halladas en autoinformes o heteroinformes en los que se utilicen rasgos de personalidad en forma de adjetivos extraídos de diccionarios (McCrae & John, 1992). Son numerosos los estudios que muestran una buena replicabilidad intercultural e interlingüística, resultados éstos a favor de la universalidad del modelo (McCrae & Costa, 1997).

Aunque este modelo resulta muy útil en la descripción de la personalidad, no parece que su capacidad para explicar el comportamiento presente una utilidad paralela. Según Eysenck (1992a, 1992b), dos son las consideraciones fundamentales que debieran determinar qué es un rasgo de personalidad. La primera de ellas supone considerar que un buen modelo teórico debe actuar como marco conceptual que permita guiar tanto la elaboración de mejores estudios de taxonomía como nuevos avances teóricos. La segunda consideración determina como necesaria la aportación de evidencias consistentes sobre la vinculación entre causación genética y organización de la conducta.

En el modelo de los Cinco Grandes, si bien están suficientemente fundamentadas las bases biológicas de las dimensiones Neuroticismo y Extroversión, muy poco se conoce acerca de las dimensiones Conciencia, Amabilidad y Apertura a la Experiencia. Por otro lado, el reactivo final en forma de adjetivos que se obtiene utilizando el criterio léxico, y que como se apuntó, se nutre del uso de diccionarios, no necesariamente es el término final en los que debiera articularse un factor de personalidad. La codificación de los rasgos de personalidad en el lenguaje natural puede no ser proporcional a la relevancia conductual de estos rasgos o a su relevancia biológica (Zuckerman & Kuhlman, 2000). No debe olvidarse que el resultado de un análisis factorial depende obviamente de la muestra de reactivos que se utiliza. Por esta razón, se considera que la definición de un rasgo básico de personalidad no puede ser resuelto simplemente mediante la ejecución de un análisis factorial sobre un conjunto de variables si éstas han sido recogidas al margen de un marco teórico y biológico que las sustente (Eysenck, 1992a, 1992b; Zuckerman, 1992).

3.2.
***El modelo
alternativo de los
Cinco Grandes***

El conocido como Modelo Alternativo de los Cinco Grandes (Zuckerman, Kuhlman, Joireman, Teta & Kraft, 1993) encaja también dentro de los llamados modelos factoriales-biológicos. El punto de partida del modelo se basa en la consideración de que los rasgos básicos de personalidad deben ser rasgos que posean una base consistente en términos biológicos-evolutivos. Es por esta razón que para obtener los ítems que constituyen el instrumento que sustenta el modelo, el Zuckerman-Kuhlman Personality Questionnaire (ZKPQ), se llevaron a cabo análisis factoriales de escalas que miden rasgos de personalidad que, particularmente, han sido utilizadas en investigación psicobiológica (Zuckerman, Kuhlman & Camac, 1988; Zuckerman, Kuhlman, Thornquist & Kiers, 1991). Como punto de partida se tomó la escala de Eysenck y sus *Grandes Tres*, y se incorporaron medidas diversas, como por ejemplo la Emotionality-Activity-Sociability-Impulsivity (EASY, Buss & Plomin, 1975), el Karolinska Scales of Personality (KSP, Klinerberg, Schalling & Magnusson, 1986), la Sensation Seeking Scale (SSS, Zuckerman, 1979) o el Strelau Temperament Inventory (STI, Strelau, 1983) entre otras medidas. Para mayor información sobre las escalas originales utilizadas, así como del proceso de obtención de los reactivos finales constitutivos del ZKPQ véase Zuckerman et al. (1988).

Como se apuntó, y a diferencia del Modelo de los Cinco Grandes, el ZKPQ incluye dimensiones de personalidad que puedan ser potencialmente replicadas entre especies diferentes, bajo la consideración de que si un rasgo se considera relevante en términos de personalidad, éste debiera poderse contextualizar en términos de su arraigo biológico y evolutivo (Zuckerman & Kuhlman, 2000).

Los rasgos básicos de personalidad recogidos en el modelo, y por tanto en el ZKPQ, así como el conjunto de elementos que comprenden cada uno de ellos son los siguientes:

- Neuroticismo-Ansiedad (N-Anx): La escala mide preocupación emocional, tensión, miedos, indecisiones, obsesión, falta de autoconfianza y especial sensibilidad a la crítica.
- Actividad (Act): Esta escala evalúa la necesidad de actividad general, dificultad para relajarse o para no hacer nada cuando es posible y preferencia por los trabajos duros, activos, de gran consumo de energía y que supongan una vida ocupada.
- Sociabilidad (Sy): Hace referencia al número de amigos, las ganas de estar con ellos y asistir a fiestas, y la preferencia por realizar actividades con otros en contraposición a realizarlas a solas.
- Impulsividad-Búsqueda de Sensaciones (ImpSS): Dicha escala evalúa una pobre capacidad para planear y la tendencia a actuar impulsivamente sin pensar en las consecuencias y búsqueda activa de experiencias nuevas que supongan excitación y riesgo.

- Agresión-Hostilidad (Agg-Host)²: La escala hace referencia a la predisposición a expresar agresividad verbal, rudeza, grosería, desconsideración hacia los demás, conducta antisocial y deseo de venganza y malicia.

La dimensión ImpSS incluye dos facetas diferentes, Impulsividad (Imp) y Búsqueda de Sensaciones (SS). La dimensión Act incluye las facetas Actividad General (GenAct) y Esfuerzo en el trabajo (WorkAct) y la dimensión Sy incluye las facetas Amigos y grupos (Parties) e Intolerancia a la soledad (Isol).

Eysenck (1967), Guilford (1975) y Zuckerman (1984) proponen que la dimensión Obertura-intelecto se entienda como un dominio separado del temperamento, lo que justifica que este factor no se incluya en el modelo. La dimensión Extroversión está dividida en dos factores diferentes identificados como Actividad y Sociabilidad. El factor Actividad emergió como un factor diferenciado según análisis factoriales de cinco factores realizados en estudios iniciales (Zuckerman et al., 1988, 1991). Como consecuencia de la identificación de esta dimensión como un rasgo del desarrollo evolutivo (véase por ejemplo, Buss & Plomin, 1984; Thomas & Chess, 1977), dicha variable merece una evaluación diferenciada como rasgo relevante de la personalidad del individuo.

La distinción entre Hostilidad y Ansiedad es también importante ya que ambos rasgos poseen diferentes bases psicobiológicas (Gray, 1982), y no debieran ser confundidas en un mismo factor de Neuroticismo, tal y como sucede en el caso del modelo de Costa y McCrae (NEO). Tanto la dimensión Impulsividad como la de Búsqueda de sensaciones disponen, además, de correlatos psicobiológicos (Zuckerman, 1983, 1984, 1991, 1994).

El ZKPQ contiene, además de las cinco escalas ya expuestas, una escala diferenciada, denominada Infrecuencia (Inf), que tiene como objeto la detección de sujetos que tiendan a responder de forma inadecuada o bien a responder prestando poca atención a la tarea. En el manual no publicado del ZKPQ (Zuckerman & Kuhlman, 1993) se expone que en muestras norteamericanas únicamente un 4% de los participantes presentaban puntuaciones superiores a 3 en esta escala y toda persona con una puntuación de 4 o más fue eliminada del estudio de factorización. En muestra española los valores en la escala de Infrecuencia tienen un rango de 0-7 ($M=1,58$, $DE=1,45$), con una moda de 1 y con un 3,7% de participantes con un valor superior a 4 (Gomà-i-Freixanet, Valero, Puntí & Zuckerman, 2004).

El ZKPQ es un instrumento que ha sido utilizado en contextos aplicados diversos. Puede hallarse investigación sobre descripción de abusadores de sustancias y predicción de evolución clínica (Ball, 1995), cuadros psicopatológicos distintos (O'Sullivan, Zuckerman & Kraft, 1996; Thronquist & Zuckerman, 1995), conductas de riesgo (Zuckerman & Kuhlman, 2000), intereses profesionales en estudiantes de medicina (Hojat & Zuckerman, 2008) entre otros ámbitos. Se han llevado a cabo adaptaciones del instrumento al alemán

² Los acrónimos de las variables de personalidad del ZKPQ se han mantenido en inglés para facilitar su identificación en las publicaciones incluidas en esta tesis.

(Ostendorf & Angleitner, 1994), japonés (Shiomi, Kuhlman, Zuckerman, Joireman, Sato & Yata, 1996), chino (Wu, Wang, Du, Li, Jiang & Wang, 2000) o italiano (De Pascalis & Russo, 2003) entre otras lenguas. El ZKPQ ha sido también aplicado o adaptado en población española (Aluja, García & García, 2004; Gutiérrez-Zotes, Ramos & Sáiz, 2001; Peñate, Ibáñez & González, 1999; Romero, Luengo, Gómez-Freguela & Sobral, 2002).

4. El objetivo de los apartados siguientes es el de describir algunas de las propiedades psicométricas básicas que a nuestro juicio el ZKPQ debiera satisfacer y, por ende, todo instrumento de evaluación de la personalidad. No se pretende ser ni exhaustivo ni pormenorizado en la descripción técnica de los términos psicométricos expuestos, ya que consideramos que para satisfacer una descripción más específica de cada uno de ellos debiera ser otra la orientación de la presente tesis. Entendemos además que son numerosos los manuales de psicometría que seguro podrán satisfacer una aproximación más precisa y extensa de todos los conceptos de interés.

4.1. Toda medida de lo psicológico siempre está dotada de cierto error. El objetivo fundamental del investigador es determinar cuál es la magnitud de dicho error con el objetivo de acotar a qué nivel de calidad asciende el instrumento y, por tanto, determinar el grado de confianza que dicho instrumento le merece. Estamos hablando del concepto psicométrico de fiabilidad y que será la primera propiedad que centrará nuestra exposición.

Si se aplicase un mismo instrumento diversas veces a un mismo sujeto con la finalidad de medir una determinada variable es casi seguro que las puntuaciones serían parecidas entre sí pero improbablemente idénticas. La discrepancia entre estos valores observados en esta hipotética situación sería atribuible al error de medida de este instrumento.

Desde la teoría clásica de los tests, y en concreto desde el modelo lineal de Spearman, se establece que la puntuación empírica de un sujeto en un test, es decir, la puntuación obtenida (y que llamaremos X), puede considerarse como una combinación lineal de dos elementos: la puntuación verdadera de este sujeto en el rasgo de interés (identificada como V) y el error de medida cometido (denominada E). Entre los tres elementos descritos se constituye la siguiente sencilla ecuación:

$$X = V + E$$

Tal y como se deduce de esta expresión, si se aplica un test a un sujeto la puntuación que obtendrá no coincidirá con el valor de la puntuación verdadera. Debería siempre tenerse en consideración el error de medida cometido ya que es el elemento fundamental que determina la fiabilidad de un instrumento. Podría definirse la fiabilidad como el grado en que una medida está libre de error o bien el grado de confianza que merece un instrumento cuando se administra a un determinado sujeto.

Desde un punto de vista matemático, la fiabilidad de un test puede definirse también como el cociente entre la varianza de las puntuaciones verdaderas y la varianza de las puntuaciones empíricas, según la siguiente expresión:

$$r = S^2V / S^2X$$

Se puede interpretar así que la fiabilidad es la proporción de la varianza de las puntuaciones empíricas de los sujetos (denominador del anterior cociente) que se debe a la varianza de las puntuaciones verdaderas (numerador del mismo cociente), o lo que es lo mismo, la proporción de varianza verdadera que hay en la varianza empírica. A medida que dicha proporción aumenta disminuye el

error de medida. Si $r=1$, el error de medida es cero y por tanto la fiabilidad es perfecta. Está claro que un resultado de este tipo, en el que el instrumento carece de error, tiene un valor puramente pedagógico. La condición de partida en investigación psicológica es que sea cual sea el instrumento de interés, éste siempre presenta un nivel de fiabilidad imperfecto.

De entre los diversos elementos que condicionan la magnitud que puede tomar la fiabilidad de un instrumento destacamos dos de ellos. Uno es el relativo a la longitud del mismo, es decir, a la cantidad de ítems que lo constituyen. El segundo tiene que ver con la variabilidad del grupo al que se administra el instrumento. Respecto al primer factor debe considerarse que existe una relación directa entre cantidad de reactivos y grado de fiabilidad estimado. Cuantos más elementos den cuenta de la variable de interés menos queda por describir de la misma y por tanto, menos error se comete en su medida. De este principio se deriva que para conseguir un aumento de la fiabilidad de un instrumento solo sería necesario aumentar la cantidad de ítems que lo constituyen, entendiéndose necesariamente que los nuevos ítems incorporados poseen una calidad paralela, es decir, comparable, a los ítems ya presentes en este instrumento.

La ecuación de Spearman-Brown da cuenta de la vinculación entre fiabilidad y longitud de un test y puede ser una herramienta útil para el investigador que se plantee un cambio en la cantidad de ítems de un test. La necesidad de cambiar la cuantía de reactivos puede justificarse bien porque el índice de fiabilidad actual no satisface las expectativas del investigador, pudiendo considerar la posibilidad de aumentar la cantidad de ítems de este instrumento, bien porque el actual instrumento dispone de demasiados ítems siendo recomendable una versión reducida del mismo. Esta ecuación se expresa de la siguiente manera:

$$R_{xx} = nr_{xx} / (1 + (nr_{xx} - r_{xx}))$$

donde:

R_{xx} =coeficiente de fiabilidad del test alargado o acortado

r_{xx} =coeficiente de fiabilidad del test inicial

n =número de veces que se ha alargado el test. Si el test se alarga n debe ser >1 , si se acorta n debe ser <1

Debe considerarse, no obstante, que aunque a medida que aumenta el número de ítems paralelos aumenta el coeficiente de fiabilidad de test, este aumento no mantiene una vinculación monotónica: a partir de un determinado valor de n no se observan aumentos de la fiabilidad sustanciales. Consecuentemente, uno de los objetivos fundamentales que deberá satisfacerse es determinar en qué medida un aumento o disminución de ítems se asocia a un valor óptimo de fiabilidad. Puede ser de gran interés la reducción de la longitud de un instrumento si esta reducción supone tener que exponer a los sujetos a un menor potencial cansancio, por ejemplo, pero solo si el nivel de fiabilidad resultante, que tenderá a ser inferior, sea esencialmente comparable al valor original o bien de una calidad suficiente.

Todos los trabajos constitutivos del compendio de publicaciones de la presente tesis han sido realizados con la versión original del ZKPQ de 99 ítems. No

zkpq

obstante, existe para el ZKPQ una propuesta de versión reducida, de la misma manera que existen versiones reducidas de otros instrumentos de evaluación de la personalidad, como por ejemplo en el caso del Eysenck Personality Questionnaire (EPQ-RS, Eysenck & Eysenck, 1991) o del NEO-PI-R (NEO-FFI, McCrae & Costa, 2004). Esta versión reducida recibe el nombre de ZKPQ-50-CC (Aluja, Rossier, García, Angleitner, Kuhlman & Zuckerman, 2006; Aluja, Rossier & Zuckerman, 2007) y está constituida por 50 ítems en total, 10 por cada uno de los cinco factores. Esta versión muestra un satisfactorio comportamiento en términos tanto de estructura factorial como de fiabilidad interna. No obstante, aunque no ponemos en duda que disponer de una versión reducida podría redundar en un instrumento quizás más eficiente en determinados contextos de evaluación, debe tenerse en consideración que la eliminación de ítems que ha justificado esta nueva versión ha sido establecida exclusivamente con población universitaria. Entendemos que el análisis del comportamiento psicométrico de esta versión en población general aportará datos relevantes sobre el grado en que puede ser aplicada también con garantías en población no universitaria.

Además de las razones metodológicas expuestas, otra consideración fundamental que ha justificado no alterar la longitud original del ZKPQ ha sido la de permitir la comparabilidad entre estudios propios y ajenos y, por supuesto, la comparabilidad interpoblacional y cultural del instrumento, elemento angular en la validación de un instrumento de personalidad. No tenemos constancia bibliográfica alguna que se utilice una versión reducida del ZKPQ en un contexto aplicado, lo que redundaría en la idea de que hasta el momento solo es posible la comparabilidad entre estudios con la versión íntegra. Por otro lado, no podemos dejar de mencionar que el ZKPQ, con sus 99 ítems, no resulta un instrumento especialmente extenso en comparación con otros instrumentos habituales de medida de la personalidad. Considérese por ejemplo el NEO-PI-R, formado por 240 ítems (Costa & McCrae, 1992) o el Temperament and Character Inventory (TCI; Cloninger, Svrakic & Przybeck, 1993) de 174 ítems. O bien en el caso de instrumentos con una orientación clínica, el Minnesota Multiphasic Personality Inventory (MMPI-II; Butcher, Dahlstrom, Graham, Tellegen & Kaemmer, 1989), que en su versión reducida dispone de 370 preguntas (Butcher & Hostetler, 1990) o el Millon Clinical Multiaxial Inventory (MCMI; Millon, 1994) constituido por 175 ítems.

Un segundo factor que condiciona la magnitud del coeficiente de fiabilidad es el relativo a la variabilidad de la muestra explorada. Cuanto menor sea la variabilidad de las puntuaciones empíricas obtenidas, es decir, cuanto menor sea la desviación típica observada, menor tiende a ser también el índice de fiabilidad asociado. Este comportamiento debe ser debidamente considerado ya que tiende a ignorarse con demasiada frecuencia cuando se estudian o se valoran las propiedades psicométricas de un instrumento: la fiabilidad de un test no es una propiedad intrínseca del instrumento. Depende de la convergencia de dos factores. Por un lado, efectivamente, el propio instrumento, por el otro, la población en la que se aplica dicho instrumento. En definitiva, un test puede presentar tantos índices de fiabilidad como muestras distintas en las que se calcule (Muñiz, 1998).

Expuestas algunas consideraciones importantes sobre la magnitud de la fiabilidad de un instrumento, cabe preguntarse sobre qué tipos de fiabilidad pueden ser de interés para un cuestionario de personalidad. Para responder a esta pregunta expondremos en primer lugar que no se tendrá en consideración el concepto de fiabilidad entre observadores, que si bien tendría una fundamental

importancia en multitud de ámbitos de evaluación, en el caso del ZKPQ resulta una propiedad irrelevante, al tratarse de un instrumento autoadministrado, aunque puede ser útil en algunos contextos tal y como expondremos más adelante. A partir de esta consideración, entendemos que son dos los tipos de fiabilidad que debieran ser satisfechos. Se trata de la fiabilidad interna, también conocida como consistencia interna y de la fiabilidad temporal, denominada también como estabilidad de la medida.

4.1.1. Fiabilidad como consistencia interna

Resultaría poco alentador observar que algunos de los ítems de un test de inteligencia administrado a una misma persona infieran de ésta un alto grado de inteligencia mientras que el resto de ítems informaran de justamente lo contrario. Nos referimos al concepto de consistencia interna para expresar el grado en que los ítems de una misma entidad conceptual (por ejemplo, los pertenecientes a un mismo factor) se interrelacionan entre sí. En el caso que nos ocupa, sería deseable que los ítems constitutivos de una misma dimensión de personalidad presentasen elevadas correlaciones absolutas entre sí.

Aunque existen diversas estrategias para el cálculo de la consistencia interna de un instrumento: Kuder-Richarson (1937), Rulon (1939), Guttman (1945), Cronbach (1951) entre otros, es éste último, el coeficiente Alfa de Cronbach, el que domina sin duda alguna el análisis de consistencia interna de la mayoría de los instrumentos publicados. Su formulación es la siguiente:

$$\alpha = \frac{I}{I-1} \left[1 - \frac{\sum S_i^2}{S_t^2} \right]$$

donde:

I es el número de ítems analizados
S_t² =varianza de la puntuación total del test
S_i² =varianza de cada uno de los ítems

Obsérvese que el índice alfa, que como todo coeficiente de fiabilidad puede tomar un valor comprendido entre 0 y 1, depende tanto de la cantidad de reactivos analizados como de las varianzas que éstos generan cuando son administrados a una determinada muestra. Tal y como se expuso con anterioridad, dos de los elementos fundamentales en la determinación de la magnitud de esta propiedad.

Los artículos *Art-1*, *Art-2*, *Art-3* y *Art-adjunto* proporcionan los correspondientes coeficientes alfa de cada escala del ZKPQ. Se han reportado los índices de consistencia interna del instrumento para cada estudio porque se ha considerado necesario hacerlo, ya que las poblaciones exploradas son diferentes y por tanto es pertinente reportar la consistencia interna para cada una de ellas. Solo el artículo relativo a los datos normativos del ZKPQ (*Art-4*) no contempla coeficiente de fiabilidad alguno. Esta circunstancia se justifica porque el objetivo de este artículo era exclusivamente reportar datos normativos sobre población general. El comportamiento relativo a las propiedades psicométricas del instrumento en esta población se recogen en la última publicación (*Art-adjunto*).

4.1.2.
Fiabilidad como
estabilidad de la
medida

Asumiendo que la personalidad de un individuo no debe cambiar sustancialmente tras el transcurso, como mínimo, de un breve período de tiempo, las puntuaciones obtenidas en estos dos momentos debieran ser esencialmente comparables. El interés radica aquí en el grado de acuerdo entre las puntuaciones obtenidas por los sujetos en aplicaciones diferentes de la misma medida.

En el análisis realizado para la satisfacción de este objetivo (*Art-1*) se obtuvieron correlaciones entre dos administraciones del ZKPQ a un subgrupo de participantes dos semanas después de una primera evaluación. Las correlaciones entre ambas administraciones fueron de 0,87 para la escala N-Anx, 0,89 para Act, 0,90 para Sy, 0,91 para ImpSS y 0,77 para Agg-Host. Estos resultados indican una satisfactoria estabilidad temporal del instrumento.

4.2.
Sobre la validez

El concepto de validez es, en comparación al concepto de fiabilidad, mucho más complejo, controvertido y especialmente relevante en investigación psicológica o psiquiátrica. Es a través de este concepto que se cuestionan muchos de los conceptos usuales en investigación comportamental, ya que no es posible estudiar la validez de un instrumento sin investigar tarde o temprano la naturaleza y el significado de las variables implicadas (Kerlinger & Lee, 2002).

Aunque no existe una forma unitaria de exponer a qué se hace referencia cuando se habla del concepto de validez, probablemente sería una pregunta del tipo ¿estamos midiendo lo que creemos que estamos midiendo?. La pregunta resulta especialmente compleja de satisfacer si se tiene en consideración la dificultad que supone medir variables psicológicas o psiquiátricas, ya que la mayoría de ellas no pueden ser observadas ni medidas directamente. Éstas variables, que reciben el nombre de constructos o variables latentes, exigen la determinación y delimitación de conductas que representen algún dominio relevante de este constructo y que se consideren indicadores del mismo. Se puede afirmar que se ha obtenido una medida del constructo cuando se dispone de una medida de las conductas seleccionadas como indicadores (Barbero, Vila & Suárez, 2003). En exceso, desgraciadamente, el proceso de construir y evaluar instrumentos de medición en psicología o psiquiatría se ha dirigido frecuentemente a la consecución de instrumentos fiables más que a la obtención de instrumentos válidos. Si bien es cierto que la frontera entre lo que es un problema de fiabilidad o de validez puede ser difuso, probablemente este tipo de error, consistente en no plantearse objetivos relativos a la validez y satisfacer exclusivamente preguntas relativas a la fiabilidad reside en la presencia, con frecuencia, de un vago conocimiento sobre las diferencias entre ambos conceptos psicométricos.

Por otro lado, afirmar que *se ha validado el instrumento X* o que *ya existe la validación del instrumento Y* es, según nuestro punto de vista, no hablar de prácticamente nada que pueda considerarse realmente relevante si no se tiene en consideración que esta *validación* suele hacer referencia exclusivamente a algún tipo concreto de validez, de entre los diversos tipos de validez que pueden ser analizados, que esta validez está únicamente (de momento) vinculada a la población en la que se analizó, y que está delimitada y condicionada siempre por los supuestos o asunciones teóricas que se tomaron en consideración para definir el diseño metodológico en el que se insertó el estudio. Y es que para referirse a la validez, como en el caso de la fiabilidad, supone un error asumir que los instrumentos son *portadores* de propiedades psicométricas como si se trataran de

cualidades intrínsecas y absolutas de los mismos, y enajenadas de las condiciones en las que el instrumento se explora.

A continuación se exponen los tipos de validez que hemos considerado de especial relevancia en el contexto de la medida de la personalidad. Para ello iniciaremos esta exposición hablando de la validez de contenido.

4.2.1. **Validez de contenido**

La validez de contenido tiene que ver con la representatividad o la adecuación del contenido de un instrumento de medición (Livingston, 1977). Cualquier propiedad psicológica dispone de un universo teórico de contenido, el cual consiste en todos los posibles elementos, también llamados reactivos, que se dicen u observan acerca de esta propiedad. Puesto que en teoría este universo posee un número infinito de reactivos, una medida con alta validez de contenido será aquella en la que teóricamente se ha producido un muestreo representativo de todos los reactivos posibles.

Por desgracia, la mayoría de veces no es posible elegir muestras aleatorias de reactivos de un universo de contenido. Es por esta razón que el proceso de enjuiciar la validez de contenido consiste, la mayoría de veces, en juzgar la calidad de los reactivos incluidos, siendo obvio que unos reactivos serán más fáciles de juzgar que otros. Aunque no existe un procedimiento específico, metodológico u estadístico, que explore esta propiedad, sí es posible recurrir a alguna estrategia de carácter parcial, centrada en el análisis del acuerdo entre juicios de especialistas (denominados *jueces*), sobre la relevancia de un reactivo en el contexto de un dominio concreto de conocimiento (Cohen, 1960; Hambleton, 1980).

Independientemente de si se aplica una aproximación más o menos sistemática y estructurada a la valoración de la validez de contenido es necesario que ésta se lleve a cabo desde un posicionamiento teórico claro y conocedor del constructo de interés. No obstante, compartimos el punto de vista de Messick (1975) cuando afirma que la especificación y representatividad del dominio son objetivos por satisfacer en la construcción de un test, pero que no son garantía de validez puesto que no proporcionan evidencia empírica para poder interpretar y extraer conclusiones. En ninguno de los artículos incluidos en esta tesis se aborda específicamente el análisis de contenido del ZKPQ. Se asume que tratándose de una prueba inspirada y constituida por ítems extraídos de instrumentos de consolidado bagaje, tanto teórico como empírico, debiera satisfacer razonablemente esta propiedad.

Existe una variante de la validez de contenido que recibe el nombre de validez aparente o de *facie*. Esta propiedad tiene que ver con la determinación del grado en que la prueba objeto de interés mide lo que aparenta medir. Individuos entrenados o no debieran observar la prueba y decidir en qué grado ésta mide lo que supuestamente debe medir. Aunque no sea exactamente ni exclusivamente éste el objetivo fundamental, el artículo incluido en esta tesis *Consensual validity parameters of the Zuckerman-Kuhlman Personality Questionnaire: Evidence from self-reports and spouse reports-(Art-2)* supone la exploración del grado en que los reactivos del ZKPQ permiten la descripción de un mismo rasgo en una misma persona según el juicio de una segunda persona. La pregunta que nos planteábamos en concreto fue: ¿en qué medida los juicios de uno de los dos individuos de la pareja respecto a la otra persona concuerdan con los juicios realizados por esta misma persona sobre sí mismo?

Las asunciones de partida son dos. La primera supone considerar que una persona próxima, en este caso la pareja, es capaz de identificar debidamente las conductas que caracterizan a la otra persona. La segunda asunción consiste en determinar que el ZKPQ está constituido por reactivos que son suficientemente claros y objetivos como para ser identificados y por tanto reproducidos respecto a la conducta de otro.

Utilizando correlaciones intraclase para cada una de las cinco dimensiones del ZKPQ, así como un análisis factorial de segundo orden tomando como elementos primarios las cinco dimensiones, tanto para las autoevaluaciones como para las heteroevaluaciones, se observa que los juicios sobre uno mismo y los juicios realizados por la pareja son altamente concordantes. Consideramos que a este efecto resulta muy ilustrativo el comportamiento del análisis de componentes principales recogido en la Tabla 6 de dicho artículo. Dicho análisis, pone en evidencia como cada uno de los pares de juicios (auto y hetero) satura claramente en un mismo factor y de forma exclusiva, siendo un total de cinco los factores que mejor resumen el comportamiento de los datos, concordando con la cantidad de dimensiones evaluadas. Más adelante, en esta misma introducción, volveremos a tratar este estudio.

4.2.2. Validez de constructo

La validez de constructo supone la más clara asociación entre los conceptos psicométricos y las prácticas psicométricas, y los conceptos teóricos. Exponiéndola en términos operativos, la validez de constructo podría definirse como el análisis de las propiedades psicológicas o de otro tipo que pueden explicar la varianza de la prueba. Supone buscar el significado de la prueba u obtener evidencia acerca de la capacidad del instrumento para medir el constructo. Tal y como indica Yela (1984) los estudios de validez de constructo se dirigen a garantizar científicamente que la variable que el test pretende medir es, efectivamente, una variable *adecuada*, cuyo concepto ofrece suficiente consistencia lógica dentro de un sistema teórico de la Psicología y que descansa en suficientes comprobaciones experimentales que lo verifican.

La tarea fundamental del investigador reside en el análisis de la relación entre el constructo y las conductas observables representativas de dicho constructo, la relación entre este constructo y otros constructos y, finalmente, la relación entre estas conductas tomadas como indicadores del constructo y las puntuaciones obtenidas por los sujetos en el instrumento (Barbero et al., 2003). Generalmente, los estudios de validez de constructo se centran en la estructura del test, tanto desde una perspectiva interna como externa. El interés descansa sobre el estudio de las interrelaciones entre las puntuaciones obtenidas por los sujetos en los distintos ítems que constituyen el test (estructura interna) y en la relación entre las puntuaciones obtenidas en el test y otras medidas que puedan considerarse relevantes para el constructo de interés (estructura externa). Entre las estrategias más frecuentemente utilizadas encontramos el análisis factorial, y por extensión todo el conjunto de técnicas englobadas dentro de esta familia de recursos estadísticos, así como el método de la matriz multimétodo-multirasgo.

Aunque no pretendemos en absoluto realizar una descripción precisa de estas dos estrategias resulta imprescindible exponer algunos de los elementos fundamentales que las constituyen. A estos elementos dedicaremos las siguientes líneas.

4.2.2.1. *Análisis factorial*

Tal y como exponen Kerlinger y Lee (2002), el análisis factorial sirve a la causa de la parsimonia científica, reduce la multiplicidad de las pruebas y medidas a una mayor simplicidad. Además, ayuda al científico a ubicar e identificar unidades o propiedades fundamentales que subyacen a pruebas y medidas.

El análisis factorial es sin duda alguna la técnica más utilizada, tanto en su vertiente exploratoria como confirmatoria, para poner a prueba las hipótesis acerca de la estructura del constructo de interés y de las relaciones del mismo con otras variables. En investigación psicológica, es en el estudio de la estructura de la personalidad donde esta técnica adquiere una importancia fundamental. Esta afirmación es así hasta el punto de considerar que todos los modelos actuales de personalidad justifican su existencia en términos de una estructura de personalidad inferida tras la ejecución de alguna estrategia factorialista. Se trata probablemente de una situación paradigmática de cómo un conjunto de técnicas estadísticas, todas ellas bajo el epígrafe de análisis factorial, sustentan y justifican todo un ámbito de conocimiento científico.

Para llevar a cabo un análisis factorial es necesario partir de un conjunto de n medidas tomadas en la misma muestra de sujetos en un conjunto de variables observables, lo que permite generar un matriz de $(n \times n)$ con las intercorrelaciones entre todas estas variables. El objetivo fundamental del análisis factorial es el de determinar en qué medida es posible resumir la red de interrelaciones entre las variables contempladas mediante el establecimiento de factores que, posteriormente, sean debidamente interpretados y etiquetados conceptualmente en forma de constructos, de variables latentes, y que se asumen como el fundamento de la prueba.

Los factores no se obtienen de forma automática, es decir, no son un producto mecanicista de la técnica. Son consecuencia de la ejecución de determinadas estrategias (como por ejemplo, la realización de un tipo de rotación u otro) y del juicio y toma de decisiones por parte del investigador sobre qué criterios pondera de entre los diversos que pueden tenerse en consideración en cualquier análisis de factorización (varianza total explicada, magnitud y distribución de las comunalidades de cada ítem, magnitud y distribución de los valores propios de cada factor, magnitud de los valores residuales, etc). Por supuesto sin olvidarse de las consideraciones necesariamente ligadas al ámbito teórico y/o clínico en el que se inserta esta investigación y que condicionaran sin duda muchas de las decisiones que deban tomarse, como por ejemplo, cuando debe interpretarse de alguna forma razonada cada una de las agrupaciones de reactivos en los diferentes factores obtenidos.

Aunque el análisis factorial pueda considerarse por muchos como una de las herramientas más poderosas diseñadas hasta el momento para el estudio de áreas complejas de interés científico del comportamiento, se trata de una técnica no exenta de críticas consistentes. Muchas de estas críticas giran entorno a la determinación de cuántos factores deben extraerse de una matriz de correlaciones y del problema de cómo rotar los factores (Gorsuch, 1997). Otras consideraciones fundamentales giran entorno a la determinación de en qué medida el análisis factorial se constituye en una máquina estadística formada por una gran cantidad de pruebas que arroja factores con escaso o nulo valor psicológico. Este tipo de argumento se basaría en la consideración de que los factores son simplemente artefactos del método, promedios que no se corresponden con realidad psicológica alguna. Además se afirmaría que, en el contexto de este mismo conjunto de argumentos, en el análisis factorial no se puede obtener más que lo que se dispone dentro de él.

No compartimos exactamente este conjunto de argumentos. Coincidimos con Kerlinger y Lee (2002) al considerar que si el análisis factorial proporciona entidades carentes de significado psicológico, y que no son más que promedios que tienen relevancia solo en el imaginario del investigador factorialista, entenderíamos que ningún constructo científico tendría significado alguno, ya que todos son, en cierta medida, promedios. Entendemos que la realidad de cualquier constructo, de cualquier factor, es su realidad empírica, científica. Es decir, dicha realidad dependería de la capacidad que tiene este factor para predecir exitosamente relaciones entre fenómenos a partir de determinadas presunciones teóricas e hipotéticas. Por otro lado, cuando se afirma que en el análisis factorial no se obtiene nada que no haya sido introducido en él, no podemos dejar de decir que estamos absolutamente de acuerdo con ello. Pero no podemos obviar que no se conoce todo lo que se pone dentro, de la misma manera que no se conocen cuáles son las fuentes de varianza común ni las relaciones entre los factores implicados. Entendemos que es en este contexto de desconocimiento en el que el análisis factorial puede aportar alguna luz y por tanto en este espacio en el que debiera enjuiciarse esta estrategia.

Hemos realizado estudios de factorización del ZKPQ en dos de los artículos presentados en esta tesis (*Art-1* y *Art-adjunto*). El primero comprende población universitaria, el segundo población general. Ambos análisis son pertinentes ya que no son necesariamente redundantes. Muchos de los trabajos de factorización realizados sobre instrumentos de evaluación de la personalidad tienden a llevarse a cabo en población universitaria. La razón probablemente es muy simple: los estudiantes universitarios son una población muy fácilmente accesible para muchos investigadores. Por otro lado, puesto que estamos hablando de rasgos de personalidad normal, se asume generalmente que si éstos existen, existen para toda la población, obviamente también para estudiantes universitarios. Se asume, además, que estos estudiantes son una muestra representativa en la que analizar satisfactoriamente las propiedades del instrumento. Consideramos que a veces así es, otras no tiene porqué serlo. Reiteramos en la consideración que los índices de fiabilidad y validez dependen de la población que se explora. Entendemos así que es señal de buena praxis psicométrica, a la vez que teórica, no asumir por defecto que las decisiones y consideraciones tomadas respecto un instrumento administrado en estudiantes universitarios son automáticamente aplicables o generalizables al resto de la población.

Estas consideraciones son las que justifican que después de un primer artículo en el que se explora principalmente la validez de constructo del instrumento en población universitaria, posteriormente se realice un segundo estudio con el objetivo de evaluar la reproductividad del modelo en población general.

4.2.2.2.
Matriz
multirasgo-
multimétodo

Esta estrategia metodológica, propuesta hace ya cincuenta años por Campbell y Fiske (1959), permite el análisis de la estructura de un test (o conjunto de tests) entorno a los conceptos de validez convergente y validez discriminante. El objetivo fundamental se dirige a la medida de un mismo constructo mediante diversos procedimientos (multimétodo) y diversos constructos (multirasgo) con idéntico método. Si las correlaciones en las puntuaciones de los sujetos entre métodos diferentes pero sobre un mismo constructo son elevadas, se estará obteniendo evidencia sobre la validez de constructo y se podrá afirmar, además, que se observa validez convergente. Si además se observan correlaciones bajas entre constructos diferentes utilizando un mismo método, se podrá afirmar que está presente cierto grado de validez discriminante.

Aunque se trata de una estrategia de gran relevancia metodológica y conceptual en investigación tanto básica como aplicada, resulta difícil o frecuentemente imposible aplicar dos o más medidas de dos o más variables con muestras elevadas de sujetos. En el caso específico de la evaluación de la personalidad podrían crearse métodos distintos para la recogida de información sobre cada dimensión considerada. Por ejemplo, generando diferentes instrumentos con formatos de respuesta también dispares o bien generando estrategias de observación diferentes u otra forma de recogida de información distinta de la estándar con formato de lápiz y papel. Sería ciertamente factible desarrollar estas estrategias metodológicas, pero con un alto coste asociado.

En el artículo compendiado *Consensual validity parameters of the Zuckerman-Kuhlman Personality Questionnaire: Evidence from self-reports and spouse reports (Art-2)*, y que fue citado con anterioridad para referirse al concepto de validez de contenido, supone también una aproximación al concepto de validez convergente, puesto que supone la utilización de dos sistemas de evaluación para referirse a un mismo objeto de estudio. Aunque se trata de un trabajo interesante del que se derivan conclusiones con un marcado valor aplicado, y así se expondrán más adelante en la discusión, es evidente que la estrategia contemplada en este estudio aborda solo de forma parcial el espacio metodológico que supone la matriz multirasgo-multimétodo.

4.2.3. **Validez** **discriminante**

Si un instrumento evalúa personalidad, éste debiera poder diferenciar debidamente entre personas dotadas de rasgos de personalidad, como mínimo, extremos. Parece razonable considerar que aquellas personas con rasgos de personalidad que se ubiquen, teóricamente, en un extremo de una dimensión, debieran obtener puntuaciones diferenciadas de aquellas personas no dotadas, también teóricamente, de estos rasgos extremos. Estamos hablando, en definitiva, de la capacidad del instrumento para discriminar debidamente entre personas ubicadas en puntos diferentes de una misma dimensión de personalidad. Y esta propiedad, entendemos, debiera ser de aplicación necesaria para cada una de las dimensiones entorno a las cuales se articula un modelo de personalidad, considerando que este análisis se desarrolla en un contexto metodológico en el que la o las poblaciones diana presentan determinados rasgos de personalidad relativamente extremos en magnitud y determinados a priori de forma independiente de la medida del propio instrumento.

A un profesional que haga uso de un modelo de personalidad para obtener, por ejemplo, una descripción del perfil de personalidad de un individuo, le puede resultar quizás muy atractivo un modelo dotado de una gran cantidad de dimensiones de personalidad, asumiendo que así podrá disponer de una descripción más detallada de este sujeto. Bajo nuestra consideración, este modelo debiera demostrar, para cada dimensión de personalidad propuesta, que además de la existencia de un cuerpo teórico más o menos razonado sobre cada uno de estos factores, existe también evidencia empírica contrastada de que cada dimensión es capaz de diferenciar poblaciones específicas, bien definidas a priori, de forma independiente y extremas en cada uno de estos atributos. De lo contrario, consideraríamos que este modelo de personalidad no demuestra que vaya más allá de las presunciones teóricas y/o de la organización consistente de ítems en forma de numerosos factores. En definitiva, no debiera confundirse la calidad de un modelo o de su instrumento con la cantidad de dimensiones que supuestamente evalúa si para cada una de estas dimensiones no se aportan datos consistentes sobre alguna forma razonada de validez externa.

Nos estamos refiriendo al concepto de validez discriminante en términos de poblaciones diana que presentan en teoría determinados rasgos de personalidad extremos. El concepto “en teoría” es de especial relevancia conceptual y metodológica ya que condiciona la validez de los resultados que puedan inferirse. Una puntuación diferenciada entre dos poblaciones de sujetos en una o más dimensiones de personalidad no aporta necesariamente información relativa a la validez discriminante del instrumento si no se asume, o mejor dicho, si no se asumió a priori, que estas dos poblaciones debían ser diferentes en términos de personalidad de acuerdo a un cuerpo teórico o empírico bien fundamentado.

Una posible estrategia para la determinación apriorística que dos poblaciones son diferentes entre sí en términos de personalidad sería administrar un instrumento de medida de la personalidad bien consolidado y determinar en qué dimensiones se expresan estas diferencias, para posteriormente testar el comportamiento del nuevo cuestionario evaluando el grado en que este nuevo instrumento reproduce las diferencias encontradas en el test de referencia. No obstante debemos considerar esta estrategia como absurda desde un punto de vista aplicado, ya que si es posible disponer de un modelo de personalidad bien consolidado, instrumentalizado mediante un test de personalidad que actúe como patrón oro, resultaría poco interesante crear un nuevo instrumento que pretendiera competir con el primero, por considerarlo no válido o insuficiente, pero que debiera recurrir a él para contrastar su calidad.

Otra estrategia, mucho más verosímil y ya comentada, consiste en administrar el instrumento de interés a dos poblaciones que desde un punto de vista teórico debieran presentar rasgos de personalidad bien diferenciados. Consideramos que el área de trabajo que brinda la clínica, y en concreto la psicopatología, se constituye en un marco ciertamente atractivo, metodológicamente hablando, en el que analizar determinadas propiedades psicométricas. Sin duda también para el análisis de la validez discriminante.

En el contexto clínico es posible formular algunos apriorismos teóricos consistentes que pueden ser de gran interés para un modelo de personalidad normal, ya que permite la determinación de poblaciones caracterizadas por rasgos de personalidad distales o excepcionales desde un punto de vista frecuencialista. Este ha sido el punto de partida en el que se ha desarrollado el estudio recogido en el artículo relativo a la validez discriminante del ZKPQ (Art-3). En concreto, se abordó el análisis del comportamiento del instrumento en población con Trastorno Límite de la Personalidad (TLP).

Este trastorno de la personalidad se caracteriza por un patrón general de inestabilidad en las relaciones interpersonales, la autoimagen y afectividad, y una notable impulsividad que comienza al principio de la edad adulta y que se expresa en diversos contextos (DSM-IV-TR, 2000). Asumimos que el análisis del comportamiento del ZKPQ en este trastorno está plenamente justificado, al tratarse de un contexto en el que los componentes de inestabilidad emocional y ansiedad, así como el de impulsividad, se constituyen en elementos angulares, tanto del trastorno como del modelo. El estudio de la validez discriminante del ZKPQ en población TLP tomará especial relevancia en la exposición de la discusión de este documento, no solo por las consideraciones estrictamente psicométricas asociadas, si no también por las consideraciones que se derivan en términos del valor potencial de los modelos dimensionales de personalidad en la fundamentación teórica de la patología psiquiátrica, especialmente del Eje II.

4.3. *Baremos*

Afirmar que una determinada persona ha obtenido una puntuación de 18 en una determinada escala carece de valor alguno si se desconoce, por ejemplo, que este valor es el valor máximo que se puede obtener, que su puntuación tipificada z adquiere un valor de 2 o que el percentil asociado a este valor es de 50. El resultado de un individuo en una escala, intrínsecamente, es informativo. Es necesario relativizar este valor, contextualizarlo, para que pueda convertirse en un valor dotado de algún tipo de significación.

Aunque no se trata intrínsecamente de una propiedad psicométrica, la baremación es un elemento clave en el proceso de elaboración de un instrumento de medida. Si se desea que un instrumento de medida de la personalidad pueda ser empleado más allá de la investigación de grupos de sujetos es necesario proporcionar a los potenciales usuarios de este instrumento algún sistema de correspondencias que permita otorgar significado a las puntuaciones de los sujetos. Para que este sistema de correspondencias se dotara de la máxima calidad éste debiera ser lo más generalizable posible, es de decir, haberse dotado de una muestra representativa de la población objeto de estudio. El tipo de selección de los sujetos, y en concreto, el tipo de muestreo llevado a cabo, se constituyen en el elemento clave en la determinación de la calidad de una propuesta de baremación.

Un muestreo aleatorio es, metodológicamente hablando, la mejor estrategia disponible para asegurar la obtención de una muestra representativa, ya que solo es en este contexto en el que cualquier individuo de la población dispone de idéntica probabilidad de formar parte de la muestra. No obstante, resulta muy difícil, por no decir simplemente imposible, poder obtener un muestreo estrictamente probabilístico en el contexto de un estudio aplicado en humanos. Son muchas las condiciones que en la práctica merman la posibilidad de ejecutar un muestreo estrictamente probabilístico: por ejemplo, la presencia de personas institucionalizadas, personas sin las aptitudes necesarias para responder a un cuestionario, otras que se niegan a colaborar, etc. Los baremos de cualquier instrumento de evaluación están siempre sujetos a circunstancias que limitan la generalización de sus valores. Consideramos que el objetivo principal del investigador debiera ser el de minimizar en la medida de lo posible estos potenciales sesgos. Se tratará de procurar que las variables que condicionan de forma adversa la generalización de los resultados no tengan un impacto relevante.

Por otro lado, debe tenerse en consideración que no existe una población ideal sobre la cual desarrollar un baremo. Puede ser población general o clínica, por ejemplo. O quizás solo clínica, pero exclusivamente correspondiente a atención primaria. La población ideal para un baremo depende del tipo de sujeto que se desea describir y, sobretodo, de los términos en los que se proponga hacerlo. Si se desea describir el comportamiento de un sujeto respecto su sintomatología depresiva mediante un determinado instrumento, una puntuación de 20 no adquiere la misma interpretación si los baremos utilizados fueron obtenidos en el contexto de atención primaria que en unidades de internamiento psiquiátrico y, por supuesto, si los baremos provienen de población general. En el caso de la medida de la personalidad normal tampoco es posible determinar qué población es la idónea para baremar un instrumento. Probablemente la mejor opción depende del uso que se le desee dar potencialmente al instrumento. En un contexto clínico, quizás un usuario de un determinado modelo de personalidad podrá describir a un paciente con mayor precisión si dispone de baremos obtenidos en contexto psiquiátrico que no si estos baremos

son relativos a población general, población a la que este profesional no está expuesto. O quizás todo lo contrario. Este profesional necesita disponer de baremos relativos a población general para poder determinar en qué medida este patrón de personalidad merece algún tipo de atención más allá del que merecía una persona fuera de un contexto asistencial. Probablemente sería idóneo poder disponer de los dos recursos, haciendo un uso diferencial según el tipo de análisis o descripción perseguido y por supuesto asumiendo, en consecuencia, que las conclusiones a las que puedan llegarse utilizando una estrategia u otra podrán ser diferentes.

Obsérvese que todas estas consideraciones tienen sentido cuando el instrumento de interés se ubica en el conjunto de instrumentos concebidos bajo un punto de vista normativo. Es decir, instrumentos en los que la valoración que se realiza sobre un sujeto en particular depende del comportamiento medio de los sujetos que le son comparables en otras variables, más o menos numerosas, y que explícitamente organizan el sistema de baremos. Otros instrumentos serán contruidos bajo una concepción criterial, en los que existe un juicio apriorístico sobre qué individuos poseen presuntamente un determinado rasgo según se supera o no un determinado punto de corte. Este último enfoque tiene un uso muy difundido en entornos educativos o clínicos, contextos en los que es habitual la consideración de poseer o no un determinado dominio curricular o en los que existe una delimitación dicotómica entre normalidad vs no normalidad o patología. Esta estrategia, no obstante, no es relevante en el contexto de los modelos de personalidad normal, caso éste del ZKPQ, en los que los sujetos se ubican en un continuo dimensional que no presenta interrupción categorial alguna.

Hasta el momento no existía ninguna publicación que proporcionara datos normativos del ZKPQ que pudieran considerarse propiamente como baremos para población general. Ciertamente, no obstante, que existen múltiples tablas publicadas en artículos diversos en los que se exponen medias y desviaciones típicas de los sujetos participantes, pero la mayoría de ellos exploran, casi de forma exclusiva, estudiantes universitarios. El estudio de Gutiérrez-Zotes (2001), que supone el desarrollo de la primera versión del ZKPQ en español, se lleva a cabo con participantes de hasta 61 años, pero atendiendo a la media de edad reportada, que es inferior a los 22 años para ambos sexos, debe considerarse que las personas de mayor edad solo suponían una pequeña parte de todos los participantes explorados. Además, la distribución de la edad resulta para el lector desconocida e inespecíficamente dispuesta, ya que solo se reportan dos grupos de edad, uno comprendido entre los 16 y 23 años y otro entre los 24 y 61 años.

El objetivo fundamental del artículo *Spanish normative data of the Zuckerman-Kuhlman Personality Questionnaire in a general population sample (Art-4)* es proporcionar datos normativos relativos a población general para el ZKPQ según sexo, para cinco intervalos de edad (18-93 años) y ajustado por niveles de estudios (primarios, secundarios y universitarios). Con la satisfacción de este objetivo el estudio se constituye en la primera publicación de estas características en cualquier cultura utilizando el ZKPQ.

5. *Discusión* Todos los artículos compendiados en esta tesis versan sobre diversos aspectos relativos a la bondad del Modelo Alternativo de los Cinco Grandes y, en concreto, de las propiedades psicométricas del instrumento que sustenta dicho modelo, el Zuckerman-Kuhlman Personality Questionnaire (ZKPQ).

El modelo de los Cinco Alternativos se constituye, bajo nuestra consideración, en un modelo consistente tanto desde un punto de vista teórico como empírico. Se trata de un modelo que inserta sus raíces teóricas en una tradición de estudio de la personalidad que se nutre de las aportaciones de Eysenck, de la psicobiología y, por supuesto, del enfoque factorialista. No obstante, no podemos dejar de señalar que el modelo de los Cinco Alternativos es un modelo que ha tenido poca difusión, con un asentamiento hoy en día todavía limitado, especialmente en comparación con modelos que gozan de una consolidación mayor en la bibliografía internacional como es el caso, por ejemplo, del NEO, el modelo de los Cinco Grandes de Costa y McCrae (1992). Parte del esfuerzo depositado en los estudios aquí recogidos reside en dar al modelo la difusión e importancia que a nuestro juicio le corresponde.

Desde el punto de vista de la validez de un modelo de personalidad, una de las primeras consideraciones que debe realizarse sobre dicho modelo es la relativa al grado en que las respuestas que articulan el instrumento reproducen la estructura conceptual que lo inspira. Hablamos de la **validez de constructo** del modelo. Tal y como se expuso con anterioridad, el análisis factorial se ha constituido en un aliado fundamental para el análisis de este tipo de validez.

Dos de las publicaciones presentadas (*Art-1* y *Art-adjunto*) llevan a cabo análisis de la estructura factorial del ZKPQ. Los resultados de ambos trabajos, realizados en muestras de 933 estudiantes universitarios y 1000 participantes recogidos de población general indican que la estructura penta factorial propuesta por el modelo de personalidad se reproduce consistentemente en población española. La solución de cinco factores es la solución más ajustada a la matriz de intercorrelaciones de los 89 ítems del ZKPQ³. En el primer estudio (*Art-1*), además, se obtienen índices de congruencia (índices Tucker de congruencia) con población norteamericana comprendidos entre 0,84 y 0,96, lo que es indicativo de una gran similitud factorial entre las dos poblaciones.

Este satisfactorio comportamiento factorial es coincidente con estudios realizados por otros equipos (Aluja et al., 2004; De Pascalis & Russo, 2003; Gutiérrez-Zotes et al., 2001; Ostendorf & Angleitner, 1994; Siomi et al., 1996; Wo et al., 2000; Zuckerman et al., 1993). El grado de congruencia obtenido permite afirmar que el modelo presenta una adecuada reproductividad entre poblaciones y culturas diferentes, elemento éste de gran valor conceptual para un modelo de personalidad y por ende, de la validez del instrumento. El segundo estudio de factorización, realizado en población general (*Art-adjunto*), posee además la especial relevancia de tratarse de la primera publicación con el ZKPQ que se lleva a cabo en este tipo de población, no solo en nuestra cultura, sino en cualquier otra.

Las varianzas totales explicadas obtenidas en estos dos estudios de factorización son enormemente coincidentes entre sí. En el caso de población universitaria este

³ Los 10 ítems de la escala de Infrecuencia se excluyen del análisis puesto que ésta no constituye dimensión de personalidad alguna.

valor asciende a un 25% mientras que en el caso de población general la cifra alcanza el 26%. Estas cifras encajan plenamente con lo observado en otras adaptaciones del ZKPQ, en las que se encontrarán valores comprendidos entre el 20% de la adaptación china y el 28% de la adaptación japonesa.

Vinculado también al comportamiento estructural del instrumento, hemos reportado para cada población en las diversas publicaciones presentadas los respectivos índices de **consistencia interna** de cada dimensión de personalidad. Todos los coeficientes obtenidos, y en todas las poblaciones, son valores ciertamente satisfactorios. Este conjunto de resultados redunda sobre la consideración de la replicabilidad del modelo en términos del grado de interrelación entre los ítems constitutivos de cada una de las dimensiones del modelo.

Aunque la valoración global de estos índices de consistencia interna es, como se apuntó, satisfactoria, avalamos la idea de que es necesario reportar e interpretar la consistencia interna de un instrumento de evaluación de forma contextualizada. Como se expuso en la introducción de este documento, los coeficientes de fiabilidad, por supuesto también los relativos a la consistencia interna, están condicionados parcialmente por el grado de variabilidad que presenta la variable objeto de interés cuando ésta es recogida en una determinada muestra. Para dar ejemplo de esta circunstancia psicométrica nos servimos del estudio en el que se exploró la validez discriminante del ZKPQ (*Art-3*). En este estudio se analizó una muestra de 74 pacientes con TLP y 148 sujetos tomados de población general. Si se observa con detalle la magnitud de las desviaciones típicas vinculadas a cada una de las cinco variables de personalidad y los correspondientes coeficientes de consistencia interna, se observará una asociación directa entre la magnitud de los coeficientes de fiabilidad y desviaciones típicas: los coeficientes de fiabilidad más elevados tienden a asociarse a desviaciones también más elevadas (página 183 de dicho artículo).

El caso paradigmático de este comportamiento es el correspondiente a la escala N-Anx. En el caso del grupo TLP la desviación típica de esta variable toma un valor de 2,76. Este valor en el caso de la muestra de sujetos de población general asciende a 4,82. Véase que, aunque en el caso de la muestra con TLP el tamaño de la muestra es la mitad del de la muestra de población general, su desviación es marcadamente inferior. Incluso para las otras cuatro variables de personalidad en la muestra de TLP se trata de una desviación especialmente reducida, ya que la desviación típica inmediatamente superior, correspondiente a la variable Sy, asciende ya a un valor de 3,46. Hipotetizamos que, muy probablemente, esta especial homogeneización de la muestra clínica en esta variable es consecuencia de seleccionar a los sujetos en base a la presencia de rasgos clínicos afines a la dimensión de personalidad neuroticismo y ansiedad, con el objetivo de formar parte del ensayo clínico del que, posteriormente, se obtendrá la muestra de sujetos que participará de la muestra analizada en este artículo. Es relevante señalar que en esta variable, y en esta muestra clínica, el coeficiente de consistencia interna toma un valor 0,73, mientras que en la muestra de población general este mismo valor asciende a 0,87.

La variable conceptualmente más relevante para el trastorno que nos ocupa (N-Anx) es en la que precisamente se observa un rendimiento en términos de consistencia interna menor. Es efectivamente un valor satisfactorio, pero es el coeficiente de menor magnitud de los cinco correspondientes a cada variable de

personalidad explorada. Se trata solo de un ejemplo de como un determinante metodológico, en concreto, del filtro que suponen los criterios de inclusión y exclusión que se aplican en una investigación y que por tanto delimitan y encuadran una determinada población de pacientes, condiciona la magnitud de la desviación típica, y a través de ésta, la magnitud de los índices de consistencia interna resultantes. Y ésta consistencia interna condicionará, inevitablemente, la precisión de las estimaciones que se realicen con este instrumento a partir de esta muestra, lo que a su vez impactará sobre la magnitud de las diversas formas de validez externa que puedan inferirse. En definitiva, no puede desvincularse la fiabilidad de un instrumento, ni por tanto su validez, de la población en la que se analizan dichas propiedades.

No podemos abandonar el concepto de consistencia interna sin exponer el comportamiento de la escala Agg-Host. Esta escala tiende sistemáticamente a presentar un resultado menos satisfactorio en las diversas muestras exploradas. Este comportamiento diferencial es coincidente con lo observado en diferentes muestras, lenguas y culturas (De Pascalis & Russo, 2003; Wu et al., 2000). Se ha sugerido que este comportamiento podría estar relacionado con diferencias culturales respecto a población de EEUU. Este aspecto deberá ser abordado en investigaciones futuras.

Un segundo tipo de fiabilidad pertinente para un instrumento como el ZKPQ es el relativo a la **estabilidad temporal**. Los resultados que reportamos en la primera de las publicaciones (*Art-1*) avalan la estabilidad temporal del instrumento ya que las correlaciones obtenidas presentan un adecuado nivel. Se trata de valores comprendidos entre 0,77 en el caso de la escala de Agg-Host y de 0,91 en el caso de la escala ImpSS.

Estamos convencidos, no obstante, que esta propiedad debe ser mejor explorada por diversas razones. La primera de ellas es debida a que únicamente ha sido explorada en una población de sujetos, en este caso población de universitarios. Como medida de fiabilidad que es, la estabilidad temporal del instrumento debiera ser analizada también en poblaciones diferentes. El objetivo último perseguido debiera ser, de nuevo, poder asumir con las debidas garantías que se trata de un resultado razonablemente generalizable y no únicamente vinculado a una determinada población. La segunda de las consideraciones que exponemos es la relativa al tiempo transcurrido entre los dos registros. En el caso del artículo propuesto, este análisis se realizó, por razones de practicidad, habiendo transcurrido solo 15 días. Atendiendo a la estabilidad del constructo que nos ocupa debiera analizarse la fiabilidad del instrumento más allá de este período temporal. Sin duda alguna se trata de otro aspecto psicométrico que debe ser explorado con mayor precisión, ampliando el tiempo de demora entre evaluaciones a la vez que analizándolo en poblaciones diferentes con el objetivo de acotar el grado de generalización de este estimador.

Volviendo a las consideraciones entorno a los criterios de validez del ZKPQ, retomamos uno de los conceptos que hemos percibido y expuesto como especialmente importante en el desarrollo de la introducción de este documento. Se trata del concepto de **validez discriminante**, elemento que consideramos de gran importancia en la validación de un instrumento de medida de la personalidad.

Una forma de analizar esta propiedad, y que a la vez se convierte en una atractiva forma de validación externa del instrumento, consiste en determinar el

grado en que el instrumento es capaz de diferenciar poblaciones que dispongan de rasgos de personalidad presumiblemente extremos o claramente diferenciados de otras poblaciones de sujetos. Por supuesto asumiendo que este tipo de análisis, centrado en determinación de variables externas e independientes del propio instrumento y apriorísticamente definidas, deberá ser combinado también con la clásica y frecuente exploración basada en el análisis de la validez interna del instrumento y que se fundamenta esencialmente en el análisis de la estructura factorial del mismo.

Consideramos que si la personalidad es un elemento fundamental del trastorno objeto de interés, éste debiera ser abordado exigiendo al instrumento ciertas garantías de diferenciación entre poblaciones específicas. Esta exigencia proponemos que, además, sea dirigida más allá de la simple búsqueda de significación estadística.

Mucha de la investigación existente respecto al análisis de las diferencias entre poblaciones en términos de dimensiones de personalidad suele descansar exclusivamente sobre la determinación de si se observa o no una discrepancia estadísticamente significativa (y por supuesto esta consideración no se restringe exclusivamente al estudio de la personalidad). Esta estrategia, si se formula exclusivamente en estos términos estadísticos, consideramos debiera relativizarse o simplemente desecharse, ya que supone asumir que aquello que es relevante en términos de significación probabilística posee automáticamente algún tipo de relevancia fuera de este contexto, y en concreto en términos de diferencias de personalidad. Hallar una discrepancia estadísticamente significativa entre dos poblaciones puede ser fruto casi exclusivamente, de la cantidad de sujetos que están incluidos en la prueba de significación. No debe olvidarse que la obtención de un efecto significativo está altamente determinado por la magnitud de las dos o más muestras implicadas en el análisis. Siempre se cumple que una discrepancia pequeña entre dos medias, minúscula incluso, puede ser estadísticamente significativa si se dota de una cantidad de sujetos suficiente. En estudios en los que se está utilizando una cantidad ingente de participantes lo difícil es no obtener algún efecto que obtenga significación estadística. Es por ello que consideramos que no es adecuado formular el concepto de *validez discriminante comprobada* entorno a la obtención simplemente de una diferencia estadísticamente significativa, por muchas correcciones que se hagan del nivel de significación, excepto que el investigador pueda inyectar alguna dosis de relevancia conceptual de rango clínico, predictivo, epidemiológico, de capacidad de clasificación, etc, y por tanto más allá de lo *mucho* o *poco* influyente que se estima el azar sobre un efecto, que es lo único que se está testando en una prueba de hipótesis estadística.

Obsérvese un ejemplo de esta apreciación: Si se toma el artículo relativo a la validez discriminante del ZKPQ en población con TLP (*Art-3*), podríamos partir del supuesto que nuestro interés fundamental descansa sobre la escala de Infrecuencia (Inf), que es la única escala que en este estudio no presenta una discrepancia estadísticamente significativa entre las dos muestras analizadas (su nivel de significación es $p=0,192$, cuando todos los demás factores disponen de una $p<0,032$). Obsérvese, además, que la discrepancia entre las dos medias, que es de 0,22 unidades, es claramente muy reducida, especialmente si se tiene en consideración que el rango teórico de esta variable es de 10 unidades o de que las respectivas desviaciones típicas son 1,12 en el caso del grupo TLP y 1,18 en el caso del grupo de población general. Podríamos calcular la cantidad de sujetos que debieran ser incluidos para determinar que una media de 1,30 en el grupo de

sujetos con TLP es significativamente diferente, estadísticamente hablando, de una media de 1,08 en esta misma variable en el grupo de sujetos correspondientes a población general. Manteniendo constantes las respectivas medias, porque se asume que los nuevos sujetos que se incorporarían serían esencialmente parecidos a los ya presentes, y asumiendo razonablemente que las nuevas correspondientes varianzas serían ligeramente inferiores, como consecuencia de aumentar la muestra y por tanto el tamaño del denominador que condicionará su cálculo, sería necesario recoger el ZKPQ de un total de 148 personas afectadas de TLP y de 300 personas de población general:

Se asume una desviación típica común de 1,1, cifra solo unas décimas inferior a la desviación típica común actual, como consecuencia de la ampliación de la muestra, y un contraste bilateral. Se asume, además, que el investigador no otorga relevancia alguna a la potencia asociada al resultado significativo (circunstancia ésta muy común en la inmensa mayoría de publicaciones en psicología o psiquiatría, no solo sobre Personalidad): se asume un potencia mínima de 0,51. El nivel de confianza se mantiene al 95%.

Considérese que esta cantidad de participantes podría ser satisfecha con relativa facilidad en una Unidad o Servicio en el que tratasen este tipo de patología psiquiátrica. Con estos dos tamaños muestrales las dos medias serían estadísticamente diferentes a un nivel de significación estándar del 5%. Como consecuencia de este efecto estadístico, ¿se afirmaría entonces que la escala de Infrecuencia posee algún tipo de atributo discriminativo? Entendemos que no.

Lo que se ha pretendido con el análisis de la validez discriminante del ZKPQ en población con TLP (*Art-3*) ha sido abordar las diferencias entre las dos muestras implicadas aportando criterios de ajuste que puedan ser sustanciales, interpretativamente hablando, a diferencia de lo que supondría reportar las diferencias en términos exclusivamente de significación estadística o incluso de valores de *d* de Cohen (tamaño del efecto). Se han utilizado parámetros diversos para enjuiciar con mejor criterio esta capacidad discriminante. Se utilizaron parámetros como la Sensibilidad (capacidad para detectar sujetos afectados), Especificidad (capacidad para detectar sujetos no afectados), Valor Predictivo Positivo (cantidad de sujetos con un resultado positivo en la prueba que son realmente positivos) y Valor Predictivo Negativo (cantidad de sujetos con un resultado negativo en la prueba que son realmente negativos). Se complementaron estos índices con valores Kappa y áreas bajo la curva. Todos los anteriores parámetros fueron calculados para cada punto de corte estudiado. Esta aproximación, que está lejos de ser novedosa en investigación epidemiológica, no suele ser frecuente, desgraciadamente, en investigación en personalidad.

El modelo multivariante obtenido constituido por las dimensiones de personalidad N-Anx, ImpSS y Act, y asumiendo un determinado punto de corte probabilístico (se trata de un modelo *logit*, en el contexto de una regresión logística) supone la obtención de una Sensibilidad del 87,8%, una Especificidad del 87,8% y un índice Kappa de 0,73. Los Valores Predictivos Positivo y Negativo, aunque dependen de la prevalencia del rasgo objeto de estudio, y por tanto están condicionados por el tipo de diseño metodológico, en concreto por la razón de sujetos entre los dos grupos explorados, toman valores ambos superiores al 80%. Estos resultados son claramente superiores a los obtenidos por otros modelos de personalidad más consolidados en la bibliografía internacional o incluso por sistemas de evaluación específicos para TLP (Fossati, Maffei, Bagnato, Donati, Fiorilli, Novella, et al., 1998; Gutiérrez, Sangorrín, Martín-Santos, Torres & Torrents, 2002).

Es relevante destacar que la variable que presenta mayor capacidad para diferenciar los dos grupos es la dimensión N-Anx, que es la que presenta el tamaño del efecto más elevado ($d=1,86$). En segundo lugar interviene la variable ImpSS ($d=0,83$) y finalmente la variable Act ($d=0,31$), ésta última en sentido contrario a las dos anteriores, es decir, asociándose una menor puntuación al grupo TLP. Cabe mencionar que, en contexto de análisis bivariante, las cinco dimensiones del ZKPQ presentaban significación estadística, pero en un contexto de análisis multivariado (aplicando una estrategia de entrada condicional) solo las anteriores tres variables se constituyen en el modelo final. Tal y como se indicó con anterioridad, la escala de Infrecuencia no presentaba discrepancia alguna entre las dos muestras, lo que puede ser interpretado como una medida de control.

El análisis fue también llevado a cabo sobre las subescalas de personalidad que proporciona el modelo, esto es, utilizando aquellos rasgos en los que determinadas dimensiones se estructuran (Act: GenAct y WorkAct; ImpSS: Imp y SS; Sy: Parties y Isol) además de las dimensiones directamente cuando éstas no disponen de subescalas (N-Anx y Agg-Host). Siguiendo la misma estrategia de análisis descrita en el caso de las dimensiones se determinó finalmente que las variables constitutivas del modelo eran N-Anx, Imp y GenAct.

Este resultado supone afirmar que el componente impulsivo (Imp), pero no el relativo a la búsqueda de sensaciones (SS), es el que presenta capacidad de diferenciación sobre el trastorno. Por otro lado, de la dimensión de actividad (Act), es exclusivamente la subescala actividad general (GenAct), en sentido inverso, pero no la relativa a esfuerzo en el trabajo (WorkAct), la que dispone de capacidad discriminante entre las dos poblaciones. La dimensión de Neuroticismo-Ansiedad (N-Anx) permanece en el modelo con un peso siempre preponderante.

De entre los dos modelos analizados, el correspondiente al análisis de las dimensiones o el correspondiente a las dimensiones y subescalas, se optó por el primero, puesto que se constituía en el modelo con mayor capacidad discriminante entre las dos poblaciones objeto de estudio. Este criterio se fundamenta sobre la base de una valoración esencialmente matemática, pero no debe menospreciarse el valor conceptual que puede suponer enfocar el Trastorno Límite de la Personalidad en términos de personas que presentan un perfil comportamental claramente impulsivo, pero no necesariamente buscadoras de sensaciones, y sin que esta impulsividad deba confundirse con un rasgo de extroversión, componente en el que muestran un comportamiento poco pronunciado e incluso inferior a la población general. Por supuesto todo ello en el contexto de una marcada tensión y preocupación emocional así como alto componente ansioso.

Estos resultados son consistentes con la consideración de que los rasgos de personalidad fundamentales que vertebran el Trastorno Límite de la Personalidad pueden ser consistentemente descritos en términos dimensionales de personalidad normal. Este enfoque contrasta con la escasa atención que ha tenido la investigación de la estructura de la personalidad normal en el contexto de los trastornos de personalidad. Algunas voces, hace más de 20 años, afirmaban ya que los trastornos de personalidad podían ser más apropiadamente representados por modelos dimensionales que no categoriales (Frances, 1980; Eysenck, 1987). No obstante el DSM-IV-TR continúa siendo un instrumento esencialmente categorial en el enfoque de los trastornos de la personalidad.

Estamos convencidos que el fuerte arraigo que tiene todavía este sistema categorial en la conceptualización de los trastornos del Eje II estriba en que resulta mucho más familiar para muchos clínicos, legitima los esfuerzos realizados hasta el momento en investigación y tratamiento, y facilita la comunicación entre colegas (Livesley & Jackson, 1992). Además es común también entre estos profesionales la percepción de que existe un pobre consenso entre los diferentes modelos dimensionales (Ball, 2001).

Respecto al primer conjunto de consideraciones compartimos el punto de vista de Clark (1993) cuando afirma que los argumentos de familiaridad, tradición o sencillez conceptual son útiles para juzgar o comparar dos sistemas que hayan mostrado equiparable validez, pero no pueden ser una estrategia argumental, ni científica ni clínica, de primera línea. Consideramos que las limitaciones en forma de excesiva comorbilidad o solapamiento entre categorías, pobre validez convergente y discriminante, puntos de corte diagnósticos arbitrarios, dilemas de clasificación, pérdida de información en los casos no prototípicos o divergencia respecto la práctica cotidiana (Ball, 2001; Verheul, 2005; Widiger, 1992; Widiger & Frances, 2002) son condiciones que evidencian que el enfoque categorial del Eje II del DSM presenta limitaciones más que relevantes. Respecto a la segunda de las reticencias, relativa a la percepción de que los modelos dimensionales no se ponen de acuerdo entre sí, debemos afirmar que no lo vemos exactamente de esta manera. Aunque ciertamente no existe una perfecta unanimidad respecto el número exacto, denominación o número de elementos específicos dentro de cada dimensión, sí existe un destacado consenso respecto la existencia de un número limitado de dimensiones de orden superior. Además, cada una de estas dimensiones, están definidas por un determinado número de rasgos de orden inferior que podrían definirse en términos de desadaptabilidad/maladaptabilidad. Quizás el mayor avance posible descansa no tanto en la potenciación de determinados instrumentos (por ejemplo ZKPQ o NEO) sino en establecer un marco teórico general que integre diversos modelos (Verheul, 2005). Tal y como indican Widiger y Simonsen (2005) sería deseable que la nueva versión del DSM, en proceso de elaboración en la actualidad, incorporara una integración de los distintos modelos dimensionales en el contexto de una estructura jerárquica en la que cada contribución y principales ventajas de cada modelo pudieran ser tomadas en consideración.

Investigadores y clínicos que estén interesados en incluir en sus sistemas de diagnóstico y evaluación una medida del funcionamiento de la personalidad normal a la vez que una descripción de rasgos de personalidad no adaptativos pueden hacer uso del ZKPQ. No obstante queda mucho camino por recorrer todavía ya que solo disponemos de una publicación que se inserte directamente y de forma explícita en investigación clínica. Se nos ocurren otras investigaciones dirigidas al análisis de cual es la capacidad predictiva del modelo respecto al pronóstico del paciente así como de la determinación del grado de especificidad del modelo matemático en términos de capacidad discriminante del ZKPQ.

Pudiera pensarse que un modelo dotado únicamente de cinco dimensiones de personalidad posee intrínsecamente una baja capacidad para mostrarse específico y discriminante ante la multitud de trastornos en los que presuntamente la personalidad juega un papel relevante. Discrepamos de esta perspectiva. En primer lugar, porque la presencia de muchas más dimensiones de personalidad no tiene por qué traducirse en una mayor validez del modelo. Apuntamos con anterioridad que no puede confundirse la cantidad de

dimensiones que proporciona un modelo con la potencial magnitud de su validez externa (convergente, discriminante, predictiva, etc). En segundo lugar, debe considerarse que la capacidad del ZKPQ para diferenciar poblaciones diferentes no depende exclusivamente de la cantidad de dimensiones constitutivas de un determinado modelo. Depende también de la combinación entre ellas, y del peso y sentido diferencial con el que cada factor participa en el modelo final. Esta multivariada y ponderada combinación de los cinco rasgos de personalidad del ZKPQ multiplica enormemente la cantidad de modelos potencialmente predictivos.

Como se apuntó, son muchas las líneas posibles de investigación, algunas de ellas dirigidas a la determinación de hasta qué punto el modelo de Zuckerman es capaz de realizar descripciones específicas ante poblaciones también específicas en términos de personalidad normal o bien se constituye únicamente como un detector inespecífico de malestar psicopatológico. Nuestro equipo ya ha comenzado a explorar otras muestras clínicas, como por ejemplo el Trastorno por Déficit de Atención en Adultos o el Trastorno Obsesivo Compulsivo, además de otras poblaciones. Los resultados de estos trabajos deberán dilucidar el sentido de la respuesta a la pregunta planteada.

En el momento en el que un instrumento de personalidad deja de ser exclusivamente objeto de interés en un contexto de investigación de grandes grupos muestrales de sujetos, de población exclusivamente universitaria o de investigación básica, y comienza a percibirse como un modelo de interés para la descripción del perfil de personalidad de personas concretas, es el momento en el que el instrumento debe proporcionar estrategias que permitan interpretar el perfil de cada una de estas personas. Parte del esfuerzo de nuestro trabajo se ha dirigido a satisfacer este objetivo, proporcionando **baremos obtenidos de población general**. Disponer de información de esta naturaleza debiera suponer para el ZKPQ dar un paso más hacia su difusión como modelo de personalidad útil en contexto aplicado.

El estudio que sustenta estos baremos (*Art-4*) se ha realizado en población comunitaria, extraída de aulas, asociaciones, en el contexto de chequeos médicos, etc. Aunque está lejos de tratarse de un muestreo aleatorio, mecanismo garante de una adecuada representatividad poblacional, la selección de la muestra propuesta en este trabajo pretende impactar sobre diversos ámbitos de la actividad cotidiana de una gran cantidad de personas. Además, esta recogida de información se lleva a cabo en circunstancias en las que las respuestas se recogen de forma anónima, favoreciendo así respuestas con una baja o nula deseabilidad o condicionamiento. Consideramos que estas condiciones, que aunque como apuntamos no son circunstancias metodológicamente óptimas, sitúan al ZKPQ en una situación menos comprometida en términos de generalización que otros instrumentos de personalidad baremados en contextos muy restrictivos o simplemente poco generalizables, o incluso en los que puede existir un potencial sesgo de las respuestas recogidas como consecuencia del contexto de supervisión en el que se obtienen. Considérese por ejemplo el NEO-PI-R en su versión española (Costa & McCrae, 1999) en el que, aunque dispone de una muestra de baremación muy elevada, está constituida en su totalidad por sujetos extraídos en el contexto de un proceso de selección de personal para optar a la obtención de un lugar de trabajo (página 77 del manual de la adaptación española de este instrumento).

En nuestro estudio de baremación (*Art-4*) se proporcionan datos obtenidos de

más de 1600 participantes divididos en cinco intervalos de edad, de los 18 a los 93 años, así como por sexo, lo que supone 10 combinaciones por cada una de las cinco dimensiones del modelo, más la relativa a la escala de Infrecuencia. Puesto que incluir tres niveles académicos en el sistema de baremación hubiese supuesto multiplicar en exceso la cantidad de combinaciones posibles se optó por minimizar su impacto en la estimación de cada una de las medias y desviaciones típicas generadas. Por esta razón el nivel académico fue ajustado matemáticamente dentro de cada combinación de edad y sexo.

Estos datos normativos están expresados en términos de medias y desviaciones típicas. Aunque somos conscientes que las puntuaciones derivadas en forma de T son puntuaciones quizás más familiares para muchos usuarios que trabajan en el campo aplicado hemos preferido reportar los baremos solo con medias y desviaciones típicas para que cualquier interesado en el modelo pueda realizar de forma sencilla y directa comparaciones con otros estudios en los que se lleve a cabo un uso del modelo, como pudiera ser, por ejemplo, en un contexto transcultural. Si se deseara una interpretación tipificada de una puntuación directa correspondiente a un sujeto, sería necesario únicamente una sencilla transformación lineal de la misma, ya sean en forma de z o de T, por ejemplo. Como es conocido, la transformación en forma de z (y por supuesto de T, ya que se trata de una puntuación derivada de z) presupone la normalidad de la distribución de la variable objeto de interés. Puesto que no se abordan en esta publicación los análisis de distribución de cada una de las 10 condiciones edad-sexo no podemos afirmar que todas ellas satisfagan este principio de normalidad. No obstante, si se observa la distribución de las cinco variables de personalidad para la muestra total (*Art-4*, páginas 326 y 327) parece razonable asumir una distribución gaussiana.

Los usuarios de estos baremos deben recordar, nuevamente, que aunque sea posible e incluso útil hacerlo, el ZKPQ no supone una evaluación clínica de la personalidad. Esto supone que, aunque en estos momentos es posible comparar el perfil de personalidad de un individuo en contexto clínico respecto a población general, el ZKPQ no puede percibirse como un instrumento detector o descriptor de patología. No hay puntos de corte ni es posible inferir categorías diagnósticas. O como mínimo no los hay desde el punto de vista de un instrumento criterial o desde una perspectiva clínica *clásica* (véase páginas 184-185, *Art-3*).

Parte de la investigación que se desarrolle en el futuro deberá proporcionar poblaciones de baremación más específicas, por ejemplo población clínica, para que cada usuario pueda disponer de una descripción más ajustada del perfil del sujeto evaluado. Además, y cuando exista una investigación debidamente contrastada que lo avale, debería proponerse mecanismos de estimación que, ahora sí, permitan determinar la probabilidad de posible trastorno, otorgando así a las variables de Personalidad un rol mucho más relevante del que hoy goza. Esto debiera convertirse en un estímulo para la mejora de los mecanismos de diagnóstico presentes así como en una mayor precisión de las predicciones de la evolución terapéutica de los pacientes evaluados.

En un contexto como éste, básicamente aplicado, toma especial relevancia el segundo de los artículos compendiados (*Art-2*). Si bien este estudio fue planteado en su origen para satisfacer un objetivo esencialmente psicométrico, y más allá del valor que tome en términos de validez aparente o de constructo, consideramos que el estudio adquiere en este momento un carácter eminentemente ecológico. Afirmar que lo descrito sobre uno mismo mediante un

instrumento de evaluación de la personalidad converge esencialmente con lo descrito por otros sobre nosotros supone una propiedad de gran interés en un contexto aplicado. Esto es así ya que permite la corroboración o contraste de la información proporcionada por la propia persona, o bien permite obtener información consistente en aquellas condiciones en las que la persona simplemente no puede aportarla con las debidas garantías. Sin duda alguna estas consideraciones toman especial interés en contexto clínico, situación ésta en la que la necesidad de contrastar información exclusivamente reportada por el paciente se constituye en una condición más que frecuente.

La satisfactoria **convergencia** hallada en este estudio entre la información autoreportada y heteroreportada es posible, entendemos, gracias al tipo de conductas que integran el listado de reactivos del ZKPQ. Solo ante un conjunto de conductas debidamente definidas, razonablemente objetivas y acotadas es posible minimizar la discrepancia entre informantes diferentes. Una de las ventajas que proporciona además esta estrategia de análisis es que los artefactos que influyen los autoinformes (por ejemplo tendencias de respuesta condicionadas por procesos de aquiescencia o de deseabilidad social) pueden considerarse esencialmente independientes de aquellos a los que se ve sometido el observador (efectos de halo, de estereotipos, etc), por lo que la presencia de acuerdo entre la dos fuentes de información se constituye en una potente fuente de validez del instrumento. Lo que en otro contexto metodológico y con otros objetivos podría ser interpretado en términos de fiabilidad entre observadores es, en el caso que nos ocupa, un dato que interpretamos en términos de validez de consenso del instrumento (McCrae & Costa, 1983).

Para finalizar esta discusión consideramos necesario recoger algunas impresiones relativas al comportamiento de la escala de Infrecuencia del ZKPQ. Se trata de una escala que, como se apuntó con anterioridad, proporciona una medida de control en términos de respuestas *inadecuadas*, cargadas de deseabilidad social, o incluso de inatención en el momento de responder al test. A diferencia de determinadas escalas comparables incluidas en otros instrumentos de personalidad (por ejemplo el caso del MCMI-III), la escala de Infrecuencia no se considera estrictamente como una medida de validez, sino más bien se trata de una escala comparable a la de sinceridad, como la presente en el EPQ de Eysenck. No existen para esta escala puntos de corte ni juicios consistentes respecto a cuando debe ser eliminado un sujeto por considerar que sus respuestas no merecen la debida credibilidad.

En otros estudios de adaptación de instrumentos de evaluación de la Personalidad en nuestro entorno cultural (Eysenck, García, Torrubia, Ávila, & Ortet, 1992; Gomà-i-Freixanet, 1995; Gomà-i-Freixanet & Wismeijer, 2002) se observa que las puntuaciones en la escala de Sinceridad del EPQ tienden a ser diferentes de las encontradas en participantes ingleses, mostrándose éstos últimos más sinceros. Este mismo patrón de respuesta se observa en la escala de Infrecuencia de la adaptación al italiano del ZKPQ (De Pascalis & Russo, 2003) o en la versión en castellano (Gutiérrez-Zotes et al., 2001), hallándose puntuaciones siempre más elevadas que en la versión americana. Como consecuencia de este comportamiento transcultural diferencial, no nos parece apropiado tomar como referencia el punto de corte propuesto por Zuckerman y Kuhlman (1993) en población americana, según el cual se podría prescindir de las valoraciones obtenidas de personas que presentan en la escala de Infrecuencia una puntuación igual o superior a 4 puntos.

Otros instrumentos de evaluación de la personalidad han resuelto con aparente eficacia esta compleja circunstancia de toma de decisión. En el caso ya comentado del MCMI-III, por ejemplo, se proporciona al usuario diversas estrategias para abordar la calidad de las respuestas obtenidas: se proporcionan escalas de validez, de sinceridad, de deseabilidad social, etc, todas ellas con el objetivo de que el usuario pueda conocer cuán válidas o creíbles resultan las respuestas de un individuo (Millon, Davis & Millon, 2007). No obstante, la determinación de la calidad de la respuesta de un individuo se determina según criterios que resultan simplemente desconocidos. En el manual de este instrumento no se reportan datos empíricos contrastados, ni siquiera juicios o parámetros bien fundamentados que puedan ser usados como criterio para enjuiciar, por ejemplo, por qué la obtención de una puntuación en una escala de sinceridad inferior a 34 o superior a 178 debe suponer considerar la respuesta del sujeto como *inválida*. Ni por qué cuando existe una puntuación superior a 75 en otra escala debe ponerse en tela de juicio la valoración del sujeto puesto que éste presenta *una tendencia a despreciarse o devaluarse, presentando unas dificultades emocionales o personales más problemáticas de las que normalmente se encontraría en una revisión objetiva* [sic] (página 129 del manual).

Somos plenamente conscientes que la determinación de lo que es presuntamente una respuesta inapropiada o simplemente no válida exige un abordaje metodológico complejo y difícil de contrastar, especialmente si se considera que, bajo nuestro juicio, esta validación debiera complementarse con la búsqueda de parámetros externos e independientes. Pero si no es así, puede caerse con relativa facilidad en la formulación de soluciones que acaban siendo tautológicas o simplemente arbitrarias por no haber sido contrastadas empíricamente, sobretodo si, además, se debe asumir que un mismo punto de corte para inferir invalidación o sospecha de escasa credibilidad tiene que servir para valorar cualquier población o cultura.

Para el ZKPQ, consideramos que es necesario aportar investigación que proponga indicadores consistentes, y especialmente razonados, que permitan validar esta escala en nuestro entorno cultural y proponer un punto de corte, como mínimo, para población general. Este punto de corte debiera permitir al usuario del instrumento disponer de una estimación consistente de la calidad de la información recogida de una determinada población sin tener que prescindir sistemáticamente de una gran cantidad de sujetos como consecuencia de aplicar un criterio excesivamente conservador. El análisis de este punto de corte, adecuadamente contrastado en nuestro entorno cultural, debiera redundar en una mejora, como mínimo, de la magnitud de la validez externa de las medidas, generando solo un número reducido de falsos positivos y negativos en un contexto de análisis grupal, o bien en la no generación de sospechas poco fundamentadas cuando un único individuo es evaluado.

6.
*Consideraciones
finales y estudios
futuros*

Los datos que se aportan en las diversas publicaciones compiladas se constituyen en elementos psicométricos que nos informan sobre la calidad del modelo que nos ocupa así como de las propiedades psicométricas del ZKPQ. Aportan información diversa con el objetivo de que cualquier usuario pueda elaborar un juicio lo más ajustado posible respecto en qué medida el modelo le ofrece razonables garantías de validez y fiabilidad.

No obstante, no puede asumirse que con esta información el modelo y el instrumento están ya validados. Nos resistimos a afirmar que el ZKPQ es un instrumento validado. Obsérvese que no hemos dado a ninguno de nuestros trabajos el título de validación del ZKPQ. Asumimos que el modelo y el instrumento están en proceso constante de validación, de la misma manera que cualquier otro modelo y cualquier otro instrumento de evaluación, sea o no de personalidad. Entendemos que se trata de un error conceptual afirmar que los instrumentos pueden considerarse validados como si la validación fuera una entidad absoluta, unidimensional, estática, independiente de las poblaciones a las que se aplica o ajena a los cambios de diversa índole tanto científica y social que van aconteciendo. No existe una sola forma, ni dos, ni tres para otorgar validez a un instrumento, lo mismo acontece para la fiabilidad. No existe una única población sobre la que analizar una propiedad, no existe una única estrategia metodológica para abordar un único propósito, ni una única técnica estadística que pueda dar fe absoluta sobre un objeto psicométrico, y lo que es más importante, no existe un único posicionamiento científico, económico, ético, cultural, desde el que enjuiciarlo. Se trata siempre de un enfoque multidimensional y de un proceso y negociación continuos.

Validar un instrumento, o más apropiadamente dicho, aportar datos sobre la validez y fiabilidad de un instrumento, consiste en tomar partida de un posicionamiento teórico (y por tanto no de otros), supone analizar y reanalizar lo realizado hasta el momento, significa contextualizar, valorar las consecuencias de la toma de determinadas decisiones, relativizar, comparar, converger, participar de un proceso dinámico que no finaliza nunca hasta que el modelo deja de considerarse adecuado u óptimo, o simplemente, desde un punto de vista funcional, hasta que se considere que el modelo o el instrumento ya no sirve o no sirve como lo hacía hasta el momento.

Los trabajos que tenemos intención de satisfacer en el futuro, y para los que en algunos casos disponemos ya de muestra recogida, tienen como objetivo aportar más datos a favor de la calidad del modelo pentafactorial de personalidad de Zuckerman. Tal y como se ha ido desarrollando en la discusión de este documento, parte importante de esta investigación se articula entorno a la satisfacción de objetivos dirigidos al análisis de la validez externa de las cinco dimensiones del modelo, explorando poblaciones diana diferentes, tanto clínicas como no clínicas. El objetivo último es que, además de proporcionar una aproximación teórica consistente y una satisfactoria validez interna en forma, también, de consistentes agrupaciones factoriales o de elevados coeficientes de fiabilidad interna, se pueda disponer de un modelo que muestre una razonable validez externa a la vez que aplicada.

7. *Referencias* Aluja, A., García, O. & García, L.F. (2004). Replicability of the three, four and five Zuckerman's personality super-factors. Exploratory and confirmatory factor analysis of the EPQ-RS, and NEO-PI-R. *Personality and Individual Differences*, 36, 1093-1108.

Aluja, A., Rossier, J., García, L.F., Angleitner, A., Kuhlman, M. & Zuckerman, M. (2006). A cross-cultural shortened version of the ZKPQ (ZKPQ-50-CC) adapted to English, French, German and Spanish languages. *Personality and Individual Differences*, 41, 619-628.

Aluja, A., Rossier, J. & Zuckerman, M. (2007). Equivalence of paper and pencil vs Internet forms of the ZKPQ-50-CC in Spanish and French samples. *Personality and Individual Differences*, 43, 2022-2032.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders, DSM-IV-TR* (4th rev. ed). Washington, DC: American Psychiatric Association.

Ball, S.A. (1995). The validity of an alternative five-factor measure of personality in cocaine abusers. *Psychological Assessment*, 7, 148-154.

Ball, S.A. (2001). Reconceptualizing personality disorders categories using personality trait dimensions: Introduction to special section. *Journal of Personality*, 69, 147-153.

Barbero, M.I., Vila, E. & Suárez, J.C. (2003). *Psicometría*. Unidades Didácticas. Madrid: UNED.

Buss, A.H. & Plomin, R. (1975). *A temperament theory of personality development*. New York: Wiley.

Buss, A.H. & Plomin, R. (1984). *Temperament: Early developing personality traits*. Hillsdale, NJ: Erlbaum.

Butcher, J.N., Dahlstrom, W.G., Graham, J.R., Tellegen, A. & Kaemmer, B. (1989). *The Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring*. Minneapolis, MN: University of Minnesota Press.

Butcher, J.N. & Hostetler, K. (1990). Abbreviating MMPI Item Administration: Past Problems and Prospect for the MMPI-2. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 2, 12-22.

Cattell, R. B. (1957). *Personality and motivation structure and measurement*. New York: World Book.

Campbell, N.R. & Fiske, D.W. (1959). Convergent and discriminant validation by multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.

Clark, L.A. (1993). Personality disorder diagnosis: Limitations of the Five-Factor Model. *Psychological Inquiry*, 4, 100-104.

Cloninger, CR., Svrakic, D.M. & Przybeck, TR. (1993) A psychobiological model of temperament and character. *Archives of General Psychiatry*, 50, 975-990.

The logo for the Zuckerman-Kuhlman Personality Questionnaire (ZKPQ), consisting of the lowercase letters 'zkpq' in a white, sans-serif font on a dark red background.

- Cohen, J.A. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.
- Costa, P.T. & McCrae, R.R. (1985). *The NEO Personality Inventory manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P.T. & McCrae, R.R. (1992). *NEO PI-R professional manual*. Odessa, FL: Psychological Assessment Resources, Inc.
- Costa, P.T. & McCrae, R.R. (1999). *NEO PI-R, Inventario de personalidad NEO revisado*. Madrid: TEA Ediciones.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- De Pascalis, V. & Russo, P. (2003). Zuckerman-Kuhlman Personality Questionnaire: preliminary results of the Italian version. *Psychological Reports*, 92, 965-974.
- Digman, J. M. (1990). Personality Structure: Emergence of the five factor model. *Annual Review of Psychology*, 41, 417-440.
- Eysenck, H.J. (1947). *Dimensions of Personality*. London, U.K.: Routledge.
- Eysenck, H.J. (1967). *The biological basis of personality*. Springfield, IL: CC Thomas.
- Eysenck, H.J. (1987). The definition of personality disorders and the criteria appropriate to their definition. *Journal of Personality Disorders*, 1, 211-219.
- Eysenck, H.L. (1992a). Four ways five factors are not basic. *Personality and Individual Differences*, 13, 667-673.
- Eysenck, H.L. (1992b). A reply to Costa and McCrae. P or A and C-The role of theory. *Personality and Individual Differences*, 13, 867-868.
- Eysenck, H.J. & Eysenck, S.B.G. (1969). *Personality Structure and Measurement*. London: Routledge.
- Eysenck, H.J. & Eysenck, S.B.G. (1991). *Eysenck Personality Scales*. London: Hodder and Stoughton.
- Eysenck, S.B.G., Garcia Sevilla, L., Torrubia, R., Ávila, C. & Ortet, G. (1992). Versió catalana de l'EPQ per a adults: un instrument per a la mesura de la personalitat. *Annals de Medicina*, 78, 223-230.
- Fossati, A., Maffel, C., Bagnato, M., Donati, D., Fiorilli, M., Novella, et al. (1998). Brief communication: Criterion validity of the personality diagnostic questionnaire (PDQ4) in mixed psychiatric sample. *Journal of Personality Disorders*, 12, 172-178.
- Frances, A. (1980). The DSM-III personality disorders section: a commentary. *American Journal of Psychiatry*, 137(9), 1050-1054.
- Gray, J.A. (1982). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. New York: Oxford University Press.

Gomà-i-Freixanet, M. (1995). Prosocial and antisocial aspects of personality. *Personality and Individual Differences*, 19, 125-134.

Gomà-i-Freixanet, M. & Wismeijer, A.A. (2002). Applying personality theory to a group of police bodyguards: A physically risky prosocial prototype?. *Psicothema*, 14, 387-392.

Gomà-i-Freixanet, M., Valero, S., Puntí, J. & Zuckerman, M. (2004). Psychometric properties of the Zuckerman-Kuhlman Personality Questionnaire in a Spanish sample. *European Journal of Psychological Assessment*, 20(2), 134-146.

Gorsuch, R.L. (1997). Exploratory Factor Analysis: Its role in item analysis. *Journal of Personality Assessment*, 68(3), 532-560.

Guilford, J.P. (1975). Factors and factors of personality. *Psychological Bulletin*, 82, 802-814.

Gutiérrez-Zotes, J. A., Ramos, J. A. & Sáiz, J. (2001). Desarrollo de la versión española del cuestionario de personalidad Zuckerman-Kuhlman (ZKPQ-III) y propiedades psicométricas. *Psiquis*, 22, 239-250.

Gutiérrez F., Sangorrín J., Martín-Santos R., Torres X. & Torrens M. (2002). Measuring the core features of personality disorders in substance abusers using the Temperament and Character Inventory (TCI). *Journal of Personality Disorders*, 16(4), 344-359.

Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, 10, 255-282.

Hambleton, R.K. (1980). True score validity and standard setting methods. En R.A. Berk (Ed.): *Criterion referenced measurement: The state of the art*. Baltimore: John Hopkins University Press.

Hojat, M. & Zuckerman, M. (2008). Personality and specialty interest in medical students. *Medical Teacher*, 30(4), 400-4006.

Kerlinger, F.N. & Lee, H.B. (2002). *Investigación del Comportamiento. Métodos de Investigación en Ciencias Sociales*. México: McGraw-Hill.

Klinterberg, A.B., Schalling, D. & Magnusson, D. (1986). *Self-report assessment of personality traits. Data from the KSP inventory on a representative sample of normal male and female subjects within a developmental project*. Reports from the Project Individual development and Adjustment, No. 64, Department of Psychology, Stockholm University, Sweden.

Kuder, G.F. & Richardson, M.W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2, 151-160.

Livesley, W.J. & Jackson, D.N. (1992). Guidelines for developing, evaluating, and revising the classification of personality disorders. *Journal of Nervous and Mental Disease*, 180, 609-618.

Livingston, S.A. (1977). Psychometric techniques for criterion-referenced testing and behavioral assessment. En J.D. Cone y R.P. Hawkins (Eds.): *Behavioral assessment: New directions in clinical psychology*. Nueva Cork: Brunner/Mazel.

- McCrae, R.R. & Costa, P.T. (1983). Join Factors in self-reports and ratings: Neuroticism, extraversion and openness to experience. *Personality and Individual Differences*, 4, 245-255.
- McCrae, R.R. & Costa P.T. (1997). Personality trait structure as a human universal. *American Psychologist*, 52, 509-516.
- McCrae, R.R. & Costa, P.T. (2004). A contemplated revision of the NEO five-factor inventory. *Personality and Individual Differences*, 36(3), 587-596.
- McCrae, R.R. & John, O. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, 60(2), 174-214.
- Messick, S. (1975). The standard problem: Meaning and values in measurement and evaluation. *American Psychologist*, 30, 955-966.
- Millon, T. (1994). *Millon Index of Personality Styles*. San Antonio: The Psychological Corporation. Traducción al español (1997). *Inventario Millon de Estilos de Personalidad*. Buenos Aires: Paidós.
- Millon, T., Davis, R. & Millon, C. (2007). *Inventario Multiaxial de Millon III*. Adaptación y baremación española. Madrid: TEA Ediciones.
- Muñiz, J. (1998). *Teoría Clásica de los Tests*. Madrid: Editorial Pirámide.
- O'Sullivan, D., Zuckerman, M. & Kraft, M. (1996). The personality of prostitutes. *Personality and Individual Differences*, 21, 445-448.
- Ostendorf, F. & Angleitner, A. (1994). A comparison of different instruments proposed to measure the Big Five. *European Review of Applied Psychology*, 44, 45-53.
- Peñate, W., Ibáñez, I. & González, M. (1999). La cuantía y naturaleza de las dimensiones básicas de personalidad: una aproximación empírica. *Análisis y Modificación de Conducta*, 25, 103-130.
- Romero, E., Luengo, M.A., Gómez-Freguela, J.A. & Sobral, J. (2002). La estructura de la personalidad en adolescentes: el modelo de cinco factores y los cinco alternativos. *Psicothema*, 14, 134-143.
- Rulon, P.J. (1939). A simplified procedure for determining the reliability of a test by split-halves. *Harvard Educational Review*, 9, 99-103.
- Shiomi, K., Kuhlman, D. M., Zuckerman, M., Joireman, J., Sato, M. & Yata, S. (1996). Examining the validity and reliability of a Japanese version of the Zuckerman-Kuhlman Personality Questionnaire (ZKPQ). *Hyogo University of Teacher Education Journal*, 2, 1-13.
- Strelau, J. (1983). *Temperament -Personality- Activity*. London: Academic Press.
- Thomas, A. & Chess, S. (1977). *Temperament and development*. New York: Brunner/Mazel.

Thornquist, M.H. & Zuckerman, M. (1995). Psychopathy, passive-avoidance learning and basic dimensions of personality. *Personality and Individual Differences*, 19, 525-534.

Verheul, R. (2005). Clinical utility of dimensional models for personality pathology. *Journal of Personality Disorders*, 19(3), 283-302.

Widiger, T.A. (1992). Categorical versus dimensional classification: Implications from and for research. *Journal of Personality Disorders*, 6, 287-300.

Widiger, T.A. & Frances, A. (2002). Toward a dimensional model for the personality disorders. In P.T. Costa and T.A. Widiger (Eds). *Personality disorders and the Five-Factor models of personality*. Washington DC: American Psychological Association.

Widiger, T.A. & Simonsen, E. (2005). Alternative dimensional models of personality disorder: Finding a common group. *Journal of Personality Disorders*, 19, 110-130.

Wu, Y., Wang, W., Du, W., Li, J., Jiang, X. & Wang, Y. (2000). Development of a Chinese version of the Zuckerman-Kuhlman Personality Questionnaire: reliabilities and gender/age effects. *Social Behavior and Personality*, 28, 241-250.

Yela, M. (1984). *Apuntes de psicometría*. Escuela de Psicología. U.C.M.

Zuckerman, M. (1979). *Sensation Seeking: Beyond the optimal level of arousal*. Hillsdale, NJ: Erlbaum.

Zuckerman, M. (1983). *Biological bases of sensation seeking, impulsivity, and anxiety*. Hillsdale, NJ: Erlbaum.

Zuckerman, M. (1984). Sensation seeking: A comparative approach to a human trait. *Behavioral and Brain Sciences*, 7, 413-417.

Zuckerman, M. (1991). *Psychobiology of Personality*. New York: Cambridge University Press.

Zuckerman, M. (1992). What is a basic factor and which factors are basic? Turtles and the way down. *Personality and Individual Differences*, 13, 675-681.

Zuckerman, M. (1994). *Behavioral expression and biosocial bases of sensation seeking*. New York: Cambridge University Press.

Zuckerman, M. & Kuhlman, D.M. (1993). *Norms for the Zuckerman-Kuhlman Personality Questionnaire (ZKPQ)*. Unpublished manuscript.

Zuckerman, M. & Kuhlman, D. M. (2000). Personality and Risk-Taking: Common Biosocial Factors. *Journal of Personality. Special Issue: Personality Perspectives on Problem Behavior*, 68(6), 999-1029.

Zuckerman, M., Kuhlman, D.M. & Camac, C. (1988). What lies beyond E and N? Factor analysis of scales believed to measure basic dimensions of personality. *Journal of Personality and Social Psychology*, 54, 96-107.

Zuckerman, M., Kuhlman, D.M., Joireman, J., Teta, P. & Kraft, M. (1993). A comparison of the three structural models for personality: the big three, the big five, and the alternative five. *Journal of Personality and Social Psychology*, 65, 757-768.

Zuckerman, M., Kuhlman, D.M., Thornquist, M. & Kiers, H. (1991). Five (or three): Robust questionnaire scale factors of personality without culture. *Personality and Individual Differences*, 12, 929-941.

8.
Apéndices



Psychometric Properties of the Zuckerman-Kuhlman Personality Questionnaire in a Spanish Sample

Montserrat Gomà-i-Freixanet¹, Sergi Valero²,
Joaquim Puntí¹, and Marvin Zuckerman³

¹Department of Health Psychology, Autonomous University of Barcelona, Catalonia, Spain

²Servei de Psiquiatria, Hospital Universitari de la Vall d'Hebron, Barcelona, Catalonia, Spain

³Department of Psychology, University of Delaware, Newark, DE, USA

Keywords: Personality traits, ZKPQ, AFFM, cultural generalizability.

Summary: An important question in trait theory is how many major traits are necessary to describe personality and exactly what traits these are. Several investigators have made attempts to answer these questions with solutions of 3, 5, and even 16 primary factors. The Zuckerman-Kuhlman Personality Questionnaire (ZKPQ) is a questionnaire aimed at the evaluation of a five-factor model, which emerged from factor analyses of scales and items. The results were five basic factors, similar in men and women: Neuroticism-Anxiety, Activity, Sociability, Impulsive Sensation-Seeking, and Aggression-Hostility. This study assesses the psychometric properties of the Catalan translation of the ZKPQ. The ZKPQ was administered to a total sample of 933 subjects with an age range from 17 to 25 years. The results obtained show good internal consistency of all the scales and good discriminant validity shown by the lack of correlation among scales. Gender differences are also in the predicted directions. Finally, the replicability of the original five-component structure was also demonstrated. The present findings show this Catalan version to be a reliable tool for research in the field of personality structure and demonstrate the cross-cultural reliability of the factor structure developed from American subjects.

Introduction

In the last two decades many efforts have been made to find the minimum number of dimensions necessary to describe personality structure (Digman, 1990). Some emphasized a three-factor model (Eysenck, 1967; Tellegen, 1985), others a five-factor model (Costa & McCrae, 1985; Goldberg 1992) or an older model of 16 factors (Cattell, Eber, & Tatsuoka, 1970). "The Big Five" model (Digman, 1990) is currently the most popular model although there is some dissent concerning its superiority to other models (Tellegen 1985; Eysenck, 1992a; Zuckerman, 1992;

Block, 1995). According to the Five-Factor Model of personality (FFM), most personality traits can be described in terms of five basic dimensions, namely: Neuroticism vs Emotional Stability (N); Extraversion or Surgency (E); Openness to Experience or Intellect, or Culture (O); Agreeableness vs Antagonism (A); and Conscientiousness (C). This model started with the early work of Fiske (1949), and Tupes and Christal (1961), and was continued by Norman (1963) and revived by Goldberg (1981). The FFM of personality was originally based on the lexical approach to personality structure which maintained that because personality traits are so central to human interactions, all important traits will have been encoded in natural

language. It was assumed that these dimensions could be found in self- or peer-ratings using trait adjectives as well as in questionnaires (McCrae & John, 1992). The FFM is most often measured by the NEO questionnaire (Costa & McCrae, 1992), which seems to show good cross-cultural replicability across the different languages and cultures studied, suggesting the universality of the FFM (McCrae & Costa, 1997).

Although this model may be very helpful in *describing* one's personality with self-reports or observer ratings, it has little utility in *explaining* behavior. As Eysenck has stated elsewhere (Eysenck, 1992a; 1992b), one way of determining which are the basic traits of personality is using a nomological network or a theoretical underpinning to guide better taxonomic studies and lead to theoretical advances. Another way is providing evidence for a biological link between genetic causation and behavioral organization. In relation to the FFM, much is known about the biological bases of N and E, but less is known about C, A, and O.

Zuckerman, Kuhlman, Joireman, Teta, and Kraft (1993) have proposed an alternative five-factor model (AFFM). The model emerged from a series of factor analyses of scales believed to measure basic dimensions of personality or temperament, particularly those used in psychobiological research (Zuckerman, Kuhlman, & Camac, 1988; Zuckerman, Kuhlman, Thornquist, & Kiers, 1991). The basic traits in the AFFM measured by the Zuckerman-Kuhlman Personality Questionnaire (ZKPQ) are: Neuroticism-Anxiety (N-Anx), Activity (Act), Sociability (Sy), Impulsive Sensation Seeking (ImpSS) and Aggression-Hostility (Agg-Host). Like Guilford (1975), Eysenck and Zuckerman view Openness-Intellect as a domain separate from temperament and so it was deliberately excluded from the AFFM. In this model, the broad Extraversion dimension is divided into the separate factors of Activity and Sociability. Activity emerged as a distinct factor in the five-factor analyses of scales in earlier studies (Zuckerman et al., 1988; Zuckerman et al., 1991). Because of its identification as a basic developmental trait (see, for example, Thomas & Chess, 1977; Buss & Plomin, 1984) activity level merits a distinctive assessment as a major trait of temperament or personality in the adult human as well as in the child. The distinction between Hostility and Anxiety is also important because both traits have different psychobiological bases (Gray, 1982), and should not be confounded within a single factor as found in the N dimension of Costa and McCrae's (NEO) model. On the other hand, Impulsivity and Sensation Seeking are closely related and, with Socialization, form a distinctive factor in five-factor analyses of scales (Zuckerman et al., 1991) and items. These two traits also have many important psychobiological correlates (Zuckerman, 1983, 1984, 1991, 1994).

The Zuckerman-Kuhlman Personality Questionnaire contains five scales measuring the five aforementioned basic traits and an additional scale named Infrequency to control for inaccurate responding or social desirability. This questionnaire has demonstrated good internal reliability and temporal stability (Zuckerman et al., 1993) and some studies have already been conducted demonstrating its validity, e.g., characteristics of drug abusers and prediction of therapy course (Ball, 1995), characteristics of psychopathology (Thornquist & Zuckerman, 1995) and risk taking (Zuckerman & Kuhlman, 2000). The aim of the present study was to test the cultural generalizability of the AFFM, i.e., if the five factors postulated by the alternative model can be replicated in a different culture. For this purpose the ZKPQ was translated into the Catalan language and administered to a Spanish sample with demographic characteristics similar to the original American one.

Materials and Methods

Subjects

The participants ($n = 1045$) were college students (freshmen and sophomores) from four different schools in Barcelona (Catalonia, Spain). Only those who responded to every one of the 99 items of the ZKPQ scales were included in the data analyses. The total sample ($n = 933$) consisted of 330 males (35.4%) and 603 females (64.6%). The ages of the respondents ranged from 17 to 25 years ($M = 19.16$; $SD = 1.19$). The ZKPQ was administered anonymously to groups in classroom settings. All subjects volunteered to take part in the study and did not receive any credit for their participation. Test-retest reliability coefficients with two weeks between testing were calculated using a second sample, similar to the previous one, of 97 university students whose ages ranged from 18 to 56 years ($M = 20.43$; $SD = 4.88$).

Translation

The adaptation of the ZKPQ was done using the back-translation method. Due to methodological concerns related to adaptation processes (Flaherty, Gaviria, Pathak, Mitchell, Wintrob, Richman, & Birz, 1988), all the researchers were psychologists, and some of them had experience with cultural studies. The first author who is proficient in both languages, translated the English version into the Catalan language, paying special attention

to the content of the items and the scale they belonged to. A second bilingual psychologist, living in the UK, translated the Catalan version back into English. Afterwards, both translators met and compared the original source and the back-translated items for equivalence of meaning. Any discrepancy was discussed until an agreement was reached. This version was then refined by paying special attention to the use of frequent and well-known words, and using a correct and easy grammar in order to ensure the items were well understood for every level of education. Then, where applicable, items were worded in a way that acknowledged the gender of the subject because the Catalan language offers this possibility. Finally, this last version was given to the University Language Advisor for final correction.

Measures and Procedure

Subjects completed the translated version of the ZKPQ, which consisted of five content scales, plus the Infrequency scale, containing a total of 99 items in a "true-false" answering format. The five scales can be described in terms of their typical content:

1. *Neuroticism-Anxiety (N-Anx, 19 items)* items describe frequent emotional upset, tension, worry, fearfulness, indecision, lack of self-confidence, and sensitivity to criticism.
2. *Activity (Act, 17 items)* items describe the need for general activity, an inability to relax and do nothing when the opportunity arises, and a preference for hard and challenging work, an active, busy life, and a high energy level.
3. *Sociability (Sy, 17 items)* items describe the number of friends one has and the amount of time spent with them, outgoingness at parties, and a preference for being with others as opposed to being alone and engaging in solitary activities.
4. *Impulsive Sensation-Seeking (ImpSS, 19 items)* items involve a lack of planning and the tendency to act impulsively without thinking and the seeking of excitement, novel experiences, and the willingness to take risks for these types of experiences. The ImpSS items are general in content and do not describe specific activities such as drinking or sex.
5. *Aggression-Hostility (Agg-Host, 17 items)* items describe a readiness to express verbal aggression; rude, thoughtless, or antisocial behavior; vengefulness and spitefulness; having a quick temper, and impatience with others.

This questionnaire also includes an *Infrequency* scale (*Infreq, 10 items*). Rather than being regarded as a scale in the normative sense, it should only be used to detect inattention to the task or simply as a validity measure for the individual test-taker. The items are mostly exaggerated, true scored, socially desirable but unlikely to be completely true statements about anyone. This scale is highly skewed, with most scores around 0 or 1.

In the unpublished manual for the ZKPQ, Zuckerman states that, in the USA samples, only 4% of the subjects have scores greater than 3 in the Infreq scale. Thus, the cut-off for the American samples was 3 and anyone with a score of 4 or more was eliminated. In the Spanish sample, the scores of the Infreq scale ranged from 0 to 7 ($M = 1.58$, $SD = 1.45$) with a mode of 1 and only 3.7% of the subjects had a score greater than 4. In our study we decided not to use the American cut-off. From other studies of questionnaire adaptations to a population similar to ours, see e.g., the Catalan adaptation of the EPQ (Eysenck, García, Torrubia, Avila & Ortet, 1992) or from our own experience (Gomà-i-Freixanet, 1995, 1997, 2001; Gomà-i-Freixanet & Wismeijer, 2002) we know that the scores of the Lie scale in our country are higher than the ones found in the UK samples. On the other hand, the same pattern has been found in the Italian adaptation (V. De Pascalis, personal communication, July, 2000) as well as in a Castilian adaptation of the ZKPQ (Gutiérrez-Zotes, Ramos & Saiz, 2001) showing higher scores on the Infreq scale than the American samples. As our aim was to test the cultural generalizability of the AFFM, we thought it would be more adequate to use the whole sample without introducing any "a priori" bias.

In order to replicate the results found in the USA sample, we used a sample of university students similar to the original one. Means and standard deviations for the scales were obtained for the total sample as well as for males and females and *t*-tests were used to compare the means as a function of sex. The internal consistencies of the scales were calculated using Cronbach's α coefficients and the intercorrelations of the scales were subsequently calculated. Finally, in order to investigate the internal structure of the ZKPQ, an exploratory factor analysis of Principal Component with Varimax rotation for males and females as well as for the total sample was carried out. As Messick (1989) has stated elsewhere, one of the most powerful procedures to validate the structure of a certain construct is its confirmation by means of factor analysis. Additionally, the coefficients of congruence (Tucker, 1951) between the loading matrices obtained in the USA sample and the Catalan sample were also computed for males and females, as well as within the Catalan sample.

Table 1. Means, standard deviations, skewness, kurtosis and α coefficients of the Catalan version of the ZKPQ.

Scales	M	SD	Total sample n = 933		Total sample n = 933	Men n = 330 Cronbach's α	Women n = 603
			Skewness	Kurtosis			
N-Anx	9.45	4.56	.09	-.91	.84	.79	.83
Act	8.05	3.55	.09	-.64	.74	.73	.75
Sy	8.95	3.41	-.41	-.26	.76	.76	.75
ImpSS	9.78	4.11	-.00	-.74	.79	.75	.80
Agg-Host	7.67	3.05	-.19	-.45	.67	.69	.66
Infreq	1.58	1.45	.96	.66	.45	.47	.39

Notes: N-Anx = Neuroticism-Anxiety, Act = Activity, Sy = Sociability, ImpSS = Impulsive Sensation Seeking, Agg-Host = Aggression-Hostility, Infreq = Infrequency.

Results

Descriptives, α Coefficients, t-Test Comparisons of Gender and Correlation Matrix

Table 1 shows the means, standard deviations, skewness, and kurtosis of the total sample, and Cronbach's α coefficients for males and females, as well as for the total sample, as indices of internal consistency. The magnitude of the coefficients α was adequate and very similar to that found in the original American version. The mean α for the ZKPQ scales for the total sample was .76, with values ranging from .67 to .84. N-Anx scale had the highest reliability (.84) and Agg-Host had the lowest (.67); Act, Sy, and ImpSS were intermediate (.74-.79). Test-retest reliabilities with two weeks between tests were N-Anx = .87, Act = .89, Sy = .90, ImpSS = .91 and Agg-Host = .77, indicating that the scales were reliable and stable.

Means and standard deviations for men and women, and differences among genders, are shown in Table 2. The genders differed significantly on the N-Anx, Sy, ImpSS, and Infreq scales. Women scored significantly higher on the N-Anx and Sy scales and men were higher on ImpSS and Infreq. There were no significant gender differences on Act and on Agg-Host scales.

Table 3 presents the correlation matrix for the total sample and for men and women separately. In the total

sample, correlations among scales ranged from -.19 to .22 with an absolute mean interscale correlation of .06. Although the absolute values were very low, almost all of the correlations were significant due to the large sample size.

Exploratory Factor Analysis

The structure of the translated version of the ZKPQ for the total sample was checked through exploratory factor analysis. Factor analysis of the 89 items remaining after excluding the items of the Infrequency scale was done with Principal Component analysis followed by normal-

Table 2. Means, standard deviations and t-test comparisons between men and women of the Catalan version of the ZKPQ scales.

Scales	Men (n = 330)		Women (n = 603)		t
	M	SD	M	SD	
N-Anx	7.39	4.00	10.57	4.46	11.14***
Act	7.81	3.47	8.17	3.59	1.49
Sy	8.26	3.53	9.32	3.29	4.53***
ImpSS	10.45	3.86	9.41	4.20	3.71***
Agg-Host	7.57	3.19	7.73	2.98	0.76
Infreq	2.03	1.60	1.33	1.29	6.75***

Notes: *** $p < .001$; N-Anx = Neuroticism-Anxiety, Act = Activity, Sy = Sociability, ImpSS = Impulsive Sensation Seeking, Agg-Host = Aggression-Hostility, Infreq = Infrequency.

Table 3. Correlations among scales of the Catalan version of the ZKPQ.

Scale	N-Anx	Act	Sy	ImpSS	Agg-Host	Infreq
N-Anx		-.01	-.06	.04	.12*	-.19
Act	-.05 (.02)		.13*	.20*	.08*	.13*
Sy	-.11* (-.12*)	.11* (.15*)		.20*	.11*	-.07*
ImpSS	.09* (.10)	.20* (.21*)	.22* (.24*)		.22*	.07*
Agg-Host	.09* (.17*)	.11* (.03)	.11* (.11)	.20* (.28*)		-.08*
Infreq	-.17* (-.06)	.15* (.14*)	-.00 (-.08)	.08 (.00)	-.12* (-.02)	

Notes: * $p < .05$; N-Anx = Neuroticism-Anxiety, Act = Activity, Sy = Sociability, ImpSS = Impulsive Sensation Seeking, Agg-Host = Aggression-Hostility, Infreq = Infrequency. In the upper-right side are reported correlations obtained for the total sample; in the lower-left side are correlations obtained for women and men (correlations for men are in parentheses).

Table 4. Result of the principal component analysis followed by a Varimax rotation of the five principal components from the ZKPQ scales for the Catalan male sample, *a priori* factor assignment, item communalities, percentage of variance explained, and congruence coefficients with the USA sample.

	<i>A priori</i> Scale	N-Anx <i>FI</i>	ImpSS <i>FI</i>	Sy <i>FI</i>	Agg-Host <i>FI</i>	Act <i>FI</i>	<i>h</i> ²
it_61	N-Anx	0.648	-0.037	-0.145	-0.002	-0.101	0.453
it_25	N-Anx	0.597	0.117	-0.144	0.236	0.077	0.452
it_30	N-Anx	0.574	0.138	-0.102	0.147	0.092	0.389
it_90	N-Anx	0.548	0.069	-0.257	0.165	-0.045	0.401
it_15	N-Anx	0.515	0.180	-0.102	0.238	-0.014	0.365
it_7	N-Anx	0.478	-0.027	-0.198	-0.078	-0.234	0.330
it_80	N-Anx	0.477	0.084	-0.028	-0.011	0.071	0.240
it_35	N-Anx	0.474	0.034	0.034	0.060	0.066	0.235
it_76	N-Anx	0.463	-0.051	0.021	0.095	0.005	0.227
it_41	N-Anx	0.429	-0.149	0.064	-0.015	0.080	0.217
it_71	N-Anx	0.411	-0.137	0.063	-0.085	-0.083	0.206
it_46	N-Anx	0.405	-0.006	-0.052	0.188	0.038	0.204
it_51	N-Anx	0.397	0.063	0.020	-0.197	-0.233	0.255
it_20	N-Anx	0.391	0.043	-0.018	0.186	0.009	0.190
it_66	N-Anx	0.388	0.106	-0.185	-0.162	0.095	0.231
it_56	N-Anx	0.381	0.106	-0.020	-0.043	0.042	0.160
it_96	N-Anx	0.347	-0.127	0.097	-0.154	-0.007	0.169
it_82	Sy	0.344	0.157	0.213	0.106	0.007	0.200
it_85	N-Anx	-0.335	0.076	-0.056	-0.116	0.000	0.134
it_99	Act	0.307	0.166	0.137	0.200	0.267	0.251
it_50	ImpSS	0.299	0.056	-0.021	0.238	-0.136	0.168
it_23	Act	0.289	-0.032	-0.132	0.037	-0.202	0.144
it_2	N-Anx	0.216	0.116	0.036	0.081	-0.086	0.075
it_39	ImpSS	-0.085	0.644	0.058	0.057	0.025	0.429
it_55	ImpSS	0.076	0.598	-0.052	0.186	0.022	0.401
it_65	ImpSS	0.032	0.555	-0.001	0.034	0.054	0.313
it_24	ImpSS	0.036	0.546	0.016	0.075	-0.009	0.305
it_34	ImpSS	-0.073	0.482	-0.118	-0.124	-0.140	0.287
it_75	ImpSS	0.002	0.476	-0.075	-0.040	0.103	0.245
it_45	ImpSS	-0.041	0.467	0.072	0.034	0.118	0.240
it_28	Act	-0.006	0.452	0.000	-0.051	0.277	0.284
it_60	ImpSS	-0.135	0.436	0.058	-0.013	0.095	0.221
it_14	ImpSS	0.178	0.435	0.100	0.190	-0.077	0.273
it_79	ImpSS	-0.011	0.406	0.064	0.057	0.003	0.173
it_84	ImpSS	0.184	0.358	0.065	-0.028	0.017	0.167
it_95	ImpSS	0.066	0.315	0.232	0.166	-0.040	0.186
it_53	Sy	0.096	0.291	0.255	0.183	-0.001	0.192
it_64	Act	0.060	0.226	-0.075	0.003	0.208	0.104
it_87	Sy	0.032	0.025	-0.607	0.086	0.068	0.382
it_17	Sy	0.013	0.202	-0.558	0.061	-0.024	0.357
it_68	Sy	0.104	0.068	-0.529	-0.014	-0.052	0.298
it_43	Sy	0.108	-0.071	-0.519	0.160	0.022	0.312
it_63	Sy	0.031	0.232	-0.516	0.012	-0.068	0.326
it_78	Sy	-0.163	0.214	0.510	-0.065	0.092	0.345
it_12	Sy	-0.009	0.050	-0.472	-0.051	0.180	0.260
it_58	Sy	-0.149	0.034	0.438	0.031	-0.093	0.225
it_22	Sy	-0.149	0.196	0.437	0.066	-0.088	0.264
it_9	Sy	-0.031	0.206	0.414	0.190	0.230	0.303
it_37	Sy	-0.139	-0.093	-0.412	0.039	-0.056	0.202
it_48	Sy	-0.075	0.397	0.397	0.010	0.217	0.369
it_92	Sy	-0.011	0.177	0.309	0.150	-0.077	0.155
it_98	Sy	0.111	0.172	0.210	0.102	0.136	0.115
it_77	Agg-Host	-0.124	0.183	-0.037	0.565	-0.049	0.372
it_86	Agg-Host	-0.034	0.141	0.081	-0.554	0.062	0.339
it_72	Agg-Host	0.003	0.131	-0.032	0.487	0.197	0.294

	<i>A priori</i> Scale	N-Anx <i>FI</i>	ImpSS <i>FII</i>	Sy <i>FIII</i>	Agg-Host <i>FIV</i>	Act <i>FV</i>	<i>h</i> ²
it_21	Agg-Host	-0.017	0.021	0.043	-0.482	-0.088	0.243
it_91	Agg-Host	0.231	0.183	-0.014	0.447	-0.024	0.287
it_62	Agg-Host	-0.090	0.114	0.177	-0.416	0.014	0.226
it_89	ImpSS	0.072	0.407	0.227	0.409	0.087	0.397
it_97	Agg-Host	-0.014	0.064	0.075	0.391	-0.033	0.163
it_70	Imp	0.168	0.322	0.217	0.355	-0.110	0.317
it_27	Sy	0.001	0.077	0.288	0.352	0.144	0.234
it_47	Agg-Host	-0.157	-0.043	-0.039	0.347	0.194	0.186
it_3	Agg-Host	0.135	0.109	0.005	0.346	0.119	0.164
it_31	Agg-Host	-0.157	0.082	0.035	-0.345	0.067	0.156
it_42	Agg-Host	-0.078	0.050	0.095	0.343	0.224	0.186
it_8	Agg-Host	0.103	-0.043	0.122	0.318	-0.010	0.129
it_57	Agg-Host	0.091	-0.017	-0.192	-0.310	0.208	0.185
it_67	Agg-Host	0.095	0.101	0.040	0.286	-0.107	0.114
it_11	Agg-Host	0.042	0.173	0.150	0.269	-0.134	0.145
it_36	Agg-Host	0.153	-0.014	-0.095	0.257	-0.113	0.111
it_44	Act	-0.071	-0.098	-0.140	-0.077	-0.600	0.401
it_33	Act	-0.070	0.138	0.202	0.021	0.558	0.377
it_74	Act	0.053	0.253	0.090	-0.061	0.526	0.356
it_38	Act	0.016	0.067	-0.110	0.008	-0.516	0.283
it_83	Act	0.020	0.219	0.005	-0.071	0.495	0.299
it_5	Act	0.002	0.100	-0.089	0.010	0.455	0.225
it_94	Act	-0.172	0.235	0.037	-0.105	0.408	0.264
it_29	ImpSS	-0.021	-0.151	-0.039	0.031	0.408	0.192
it_1	ImpSS	-0.002	0.194	-0.002	0.088	-0.401	0.206
it_59	Act	0.052	0.152	0.309	-0.190	0.357	0.284
it_54	Act	0.171	0.096	-0.090	0.165	0.349	0.196
it_18	Act	0.000	0.249	-0.269	-0.204	0.320	0.279
it_13	Act	0.048	0.115	-0.034	0.202	0.306	0.151
it_16	Agg-Host	-0.040	-0.090	-0.237	-0.120	0.264	0.150
it_88	Act	0.068	0.007	0.082	-0.106	-0.258	0.089
it_19	ImpSS	-0.159	0.188	0.061	-0.035	-0.244	0.125
it_49	Act	-0.028	0.006	0.108	0.027	-0.233	0.067
it_6	ImpSS	-0.033	-0.128	0.068	-0.152	0.191	0.082
% variance		5.58	5.37	4.68	4.41	4.35	
Congruence		0.90	0.84	0.92	0.85	0.91	

Notes: N-Anx = Neuroticism-Anxiety, ImpSS = Impulsive Sensation Seeking, Sy = Sociability, Agg-Host = Aggression-Hostility, Act = Activity, *h*² = communality.

ized Varimax rotation. Figure 1 shows the plot of the first 40 eigenvalues in order to evaluate more precisely the course of the eigenvalues.

Eigenvalues began to level off after five factors and factors beyond the sixth showed little change. The first seven eigenvalues were as follows: 6.98, 6.69, 4.24, 4.05, 3.13, 2.14, and 2.04. Therefore, the first five factors of the total sample explained 25.09% of the variance. Additionally, factor analyses were conducted separately for males and females and five factors were extracted and rotated. Factor loadings of each item in the rotated components retained are presented in Tables 4 and 5, as well as item communalities and percentage of explained variance by each component. In the male sample (Table 4), the first two factors were N-Anx and ImpSS, and accounted for 5.6% and 5.4% of the explained variance, respectively. The next three factors were Sy, Agg-Host,

and Act, which accounted for 4.7%, 4.4%, and 4.4% of the explained variance, respectively. In the female sample (Table 5), the first two factors were N-Anx and ImpSS, which accounted for 6.4% and 5.9% of the explained variance, respectively. The next three factors were Act, Sy, and Agg-Host, which accounted for 4.9%, 4.4%, and 3.8% of the explained variance, respectively.

Additionally, Tucker's congruence coefficients between the American sample and the Catalan sample for males and females were calculated. Two components are identical if Tucker's ϕ denotes 0.85 or higher (Haven & ten Berge, 1977). This coefficient is taken as a good index of factor similarity, thus evidencing recoverableness of the corresponding components. We found ϕ coefficients ranging from 0.84 to 0.96 (see Tables 4 and 5) and generally being in the nineties, indicating excellent factor replicability. Thus, recoverability is demonstrated

Table 5. Result of the principal component analysis followed by a Varimax rotation of the five principal components from the ZKPO scales for the Catalan female sample, *a priori* factor assignment, item communalities, percentage of variance explained and congruence coefficients with the USA sample.

	<i>A priori</i> Scale	N-Anx <i>F</i>	ImpSS <i>F</i>	Act <i>F</i>	Sy <i>F</i>	Agg-Host <i>F</i>	<i>h</i> ²
it_61	N-Anx	0.636	-0.076	-0.134	0.008	-0.118	0.442
it_30	N-Anx	0.626	0.192	0.041	-0.052	0.072	0.438
it_90	N-Anx	0.612	0.161	-0.021	-0.205	0.094	0.452
it_25	N-Anx	0.578	0.221	0.144	-0.094	0.156	0.437
it_80	N-Anx	0.562	0.128	-0.018	-0.080	-0.045	0.341
it_15	N-Anx	0.561	0.210	0.101	-0.045	0.054	0.375
it_7	N-Anx	0.553	-0.041	-0.159	0.000	-0.199	0.372
it_51	N-Anx	0.529	-0.024	-0.110	0.020	-0.193	0.330
it_35	N-Anx	0.518	-0.136	-0.118	0.110	0.057	0.316
it_96	N-Anx	0.518	-0.101	0.029	0.042	-0.074	0.286
it_71	N-Anx	0.477	-0.013	-0.108	0.051	-0.075	0.248
it_46	N-Anx	0.436	0.050	-0.040	0.032	0.133	0.213
it_85	N-Anx	-0.433	0.116	0.058	0.050	-0.237	0.263
it_66	N-Anx	0.427	-0.011	0.132	-0.040	0.115	0.215
it_76	N-Anx	0.425	-0.002	-0.022	-0.172	0.184	0.245
it_41	N-Anx	0.384	-0.119	-0.125	0.188	0.155	0.237
it_99	Act	0.384	0.169	0.234	0.009	0.320	0.333
it_20	N-Anx	0.381	-0.004	0.089	-0.061	-0.014	0.157
it_50	ImpSS	0.327	0.181	-0.092	-0.047	0.159	0.176
it_56	N-Anx	0.316	-0.003	0.105	-0.010	0.073	0.117
it_2	N-Anx	-0.242	0.085	-0.009	-0.025	-0.021	0.067
it_36	Agg-Host	0.148	0.014	-0.060	0.011	0.141	0.046
it_39	ImpSS	-0.044	0.681	0.087	-0.041	0.051	0.478
it_84	ImpSS	0.038	0.557	0.075	0.025	0.120	0.332
it_14	ImpSS	0.102	0.538	-0.036	0.087	0.256	0.374
it_65	ImpSS	-0.009	0.522	0.137	-0.013	0.003	0.291
it_24	ImpSS	-0.052	0.519	0.120	-0.047	0.048	0.291
it_55	ImpSS	0.062	0.510	0.103	-0.053	0.101	0.287
it_89	ImpSS	0.051	0.491	0.071	0.051	0.346	0.371
it_1	ImpSS	0.014	0.476	-0.225	-0.026	-0.037	0.280
it_34	ImpSS	-0.012	0.474	0.003	-0.133	-0.064	0.247
it_29	ImpSS	-0.026	-0.474	0.208	0.006	0.134	0.286
it_95	ImpSS	0.085	0.452	0.039	0.269	0.055	0.289
it_6	ImpSS	-0.021	-0.430	0.060	-0.010	-0.036	0.190
it_70	ImpSS	0.010	0.428	-0.070	0.172	0.134	0.236
it_75	ImpSS	-0.024	0.414	0.225	-0.101	-0.120	0.247
it_79	ImpSS	0.026	0.384	0.096	0.087	-0.034	0.166
it_48	Sy	-0.185	0.361	0.190	0.324	-0.012	0.306
it_28	Act	-0.234	0.340	0.291	-0.051	0.151	0.280
it_19	ImpSS	-0.094	0.333	-0.184	-0.017	-0.039	0.155
it_60	ImpSS	-0.043	0.304	0.196	0.143	0.024	0.154
it_45	ImpSS	-0.024	0.284	0.235	0.068	0.048	0.144
it_92	Sy	0.125	0.254	0.023	0.217	-0.001	0.128
it_53	Sy	0.002	0.244	0.056	0.115	0.037	0.077
it_27	Sy	-0.110	0.237	0.203	0.227	0.117	0.175
it_83	Act	0.072	0.070	0.655	-0.040	-0.004	0.440
it_33	Act	-0.005	0.055	0.644	0.102	-0.011	0.428
it_44	Act	-0.004	0.032	-0.638	-0.096	0.036	0.419
it_94	Act	-0.190	0.045	0.568	0.072	0.047	0.368
it_74	Act	-0.038	0.079	0.556	0.075	-0.044	0.325
it_38	Act	0.022	0.138	-0.516	0.020	0.118	0.300
it_59	Act	-0.098	0.151	0.492	0.033	-0.154	0.299
it_5	Act	0.014	-0.029	0.465	-0.016	-0.013	0.218
it_54	Act	0.231	0.015	0.368	-0.043	0.072	0.196
it_13	Act	0.081	0.095	0.365	-0.021	0.056	0.152

	<i>A priori</i> Scale	N-Anx <i>FI</i>	ImpSS <i>FI</i>	Act <i>FI</i>	Sy <i>FI</i>	Agg-Host <i>FI</i>	<i>h</i> ²
it_18	Act	-0.101	0.100	0.321	-0.272	0.029	0.198
it_23	Act	0.282	0.028	-0.289	-0.099	-0.176	0.204
it_64	Act	0.053	0.149	0.220	-0.080	0.019	0.080
it_88	Act	0.066	-0.031	-0.183	0.104	-0.076	0.055
it_58	Sy	-0.036	0.033	0.004	0.662	-0.014	0.441
it_87	Sy	0.119	-0.072	0.050	-0.633	0.040	0.424
it_22	Sy	-0.017	0.083	-0.096	0.600	0.058	0.380
it_12	Sy	0.019	-0.012	0.114	-0.560	0.005	0.327
it_63	Sy	-0.026	0.259	0.006	-0.545	-0.103	0.375
it_68	Sy	0.114	0.061	0.056	-0.531	-0.031	0.303
it_17	Sy	-0.048	0.279	0.005	-0.523	-0.102	0.364
it_43	Sy	0.215	-0.141	-0.019	-0.467	0.056	0.287
it_78	Sy	-0.163	0.174	0.279	0.402	-0.040	0.298
it_37	Sy	-0.060	-0.162	-0.149	-0.389	0.009	0.203
it_9	Sy	-0.167	0.215	0.298	0.302	0.112	0.266
it_82	Sy	0.217	0.046	0.010	0.279	0.068	0.132
it_49	Act	0.038	0.074	-0.194	0.198	-0.046	0.086
it_98	Sy	0.023	0.182	0.131	0.184	-0.044	0.087
it_97	Agg-Host	-0.029	-0.023	-0.003	0.111	0.535	0.300
it_67	Agg-Host	0.019	0.092	0.068	-0.055	0.535	0.303
it_77	Agg-Host	-0.063	0.220	0.040	0.094	0.520	0.333
it_72	Agg-Host	-0.013	0.120	0.199	-0.070	0.519	0.329
it_91	Agg-Host	0.187	0.080	0.161	-0.010	0.451	0.271
it_86	Agg-Host	-0.020	0.042	0.036	0.002	-0.442	0.199
it_42	Agg-Host	-0.293	0.073	0.126	0.068	0.430	0.297
it_21	Agg-Host	-0.054	-0.014	0.045	0.149	-0.418	0.202
it_8	Agg-Host	0.162	0.149	0.049	0.043	0.391	0.205
it_57	Agg-Host	-0.020	-0.127	0.155	-0.029	-0.361	0.172
it_62	Agg-Host	-0.002	0.128	0.068	0.087	-0.315	0.128
it_11	Agg-Host	0.061	0.000	-0.226	0.101	0.276	0.141
it_31	Agg-Host	-0.139	0.100	0.029	-0.038	-0.273	0.106
it_47	Agg-Host	-0.201	0.043	0.089	0.180	0.220	0.131
it_16	Agg-Host	-0.051	-0.040	-0.012	-0.085	-0.181	0.044
it_3	Agg-Host	0.004	-0.049	-0.012	-0.079	0.094	0.018
% variance		6.44	5.93	4.90	4.44	3.82	
Congruence		0.96	0.94	0.93	0.93	0.88	

Notes: N-Anx = Neuroticism-Anxiety, ImpSS = Impulsive Sensation Seeking, Act = Activity, Sy = Sociability, Agg-Host = Aggression-Hostility, *h*² = communality.

throughout the four groups of subjects with only one exception, Catalan males compared with USA males for the ImpSS scale. However, the coefficient obtained is extremely close to Haven & ten Berge's criterion (0.844). Therefore, the results obtained in this study clearly indicate the strong congruence with the original American factor structure in the Catalan sample.

Table 6 shows the factorial solution for the total sample. The first two factors were N and ImpSS, and the next three factors were Act, Sy, and Agg-Host. Of the items with loadings equal or higher than .30, two items (48 and 28) loaded highest on a different component having the second highest loading on the target component (the magnitude of these loadings being practically equal) and only 1 item failed to load on the target component (Item 99 "other people often urge me to take it easy"), which

had a primary loading on N-Anx instead of Act. This item was originally assigned to the Activity scale, but its content could also be easily understood in the sense of worrisomeness and anxious tension.

In order to replicate the original five-component structure of the AFFM within the present study, Tucker's congruence coefficients between Catalan males and females were calculated. The obtained coefficients were N-Anx = .95, Act = .91, Sy = .91, ImpSS = .90, and Agg-Host = .85. Thus, the criterion was met for all of the five scales for the Catalan males and females. If we adhere to Haven and ten Berge's (1977) demonstration, the obtained results show the replicability of the original five-component structure of the AFFM, both within the present study for males and females, as well as between the present study and the original one.

Table 6. Result of the principal component analysis followed by a Varimax rotation of the five principal components from the ZKPQ scales for the total Catalan sample, *a priori* factor assignment, item communalities, and percentage of variance explained.

	<i>A priori</i> Scale	N-Anx <i>FI</i>	ImpSS <i>FII</i>	Act <i>FIII</i>	Sy <i>FIV</i>	Agg-Host <i>FV</i>	<i>h</i> ²
it_61	N-Anx	0.654	-0.060	-0.127	0.052	-0.080	0.457
it_25	N-Anx	0.615	0.184	0.113	0.110	0.168	0.465
it_30	N-Anx	0.612	0.189	0.034	0.088	0.076	0.425
it_80	N-Anx	0.585	0.070	0.015	0.022	-0.042	0.350
it_90	N-Anx	0.585	0.146	-0.050	0.243	0.097	0.434
it_15	N-Anx	0.552	0.219	0.049	0.085	0.096	0.371
it_35	N-Anx	0.541	-0.104	-0.030	-0.107	0.055	0.319
it_7	N-Anx	0.537	-0.027	-0.194	0.072	-0.165	0.359
it_41	N-Anx	0.469	-0.171	-0.036	-0.194	0.093	0.297
it_51	N-Anx	0.465	0.032	-0.167	0.004	-0.206	0.287
it_46	N-Anx	0.460	0.016	-0.017	-0.014	0.151	0.235
it_85	N-Anx	-0.455	0.127	0.019	0.026	-0.173	0.254
it_71	N-Anx	0.452	-0.041	-0.121	-0.051	-0.091	0.232
it_76	N-Anx	0.446	-0.016	-0.025	0.113	0.154	0.236
it_96	N-Anx	0.440	-0.079	-0.020	-0.044	-0.126	0.218
it_66	N-Anx	0.432	0.009	0.127	0.078	-0.009	0.209
it_20	N-Anx	0.428	-0.011	0.075	0.010	0.079	0.196
it_99	Act	0.411	0.144	0.255	-0.071	0.256	0.326
it_56	N-Anx	0.395	-0.012	0.109	-0.026	0.036	0.170
it_82	Sy	0.298	0.073	0.032	-0.265	0.089	0.173
it_2	N-Anx	-0.295	0.138	-0.059	0.052	0.017	0.113
it_50	ImpSS	0.290	0.176	-0.129	0.067	0.177	0.168
it_39	ImpSS	-0.071	0.663	0.098	0.025	0.047	0.457
it_55	ImpSS	0.048	0.535	0.096	0.081	0.137	0.323
it_14	ImpSS	0.094	0.524	-0.045	-0.055	0.219	0.337
it_24	ImpSS	-0.020	0.522	0.113	0.032	0.059	0.290
it_65	ImpSS	0.032	0.506	0.158	-0.002	0.012	0.282
it_84	ImpSS	0.066	0.504	0.058	-0.020	0.060	0.266
it_89	ImpSS	0.072	0.455	0.096	-0.110	0.366	0.368
it_95	ImpSS	0.053	0.440	0.003	-0.219	0.082	0.251
it_34	ImpSS	-0.000	0.439	0.008	0.107	-0.084	0.211
it_1	ImpSS	-0.087	0.436	-0.308	0.084	-0.006	0.300
it_70	ImpSS	0.048	0.429	-0.082	-0.165	0.210	0.264
it_29	ImpSS	0.065	-0.416	0.305	-0.057	0.102	0.284
it_79	ImpSS	-0.028	0.413	0.060	-0.040	-0.017	0.177
it_75	ImpSS	0.013	0.408	0.225	0.084	-0.100	0.235
it_48	Sy	-0.123	0.360	0.230	-0.354	-0.013	0.323
it_6	Sy	-0.019	-0.349	0.117	-0.033	-0.061	0.141
it_28	Act	-0.154	0.339	0.334	0.030	0.092	0.259
it_60	Act	-0.004	0.323	0.190	-0.113	0.014	0.153
it_45	Act	0.028	0.317	0.237	-0.079	0.040	0.165
it_19	Act	-0.169	0.310	-0.210	0.032	-0.030	0.171
it_53	Sy	0.012	0.286	0.038	-0.141	0.086	0.111
it_92	Sy	0.083	0.263	-0.023	-0.234	0.025	0.132
it_98	Sy	0.003	0.223	0.105	-0.145	0.006	0.082
it_33	Act	0.016	0.062	0.619	-0.153	-0.005	0.411
it_83	Act	0.119	0.076	0.619	-0.009	-0.030	0.403
it_44	Act	-0.026	-0.019	-0.613	0.094	0.003	0.386
it_74	Act	0.014	0.118	0.558	-0.082	-0.052	0.335
it_94	Act	-0.145	0.071	0.536	-0.088	-0.006	0.321
it_38	Act	0.047	0.092	-0.490	-0.004	0.089	0.259
it_5	Act	-0.009	0.019	0.449	0.076	-0.015	0.208
it_59	Act	-0.106	0.179	0.426	-0.082	-0.164	0.258
it_54	Act	0.218	0.040	0.348	0.070	0.101	0.185
it_18	Act	-0.089	0.122	0.337	0.286	-0.055	0.221
it_13	Act	0.057	0.114	0.329	0.054	0.103	0.138

	<i>A priori</i> Scale	N-Anx <i>FI</i>	ImpSS <i>FI</i>	Act <i>FI</i>	Sy <i>FI</i>	Agg-Host <i>FI</i>	<i>h</i> ²
it_23	Act	0.246	0.047	-0.288	0.147	-0.107	0.179
it_49	Act	-0.039	0.093	-0.232	0.129	-0.019	0.081
it_64	Act	0.069	0.160	0.226	0.075	0.004	0.087
it_88	Act	0.088	-0.025	-0.195	-0.108	-0.082	0.065
it_87	Sy	0.031	-0.040	0.040	0.649	0.069	0.430
it_58	Sy	-0.060	0.061	-0.041	-0.579	-0.010	0.344
it_68	Sy	0.058	0.064	0.007	0.554	-0.019	0.315
it_17	Sy	-0.084	0.248	-0.009	0.553	-0.038	0.376
it_63	Sy	-0.056	0.236	-0.013	0.550	-0.044	0.364
it_22	Sy	-0.038	0.131	-0.081	-0.549	0.050	0.329
it_12	Sy	-0.037	-0.003	0.136	0.543	-0.020	0.315
it_43	Sy	0.135	-0.104	-0.024	0.504	0.087	0.291
it_78	Sy	-0.163	0.202	0.222	-0.436	-0.054	0.310
it_37	Sy	-0.080	-0.166	-0.102	0.376	0.024	0.186
it_9	Sy	-0.130	0.235	0.272	-0.324	0.133	0.268
it_27	Sy	-0.082	0.210	0.166	-0.229	0.200	0.171
it_77	Agg-Host	-0.069	0.206	0.020	-0.051	0.542	0.344
it_72	Agg-Host	0.065	0.071	0.231	0.014	0.516	0.329
it_97	Agg-Host	0.029	-0.020	0.016	-0.133	0.495	0.264
it_86	Agg-Host	-0.028	0.060	0.062	-0.032	-0.488	0.247
it_67	Agg-Host	0.008	0.116	0.004	0.046	0.442	0.211
it_91	Agg-Host	0.157	0.157	0.077	0.063	0.441	0.253
it_21	Agg-Host	-0.017	-0.028	0.015	-0.137	-0.428	0.203
it_42	Agg-Host	-0.221	0.060	0.168	-0.085	0.411	0.256
it_8	Agg-Host	0.183	0.079	0.026	-0.098	0.358	0.178
it_62	Agg-Host	-0.080	0.145	0.037	-0.083	-0.352	0.160
it_57	Agg-Host	0.026	-0.101	0.175	0.088	-0.352	0.173
it_47	Agg-Host	-0.116	-0.033	0.144	-0.142	0.296	0.143
it_31	Agg-Host	-0.193	0.106	0.031	0.041	-0.293	0.137
it_11	Agg-Host	0.060	0.057	-0.175	-0.120	0.292	0.137
it_3	Agg-Host	0.004	0.052	0.033	0.090	0.223	0.061
it_36	Agg-Host	0.116	0.028	-0.107	0.055	0.193	0.066
it_16	Agg-Host	0.007	-0.113	0.103	0.106	0.143	0.055
% variance		6.46	5.66	4.73	4.47	3.77	

Notes: N-Anx = Neuroticism-Anxiety, ImpSS = Impulsive Sensation Seeking, Act = Activity, Sy = Sociability, Agg-Host = Aggression-Hostility, *h*² = communality.

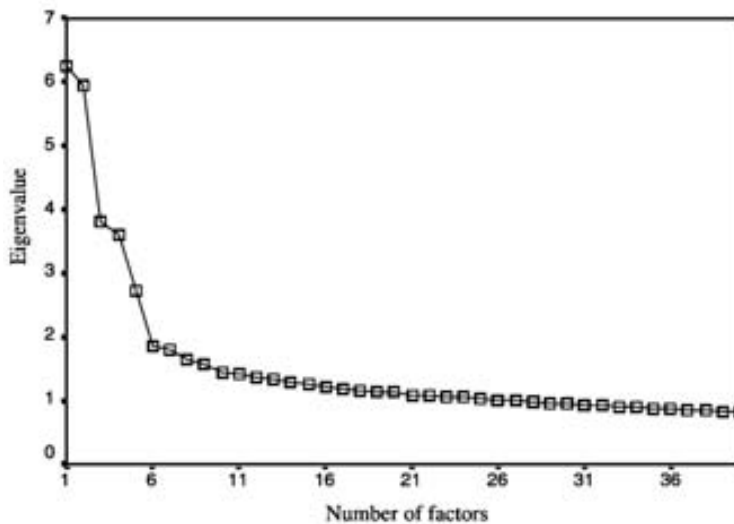


Figure 1. Scree plot of the first 40 eigenvalues for the factor analysis of the ZKPQ.

Discussion

This study investigated the psychometric properties of the Catalan adaptation of the ZKPQ. The magnitude of the α coefficients in the Catalan ZKPQ scales was practically equal to the coefficient estimates for the original USA sample, except for the Agg-Host scale, which appeared to be somewhat less internally consistent. In our sample the N-Anx scale was the most reliable, and the Agg-Host scale was the least reliable. The same general trend for the α reliabilities was found in other translations of the ZKPQ questionnaire to different languages: Italian (V. De Pascalis, personal communication, July, 2000), German (Ostendorf & Angleitner, 1994), Chinese (Wu, Wang, Du, Li, Jiang, & Wang, 2000), Japanese (D. M. Kuhlman, personal communication, July 2002), and in two different Castilian versions (Peñate, Ibáñez, & González, 1999; Gutiérrez-Zotes, Ramos, & Saiz, 2001), i.e., N-Anx being the highest and Agg-Host the lowest. All these studies, along with ours, conducted in six different languages, provide support for cross-language convergence of the ZKPQ at the scale level. Nevertheless, the fact that the Agg-Host scale has the lowest internal consistency in all of these studies – with different samples, languages, and cultures – may indicate either a translation or a cultural problem.

The means of the scales of the Catalan version of the ZKPQ were very similar to those of the American version (Zuckerman et al., 1993) and they follow the general trend found in the gender difference literature, i.e., women score significantly higher on N-Anx and Sy, and men on ImpSS. However, unlike the results of the original study, there were no significant gender differences on Act and on Agg-Host scales. Women were significantly higher than men on N-Anx in every national sample tested and women also scored significantly higher on Sy in the German and Japanese sample. Other gender differences found in the American sample such as men being higher on Act and Agg-Host were not found in other groups and in the German sample the university women were actually higher than men on Agg-Host (F. Ostendorf, personal communication, January, 2002). Perhaps the lack of gender differences in some of the scales in European countries may have to do with the type of women who elect to take higher education studies.

The ZKPQ also has good discriminant validity shown by the low intercorrelations among scales, thus empirically replicating the relative independence of scales derived from factor analysis. An interesting correlational pattern was found for the Infrequency scale. If this scale had to be considered as measuring social desirability, as the Lie scale of the EPQ does, the pattern of correlations should be totally different, showing high and positive

correlations with the more desirable traits (i.e., Act and Sy) and high but negative with the less desirable traits (i.e., N-Anx, ImpSS and Agg-Host). However, this pattern was not found, therefore, the Infrequency scale of the ZKPQ seems to detect careless responding or inattention to the task, rather than social desirability.

The results of the Principal Component analysis, using orthogonal Varimax rotation, demonstrated the replicability of the original five-component structure of the AFFM of personality in the Catalan language. For comparative purposes we performed two principal component analyses, one with the whole sample and another using the cut-off point of 3 in the Infrequency scale, as in the American sample, and the loadings of the items on the factors were very similar. Unlike the American study, which did not include data from the subjects exceeding the validity scale cut-off, in the Catalan study we did, but since the results were practically the same, we can conclude that the inclusion of subjects with high scores on the Infrequency scale apparently does not interfere much with the replicability of the AFFM. The percentage of explained variance of the present translation is not different from the ones found in other translations, which range from 20.75% (Wu et al., 2000) to 28% (in the Japanese sample).

Additionally, Tucker's coefficients of congruence with the original sample, as well as within the Catalan sample were calculated. Haven and ten Berge (1977) have demonstrated that when Tucker's coefficients of congruence "are as high as .85, then the interpretations of the corresponding components are judged equal by factor analysis experts. Therefore, one may adopt congruence of .85 or higher as evidence of recoverableness." All the coefficients were equal or higher than .85, except for the ImpSS scale when comparing USA males with Catalan males. If we adhere to Haven and ten Berge's demonstration, the congruence coefficients obtained in the Catalan adaptation of the ZKPQ are in favor of the replicability of the original five-component structure of the AFFM both within the Catalan males and females as well as between the Catalan sample and the original USA sample.

All in all, the results obtained from the Catalan adaptation of the Zuckerman-Kuhlman Personality Questionnaire showed good internal consistency of its scales, gender differences in the expected directions, independence among the scales derived from factor analysis and recoverableness of the original five-component structure. This questionnaire has proved useful as a personality tool to determine the characteristics of drug abusers (Ball, 1995) and other related psychopathologies (Thornquist & Zuckerman, 1995). As such, the ZKPQ has not only theoretical value but also applied value and research purposes as well.

Author's Notes

This research was supported by the Autonomous University of Barcelona (Spain) with grant EMER99-18. We especially thank D. Michael Kuhlman (University of Delaware), Jos M.F. ten Berge (University of Groningen), and Eduardo Doval (Autonomous University of Barcelona). Comments by the anonymous referees are gratefully acknowledged.

A copy of the Catalan ZKPQ questionnaire can be requested from the first author.

References

- Ball, S.A. (1995). The validity of an alternative five factor measure of personality in cocaine abusers. *Psychological Assessment, 7*, 148-154.
- Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin, 117*, 187-215.
- Buss, A.H., & Plomin, R. (1984). *Temperament: Early developing personality traits*. Hillsdale, NJ: Erlbaum.
- Cattell, R.B., Eber, H.W., & Tatsuoka, M.M. (1970). *The handbook for the Sixteen Personality Factor Questionnaire*. Champaign, IL: Institute for Personality and Ability Testing.
- Costa, P.T., Jr., & McCrae, R.R. (1985). *The NEO Personality Inventory*. Odessa, FL: Psychological Assessment Resources.
- Costa, P.T., Jr., & 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.
- Digman, J.M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*, 417-440.
- Eysenck, H.J. (1967). *The biological basis of personality*. Springfield, IL: CC Thomas.
- Eysenck, H.J. (1992a). Four ways five factors are not basic. *Personality and Individual Differences, 13*, 667-673.
- Eysenck, H.J. (1992b). A reply to Costa and McCrae. P or A and C - The role of theory. *Personality and Individual Differences, 13*, 867-868.
- Eysenck, S.B.G., García Sevilla, L., Torrubia, R., Avila, C., & Ortet, G. (1992). Versió catalana de l'EPQ per a adults: un instrument per a la mesura de la personalitat [Catalan version of the EPQ for adults: A tool for personality measurement]. *Annals de Medicina, 78*, 223-230.
- Fiske, D.W. (1949). Consistency of the factorial structures of personality ratings from different sources. *Journal of Abnormal and Social Psychology, 44*, 329-344.
- Flaherty, J.A., Gaviria, F.M., Pathak, D., Mitchell, T., Wintrob, R., Richman, J.A., & Birz, S. (1988). Developing instruments for cross-cultural psychiatry research. *Journal of Nervous and Mental Disease, 176*, 257-263.
- Goldberg, L.R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed.), *Review of personality and social psychology* (pp. 141-165). Beverly Hills, CA: Sage.
- Goldberg, L.R. (1992). The development of markers for the big-five factor structure. *Psychological Assessment, 4*, 26-42.
- Gomà-i-Freixanet, M. (1995). Prosocial and antisocial aspects of personality. *Personality and Individual Differences, 19*, 125-134.
- Gomà-i-Freixanet, M. (1997). Consensual validity of the EPQ: Self-report and spouse-reports. *European Journal of Psychological Assessment, 13*, 179-185.
- Gomà-i-Freixanet, M. (2001). Prosocial and antisocial aspects of personality in women: A replication study. *Personality and Individual Differences, 30*, 1401-1411.
- Gomà-i-Freixanet, M., & Wismeijer, A.A. (2002). Applying personality theory to a group of police bodyguards: A physically risky prosocial prototype? *Psicothema, 14*, 387-392.
- Gray, J.A. (1982). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. New York: Oxford University Press.
- Guilford, J.P. (1975). Factors and factors of personality. *Psychological Bulletin, 82*, 802-814.
- Gutiérrez-Zotes, J.A., Ramos, J.A., & Saiz, J. (2001). Desarrollo de la versión española del cuestionario de personalidad Zuckerman-Kuhlman (ZKPQ-III) y propiedades psicométricas [Development of the Spanish version of the ZKPQ questionnaire and psychometric properties]. *Psiquis, 22*, 239-250.
- Haven, S., & ten Berge, J.M.F. (1977). *Tucker's coefficient of congruence as a measure of factorial invariance: An empirical study*. Unpublished manuscript, University of Groningen, The Netherlands (cited in J.M.F. ten Berge, 1986).
- McCrae, R.R., & Costa, P.T., Jr. (1997). Personality trait structure as a human universal. *American Psychologist, 52*, 509-516.
- McCrae, R.R., & John, O.P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality, 60*, 175-215.
- Messick, S. (1989). Validity. In R.L. Lynn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York: Macmillan.
- Norman, W.T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology, 66*, 574-583.
- Ostendorf, F., & Angleitner, A. (1994). A comparison of different instruments proposed to measure the Big Five. *European Review of Applied Psychology, 44*, 45-53.
- Peñate, W., Ibáñez, I., & González, M. (1999). La cuantía y naturaleza de las dimensiones básicas de personalidad: una aproximación empírica [The quantity and nature of basic personality dimensions: An empirical approach]. *Análisis y Modificación de Conducta, 25*, 103-130.
- Tellegen, A. (1985). Structures of mood and personality and their relevance to assessing anxiety with an emphasis on self-report. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorders* (pp. 681-706). Hillsdale, NJ: Erlbaum.
- ten Berge, J.M.F. (1986). Rotation to perfect congruence and the cross-validation of component weights across populations. *Multivariate Behavioral Research, 21*, 41-64.
- Thomas, A., & Chess, S. (1977). *Temperament and development*. New York: Brunner/Mazel.
- Thorndike, M.H., & Zuckerman, M. (1995). Psychopathy, passive-avoidance learning, and basic dimensions of personality. *Personality and Individual Differences, 19*, 525-534.
- Tucker, L.R. (1951). *A method for synthesis of factor analysis studies* (Personnel Research Section Report No. 984). Washington, DC: Dept of the Army (cited in J.M.F. ten Berge, 1986).
- Tupes, E.C., & Christal, R.E. (1961). *Recurrent personality factors based on trait ratings*. U.S. Air Force, ASD Technical Reports (No. 61-79).
- Wu, Y., Wang, W., Du, W., Li, J., Jiang, X., & Wang, Y. (2000). Development of a Chinese version of the Zuckerman-Kuhlman

- personality questionnaire: Reliabilities and gender/age effects. *Social Behavior and Personality*, 28, 241–250.
- Zuckerman, M. (1983). *Biological bases of sensation seeking, impulsivity, and anxiety*. Hillsdale, NJ: Erlbaum.
- Zuckerman, M. (1984). Sensation seeking: A comparative approach to a human trait. *Behavioral and Brain Sciences*, 7, 413–417.
- Zuckerman, M. (1991). *Psychobiology of personality*. New York: Cambridge University Press.
- Zuckerman, M. (1992). What is a basic factor and which factors are basic? Turtles all the way down. *Personality and Individual Differences*, 13, 675–681.
- Zuckerman, M. (1994). *Behavioral expressions and biosocial bases of sensation seeking*. New York: Cambridge University Press.
- Zuckerman, M., & Kuhlman, D.M. (2000). Personality and risk-taking: Common biosocial factors. *Journal of Personality*, 68, 999–1029.
- Zuckerman, M., Kuhlman, D.M., & Camac, C. (1988). What lies beyond E and N? Factor analyses of scales believed to measure basic dimensions of personality. *Journal of Personality and Social Psychology*, 54, 96–107.
- Zuckerman, M., Kuhlman, D.M., Joireman, J., Teta, P., & Kraft, M. (1993). A comparison of three structural models for personality: the big three, the big five, and the alternative five. *Journal of Personality and Social Psychology*, 65, 757–768.
- Zuckerman, M., Kuhlman, D.M., Thornquist, M., & Kiers, H. (1991). Five (or three): Robust questionnaire scale factors of personality without culture. *Personality and Individual Differences*, 12, 929–941.

Montserrat Gomà-i-Freixanet, PhD.
Department of Health Psychology
Autonomous University of Barcelona
Edifici B,
E-08193 Bellaterra
Catalonia, Spain
Tel. +34 93 581-3854
Fax +34 93 581-2125
E-mail montserrat.goma@uab.es

Consensual Validity Parameters of the Zuckerman–Kuhlman Personality Questionnaire: Evidence From Self-Reports and Spouse Reports

Montserrat Gomà-i-Freixanet and Andreas A. J. Wismeijer

*Department of Health Psychology
Autonomous University of Barcelona*

Sergi Valero

*Department of Psychiatry
Hospital Universitari de la Vall d'Hebron
Barcelona, Catalonia, Spain*

One way of validating questionnaire responses is correlating them with ratings made by external assessors who know the ratee well: This is known as consensual validity. In this study, we assessed the consensual validity of the Zuckerman–Kuhlman Personality Questionnaire (ZKPQ; Zuckerman, 2002; Zuckerman, Kuhlman, Joireman, Teta, & Kraft, 1993). A multitrait-multimethod matrix of self-reported and spouse reported personality ratings was used to establish convergent and discriminant validity by means of Campbell & Fiske's (1959) evaluative criteria. Self-reports of 86 men and 85 women were correlated with their spouses' reports. Intraclass correlations ranged from .47 to .63 for the 5 dimensions, providing strong evidence of convergent and discriminant validity. The results obtained favor the use of the ZKPQ as a valid self-report measure of personality traits.

The alternative Five-factor model (AFFM) proposed and described by Zuckerman, Kuhlman, Thornquist, and Kiers (1991) originated as an alternative to the Five-factor model (FFM; Digman, 1990) to make up for the latter's supposed lack of explanatory power. The origins of the FFM and the AFFM were quite different. The former originated in research on the lexical properties of adjectives in the language pertaining to personality descriptors. Costa, McCrae, and Arenberg (1980) started with a three-factor model (Extraversion, Neuroticism, and Openness to Experience), resembling Eysenck's (1967) model at least in the first two factors. Costa and McCrae (1985) later added two more factors (Agreeableness and Conscientiousness) to bring the model closer to the five factors identified in lexical analyses. The model is essentially a descriptive one designed to tap those dimensions of personality identified in the lexical analyses of dictionaries and factor analyses of reduced subsets of the personality relevant words. It is essentially atheoretical in its origin and descriptive of traits that are uniquely human and do not translate easily into compara-

tive descriptions of animal behavior traits (e.g., Conscientiousness, Agreeableness).

The AFFM was being developed at the same time as Costa and McCrae (1985) were developing their own model. In preparation for his book on the *Psychobiology of Personality*, Zuckerman (1991) began looking for a framework to describe personality traits with biological-evolutionary roots and a potential for comparative analyses. Zuckerman et al. (1991) used questionnaire scales that had already been widely used in human psychobiological research and/or in studies of temperament in children and adults. Factor analyses using several markers for each hypothesized trait consistently yielded five factors reliably identifiable across genders (Zuckerman, Kuhlman, & Camac, 1988; Zuckerman et al., 1991). Both Eysenck (1992a, 1992b) and Zuckerman (1992) pointed out that one way of determining which are the basic traits of personality is using a nomological network or a theoretical underpinning to guide taxonomic studies and lead to theoretical advances and that psychobiological studies of personality provide data for the understanding of the

neurobiological and genetic underpinnings of personality. Relying only on the encoding of personality traits in language is treacherous, as this encoding probably reflects the observability of these traits in social interactions and may not necessarily mirror the proportional biological relevance of the traits (Zuckerman, 1992; Zuckerman, Kuhlman, Joireman, Teta, & Kraft, 1993). Therefore, using psychobiological data instead allows researchers to explore the biological origins of personality (Eysenck, 1992a).

The Zuckerman-Kuhlman Personality Questionnaire (ZKPQ; Zuckerman, 2002; Zuckerman et al., 1993) was developed to measure the dimensions that constitute the AFFM, namely, Neuroticism-Anxiety (N-Anx), Activity (Act), Sociability (Sy), Impulsive Sensation-Seeking (ImpSS), and Aggression-Hostility (Agg-Host). Since its first publication in 1993, the ZKPQ has undergone extensive psychometric testing, targeting both reliability and validity parameters. Concerning criterion validity, the ZKPQ have been tested in different populations such as team sportsmen and sportswomen (O'Sullivan, Zuckerman, & Kraft, 1998), prostitutes (O'Sullivan, Zuckerman, & Kraft, 1996), and cocaine abusers (Ball, 1995). Several cross-cultural studies have been performed as well using translations in Germany (Ostendorf & Angleitner, 1994), Japan (Shiomi et al., 1996), Israel (I. Montag, 2001), China (Wu et al., 2000), Italy (De Pascalis & Russo, 2003), and Spain (Gomà-i-Freixanet, Valero, Puntí, & Zuckerman, 2004; Kuhlman, Zuckerman, Gomà-i-Freixanet, & Shiomi, 2003).

Our aim of this study was to provide data on the consensual validity of the ZKPQ using the Catalan translation by Gomà-i-Freixanet et al. (2004). The psychometric evaluation of this version showed good internal consistency and discriminant validity of scales. The mean Cronbach's alpha for the ZKPQ scales was .76, with values ranging from .67 to .84, indicating adequate internal consistency within scales. Also, the correlations among scales ranged from $-.19$ to $.22$, with an absolute mean interscale correlation of $.06$, indicating adequate discriminant validity. Furthermore, gender differences in means were in accordance with the original U.S. sample, and the original U.S. factor structure was largely replicated with Tucker's congruence coefficients generally being in the 90s (Gomà-i-Freixanet et al., 2004).

The construct validity of a test is defined by the degree to which the test measures the variable(s) that it is designed to measure. This is easy to establish when one has a criterion against which the test can be evaluated, but unfortunately, this is not always the case. One way of overcoming this problem could be what is known as "consensual validation" (McCrae, 1982), meaning that one can validate questionnaire responses by correlating them with ratings made by external assessors who know the ratee well. This method thus provides both reports from observers who can interpret specific overt behaviors as evidence of underlying traits and self-reports that provide unique access to the private thoughts of the rated individual. One of the advantages of this

approach (McCrae & Costa, 1983) is that the artifacts that may influence self-reports (e.g., acquiescence, social desirability, and defensive responding) in general may be considered to be independent from those that influence observer reports (e.g., halo effects or effects of stereotypes). An agreement between self-reports and observer reports therefore constitutes powerful evidence of consensual validation. In a former article, Gomà-i-Freixanet (1997) gave a detailed overview of factors affecting the accuracy or agreement between self-reports and observer reports. Some of the most important conditions leading to improved accuracy of reports are the instrument itself and the acquaintance between the target and the rater. McCrae & Costa (1983) stated that one could increase the correspondence between the two sources by employing the same instrument for both the target individual and the rater. Regarding the target-rater acquaintance, it is known that self-peer convergence and interrater reliability increase with the degree of acquaintance between the judge and the target. Norman and Goldberg (1966) pointed out that, judged against the external criteria of self-reports, raters who have had longer acquaintances with the targets give consistently more accurate ratings. Several authors (e.g., McCrae, 1982; Watson & Clark, 1991) found that spouses are more accurate in assessing personality traits than friends, and friends are more accurate than strangers.

This research is the first that aimed to assess the validity of the ZKPQ by means of the consensual validation method using the multitrait-multimethod matrix (Campbell & Fiske, 1959). This approach allows us to assess the convergent and discriminant validity of the ZKPQ scales. By convergent validity, it is meant that measures of the same trait should be in agreement even if they are measured with different methods, and discriminant validity refers to different traits that should be distinguished from one another even if they are measured with the same method. As one can see from the previous reported literature, self-peer agreement increases with the degree of acquaintance between the rater and the ratee. Therefore, we used peer raters well acquainted with their targets: spouses. Thus, the same questionnaire was answered by both members of the couple, spouses were well acquainted with their targets, and using both self-reports and observer reports (from now on referred to as "spouse reports"), we controlled for the artifacts affecting reports in general.

METHOD

Participants

Participants were friends, parents, and relatives of undergraduate and postgraduate students they recruited to participate in the study. Also, some participants were recruited by one of the authors himself (A. Wismeijer), although they did not know the author. The primary convenience sample consisted of 88 couples. As occasionally some couples returned

self-reports or spouse reports for one person only or because reports were not answered properly (defined by us as having $\geq 10\%$ of double or missing answers), data of 86 men and 85 women were eligible for analysis. Age ranged from 19 to 75 years ($M = 38.21$, $SD = 13.14$). Mean amount of years the couples are living together was 13.33 years, with a range of .08 to 44 years ($SD = 11.66$). The mode of level of education was high school and the socioeconomic status was middle class. All respondents participated without inducements and voluntarily in the study.

Materials

Data were obtained using the Catalan translation of the ZKPQ (Gomà-i-Freixanet et al., 2004). The ZKPQ consists of 99 dichotomous items (in sentence format and true-false response set) covering five scales and an additional Infrequency (Infreq; 10 items) scale that allows eliminating participants with careless responding. The dimensions measured are N-Anx (19 items), Act (17 items), Sy (17 items), ImpSS (19 items), and Agg-Host (17 items).

Procedure

The study took place in the province of Barcelona, Catalonia, Spain. All participants spoke Catalan as their primary language. Catalan is one of the four official languages in Spain spoken by more than 7 million inhabitants in different countries (Andorra, France, and Italy). Each individual participant was provided with an envelope containing an introductory letter, the ZKPQ, written instructions, and two answering forms, one for himself or herself and the other to be answered as spouse (i.e., the participants were instructed to answer the ZKPQ as how they see themselves and how they see their spouse). The introductory letter explained globally the goals of the study ("the study you will collaborate in attempts to evaluate the functioning of the Catalan translation of an American questionnaire"), with no reference being made to the consensual agreement component of the study. The letter was signed by the principal investigators (M. Gomà-i-Freixanet and A. Wismeijer). The written instructions stressed to the participants to answer the questionnaire alone and without help of the spouse and instructed the participant to put the answered forms after completion in the provided envelope and seal it immediately. The instructions (written in Catalan) followed the following format:

- Please answer the questionnaire when you are alone and without help of your partner.
- First fill out the questionnaire as how you see yourself and only then fill out again the questionnaire as how you see your partner (i.e., how you see your partner and NOT how you think your partner will fill out the questionnaire).
- It is very important to answer the questionnaire honestly.

- When you filled out both answering forms, please put the questionnaire and the answering forms in the envelope and close it to guarantee your privacy.

Each envelope was precoded as well as the answering forms; therefore, all questionnaires were answered anonymously and confidentiality was guaranteed. One member of the couple returned the sealed envelopes to the investigator or the student that recruited the couple, in which case the student subsequently returned the envelopes to the investigators. No envelope was returned opened or with the seal damaged. Thus, from each couple, two sealed envelopes were returned containing in total four answering forms. Apart from the written instructions each participant received, students were first well instructed on how to give the instructions to fill in the questionnaires before recruiting their family members and/or friends so they could repeat the instructions if any doubts would arise after reading the written instructions. Telephone numbers of the principal authors (M. Gomà-i-Freixanet and A. Wismeijer) were included in the letter as well to resolve any doubts. No participant contacted us for questions or doubts. As the study was not intrusive in any sort, neither informed consent waivers nor participant debriefing following participation were required.

Analyses

A multitrait-multimethod matrix was used to determine the consensual validity by comparing the self-reports and spouse reports and establishing the degree of agreement on the different scales of the questionnaire. This approach is based on the assumption that if our measures are valid, we should expect to see a certain pattern emerge among the correlations in the multitrait-multimethod matrix. Specifically, correlations in the principal validity diagonal should be larger than the ones in the same row and column. The validity diagonal values show the correlation of the same trait across different methods. In this way, we could determine whether self-ratings differed from ratings the participants received by their spouses. Therefore, intraclass correlations (ICCs) between self-reports and spouse reports were computed for each scale, for the total sample as well as for both genders to obtain additional validity data and to ascertain if there was any differential pattern in both genders. Pearson correlation measures the intensity of the linear association between two variables but does not give information on the observed agreement, thus ignoring differences in rater's levels of response. Contrarily, the ICC coefficient is a more accurate statistic than Pearson's zero order correlation, as the ICC is sensitive to magnitudinal differences between the variables (Bland & Altman, 1986; Prieto, Lamarca, & Casado, 1998). Using traditional zero order correlations, a high correlation between self-reports and spouse reports on a given scale might be found, even though both raters rate each other at a different level. As long as this difference is consistent, it will

not lower the Pearson's correlation coefficient. The ICC solves this problem by taking the differences in magnitude between the scores into account. Using this technique, high yet consistent differences between variables will therefore not lead to an inflated correlation.

The Type I error rate per comparison was set by default to .05. As the ZKPQ has a total of six scales, and age was included in the analysis, a Bonferroni correction for multiple tests of $.05/7 = .007$ was used for the between-participant unpaired *t* tests. A correction of $.05/5 = .01$ was used for the self-reports and spouse reports paired *t* tests, as age and Infreq were not included in these analyses. Finally, an additional principal components analysis (PCA) of the 10 ZKPQ scales scores of self-ratings and spouse ratings was executed. We used this statistical technique, as it fitted optimally to the number of actual observations. Our sample size prevented us from conducting confirmatory factor analysis via structural equation modeling. We forced a five-factor solution and subsequently used the Varimax method for rotation. We predicted that if the five factors of the ZKPQ (Infreq scale was left out) would genuinely show consensual validity, each factor in the rotated matrix would be heavily loaded by both the self-ratings and spouse ratings of the same scale.

RESULTS

Descriptive Statistics

Descriptive statistics from the sample are shown in Table 1. Men and women did not differ significantly on age, although the former were on average 3.20 years older. Means of the scales of the ZKPQ are very similar to the original U.S. version and to the Catalan version (Gomà-i-Freixanet et al., 2004), and they follow the general trend found in men and women, that is, women score higher on N-Anx and Sy and lower on Imp-SS, and Infreq, although these differences only reached significance on the N-Anx scale.

In examining relations between self-data and spouse data, we compared the means across the two methods of data obtainment. As was discussed earlier (Gomà-i-Freixanet, 1997), two different patterns might be expected on theoretical grounds. First, researchers who emphasize the biased nature of self-report data would predict that self-raters will respond in a more socially desirable manner (i.e., higher levels on Sy and generally lower levels on N-Anx, ImpSS, Agg-Host and Infreq) than their spouses. Conversely, those who emphasize the biased nature of spouse-report data would predict that spouses will assign generally lower levels of neuroticism than the self-raters because this trait is less externally observable (Johnson, 1997). Table 2 presents the means and standard deviations for the ZKPQ dimensions for self-reports and spouse reports. The data did not support either contention, as we did not find any significant difference between self-reports and spouse reports.

TABLE 1
Means and Standard Deviations
for Self-Reported ZKPQ Scales for Men
and Women and *t* Test Comparisons

Variable	Men ^a		Women ^b		<i>t</i>	Cohen's <i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Age (years)	39.84	13.50	36.64	12.67	1.57	.25
ZKPQ						
N-Anx	5.92	4.06	9.84	4.88	-5.73*	-.88
Act	8.35	3.55	7.56	3.79	1.40	.22
Sy	6.58	3.63	6.89	3.51	-0.59	-.09
ImpSS	7.58	4.09	7.08	4.16	0.79	.12
Agg-Host	7.12	3.33	6.41	2.79	1.50	.23
Infreq	2.13	1.57	1.66	1.48	2.01	.31

Note. ZKPQ = Zuckerman-Kuhlman Personality Questionnaire; N-Anx = Neuroticism-Anxiety; Act = Activity; Sy = Sociability; ImpSS = Impulsive Sensation Seeking; Agg-Host = Aggression-Hostility; Infreq = Infrequency. ^a*n* = 86, ^b*n* = 85.

**p* < .001, two-tailed.

TABLE 2
Means and Standard Deviations for ZKPQ
Scales for Self-Reports and Spouse Reports
and *t* Test Comparisons With Paired Data

Scale	Self-Reports ^a		Spouse Reports ^b		<i>t</i>	Cohen's <i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
ZKPQ						
N-Anx	7.83	4.90	7.82	4.80	0.04	.00
Act	7.92	3.70	7.95	3.72	-0.10	-.01
Sy	6.77	3.55	7.15	4.14	-1.34	-.07
ImpSS	7.37	4.11	7.29	4.19	0.28	.02
Agg-Host	6.76	3.06	6.74	3.75	0.07	.00

Note. ZKPQ = Zuckerman-Kuhlman Personality Questionnaire; N-Anx = Neuroticism-Anxiety; Act = Activity; Sy = Sociability; ImpSS = Impulsive Sensation-Seeking; Agg-Host = Aggression-Hostility.

^a*n* = 171, ^b*n* = 171.

To examine the internal consistency of self-reports and spouse reports, we calculated Cronbach's alpha for both. The mean self-reported α for the ZKPQ scales was .78, with values ranging from .69 to .87. Table 3 shows that N-Anx has the highest internal consistency and Agg-Host the lowest. These coefficients are adequate and very similar to those found in the original U.S. version and even slightly higher than those of the Catalan version. The mean spouse-reported α for the same scales was .81, with values ranging from .77 to .86.

Heteromethod Correlations

Subsequently, ICCs were computed to determine the convergent and discriminant validity of the ZKPQ scales. Table 4 shows the multitrait-multimethod matrix with the heteromethod correlations between the self-reports and spouse reports. The principal diagonal of this table, underlined, con-

TABLE 3
Internal Consistency for Self-Reported and Spouse Rated ZKPQ Scales

Scale	Alpha Coefficients	
	Self-Reports	Spouse Reports
ZKPQ		
N-Anx	.87	.86
Act	.76	.77
Sy	.78	.84
ImpSS	.80	.80
Agg-Host	.69	.79

Note. ZKPQ = Zuckerman-Kuhlman Personality Questionnaire; N-Anx = Neuroticism-Anxiety; Act = Activity; Sy = Sociability; ImpSS = Impulsive Sensation-Seeking; Agg-Host = Aggression-Hostility.

TABLE 4
Multitrait-Multimethod Matrix for Self-Reports and Spouse Reports

Scale	Spouse Reports				
	N-Anx	Act	Sy	ImpSS	Agg-Host
Self-Reports					
N-Anx	<u>.63***</u>	-.11	-.07	.04	.16*
Act	-.02	<u>.47***</u>	.24**	.28***	.16*
Sy	-.15	.16*	<u>.54***</u>	.34***	.20**
ImpSS	-.01	.10	.34***	<u>.63***</u>	.22**
Agg-Host	.00	-.14	.08	.10	<u>.53***</u>

Note. Convergent correlations are underlined. N-Anx = Neuroticism-Anxiety; Act = Activity; Sy = Sociability; ImpSS = Impulsive Sensation-Seeking; Agg-Host = Aggression-Hostility.

* $p < .05$, two-tailed. ** $p < .01$, two-tailed. *** $p < .001$, two-tailed.

tains the convergent validity coefficients for the five scales. Overall, the results indicate a clear convergent and discriminant pattern. In terms of convergent validity, all scales showed a significant level of self-peer agreement.

Total sample convergent correlations ranged from .47 to .63 (absolute mean interscale correlation = .56). Clearly, all convergent correlations were far higher than the so-called ".3 barrier" (Mischel, 1968). In contrast, all but two of the off-diagonal values assessing discriminant validity stayed below the .3 barrier. Using the criterion that convergent correlations should be higher than any other values in its row or column of the heteromethod matrix (Campbell & Fiske, 1959), we found that all the scales displayed a good level of discriminant validity. These results support the construct validity of these specific scales. Self-reports and spouse reports showed a significant level of agreement on each of the scales.

Table 5 shows convergent correlations between self-reports and spouse reports for the total sample as well as for men and women separately to analyze any gender-based differences in ratings, that is, whether gender can influence the convergence of spouse ratings on any scale. Note that the correlations provided in the first column of Table 5 are the same as the principal diagonal correlations underlined in Table 4. The

95% confidence intervals are provided for these coefficients to show the estimation of population parameters. Convergent correlations ranged from .48 to .60 for men (absolute mean interscale correlation = .54) and .45 to .71 for women (absolute mean interscale correlation = .55).

The results for men and women followed the general pattern found in the total sample; there were, however, differences in absolute values between men and women in the magnitude of correlations in two scales: Sy and ImpSS. Agreement between self-reports and spouse reports was higher for Sy when women assessed men, whereas for the ImpSS scale, agreement was higher when men assessed women.

To evaluate the possible influence of the years the partners were living together on the level of agreement between self-reports and spouse reports, we generated a new variable named *consensus*, which consisted of subtracting for each scale the self-scores from the spouse scores. This new variable gave us a measure of the difference between self-reports and spouse-reports scores. None of the correlations between consensus and years living together, which ranged from 0 to 44 years, were significant for any of the scales: N-Anx ($r = .13$), Act ($r = .04$), Sy ($r = .05$), ImpSS ($r = .04$), Agg-Host ($r = .07$), and Infreq ($r = .09$).

PCA

Finally, Table 6 shows the results of the factor analysis of the self-reported and spouse-reported scales (excluding the Infreq scale) using a PCA followed by normalized varimax rotation of the five scales of the ZKPQ. The rationale behind this methodological strategy was that if the ZKPQ showed a high degree of consensual validity, then the responses made by both assessors should be highly consistent among them for each one of the scales. A PCA that contemplates the assessments of both assessors should consistently gather in each factor self-reports

TABLE 5
Intraclass Correlations With 95% Confidence Interval Between Self-Reports and Spouse Reports for the Total Sample and for Men and Women Separately

Scale	Entire Sample ^a		Men ^b		Women ^c	
	ICC	CI	ICC	CI	ICC	CI
ZKPQ						
N-Anx	<u>.63</u>	.53 to .71	.54	.37 to .67	.59	.44 to .72
Act	.47	.34 to .58	.48	.29 to .62	.46	.28 to .62
Sy	.54	.42 to .64	.60	.45 to .72	.45	.26 to .61
ImpSS	.63	.53 to .71	.56	.39 to .69	.71	.58 to .80
Agg-Host	.53	.41 to .63	.54	.37 to .68	.53	.36 to .67

Note. All the correlations were statistically significant at $p < .001$, two-tailed. ICC = intraclass correlations; CI = 95% confidence interval; ZKPQ = Zuckerman-Kuhlman Personality Questionnaire; N-Anx = Neuroticism-Anxiety; Act = Activity; Sy = Sociability; ImpSS = Impulsive Sensation-Seeking; Agg-Host = Aggression-Hostility.
^a $n = 171$, ^b $n = 86$, ^c $n = 85$.

TABLE 6
Result of the Principal Components
Analysis Followed by a Varimax Rotation
of the Five ZKPQ Scales

Scale	Rotated Factor Loadings				
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
ImpSS self	.89	.04	.21	.14	.08
ImpSS spouse	.81	-.01	.04	.31	.19
N-Anx spouse	-.13	.92	.02	-.04	.06
N-Anx self	.17	.86	.12	-.11	-.17
Agg-Host self	.14	.01	.89	-.01	-.06
Agg-Host spouse	.10	.16	.82	.21	.06
Sy spouse	.22	-.01	.04	.85	.16
Sy self	.18	-.15	.15	.82	.06
Act spouse	.03	-.02	-.19	.16	.85
Act self	.23	-.08	.20	.06	.81
% variance	16.49	16.31	16.21	15.89	14.84

Note. Factorial weights corresponding to pairs of self-ratings and spouse ratings are underlined. ZKPQ = Zuckerman-Kuhlman Personality Questionnaire; N-Anx = Neuroticism-Anxiety; Act = Activity; Sy = Sociability; ImpSS = Impulsive Sensation-Seeking; Agg-Host = Aggression-Hostility.

and spouse reports. It can be seen that the five rotated components are made up of the combinations of the self-reported and the corresponding spouse-reported ratings of the same scale. The five-factor solution together explained 79.74% of the variance. It is noteworthy to mention the magnitude of the factorial weights that corresponded to each pair of self-ratings versus spouse ratings for each scale that were characterized by highly and very similar factorial weights. Furthermore, there was a high discrepancy between these pairs of assessments and the rest of factorial weights in each factor. These results add additional data about the convergent and discriminant validity of ZKPQ scales.

DISCUSSION

The data obtained with this sample provide strong evidence for the satisfying psychometric properties of the Catalan version of the ZKPQ in general and its consensual validity in particular. With respect to the psychometric properties, gender differences in means among the scales were in the same direction and sense of those found in the original U.S. sample (Zuckerman & Kuhlman, 1993), in a different sample of university students (Gomà-i-Freixanet et al., 2004), and in other cross-cultural samples (e.g., Kuhlman et al., 2003). In general, women score higher on N-Anx and Sy, lower on ImpSS and Infreq, and similar to men, on Act and Agg-Host. With regard to the internal consistency, the results are in accordance to those previously found in other studies and slightly higher than those previously found in another sample with university students (Gomà-i-Freixanet et al., 2004).

In relation to the consensual validity parameters, the data obtained provide clear evidence for the consensual validity

of the personality dimensions assessed by the ZKPQ. Convergent correlations well above the .3 barrier were found. These results seem unlikely to stem from the artifacts of social desirability, acquiescence, extreme responding, or shared stereotypes because these sources of variance are generally not found simultaneously in self-reports and observer reports. Furthermore, as Eysenck and Eysenck (1985) pointed out, although such factors are not always entirely absent, they play only a relatively small part in such personality questionnaires, except under special conditions of motivation in which dissimulation may assume a more prominent role. These conditions generally are present in selection processes, forensic contexts, or in the case of not obtaining data anonymously. Neither condition was present in our study. Additional data come from the comparison of the mean scores for self-reports and spouse reports in that they did not show significant differences.

The magnitude of the convergent correlations—generally .5 to .6—was larger than typically has been reported (e.g., Borgatta, 1964; McCrae & Costa, 1987). Two factors could probably have contributed to the relatively higher correlations: the psychometric adequacy of the instrument itself and/or the nature of the raters. With reference to the instrument, these results seem to add data on the sound psychometric properties of the ZKPQ, specifically on validity. With respect to the second factor, it has been well established (e.g., Norman & Golberg, 1966; Watson & Clark, 1991) that the choice of well-qualified raters enhances the accuracy of personality assessments. On the whole, spouses were well acquainted with the participants they rated, most over a period of many years. Our data also address another issue pointed out by McCrae & Costa (1989). When reliable and valid measures are used, the correlations considerably exceed the .3 barrier, being better characterized as facing the ".6 barrier." It seems likely that the correlations found are near the ceiling for self-other agreement. Raters will always diverge to some extent from the individual's phenomenological view of himself or herself, and indeed, it would be disconcerting to think others could know us as intimately as we know ourselves.

Another issue that comes from our results and merits some comment is the lack of correlation between the years the partners were living together and the degree of agreement between self-reports and spouse reports. This lack of correlation corroborates some previous studies with the Eysenck Personality Questionnaire (Gomà-i-Freixanet, 1997) and with other instruments (Buss, 1984; Caspi, Herbener, & Ozer, 1992). It has already been mentioned that self-peer convergence increases with the degree of acquaintance between the judge and the target and that spouses are more accurate in assessing personality traits than friends and friends more accurate than strangers. Therefore, it seems as if there is a differential degree of agreement related to the degree of acquaintance, but the degree of agreement does not increase with the years of living together in this sample.

Additional data about convergent and discriminant validity of the ZKPQ scales stem from the PCA executed with self-reports and spouse reports scales scores. Convergent validity indexes are characterized by high and highly similar loadings of the self-reports and spouse reports of the same scale on a given factor. Discriminant validity indexes are characterized by relatively low loadings of the remaining scales on that given factor. The results unambiguously show this pattern of behavior for each pair of reports of the same scale and in every factor.

To summarize, the magnitude of the convergent correlations found in this study among self-reports and spouse reports were larger than typically reported and near the ceiling for self-other agreement (.6 barrier), and the data obtained from the PCA indicate that the consensual validity parameters of the ZKPQ are adequate, thus advocating the use of the self-reported ZKPQ as a valid instrument for personality assessment. Moreover, the results of this study when discussed in the context of previous research undergone in other cultures provide support for the cross-cultural validity of the ZKPQ and the AFFM as a method of evaluating personality functioning.

Finally, the use of self-reports and spouse reports obtained with reliable and valid instruments in personality assessment is not only useful for research purposes but for applied purposes as well. For example, it could be useful in assisting for diagnosis, tailoring treatment techniques, and predicting compliance and success (Mutén, 1991). Comparisons among self-ratings and spouse ratings could also be useful in some psychotherapeutic formulations and in some assessment contexts that clearly encourage biased responding such as in selection settings or in forensic contexts. It would also be a useful alternative in cases in which the ability to report accurately is impaired by cognitive deficits or by psychiatric illness. Thus, using spousal ratings to assess personality in general is a potentially fruitful avenue for future research.

ACKNOWLEDGMENTS

This research was supported by the Autonomous University of Barcelona, Spain with Grant EMER99-18. We thank the students and participants that cooperated in the study and Carmina Puig Sobrevals for her logistical help. Comments by Marvin Zuckerman and by the anonymous reviewers are gratefully acknowledged.

REFERENCES

Ball, S. (1995). The validity of an alternative Five-factor measure of personality in cocaine abusers. *Psychological Assessment*, 7, 148-154.
 Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 1, 307-310.
 Borgama, E. F. (1964). The structure of personality characteristics. *Behavioral Science*, 9, 8-17.

Bus, D. M. (1984). Marital assortment for personality dispositions: Assessment with three different data systems. *Behavior Genetics*, 14, 111-123.
 Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
 Caspi, A., Herbener, E. S., & Ozer, D. J. (1992). Shared experiences and the similarity of personalities: A longitudinal study of married couples. *Journal of Personality and Social Psychology*, 62, 281-291.
 Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO Personality Inventory*. Odessa, FL: Psychological Assessment Resources.
 Costa, P. T., Jr., McCrae, R. R., & Arnberg, D. (1980). Enduring dispositions in adult males. *Journal of Personality and Social Psychology*, 38, 793-800.
 De Pascalis, V., & Russo, P. (2003). Zuckerman-Kuhlman Personality Questionnaire: Preliminary results of the Italian version. *Psychological Reports*, 92, 965-974.
 Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41, 417-440.
 Eysenck, H. J. (1967). *The biological basis of personality*. Springfield, IL: Thomas.
 Eysenck, H. J. (1992a). Four ways five factors are not basic. *Personality and Individual Differences*, 13, 667-673.
 Eysenck, H. J. (1992b). A reply to Costa and McCrae. P or A and C—The role of theory. *Personality and Individual Differences*, 13, 867-868.
 Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences*. London: Plenum.
 Gomà-i-Freixanet, M. (1997). Consensual validity of the EPQ: Self-Reports and Spouse-Reports. *European Journal of Psychological Assessment*, 13, 179-185.
 Gomà-i-Freixanet, M., Valero, S., Puntí, J., & Zuckerman, M. (2004). Psychometric properties of the Zuckerman-Kuhlman Personality Questionnaire in a Spanish sample. *European Journal of Psychological Assessment*, 20, 134-146.
 Johnson, J. A. (1997). Units of analysis for the description and explanation of personality. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 73-93). San Diego, CA: Academic.
 Kuhlman, D., Zuckerman, M., Gomà-i-Freixanet, M., & Shiomu, K. (2003). *Cultural and gender differences in personality within three cultures*. Manuscript submitted for publication.
 McCrae, R. R. (1982). Consensual validation of personality traits: Evidence from self-reports and ratings. *Journal of Personality and Social Psychology*, 43, 293-303.
 McCrae, R. R., & Costa, P. T., Jr. (1983). Joint factors in self-reports and ratings: Neuroticism, extraversion and openness to experience. *Personality and Individual Differences*, 4, 245-255.
 McCrae, R. R., & Costa, P. T., Jr. (1987). Validation of the Five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52, 81-90.
 McCrae, R. R., & Costa, P. T., Jr. (1989). Different points of view: Self-reports and ratings in the assessment of personality. In J. P. Forgas & M. J. Innes (Eds.), *Recent advances in social psychology: An international perspective* (pp. 429-439). Amsterdam: Elsevier.
 Mischel, W. (1968). *Personality and assessment*. New York: Wiley.
 Montag, I. (2001). *The Hebrew ZKPQ: Preliminary reliability and validity results*. Unpublished manuscript.
 Mutén, E. (1991). Self-reports, spouse ratings, and psychophysiological assessment in a behavioural medicine program: An application of the Five-factor model. *Journal of Personality Assessment*, 57, 449-464.
 Norman, W. T., & Goldberg, L. R. (1966). Raters, rates, and randomness in personality structure. *Journal of Personality and Social Psychology*, 4, 681-691.
 Ostendorf, F., & Angleitner, A. (1994). A comparison of different instruments proposed to measure the Big Five. *European Review of Applied Psychology*, 44, 45-53.
 O'Sullivan, D., Zuckerman, M., & Kraft, M. (1996). The personality of prostitutes. *Personality and Individual Differences*, 21, 445-448.

- O'Sullivan, D., Zuckerman, M., & Kraft, M. (1998). Personality characteristics of male and female participants in team sports. *Personality and Individual Differences*, 25, 119-128.
- Prieto, L., Lamarca, R., & Casado, A. (1998). La evaluación de la fiabilidad en las observaciones clínicas: El coeficiente de correlación intraclass [The assessment of reliability in clinical observations: The intraclass correlation coefficient]. *Medicina Clínica*, 110, 142-145.
- Shiomi, K., Kuhlman, D., Zuckerman, M., Joireman, J., Sato, M., & Yata, S. (1996). Examining the validity and reliability of a Japanese version of the Zuckerman-Kuhlman Personality Questionnaire (ZKPQ). *Hyogo University of Teacher Education Journal*, 2, 1-13.
- Watson, D., & Clark, L. A. (1991). Self- versus peer ratings of specific emotional traits: Evidence of convergent and discriminant validity. *Journal of Personality and Social Psychology*, 60, 927-940.
- Wa, Y., Wang, W., Du, W., Li, J., Jiang, X., & Wang, Y. (2000). Development of a Chinese version of the Zuckerman-Kuhlman Personality Questionnaire: Reliabilities and gender/age effects. *Social Behaviour and Personality*, 28, 241-250.
- Zuckerman, M. (1991). *Psychobiology of personality*. New York: Cambridge University Press.
- Zuckerman, M. (1992). What is a basic factor and which factors are basic? Turtles all the way down. *Personality and Individual Differences*, 13, 675-681.
- Zuckerman, M. (2002). Zuckerman-Kuhlman Personality Questionnaire (ZKPQ): An alternative five-factorial model. In B. De Raad & M. Perugini (Eds.), *Big Five assessment* (pp. 377-396). Seattle, WA: Hogrefe & Huber.
- Zuckerman, M., & Kuhlman, D. (1993). *Norms for the Zuckerman-Kuhlman Personality Questionnaire (ZKPQ)*. Unpublished manuscript.
- Zuckerman, M., Kuhlman, D., & Camac, C. (1988). What lies beyond E and N? Factor analyses of scales believed to measure basic dimensions of personality. *Journal of Personality and Social Psychology*, 54, 96-107.
- Zuckerman, M., Kuhlman, D., Joireman, J., Teta, P., & Kraft, M. (1993). A comparison of three structural models for personality: The Big Three, the Big Five, and the Alternative Five. *Journal of Personality and Social Psychology*, 65, 757-768.
- Zuckerman, M., Kuhlman, D., Thornquist, M., & Kiers, H. (1991). Five (or three): Robust questionnaire scale factors of personality without culture. *Personality and Individual Differences*, 12, 929-941.

Montserrat Gomà-i-Freixanet
 Department of Health Psychology
 Autonomous University of Barcelona
 Edifici B
 E-08193 Bellaterra
 Catalonia, Spain
 Email: montserrat.goma@uab.es

Received April 27, 2004
 Revised September 23, 2004

DISCRIMINANT VALIDITY OF THE ZKPQ IN A SAMPLE MEETING BPD DIAGNOSIS VS. NORMAL-RANGE CONTROLS

Montserrat Gomà-i-Freixanet, PhD, Joaquim Soler, PsyD,
Sergi Valero, PsyD, Juan Carlos Pascual, MD,
and Victor Pérez Sola, MD, PhD

Widiger and Simonsen (2005) state that given the limitations of the categorical model of Personality Disorders classification proposals are to be expected for dimensional classifications. The purpose of this paper is to test the alternative five factorial model (AFFM) of personality in a sample with PDs. Subjects were administered the ZKPQ to test the discriminant capacity of the AFFM in classifying subjects diagnosed with BPD ($n = 74$) vs normal-range controls ($n = 148$) paired by age and sex, and identifying sensitive and/or specific dimensions that can be of help in diagnosing BPD. The results showed that high scores on N-Anx and Imp-SS, and low scores on Act are prognostic factors for being diagnosed with BPD. Likewise, this model correctly classified 88% of subjects with a kappa index of 0.73. The AFFM of personality appears to have a substantial power for predicting SCID-II interview-based BPD diagnosis.

The classification of personality disorders proposed in the different versions of the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., rev. ed.; *DSM-IV*; American Psychiatric Association, 2000) continues to be the subject of discussion and debate. Two of the cardinal points of debate are the categorical nature of the *DSM* system and the contents of its classification. Although some improvements have been introduced in the different versions of the *DSM*, descriptions of personality disorders are still largely the result of committee deliberations with limited empirical support (Livesley, 1987; Widiger & Frances, 1987) and little attention to conceptions derived from the study of normal personality structure. Since the early claims (Eysenck, 1987; Frances, 1980) that personality disorders

From Department of Health Psychology, Universitat Autònoma de Barcelona, Catalonia, Spain (M. G.-F.); Department of Psychiatry, Hospital de la Santa Creu i St. Pau, CIBER-SAM, Barcelona, Catalonia, Spain (J. S., J. C. P., V. P. S.); and Department of Psychiatry, Hospital Universitari Vall d'Hebron, Barcelona, Catalonia, Spain (S. V.).

Address correspondence to Montserrat Gomà-i-Freixanet, PhD., Department of Health Psychology, Universitat Autònoma de Barcelona, Edifici B, E-08193 Bellaterra, Catalonia, Spain; E-mail: montserrat.goma@uab.es, www.zkpq.com

(PDs) could be more appropriately represented by a dimensional model than by a categorical one, there is growing unequivocal evidence giving empirical support to alternative dimensional models (e.g., Clark, 1993; Dyce & O'Connor, 1998; Livesley, 1998; Widiger & Costa, 1994; Wiggins & Pincus, 1989). Although the diagnostic approach of the present version of the *DSM* is categorical in nature, it simultaneously stresses that enduring and maladaptive personality characteristics only constitute a PD when they cause significant functional impairment or subjective distress: a view consistent with dimensional approaches that conceptualize PDs as extreme variants of normal personality traits (e.g., Livesley, Schroeder, Jackson, & Jang, 1994; Widiger, 1992, 1993).

In a recent seminal paper, Widiger and Simonsen (2005) state that given the recognition of the many limitations of the categorical model of PDs classification, "one expected response is proposals for dimensional classifications" (p. 110). It seems that for a future *DSM* edition the best bet for PDs diagnosis could be an integration of the different dimensional models into a hierarchical structure where the contributions and potential advantages of each one of the models could be taken into account. Therefore, they expect that a common structure is likely to be found among these models as they come from the study of the normal personality structure and have the common aim of identifying the fundamental dimensions that underlie normal and maladaptive personality. Among these dimensional models, Widiger and Simonsen (2005) cite the model of Zuckerman, Kuhlman, Joireman, Teta, and Kraft (1993). This model, described as the alternative five factorial model (AFFM), emerged from a series of factor analyses of scales believed to measure basic dimensions of personality or temperament, particularly those used in psychobiological research (Zuckerman, Kuhlman, & Camac, 1988; Zuckerman, Kuhlman, Thornquist, & Kiers, 1991). The basic traits of the AFFM are Neuroticism-Anxiety (N-Anx), Activity (Act), Sociability (Sy), Impulsive Sensation Seeking (ImpSS), and Aggression-Hostility (Agg-Host) and are measured by the Zuckerman-Kuhlman Personality Questionnaire (ZKPQ; Zuckerman & Kuhlman, 1993). In this model, no measures of cultural interests or intellectual styles were included because of Zuckerman's (1984) conception that basic traits should be easily comparable to traits in other species and found throughout the human lifespan. In the same line of thinking, Aggression rather than Agreeableness, and Impulsive Sensation Seeking rather than Conscientiousness were included. Furthermore, the broad dimension of Extraversion was divided into the separate factors of Activity and Sociability (Zuckerman, 2002) because in earlier studies Act emerged as a distinct factor at the five-factor analyses of scales (Zuckerman et al., 1988; Zuckerman et al., 1991). Due to its identification as a basic developmental trait (see, for example, Buss & Plomin, 1984; Thomas & Chess, 1977) activity level merits a distinctive assessment as a major trait of temperament in the child as well as of personality in the adult human. Moreover, the distinction between Hostility and Anxiety is also important because both

traits have different psychobiological bases (Gray, 1982), and should not be confounded within a single factor. On the other hand, Imp and SS are conceptually closely related and have many important psychobiological correlates (Zuckerman, 1983, 1984, 1991, 1994). Together with Socialization they form a distinctive factor in five-factor analyses of scales (Zuckerman et al., 1991) and of items. Finally, a measure of social desirability was also included in the questionnaire (Infrequency scale) to ensure none of the basic traits were affected by this response set bias. The ZKPQ questionnaire developed to assess this alternative model has demonstrated good internal reliability, temporal stability, validity and cross-cultural replicability (Gomà-i-Freixanet, Valero, Puntí, & Zuckerman, 2004; Gomà-i-Freixanet, Wismeijer, & Valero, 2005; Zuckerman, 2002).

As Widiger and Simonsen (2005) recognize, there is a need for supportive empirical data on each proposed alternative dimensional model of PD in making decisions regarding which specific component of each model should be included within this proposed integrative hierarchical structure. Therefore, the aim of the present study was (1) to investigate the characteristic ZKPQ profile of Borderline Personality Disorder (BPD) patients when compared with demographically matched normal-range controls (i.e., which is the dimensional profile of BPD patients on dimensions derived from the AFFM) and (2) to test the discriminant capacity of the ZKPQ in classifying subjects diagnosed with BPD vs. normal-range controls paired by age and sex, and identify sensitive and/or specific dimensions that can be of help in diagnosing BPD. Specifically, the ability of the ZKPQ in predicting a categorical PD diagnosis was assessed. The present study provides first time evidence of the specific ZKPQ profile of patients meeting BPD as assessed by the SCID-II.

METHOD

PARTICIPANTS

For the purpose of this study, we used two samples matched by age and sex. The age range for both samples was from 19 to 43 years, with a proportion of 87.8% being females. The clinical sample consisted of a convenience sample of 74 outpatients from the Department of Psychiatry at the Hospital de la Santa Creu i Sant Pau from the Autonomous University of Barcelona, Spain. This sample, fulfilling BPD diagnosis, consisted of 65 women and 9 men (mean \pm SD age 27.32 ± 5.32 years) with 55.6% of them having completed high school studies. The frequencies of *DSM-IV* co-occurring Axis II diagnoses (including PDs from the *DSM-IV* appendix) observed were: (a) schizoid, 1.4%; (b) schizotypal, 2.9%; (c) paranoid, 24.6%; (d) antisocial, 10.1%; (e) histrionic, 8.7%; (f) narcissistic, 2.9%; (g) dependent, 13%; (h) avoidant, 13%; (i) obsessive-compulsive, 14.5%; (j) depressive, 44.9%; and (k) negativistic, 26.1%.

To test for the clinical specificity of the dimensional personality profile, the BPD sample was matched by sex and age with a normal-range sample which acted as a control group. A case-control strategy was used, randomly selecting two controls for each case. In total, the control sample comprised of 148 subjects, 130 women and 18 men (mean \pm SD age 27.32 \pm 5.32 years) with 63.9% of them having completed college studies. This control group was extracted from a comprehensive general population sample pool of 1,169 subjects, which matched the IDESCAT Census Projections for the year 2000 in the distribution of age and sex groups (Gomà-i-Freixanet et al., 2003). In line with other studies (Pukrop, 2002), BPD patients attained significantly lower levels of education (Chi-square = 31.79, $p = 0.0005$) than controls.

ASSESSMENT INSTRUMENTS

The *Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II; First, Gibbon, Spitzer, Williams, & Benjamin, 1997)* was used as external diagnostic standard for DSM-IV PD assessment. Although there is a lack of an agreed-upon "gold standard" for psychiatric diagnosis, the previous version of this instrument has shown to have adequate interrater reliability and to be useful in providing fine discriminations among the Axis II PDs (e.g., First, Spitzer, Gibbon, & Williams, 1995). The Spanish translation of the *DSM-III-R* version of the SCID-II gave an overall kappa of 0.85 (Gómez Beneyto et al., 1994).

The *Zuckerman-Kuhlman Personality Questionnaire (ZKPQ; Zuckerman & Kuhlman, 1993)* consists of 5 content scales. Additionally, two facet scores can be obtained from the Act scale: Need for General Activity, impatience and restlessness (GenAct); and need for Work Activity (WorkAct), and from the Sy scale: Parties and friends (Parties) and Isolation intolerance (Isol). In total it has 99 dichotomous items, in sentence format and true-false response set. The ZKPQ also includes an *Infrequency* scale (In-freq, 10 items) detecting inattention to the task. The items are mostly exaggerated, true scored, socially desirable but unlikely to be completely true statements about anyone. This scale is highly skewed, with most scores around 0 or 1. In Spanish samples, scores higher than 4 in this scale are considered to indicate questionable validity for that record (Gomà-i-Freixanet et al., 2004).

Participants from both samples completed the Spanish version of the ZKPQ (Gomà-i-Freixanet et al., 2004). This instrument has shown good psychometric properties in Spanish samples, with test-retest reliability coefficients over a 2-week period ranging from 0.77 to 0.91, internal consistency α coefficients ranging from 0.67 to 0.84 and convergent, discriminant, and consensual validity. The factorial structure has also been replicated in Spanish samples with congruence coefficients ranging from 0.88 to 0.96, and from 0.84 to 0.92 in a female and male sample respectively (Gomà-i-Freixanet et al., 2004; Gomà-i-Freixanet et al., 2005).

PROCEDURE

Subjects fulfilling *DSM-IV* diagnostic criteria for BPD (American Psychiatric Association, 1994) and the rest of inclusion/exclusion criteria were accepted to participate in a single-centre, randomized, double-blind, and placebo-controlled clinical trial that was including pharmacological therapy. The inclusion criteria to participate in were: (1) meeting *DSM-IV* diagnostic criteria for BPD as assessed by the SCID-II for *DSM-IV* Axis II PDs; (2) having an age range from 18 to 45 years; (3) no current episodes of Axis I disorders or unstable symptomatology of the comorbid Axis I disorders; (4) having a Clinical Global Impression (CGI; Guy, 1976) severity of illness score ≥ 4 ; (5) not receiving psychotherapy; and (6) for female subjects, using medically accepted contraception. Subjects with organic brain syndrome, schizophrenia, drug-induced psychosis, alcohol, or other substance dependence, bipolar disorder, mental deficiency, or major depressive disorder were not included in the study. Participants also underwent complete physical examination, and laboratory and pregnancy tests before entering to the study. The preceding clinical history and diagnostic interviews were conducted by two clinical psychologists from the BPD unit with wide experience using diagnostic interviews for Axis II, and were administered and scored blind to the ZKPQ scores. Subjects were administered the ZKPQ before the pharmacological treatment started. The Ethical Committee of the institution approved the protocol and all subjects gave their written informed consent before participating in this study.

As stated above, the control sample was a random subsample of a much larger sample. The questionnaires were answered anonymously and only demographic data such as sex, age, education level, and place of residence were recorded. All respondents participated voluntarily in the study and did not receive any emolument for their participation. As the study was not intrusive in any way, neither informed consent waivers nor participant debriefing following participation were required from the controls.

STATISTICAL ANALYSIS

The *DSM-IV* cut-off points were used for categorical diagnosis of BPD (First et al., 1997). The data analyses followed 3 steps: descriptive analyses, predictive analyses, and model adjustment. In the first step, differences between groups were tested using two-tailed independent Student's *t*-test, and Pearson's correlation coefficients among ZKPQ scales for both groups were carried out. Subsequently, we performed logistic regression analyses to study the independent contribution of each ZKPQ dimension and facet to the prediction of the categorical diagnosis of BPD. We opted for logistic regression analysis because, contrary to Pearson's correlation, logistic regression analysis adjusts associations among variables. Finally, for adjusting the model and checking the overall discriminatory power of the ZKPQ, different cut-off points for ZKPQ dimensions and facets were ob-

tained, and concordance was then studied by kappa indexes. Sensitivity and specificity indexes, hit rates (true positives + true negatives)/*n*, along with positive and negative predictive values were also obtained. Moreover, the Receiver Operating Characteristic (ROC) curve was produced with the selected cut-off point. The ROC curve summarizes all pairs of sensitivity and specificity values which can be achieved when the threshold is changed from low to high scores. The area under the ROC curve represents the ability of a scale to identify a particular PD; precisely in this case, the ability of the ZKPQ to identify a BPD diagnosis. An area under the ROC curve of 0.5 indicates an inaccurate test, whereas an area of 1.0 indicates perfect diagnostic accuracy.

RESULTS

COMPARISON OF BPD PATIENTS WITH NORMAL-RANGE CONTROLS

Means, standard deviations, *t*-test differences, Cohen's *d* and Cronbach's alphas of the ZKPQ dimensions in BPD and control groups are shown in Table 1. Both groups differed significantly on all scales but on the Infreq scale: N-Anx ($t_{216} = 15.50$, $p = 0.0005$), Act ($t_{220} = 2.15$, $p = 0.032$), Sy ($t_{220} = 3.49$, $p = 0.001$), ImpSS ($t_{220} = 5.82$, $p = 0.0005$), Agg-Host ($t_{220} = 5.76$, $p = 0.0005$), and Inf ($t_{220} = 1.31$, $p = 0.192$). The BPD group scored significantly higher on N-Anx, ImpSS and Agg-Host, and significantly lower on Act and Sy. The absence of significant differences between both groups on the Infreq scale (measuring inattention to the task) and the low scores (around 0 or 1) gives additional validity to the obtained data in the BPD group. Also worth mentioning is the high homogeneity of the scores on the N-Anx scale in the BPD group ($SD = 2.76$). Internal consistencies ranged from 0.73 to 0.79 in the BPD sample, and from 0.66 to 0.87 in the control group sample. Internal reliabilities of the control sample are similar to those found in other groups with the same cultural background and similar age ranges (Gomà-i-Freixanet et al., 2004; Gomà-i-Freixanet et al., 2005). Reliabilities

TABLE 1. Mean Differences on the ZKPQ Dimensions Between BPD and Normal Control Groups, Cronbach's Alphas and Cohen's *d*

Scale	BPD (n = 74)			Control (n = 148)			<i>t</i>	<i>p</i>	Cohen's <i>d</i>
	<i>M</i>	<i>SD</i>	α	<i>M</i>	<i>SD</i>	α			
ZKPQ									
N-Anx	16.04	2.76	0.73	8.13	4.82	0.87	15.50	.0005	1.86
Act	7.28	3.56	0.78	8.36	3.48	0.73	2.15	.032	0.31
Sy	6.89	3.46	0.75	8.56	3.31	0.76	3.49	.001	0.50
ImpSS	12.16	4.08	0.79	8.35	4.83	0.86	5.82	.0005	0.83
Agg-Host	10.35	3.52	0.79	7.76	2.96	0.66	5.76	.0005	0.82
Infreq	1.30	1.12	—	1.08	1.18	—	1.31	.192	0.18

Note. ZKPQ = Zuckerman-Kuhlman Personality Questionnaire; N-Anx = Neuroticism-Anxiety; Act = Activity; Sy = Sociability; ImpSS = Impulsive Sensation Seeking; Agg-Host = Aggression-Hostility; Infreq = Infrequency.

of the BPD sample are similar to those found in other clinical samples (Ball, 1995).

CORRELATION ANALYSES

For comparison, Table 2 shows Pearson's correlation coefficients among ZKPQ scales for both groups. In general, as in a university student sample (Gomà-i-Freixanet et al., 2004) and in a general population sample (Gomà-i-Freixanet et al., 2005) the correlation coefficients for the control group are low in magnitude, although they are statistically significant due to the sample size. In BPD patients, N-Anx was significantly and positive correlated with both ImpSS and Agg-Host dimensions, while there was no such correlation in the general population sample.

PREDICTIVE POWER OF THE ZKPQ AT THE LEVEL OF DIMENSIONS

In order to test the ability of the ZKPQ dimensions to predict a categorical BPD diagnosis, we performed a logistic regression analysis using the stepwise method for entering the variables. Given the small number of predictors, alpha was set at 0.05 for entry into the equation. All five ZKPQ dimensions (Infreq scale is not considered a dimension in the normative sense) were included in the logistic regression model where the group of origin (either BPD or control) was the dependent variable. The groups were coded into the model as 1 per BPD group and 0 per control group, and the resulting final model was statistically significant ($\chi^2 = 149.75$, $p = 0.0005$). Table 3 shows that 3 out of 5 dimensions entered into the model: N-Anx, ImpSS and Act. The Wald statistic informs on the impact of each of the predictive dimensions for the categorical diagnoses of BPD. The obtained model shows that having high scores on N-Anx and ImpSS, and low scores on Act is a prognostic factor with high probability

TABLE 2. Correlations Between ZKPQ Dimensions in BPD Group (n = 74) and Control Group (n = 148) Separately

	Control					
	N-Anx	Act	Sy	ImpSS	Agg-Host	Infreq
BPD						
N-Anx		.072	-.167*	-.009	.095	-.078
Act	.148		.170*	.364***	.260**	.236**
Sy	-.111	.337**		.308***	.190*	.061
ImpSS	.303**	.317**	.334**		.347***	.157
Agg-Host	.473***	.236*	.164	.477***		-.055
Infreq	-.088	-.021	.030	.229*	-.204	

Note. In the upper-right side correlations for the control group are reported; in the lower-left side correlations for the BPD group.

N-Anx = Neuroticism-Anxiety; Act = Activity; Sy = Sociability; ImpSS = Impulsive Sensation Seeking; Agg-Host = Aggression-Hostility; Infreq = Infrequency.

* $p < .05$, ** $p < .01$, *** $p < .001$, two-tailed tests.

**TABLE 3. Logistic Regression Analysis
Output of ZKPQ Dimensions**

Scale	B	Wald	Sig.	Exp(B)
N-Anx	.452	44.75	<.0005	1.57
ImpSS	.196	13.21	<.0005	1.22
Act	-.220	10.50	.001	0.80

Note. BPD = 1; control group = 0
N-Anx = Neuroticism-Anxiety; ImpSS = Impulsive
Sensation Seeking; Act = Activity.

of being endorsed with a BPD diagnosis. Using the following algorithm,

$$p = \frac{1}{1 + e^{6.77 - .452*N-Anx - .196*ImpSS - .220*Act}}$$

estimations about the probability of being endorsed with a BPD diagnosis can be made. Using the default probabilistic cut-off point of 0.5, the obtained model had a sensitivity (the probability that a test will be positive given a patient with the condition) of 82.4%; a specificity (the probability that a test will be negative given a patient without a condition) of 89.9%; a positive predictive value (PPV: the probability that a patient will have a condition given a positive test result) of 80.3%; a negative predictive value (NPV: the probability that a patient will not have a condition given a negative test result) of 91.1%; a hit rate of 87.4% and a kappa index of 0.72 ($p < 0.0005$).

ACCURACY OF THE MODEL AT THE LEVEL OF DIMENSIONS

Accuracy can be improved if different cut-off points are chosen. The accuracy of the model exemplified by concordance and predictive indexes can be seen in Table 4. Analyses showed that, predictably, sensitivity was increased and specificity decreased or vice versa when lower or higher cut-off points were used. It seemed to us that the optimal adjustment was observed using the probabilistic cut-off point of 0.4 where we found a high and balanced sensitivity (87.8%) and specificity (87.8%), with a high PPV (78.3%) and NPV (93.5%). Likewise, this cut-off point showed the highest hit rate, with 87.8% of the participants correctly classified and the highest

**TABLE 4. Cut-Off Points and Model Adjustment for ZKPQ Dimensions
and Concordance with BPD Diagnosis**

Cut-off points	%						
	Sensitivity	Specificity	Hit Rate	PPV ¹	NPV ²	Kappa	p
.3	90.5	84.5	86.5	74.4	94.7	.712	.0005
.4	87.8	87.8	87.8	78.3	93.5	.734	.0005
.5	82.4	89.9	87.4	80.3	91.1	.718	.0005
.6	77.0	92.6	87.4	83.8	88.9	.710	.0005
.7	68.9	95.9	86.9	89.4	86.0	.688	.0005

Note. ¹PPV = Positive Predictive Value (probability of true BPD after a positive result in predictor); ²NPV = Negative Predictive Value (probability of true non-BPD after a negative result in predictor).

kappa index of 0.73 ($p < 0.0005$). This kappa index, the usual metric of categorical agreement, informs of the accuracy of the model, i.e., of the standardized difference between the observed and the expected agreement. In the obtained model, the accuracy of the ZKPQ is not only statistically significant but high in magnitude. The obtained area under the ROC curve with a cut-off point of 0.4 was 0.94 (CI 95% 0.90–0.97), indicating almost perfect diagnostic accuracy.

PREDICTIVE POWER OF THE ZKPQ AND ACCURACY OF THE MODEL AT THE LEVEL OF FACETS

Given that the ZKPQ can offer facet scores in 3 out of the 5 scales, we performed a second logistic regression analysis using the stepwise method for entering the variables. We entered into the equation 2 dimensions (N-Anx and Agg-Host) plus 6 facets (GenAct, WorkAct, Parties, Isol, Imp, and SS). The resulting final model was statistically significant ($\text{Chi-square}_3 = 157.86$, $p = 0.0005$). Table 5 shows that 3 out of 8 scales entered into the model: N-Anx, Imp, and GenAct. The obtained model shows that having high scores on N-Anx and Imp, and low scores on GenAct is a prognostic factor with high probability of being endorsed with a BPD diagnosis. The optimal adjustment seemed to be observed using the probabilistic cut-off point of 0.5 where we found a high sensitivity (79.7%) and specificity (93.2%), with a high PPV (85.5%) and NPV (90.1%). Likewise, this cut-off point showed the highest hit rate, with 88.7% of the participants correctly classified and the highest kappa index of 0.74 ($p < 0.0005$).

DISCUSSION

The overall results show that BPD can be accurately described in terms of dimensional trait profiles. Subjects meeting BPD diagnosis scored higher than controls on N-Anx, ImpSS, AggHost, and lower on Sy and Act. These results are in line with data obtained with subjects meeting the same diagnosis but assessed with different instruments (e.g., when comparing BPD with controls controlling for age and sex and using the NEO, Pukrop (2002) found that BPD also differed on Neuroticism, Agreeableness, and Extraversion). When we tested for the discriminant capacity of the ZKPQ, the dimensions that best discriminated were N, ImpSS, and Act. Although

**TABLE 5. Logistic Regression Analysis
Output of ZKPQ Facets**

Scale	B	Wald	Sig.	Exp(B)
N-Anx	.428	37.91	<.0005	1.53
Imp	.485	18.22	<.0005	1.62
GenAct	-.231	5.33	.021	0.79

Note. BPD = 1; control group = 0.
N-Anx = Neuroticism-Anxiety; Imp = Impulsivity; GenAct = General Activity.

AggHost and Sy were significantly different at the level of bivariate analysis, these two dimensions recapitulated at the multivariate analysis due to the statistical significance of those that remained in the model. The kappa of 0.73 obtained with the ZKPQ is higher than those found in the literature with SCID-II BPD diagnosis and other instruments such as the PDQ-4+ (Fossati et al., 1998; $k = 0.19$) and the TCI (Gutiérrez, Sangorrin, Martín-Santos, Torres, & Torrens, 2002; $k = 0.53$). Furthermore, to make a more detailed analysis of the BPD profile, we introduced the facets, whenever possible, and the variables that entered into the model were N-Anx, Imp, and GenAct. Although for diagnosis purposes it is easier for the clinician to work at the level of dimensions, for research purposes it is interesting to know that the component of Imp (and not that of SS) best discriminates among groups. In the same manner, in relation to the E dimension, it is interesting to differentiate among the different components of that dimension. It is the Act dimension rather than that of Sy, that predicts BPD diagnosis, and specifically within the Act dimension itself, it is GenAct rather than WorkAct that best discriminates.

The present study sought to test specifically the AFFM in BPD patients for several reasons. First, according to Livesley (2005), BPD is one of the most valid *DSM-IV* diagnoses along with the antisocial, schizoid-avoidant, and obsessive compulsive PDs. Thus, data obtained using this diagnosis provide an additional guarantee regarding its validity. Second, BPD symptomatology seems to be well characterized by the AFFM. Although this instrument was not designed specifically to measure personality pathology, the ZKPQ provides data on basic personality traits that may be reflected in a wide range of adaptive and maladaptive behaviors as well as of habits and attitudes (e.g., two of the cardinal symptoms of BPD namely affective dysregulation and behavioral dyscontrol could well be tapped by N-Anx and ImpSS, respectively). Third, there are some controversies about which dimensions best describe BPD. According to Widiger and Trull (1997), subtle distinctions exist in relation to how personality domains are conceptualized, thus provoking a disagreement at lower-order levels about the selection and placement of particular facets into particular scales. Livesley (2005) states that this disagreement stems partly from the lack of clarity about the exact meaning of secondary and primary traits; what is a primary trait to one theorist is a secondary trait to others, e.g., Impulsivity trait placement in the alternative models of Eysenck (1991), Costa and McCrae (1998), and Zuckerman (2002). And fourth, Borderline diagnosis seems to be characterized by a core phenotype consisting of affective dysregulation, behavioral dyscontrol, and disturbed interpersonal relationships. These dimensions have strong genetic influences on traits that underlie them: N-Anx is associated with increased responsivity of cholinergic systems, and Impulsive Aggression with reduced serotonergic activity in the brain (Skodol et al., 2002). Therefore, biologically rooted models such as the AFFM connect temperamental dispositions with biological markers such as neurotransmitters or hormones, that in turn can serve as genetic

criteria to supplement the usual statistical criteria used to determine the number and content of secondary constructs (Livesley, 2005).

CONCLUSIONS

The data obtained from this study is the first attempt to provide validation of the ZKPQ in a clinical sample meeting BPD criteria by examining its psychometric properties and its discriminant and predictive power. Nevertheless, it is important to acknowledge several important issues regarding the interpretation of the present results. First, our results could be different with the use of an alternative model of PD diagnosis. In addition, the probabilistic cut-off points chosen must be considered as preliminary until they have been cross-validated in independent samples. Furthermore, the sole reliance on patient's reports to assess personality traits could produce an underestimation of traits with low social desirability. However, since we did not find significant differences on the Infreq scale between controls and patients, it seems unlikely that our results lose relevance because of this positive response set bias.

In summary, the AFFM of personality with its instrument of measurement, the ZKPQ, appears to have a substantial power for predicting SCID-II interview-based BPD diagnosis as compared to other trait systems of assessment. Thus, the measures obtained from the ZKPQ could indicate temperamental vulnerability to BPD that can be triggered by developmental events. Researchers and clinicians whose interest includes general personality functioning as well as maladaptive personality traits might be well served by using this instrument as a useful tool for diagnosing, case conceptualization, differential treatment planning, and predict response to treatment (Widiger, 1993). Accordingly, results of the current study could add to the growing base of knowledge on the utility of dimensional models of personality in the advance of PDs conceptualization.

REFERENCES

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., rev. ed.). Washington, DC: Author.
- Ball, S. A. (1995). The validity of an alternative five-factor measure of personality in cocaine abusers. *Psychological Assessment*, 7, 148-154.
- Buss, A. H., & Plomin, R. A. (1984). *Temperament: Early developing personality traits*. Hillsdale, NJ: Erlbaum.
- Clark, L. A. (1993). *The Schedule for Non-adaptive and Adaptive Personality (SNAP): Manual for administration and scoring*. Minneapolis: University of Minnesota Press.
- Costa, P. T., Jr., & McCrae, R. R. (1998). Six approaches to the explication of facet-level traits: Examples from conscientiousness. *European Journal of Personality*, 12, 117-134.
- Dyce, J. A., & O'Connor, B. P. (1998). Personality disorders and the five-factor model: A test of facet-level predictions. *Journal of Personality Disorders*, 12, 31-45.

- Eysenck, H. J. (1987). The definition of personality disorders and the criteria appropriate to their definition. *Journal of Personality Disorders, 1*, 211-219.
- Eysenck, H. J. (1991). Dimensions of personality: 16, 5 or 3? Criteria for a taxonomic paradigm. *Personality and Individual Differences, 12*, 773-790.
- First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J.B.W. (1995). *Structured clinical interview for Axis I DSM-IV disorders, non-patient edition (SCID/NP, Version 2.0)*. New York: New York State Psychiatric Institute.
- First, M. B., Gibbon, M., Spitzer, R. L., Williams, J.B.W., & Benjamin, L. S. (1997). *Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II)*. Washington, DC: American Psychiatric Association.
- Fossati, A., Maffei, C., Bagnato, M., Donati, D., Donini, M., Fiorilli, M., et al. (1998). Brief communication: Criterion validity of the personality diagnostic questionnaire 4+(PDQ4+) in a mixed psychiatric sample. *Journal of Personality Disorders, 12*, 172-178.
- Frances, A. J. (1980). The DSM-III personality disorders section: A commentary. *American Journal of Psychiatry, 137*, 1050-1054.
- Gomà-i-Freixanet, M., Valero, S., Puntí, J., Comas, J., Salcedo, A., & Zuckerman, M. (2003). Psychometric appraisal of the ZKPQ in a general population. *11th Biennial Meeting of the International Society for the Study of Individual Differences*. Graz, Austria.
- Gomà-i-Freixanet, M., Valero, S., Puntí, J., & Zuckerman, M. (2004). Psychometric properties of the Zuckerman-Kuhlman Personality Questionnaire in a Spanish sample. *European Journal of Psychological Assessment, 20*, 134-146.
- Gomà-i-Freixanet, M., Wismeyer, A. J., & Valero, S. (2005). Consensual validity parameters of the Zuckerman-Kuhlman Personality Questionnaire: Evidence from self-reports and spouse reports. *Journal of Personality Assessment, 84*, 279-286.
- Gómez Beneyto, M., Villar, M., Renovell, M., Pérez, F., Hernández, M., Leal, C., et al. (1994). The diagnosis of personality disorder with a modified version of the SCID-II in a Spanish clinical sample. *Journal of Personality Disorders, 8*, 104-110.
- Gray, J. A. (1982). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. New York: Oxford University Press.
- Gutiérrez, F., Sangorrín, J., Martín-Santos, R., Torres, X., & Torrens, M. (2002). Measuring the core features of personality disorders in substance abusers using the Temperament and Character Inventory (TCI). *Journal of Personality Disorders, 16*, 344-359.
- Guy, W. (Ed.) (1976). *ECDEU Assessment Manual for Psychopharmacology: Publication ADM 76-338*. (pp. 218-222). Washington, DC: US Department of Health, Education, and Welfare.
- Livesley, W. J. (1987). A systematic approach to the delineation of personality disorders. *American Journal of Psychiatry, 144*, 772-777.
- Livesley, W. J. (1998). Suggestions for a framework for an empirically based classification of personality disorder. *Canadian Journal of Psychiatry, 43*, 137-147.
- Livesley, W. J. (2005). Behavioral and molecular genetic contributions to a dimensional classification of personality disorder. *Journal of Personality Disorders, 19*, 131-155.
- Livesley, W. J., Schroeder, M. L., Jackson, D. N., & Jang, K. L. (1994). Categorical distinctions in the study of personality disorder: Implications for classification. *Journal of Abnormal Psychology, 103*, 6-17.
- Pukrop, R. (2002). Dimensional personality profiles of borderline personality disorder in comparison with other personality disorders and healthy controls. *Journal of Personality Disorders, 16*, 135-147.
- Skodol, A. E., Siever, L. J., Livesley, W. J., Gunderson, J. G., Pfohl, B., & Widiger, T. A. (2002). The borderline diagnosis II: Biology, genetics and clinical course. *Biological Psychiatry, 51*, 951-963.
- Thomas, A., & Chess, S. (1977). *Temperament and development*. New York: Brunner/Mazel.
- Widiger, T. A. (1992). Categorical versus dimensional classification: Implications from and for research. *Journal of Personality Disorders, 6*, 287-300.
- Widiger, T. A. (1993). The DSM-III-R categorical personality disorder diagnoses: A critique and an alternative. *Psychological Inquiry, 4*, 75-90.

- Widiger, T. A., & Costa, P. T. (1994). Personality and personality disorders. *Journal of Abnormal Psychology, 103*, 78-91.
- Widiger, T. A., & Frances, A. J. (1987). Interviews and inventories for the measurement of personality disorders. *Clinical Psychology Review, 7*, 49-75.
- Widiger, T. A., & Simonsen, E. (2005). Alternative dimensional models of personality disorder: Finding a common ground. *Journal of Personality Disorders, 19*, 110-130.
- Widiger, T. A., & Trull, T. J. (1997). Assessment of the Five Factor Model of personality. *Journal of Personality Assessment, 68*, 228-250.
- Wiggins, J. S., & Pincus, A. L. (1989). Conceptions of personality disorders and dimensions of personality. *Psychological Assessment, 1*, 305-316.
- Zuckerman, M. (1983). *Biological bases of sensation seeking, impulsivity, and anxiety*. Hillsdale, NJ: Erlbaum.
- Zuckerman, M. (1984). Sensation seeking: A comparative approach to a human trait. *Behavioral and Brain Sciences, 7*, 413-417.
- Zuckerman, M. (1991). *Psychobiology of personality*. New York: Cambridge University Press.
- Zuckerman, M. (1994). *Behavioral expressions and biosocial bases of sensation seeking*. New York: Cambridge University Press.
- Zuckerman, M. (2002). Zuckerman-Kuhlman Personality Questionnaire (ZKPQ): An alternative five-factorial model. In B. De Raad & M. Perugini (Eds.), *Big Five assessment* (pp. 377-396). Seattle, WA: Hogrefe & Huber Publishers.
- Zuckerman, M., & Kuhlman, D. (1993). *Norms for the Zuckerman-Kuhlman Personality Questionnaire (ZKPQ)*. Unpublished manuscript.
- Zuckerman, M., Kuhlman, D., & Carnac, C. (1988). What lies beyond E and N? Factor analyses of scales believed to measure basic dimensions of personality. *Journal of Personality and Social Psychology, 54*, 96-107.
- Zuckerman, M., Kuhlman, D., Joireman, J., Teta, P., & Kraft, M. (1993). A comparison of three structural models for personality: The Big Three, the Big Five, and the Alternative Five. *Journal of Personality and Social Psychology, 65*, 757-768.
- Zuckerman, M., Kuhlman, D., Thornquist, M., & Kiers, H. (1991). Five (or three): Robust questionnaire scale factors of personality without culture. *Personality and Individual Differences, 12*, 929-941.

Spanish normative data of the Zuckerman-Kuhlman Personality Questionnaire in a general population sample

Montserrat Gomà-i-Freixanet and Sergi Valero Ventura*

Autonomous University of Barcelona, and * Hospital Universitari Vall d'Hebron de Barcelona

The Zuckerman-Kuhlman Personality Questionnaire (ZKPQ), a tool designed to measure the alternative five-factor model of personality, was translated and adapted into the Spanish language. To date, there appears to be no normative data for the ZKPQ in general population samples in any culture. The aim of this study was to obtain Spanish normative data for the scales and subscales in a community sample of the general population. The questionnaire was administered to 1,678 participants of both genders (55.8 % females) with ages ranging from 18 to 93 years. The magnitude of the sample, as well as an adequate representativity of both genders and the wide age range being considered, provides a satisfactory description of each scale and subscale of this instrument. The normative data obtained showed a normal distribution of the scales. The norms (means and standard deviations) are presented by gender and five age ranges and are adjusted according to the level of education attained. Gender differences are in accordance with data obtained with instruments measuring similar personality traits. This is the first time the normative data of the ZKPQ is reported in a general community sample in any culture.

Datos normativos españoles del Zuckerman-Kuhlman Personality Questionnaire en una muestra de población general. El Zuckerman-Kuhlman Personality Questionnaire (ZKPQ), un instrumento diseñado para la evaluación del modelo alternativo de los cinco factores de personalidad, ha sido traducido y adaptado a la lengua española. Hasta la fecha, no existen datos normativos en población general en ninguna cultura. El objetivo del presente estudio fue obtener datos normativos españoles de las escalas y subescalas en una muestra de la población general. El cuestionario fue administrado a 1.678 participantes de ambos géneros (55.8 % mujeres) con edades comprendidas entre los 18 y 93 años. La magnitud de la muestra, así como una adecuada representatividad de ambos géneros y el amplio rango de edad considerado, proporcionan una descripción satisfactoria de cada escala y subescala de dicho instrumento. Los datos normativos obtenidos mostraron una distribución normal de las escalas. Las medias y desviaciones estándar son presentadas por género y cinco rangos de edad, y ajustadas por el nivel de educación alcanzado. Las diferencias de género van en la misma dirección que las obtenidas con instrumentos que evalúan rasgos de personalidad similar. Esta es la primera vez que se reflejan datos normativos del ZKPQ en una muestra de población general.

Zuckerman, Kuhlman, Joireman, Teta and Kraft (1993) have proposed an alternative five factor model (AFFM) of personality. This model emerged from a series of factor analyses of scales believed to measure basic dimensions of personality or temperament, particularly those used in psychobiological research (Zuckerman, Kuhlman, & Camac, 1988; Zuckerman, Kuhlman, Thornquist, & Kiers, 1991). Eysenck (1992) and Zuckerman (1992) pointed out that psychobiological studies of personality provide a better understanding of the neurobiological and genetic underpinnings of personality. Relying only on the encoding of personality traits in language may be treacherous, as this encoding

probably reflects the observability of these traits in social interactions and may not necessarily mirror the proportional biological relevance of the traits (Zuckerman, 1992; Zuckerman et al., 1993). Therefore, in contrast to the lexical approach, using psychobiological data allows researchers to explore the causal, biological origins of personality (Eysenck, 1992).

The Zuckerman-Kuhlman Personality Questionnaire (ZKPQ) was developed to measure the dimensions that constitute the AFFM (Zuckerman, 2002; Zuckerman et al., 1993). The basic traits in the AFFM and measured by the Zuckerman-Kuhlman Personality Questionnaire are: Neuroticism-Anxiety (N-Anx), Activity (Act), Sociability (Sy), Impulsive Sensation Seeking (ImpSS) and Aggression-Hostility (Agg-Host). In this model, no measures of cultural interests or intellectual styles were included because of Zuckerman's (1984) conception that basic traits should be easily comparable to traits in other species and found throughout the human lifespan. Similarly, Aggression rather than Agreeableness, and Impulsive Sensation Seeking rather than

Fecha recepción: 9-4-07 • Fecha aceptación: 25-6-07
Correspondencia: Montserrat Gomà-i-Freixanet
Faculty of Psychology
Autonomous University of Barcelona
08193 Bellaterra (Spain)
E-mail: montserrat.goma@uab.es / www.zkpq.com

Conscientiousness were included. Furthermore, the broad dimension of Extraversion was divided into the separate factors of Activity and Sociability (Zuckerman, 2002) because in earlier studies Activity emerged as a distinct factor at the five-factor analyses of scales (Zuckerman et al., 1988; Zuckerman et al., 1991).

Due to its identification as a basic developmental trait (see, for example, Buss & Plomin, 1984; Thomas & Chess, 1977) activity level merits a distinctive assessment as a major trait of temperament in the child as well as of personality in the adult human. Moreover, the distinction between Hostility and Anxiety is also important because both traits have different psychobiological bases (Gray, 1982) and should not be confounded within a single factor. On the other hand, Impulsivity and Sensation Seeking are conceptually closely related and have many important psychobiological correlates (Zuckerman, 1983, 1984, 1991, 1994). Together with Socialization they form a distinctive factor in five-factor analyses of scales (Zuckerman et al., 1991) and of items. Finally, a measure of social desirability was also included in the questionnaire, named Infrequency scale (Infreq) to ensure none of the basic traits were affected by this response set bias and to control for inaccurate responding.

Since its first publication in 1993, the ZKPQ has undergone extensive psychometric testing, demonstrating good internal reliability, temporal stability, validity and cross-cultural replicability (Zuckerman, 2002). Concerning concurrent validity, for example, some studies have already been conducted, e.g., characteristics of drug abusers and prediction of therapy course (Ball, 1995), characteristics of psychopathology (Thornquist & Zuckerman, 1995; O'Sullivan, Zuckerman, & Kraft, 1996) and risk taking (O'Sullivan, Zuckerman, & Kraft, 1998; Zuckerman & Kuhlman, 2000). Several cross-cultural studies have also been performed as well using translations into other languages, e. g. German (Ostendorf & Angleitner, 1994), Japanese (Shiomi, Kuhlman, Zuckerman, Joireman, Sato, & Yata, 1996), Chinese (Wu, Wang, Du, Li, Jiang, & Wang, 2000), or Italian (De Pascalis & Russo, 2003) among others.

The ZKPQ has also been used with Spanish samples (Aluja, García, & García, 2004; Gomà-i-Freixanet, Valero, Puntí, & Zuckerman, 2004; Gutiérrez-Zotes, Ramos, & Sáiz, 2001; Peñate, Ibáñez, & González, 1999; Romero, Luengo, Gómez, & Sobral, 2002). Gomà-i-Freixanet et al. (2004) have reported that the ZKPQ has shown good psychometric properties in non-clinical samples, with test-retest reliability coefficients over a 2-week period ranging from 0.77 to 0.91, internal consistency alpha coefficients ranging from 0.67 to 0.84, and convergent, discriminant and consensual validity (Gomà-i-Freixanet, Wismeijer, & Valero, 2005). The factorial structure has also been replicated with congruence coefficients ranging from 0.88 to 0.96, and from 0.84 to 0.92 in a female and male sample respectively (Gomà-i-Freixanet et al., 2004). This questionnaire has also been used with clinical samples (Gomà-i-Freixanet et al., in press) showing good discriminant properties.

However, even though this questionnaire has undergone much psychometric testing most of the data have been obtained with student's populations. Up to this point, there appears to be no normative data on the ZKPQ in general population samples in any culture. Such normative data are necessary to interpret the figures obtained with the ZKPQ, but are also of interest to address some cross-cultural issues about the assessment of personality in various

countries, e.g. the US compared to European countries. Thus, the primary objective of this study was to obtain Spanish normative data for the scales and subscales of the ZKPQ in a community sample of the general population. These data are fundamentally important to interpret the scores obtained by one individual compared to his own group of reference, either in clinical or non-clinical samples.

Method

Participants

Our community sample consisted of 1,678 participants, 741 males (44.2%) and 937 females (55.8%). The age of the respondents ranged from 18 to 88 yrs for males and 18 to 93 yrs for females, with a total mean age of 40.26 yrs ($SD= 18.84$).

Instrument

The Zuckerman-Kuhlman Personality Questionnaire (ZKPQ; Zuckerman & Kuhlman, 1993) consists of 5 content scales, plus an Infrequency scale that allows eliminating subjects with careless responding. The ZKPQ has in total 99 dichotomous items (in sentence format and true-false response set). The five scales can be described in terms of their typical content:

1. *Neuroticism-Anxiety (N-Anx, 19 items)* items describe frequent emotional upset, tension, worry, fearfulness, obsessive indecision, lack of self-confidence and sensitivity to criticism.
2. *Activity (Act, 17 items)* items describe the need for general activity, an inability to relax and do nothing when the opportunity arises, and a preference for hard and challenging work, an active busy life and high energy level. Two facet scores can be obtained from this scale: Need for General Activity, impatience and restlessness (*GenAct, 9 items*) and need for Work Activity (*WorkAct, 8 items*).
3. *Sociability (Sy, 17 items)* items describe the number of friends one has and the amount of time spent with them, outgoingness at parties and a preference for being with others as opposed to being alone and engaging in solitary activities. Two facet scores can also be obtained: Parties and friends (*Parties, 9 items*) and Isolation intolerance (*Isol, 8 items*).
4. *Impulsive Sensation-Seeking (ImpSS, 19 items)* items involve a lack of planning and the tendency to act impulsively without thinking and the seeking of excitement, novel experiences, and the willingness to take risks for these types of experiences. The ImpSS items are general in content and do not describe specific activities such as risky sports, drinking, drugs, or having sex. These items were eliminated to avoid confounding in studies of persons who actually engage in one or another of these activities. Two facet scores can be obtained from this scale: Impulsivity (*Imp, 8 items*) and Sensation Seeking (*SS, 11 items*).
5. *Aggression-Hostility (Agg-Host, 17 items)* items describe a readiness to express verbal aggression; rude, thoughtless or antisocial behaviour; vengefulness and spitefulness; having a quick temper and impatience with others.

The ZKPQ also includes an *Infrequency* scale (*Infreq. 10 items*). Rather than being regarded as a scale in the normative sense, it should only be used to detect inattention to the task or simply a validity measure for the individual test-taker. The items are mostly exaggerated, true scored, socially desirable but unlikely to be completely true statements about anyone. This scale is highly skewed, with most scores around 0 or 1.

Procedure

For the purpose of this study, subjects were contacted from very different sites (classroom, home, while waiting for a yearly health check, leisure associations, etc.) and asked to answer the questionnaire. The questionnaire was provided with written instructions and an introductory letter explaining globally the goal of the study («the study you will collaborate in attempts to evaluate the functioning of an American questionnaire in our culture»). Most of the questionnaires were administered in a group situation and others individually. In this latter situation, the subject also received a prepaid envelope which had to be posted. Around 90% of the respondents correctly completed the questionnaire. The ZKPQ was administered anonymously, and demographic variables such as age (5 age-ranges), gender and level of education (primary studies, secondary studies, higher education) were obtained. As the study was not intrusive in any way, neither informed consent waivers nor participants debriefing following participation were required. Subjects participated voluntarily in the study and did not receive any emolument for their collaboration.

Statistical analysis

Age and gender effects on scales and subscales were checked by means of an ANOVA. Bonferroni correction for multiple comparisons was not applied because multiple comparisons were not the primary goal of the study, but the behaviour of each single scale regarding age and gender. The effect of the variable «level of

education» on the different scales and subscales of the ZKPQ was checked by means of subsequent ANOVAs. Partial correlation coefficients among ZKPQ scales controlling for age, gender and level of education, were used to assess the independence of the dimensions. Finally, an ANCOVA was performed including as principal effects age and gender, and the level of education as a covariant. With this statistical strategy we have a more adjusted estimation of the normative data. Means and standard deviations for both genders and by the 5 age-ranges were calculated for all the scales and subscales.

Results

The distribution of the sample by the 5 age-ranges and for both genders is shown in Table 1. The cut-off points for the 5 age-ranges were decided with the aim of obtaining a precise description of the variable age taking into account that the number of participants in each range would be sufficiently high to generate adequate estimations. In considering 5 age-ranges and 2 genders, 10 different conditions were generated with a sufficient number of subjects in each condition.

The distributions of the scores for the 6 scales are shown in Figures 1 to 6. The histograms illustrate that the scales show an

Age range	Men	Women	Total
18-25	178	325	503
26-40	216	228	442
41-55	188	192	380
56-70	77	92	169
71-93	82	102	184
Total	741	937	1678

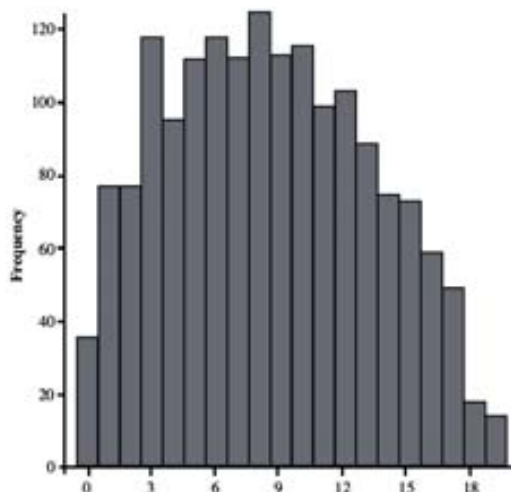


Figure 1. Histogram of the Neuroticism-Anxiety scale

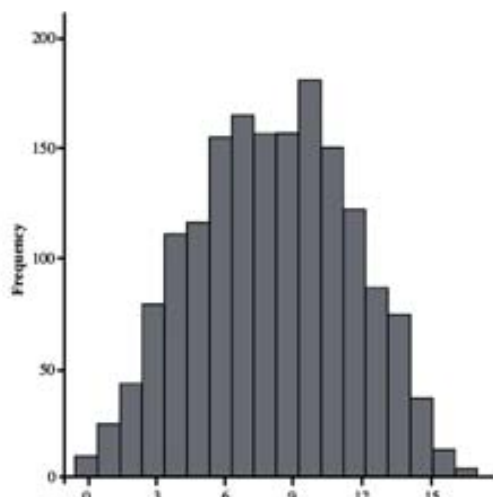


Figure 2. Histogram of the Activity scale

essentially normal distribution, except for the Infreq scale that shows a decreasing monotonic distribution.

In order to assess the effect of the variable age (5 categories) and gender in all the scales and subscales, we performed an ANOVA with these two factors (see Table 2). Both age and gender were significant in almost all conditions, except and only for the variable gender in the Act scale and its GenAct subscale, in the Parties subscale of the Sy scale, and in the Imp subscale of the ImpSS scale.

Also, in order to assess the effect of the variable level of education, subsequent ANOVAs revealed statistical significant differences among scales and subscales. Therefore, it seemed that this variable should be taken into account in the calculation of the normative data. Nevertheless, this variable has 3 categories, age has 5 categories and gender has 2 categories, giving in total 30 (3x5x2) categories. Thus, and because of the categorical nature of this variable, we have generated two dummy variables taking as a

reference the category of «primary studies». The first dummy variable was the contrast between the category «primary studies» vs «secondary studies». The second dummy variable generated was the contrast between the category «primary studies» vs «higher education». These dummy variables were also taken into account for the calculation of the correlation coefficients and the normative data. Table 3 shows ANOVAs for scales and subscales analysing the effect of level of education.

Table 4 shows the correlation matrix among the different scales of the ZKPQ. The correlation coefficients presented are partial correlations controlling for age, gender and level of education in order to obtain better adjusted estimations. As it can be observed, there exists a low level of covariation among the dimensions with correlations ranging from $r = -.01$ to $r = .25$, with an absolute mean interscale correlation of $r = .14$. These data show the independence of the dimensions although almost all of the correlations were significant due to the large sample size.

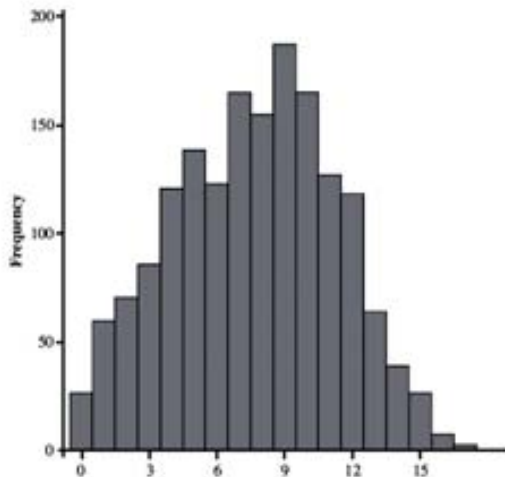


Figure 3. Histogram of the Sociability scale

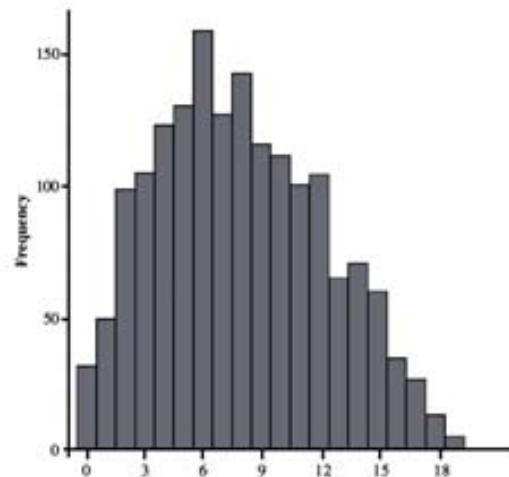


Figure 4. Histogram of the ImpulsiveSensationSeeking scale

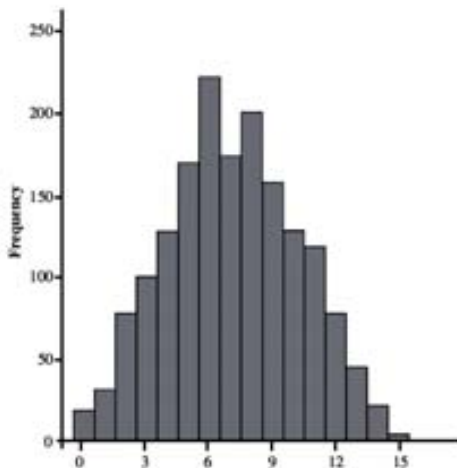


Figure 5. Histogram of the Aggression-Hostility scale

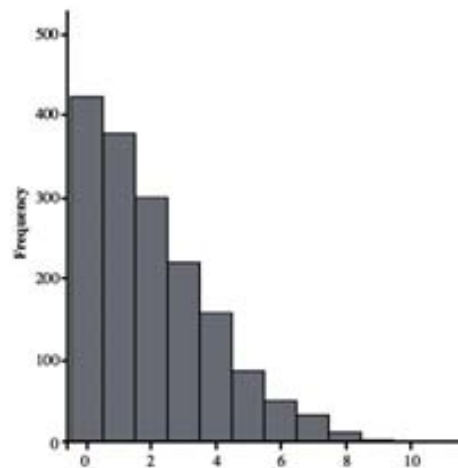


Figure 6. Histogram of the Infrequency scale

Normative data by gender and age are presented in Table 5. The descriptive statistics shown are estimated due to the adjustment of the variable level of education. For each scale and subscale, we performed an ANCOVA taking as principal effects age and gender, and level of education (2 dummy variables) as a covariant. With this statistical strategy, normative means and standard deviations are more precisely estimated.

Discussion

The normative data obtained from a community sample of the general population showed a normal distribution of the scales. This implies that the items on each scale distribute the participants along the continuum on each dimension.

The pertinence of providing norms by gender is supported by the significant differences obtained in all but one scale (Act),

women scoring higher on N-Anx and Sy, and men higher on ImpSS, Agg-Host and Infreq. These results are in accordance with data obtained with instruments measuring similar personality traits, as the EPQ (Lynn & Martin, 1997) or the NEO-PI-R (Costa, Terracciano, & McCrae, 2001). The same applies for providing normative data by different age-ranges. Age was significantly different on all scales and subscales which imply that these dimensions may be influenced by subsequent biological and social factors. Nonetheless, the similarity with findings by Eysenck & Eysenck (1976) made 30 years ago suggests that the age related changes found are independent from potential cohort effects.

Finally, the fact of giving normative data adjusted by level of education is supported by the fact that all but one scale (ImpSS) show significant differences according to the level of education reached. This variable has been seldom considered in studies on normative data. We believe that this variable is worth considering in future publications of normative data because life expectancy is increasing and the older generations attained lower levels of education. Also, these results suggest that specific cross-cultural

Table 2
Two factors (age¹ and gender) ANOVAs for scales and subscales

Scale	F	p
ZKPQ		
N-Anx		
	Age	22.35
	Gender	121.43
		.0005
Act		
	Age	6.88
	Gender	1.56
		.212
GenAct		
	Age	22.78
	Gender	2.48
		.115
WorkAct		
	Age	6.18
	Gender	23.42
		.0005
Sy		
	Age	41.01
	Gender	10.44
		.001
Parties		
	Age	13.63
	Gender	1.57
		.211
Isol		
	Age	53.51
	Gender	16.19
		.0005
ImpSS		
	Age	25.38
	Gender	4.31
		.038
Imp		
	Age	4.99
	Gender	1.65
		.199
SS		
	Age	41.71
	Gender	17.25
		.0005
Agg-Host		
	Age	22.69
	Gender	10.74
		.001
Infreq		
	Age	134.90
	Gender	5.56
		.018

Note: ZKPQ= Zuckerman-Kuhlman Personality Questionnaire; N-Anx= Neuroticism-Anxiety; Act= Activity; Sy= Sociability; ImpSS= Impulsive Sensation Seeking; Agg-Host= Aggression-Hostility; Infreq= Infrequency.
¹ The variable Age has 5 categories

Table 3
ANOVAs for scales and subscales analysing the effect of level of education

	F	p
ZKPQ		
N-Anx	20.21	.0005
Act	4.93	.007
GenAct	23.87	.0005
WorkAct	7.26	.001
Sy	22.77	.0005
Parties	3.37	.035
Isol	40.01	.0005
ImpSS	.78	.459
Imp	13.83	.0005
SS	8.81	.0005
Agg-Host	5.42	.005
Infreq	116.03	.0005

Note: ZKPQ= Zuckerman-Kuhlman Personality Questionnaire; N-Anx= Neuroticism-Anxiety; Act= Activity; Sy= Sociability; ImpSS= Impulsive Sensation Seeking; Agg-Host= Aggression-Hostility; Infreq= Infrequency

Table 4
Partial correlations among scales of the ZKPQ controlling for the effect of age, gender and level of education¹

Scale	N-Anx	Act	Sy	ImpSS	Agg-Host
ZKPQ					
Act	-.009				
Sy	-.052*	.172**			
ImpSS	.157**	.239**	.239**		
Agg-Host	.217**	.097**	.100**	.253**	
Infreq	-.037	.208**	.135**	.153**	-.098**

Note: ZKPQ= Zuckerman-Kuhlman Personality Questionnaire; N-Anx= Neuroticism-Anxiety; Act= Activity; Sy= Sociability; ImpSS= Impulsive Sensation Seeking; Agg-Host= Aggression-Hostility; Infreq= Infrequency.
*p<.05, two-tailed; **p<.0005, two-tailed.
¹ The variable level of education has been considered by their 2 dummy variables

Table 5
Normative data by gender and age adjusted by level of education

Scale	Gender	Age ranges	Mean	SD
N-Anx	Men	18-25	8.47	4.46
		26-40	5.75	4.41
		41-55	6.02	4.39
		56-70	7.30	4.42
		71-93	8.07	4.54
	Women	18-25	10.44	4.78
		26-40	8.46	4.52
		41-55	9.47	4.45
		56-70	9.93	4.55
		71-93	10.31	4.77
Act	Men	18-25	8.03	3.55
		26-40	8.47	3.51
		41-55	8.59	3.50
		56-70	8.86	3.53
		71-93	9.40	3.61
	Women	18-25	7.78	3.80
		26-40	8.29	3.54
		41-55	7.76	3.54
		56-70	9.00	3.62
		71-93	9.16	3.80
GenAct	Men	18-25	3.83	2.47
		26-40	3.98	2.44
		41-55	4.24	2.43
		56-70	4.90	2.46
		71-93	5.47	2.51
	Women	18-25	3.93	2.65
		26-40	4.27	2.45
		41-55	4.44	2.47
		56-70	5.38	2.52
		71-93	5.30	2.65
WorkAct	Men	18-25	4.20	1.72
		26-40	4.49	1.70
		41-55	4.36	1.70
		56-70	3.96	1.71
		71-93	3.93	1.75
	Women	18-25	3.86	1.84
		26-40	4.02	1.71
		41-55	3.33	1.72
		56-70	3.62	1.76
		71-93	3.87	1.84
Sy	Men	18-25	8.61	3.47
		26-40	7.27	3.44
		41-55	6.29	3.42
		56-70	5.78	3.45
		71-93	7.52	3.53
	Women	18-25	9.32	3.73
		26-40	8.02	3.44
		41-55	6.33	3.47
		56-70	6.66	3.54
		71-93	7.97	3.72
Parties	Men	18-25	3.38	1.98
		26-40	2.97	1.95
		41-55	2.61	1.96
		56-70	2.49	1.97
		71-93	3.23	2.01
	Women	18-25	3.38	2.13
		26-40	3.06	1.97
		41-55	2.42	1.98
		56-70	2.79	2.01
		71-93	3.58	2.12

Table 5 (continuation)
Normative data by gender and age adjusted by level of education

Scale	Gender	Age ranges	Mean	SD
Incl	Men	18-25	5.23	2.17
		26-40	4.30	2.15
		41-55	3.68	2.13
		56-70	3.29	2.16
		71-93	4.29	2.21
	Women	18-25	5.93	2.33
		26-40	4.96	2.16
		41-55	3.91	2.18
		56-70	3.87	2.22
		71-93	4.38	2.32
ImpSS	Men	18-25	10.16	4.24
		26-40	8.60	4.20
		41-55	7.08	4.17
		56-70	6.12	4.21
		71-93	6.92	4.31
	Women	18-25	9.51	4.54
		26-40	7.72	4.21
		41-55	6.19	4.23
		56-70	6.18	4.33
		71-93	6.80	4.53
Imp	Men	18-25	3.55	2.13
		26-40	2.87	2.12
		41-55	2.61	2.11
		56-70	2.17	2.12
		71-93	2.70	2.17
	Women	18-25	3.36	2.29
		26-40	2.98	2.12
		41-55	2.57	2.14
		56-70	2.61	2.18
		71-93	3.09	2.28
SS	Men	18-25	6.67	2.85
		26-40	5.73	2.82
		41-55	4.47	2.80
		56-70	3.95	2.83
		71-93	4.21	2.90
	Women	18-25	6.16	3.06
		26-40	4.74	2.83
		41-55	3.61	2.85
		56-70	3.57	2.91
		71-93	3.70	3.05
Agg-Host	Men	18-25	8.15	3.06
		26-40	6.84	3.03
		41-55	6.92	3.01
		56-70	7.11	3.04
		71-93	6.22	3.12
	Women	18-25	8.48	3.28
		26-40	7.28	3.04
		41-55	6.36	3.06
		56-70	5.92	3.12
		71-93	4.51	3.27
Infreq	Men	18-25	1.51	1.65
		26-40	1.94	1.63
		41-55	2.27	1.63
		56-70	3.10	1.64
		71-93	3.90	1.58
	Women	18-25	1.08	1.77
		26-40	1.43	1.64
		41-55	1.88	1.65
		56-70	2.92	1.69
		71-93	4.16	1.77

Note: N-Anx= Neuroticism-Anxiety; Act= Activity; Sy= Sociability; ImpSS= Impulsive Sensation Seeking; Agg-Host= Aggression-Hostility; Infreq= Infrequency

normative data should be taken into account for the interpretation of non-clinical and clinical scores derived from translated versions of questionnaires.

This is the first time the normative data of the Spanish version of the ZKPQ in a general community sample is reported. The magnitude of the sample as well as an adequate representativity of both genders and the wide range of age being considered, gives a satisfactory description of each scale and subscales of this instrument.

We are confident that the availability of the norms as well as the psychometric properties of the ZKPQ will stimulate its use not only in normal populations but in clinical settings. Researchers and clinicians whose interest also includes general personality functioning as well as maladaptive personality traits might be well served by using this instrument as a useful adjunct for diagnosing, case conceptualization, differential treatment planning and predicting response to treatment.

References

- Aluja, A., García, O., & García, L.F. (2004). Replicability of the three, four and five Zuckerman's personality super-factors: Exploratory and confirmatory factor analysis of the EPQ-RS, ZKPQ, and NEO-PI-R. *Personality and Individual Differences*, *36*, 1093-1108.
- Ball, S. (1995). The validity of an alternative five-factor measure of personality in cocaine abusers. *Psychological Assessment*, *7*, 148-154.
- Buss, A.H., & Plomin, R. (1984). *Temperament: Early developing personality traits*. Hillsdale, NJ: Erlbaum.
- Costa, P.T., Terracciano, A., & McCrae, R.R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, *81*, 322-331.
- De Pascalis, V., & Russo, P. (2003). Zuckerman-Kuhlman Personality Questionnaire: Preliminary results of the Italian version. *Psychological Reports*, *92*, 965-974.
- Eysenck, H.J. (1992). Four ways five factors are not basic. *Personality and Individual Differences*, *13*, 667-673.
- Eysenck, H.J., & Eysenck, S.B.G. (1976). *Psychoticism as a dimension of personality*. London: Hodder and Stoughton.
- Gomà-i-Freixanet, M., Valero, S., Puntí, J., & Zuckerman, M. (2004). Psychometric properties of the Zuckerman-Kuhlman Personality Questionnaire in a Spanish sample. *European Journal of Psychological Assessment*, *20*, 134-146.
- Gomà-i-Freixanet, M., Wismeijer, A.J., & Valero, S. (2005). Consensual validity parameters of the Zuckerman-Kuhlman Personality Questionnaire: Evidence from self-reports and spouse reports. *Journal of Personality Assessment*, *84*, 279-286.
- Gomà-i-Freixanet, M., Soler, J., Valero, S., Wismeijer, A.J., Pascual, J.C., & Pérez Sola, V. (in press). Discriminant validity of the ZKPQ in a sample meeting BPD diagnosis vs normal-range controls. *Journal of Personality Disorders*.
- Gray, J.A. (1982). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. New York: Oxford University Press.
- Gutiérrez-Zotes, J.A., Ramos, J.A., & Sáiz, J. (2001). Desarrollo de la versión española del cuestionario de personalidad Zuckerman-Kuhlman (ZKPQ-III) y propiedades psicométricas (Development of the Spanish version of the ZKPQ questionnaire and psychometric properties). *Psiquis*, *22*, 239-250.
- Lynn, R., & Martin, T. (1997). Gender differences in extraversion, Neuroticism and Psychoticism in 37 nations. *The Journal of Social Psychology*, *137*, 369-373.
- Ostendorf, F., & Angleitner, A. (1994). A comparison of different instruments proposed to measure the Big Five. *European Review of Applied Psychology*, *44*, 45-53.
- O'Sullivan, D., Zuckerman, M., & Kraft, M. (1996). The personality of prostitutes. *Personality and Individual Differences*, *21*, 445-448.
- O'Sullivan, D., Zuckerman, M., & Kraft, M. (1998). Personality characteristics of male and female participants in team sports. *Personality and Individual Differences*, *25*, 119-128.
- Peñate, W., Ibáñez, I., & González, M. (1999). La cuantía y naturaleza de las dimensiones básicas de personalidad: una aproximación empírica (The number and nature of the basic personality dimensions: An empirical approach). *Análisis y Modificación de Conducta*, *25*, 103-130.
- Romero, E., Luengo, M.A., Gómez-Freguela, J.A., & Sobral, J. (2002). La estructura de los rasgos de personalidad en adolescentes: el modelo de cinco factores y los cinco alternativos. *Psicothema*, *14*, 134-143.
- Shiomi, K., Kuhlman, D., Zuckerman, M., Joireman, J., Sato, M., & Yata, S. (1996). Examining the validity and reliability of a Japanese version of the Zuckerman-Kuhlman Personality Questionnaire (ZKPQ). *Hyo-go University of Teacher Education Journal*, *2*, 1-13.
- Thomas, A., & Chess, S. (1977). *Temperament and development*. New York: Bruner/Mazel.
- Thornquist, M.H., & Zuckerman, M. (1995). Psychopathy, passive-avoidance learning and basic dimensions of personality. *Personality and Individual Differences*, *19*, 525-534.
- Wu, Y., Wang, W., Du, W., Li, J., Jiang, X., & Wang, Y. (2000). Development of a Chinese version of the Zuckerman-Kuhlman Personality Questionnaire: Reliabilities and gender/age effects. *Social Behaviour and Personality*, *28*, 241-250.
- Zuckerman, M. (1983). *Biological bases of sensation seeking, impulsivity and anxiety*. Hillsdale, NJ: Erlbaum.
- Zuckerman, M. (1984). Sensation seeking: A comparative approach to a human trait. *Behavioral and Brain Sciences*, *7*, 413-417.
- Zuckerman, M. (1991). *Psychobiology of personality*. New York: Cambridge University Press.
- Zuckerman, M. (1992). What is a basic factor and which factors are basic? Turtles all the way down. *Personality and Individual Differences*, *13*, 675-681.
- Zuckerman, M. (1994). *Behavioral expressions and biosocial bases of sensation seeking*. New York: Cambridge University Press.
- Zuckerman, M. (2002). Zuckerman-Kuhlman Personality Questionnaire (ZKPQ): An alternative five-factorial model. In B. De Raad & M. Perugini (Eds.), *Big Five Assessment* (pp. 377-396). Seattle, WA: Hogrefe & Huber Publishers.
- Zuckerman, M., & Kuhlman, D.M. (1993). *Norms for the Zuckerman-Kuhlman Personality Questionnaire (ZKPQ)*. Unpublished manuscript.
- Zuckerman, M., & Kuhlman, D.M. (2000). Personality and risk-taking: Common biosocial factors. *Journal of Personality*, *68*, 999-1029.
- Zuckerman, M., Kuhlman, D.M., & Carnac, C. (1988). What lies beyond E and N? Factor analyses of scales believed to measure basic dimensions of personality. *Journal of Personality and Social Psychology*, *54*, 96-107.
- Zuckerman, M., Kuhlman, D.M., Joireman, J., Teta, P., & Kraft, M. (1993). A comparison of three structural models for personality: The big three, the big five and the alternative five. *Journal of Personality and Social Psychology*, *65*, 757-768.
- Zuckerman, M., Kuhlman, D.M., Thornquist, M., & Kiers, H. (1991). Five (or three): Robust questionnaire scale factors of personality without culture. *Personality and Individual Differences*, *12*, 929-941.

ZUCKERMAN-KUHLMAN PERSONALITY QUESTIONNAIRE:
PSYCHOMETRIC PROPERTIES IN A SAMPLE
OF THE GENERAL POPULATION¹

MONTSERRAT GOMÀ-I-FREIXANET

*Department of Health Psychology
School of Psychology
Autonomous University of Barcelona*

SERGI VALERO

*Department of Psychiatry
Hospital Universitari Vall d'Hebron de Barcelona
Autonomous University of Barcelona*

ANNA MURO AND SILVIA ALBIOL

*Department of Health Psychology, School of Psychology
Autonomous University of Barcelona*

Summary.—The Zuckerman-Kuhlman Personality Questionnaire, a tool designed to measure the alternative five-factor model of personality, was translated and adapted into the Spanish language. Some psychometric results with students have already been reported, but to date, there appears to be no published data on the psychometric properties of the questionnaire in a general population sample in any culture. The aim of the present study was to provide psychometric properties for a large Spanish community sample with a wide representation of adult age-range, a representative distribution of sexes, and a balanced distribution of education within the sample. Participants volunteered and were selected on a random basis. The questionnaire was administered to 1,000 participants of both sexes (51.4% women) with ages ranging from 18 to 88 years ($M=45.4$, $SD=18.7$). Analysis gave a normal distribution of all scales, good internal consistencies, and discriminant validity. Sex differences were in accord with data from tests measuring similar personality traits. Finally, the original five-component structure was also replicated.

The Zuckerman-Kuhlman Personality Questionnaire was developed to measure dimensions of an alternative five-factor model. This model (Zuckerman, Kuhlman, Joireman, Teta, & Kraft, 1993; Zuckerman, 2002) emerged from a series of factor analyses of scales believed to measure basic dimensions of personality or temperament, particularly those used in psychobiological research (Zuckerman, Kuhlman, & Camac, 1988; Zuckerman, Kuhlman, Thornquist, & Kiers, 1991). This alternative theoretical model, influenced by previous research on personality by Eysenck (1991) and Gray (1982), is mainly focused on temperamental characteristics of personality, placing great emphasis on the psychobiological aspects of personality dimensions. Eysenck (1992) and Zuckerman (1992) suggested that psychobiological studies of personality provide a better understanding of the neurobiological and genetic

¹Address correspondence to Montserrat Gomà-i-Freixanet, Ph.D., Department of Health Psychology, Autonomous University of Barcelona, E-08193 Bellaterra, Catalonia, Spain or e-mail (montserrat.goma@uab.cat) or (www.zkpq.com).

underpinnings of personality. Relying only on the encoding of personality traits in language may be inaccurate, as this encoding probably reflects the observability of these traits in social interactions and may not necessarily mirror the proportional biological relevance of the traits (Zuckerman, 1992; Zuckerman, *et al.*, 1993). Therefore, in contrast to the lexical approach, using psychobiological data allows researchers to explore the causal, biological origins of personality (Eysenck, 1992).

Since its first publication in 1993, the Zuckerman-Kuhlman Personality Questionnaire has undergone extensive psychometric testing, demonstrating good internal reliability, temporal stability, validity, and cross-cultural replicability (Zuckerman, 2002; Joireman & Kuhlman, 2004). Concerning validity, the questionnaire has shown adequate consensual validity (Gomà-i-Freixanet, Wismeijer, & Valero, 2005), known as validation of responses by correlating these with ratings made by external assessors who know the rate well. Regarding concurrent validity, some studies have been conducted, e.g., characteristics of drug abusers and prediction of course of therapy (Ball, 1995), characteristics of psychopathology (Thornquist & Zuckerman, 1995; O'Sullivan, Zuckerman, & Kraft, 1996; Gomà-i-Freixanet, Soler, Valero, Pascual, & Pérez Sola, 2008), and risk taking (O'Sullivan, Zuckerman, & Kraft, 1998; Zuckerman & Kuhlman, 2000). Several cross-cultural studies have also been performed using translations into other languages, e.g., German (Ostendorf & Angleitner, 1994), Japanese (Shiomi, Kuhlman, Zuckerman, Joireman, Sato, & Yata, 1996), Chinese (Wu, Wang, Du, Li, Jiang, & Wang, 2000), and Italian (De Pascalis & Russo, 2003), among others.

However, most of the published data on psychometric properties of the Zuckerman-Kuhlman Personality Questionnaire have been obtained from college students. These populations have at least three limitations. Namely, the entire adult age-range is not represented (in general the age range is 19 to 23 years), the distribution of sexes is not always balanced (in general women are overrepresented), and amount of education (college education) is not representative of the general population. Therefore, before examining higher inference with the results obtained with this questionnaire, it is appropriate to test the psychometric properties in a general population sample with a balanced distribution of adults in terms of age, sex, and amount of education.

The aim of the present study was to provide, for the first time, the psychometric properties of the Zuckerman-Kuhlman Personality Questionnaire in a large population sample with a wide representation of adults in terms of age, a representative distribution of sexes, and a balanced distribution of amount of education within the sample.

METHOD

Participants

The general community sample consisted of 1,000 participants, 514

women (51.4%) and 486 men (48.6%). The sample matched the IDESCAT (Statistical Institute of Catalonia) census projections for the year 2000 in distribution of sex and age groups. The mean age of the total sample was 45.4 ($SD=18.7$), with the range being 18 to 88 years. The distribution of education in this sample was 30.3% elementary school, 30.8% high school, and 37.4% college education.

Measure

Zuckerman-Kuhlman Personality Questionnaire (Zuckerman & Kuhlman, 1993).—This questionnaire has five content scales, plus an Infrequency scale. There are 99 dichotomous items in sentence format and true-false response set. The five scales can be described in terms of their typical content. (1) Neuroticism-Anxiety has 19 items which describe frequent emotional upset, tension, worry, fearfulness, indecision, lack of self-confidence, and sensitivity to criticism. (2) Activity has 17 items which describe the need for general activity, an inability to relax and do nothing when the opportunity arises, and a preference for hard and challenging work, an active busy life, and high energy. (3) Sociability has 17 items which describe the number of friends one has and the amount of time spent with them, outgoingness at parties, and a preference for being with others as opposed to being alone and engaging in solitary activities. (4) Impulsive Sensation Seeking has 19 items which involve a lack of planning and a tendency to act impulsively without thinking, plus seeking excitement, novel experiences, and willingness to take risks for these types of experiences. The items are general in content and do not describe specific activities such as drinking or sex. (5) Aggression-Hostility has 17 items which describe a readiness to express verbal aggression, rude, thoughtless or antisocial behaviour, vengefulness and spitefulness, having a quick temper, and impatience with others.

This questionnaire also includes an Infrequency scale of 10 items. Rather than being regarded as a scale in the normative sense, it should only be used to detect inattention to the task or be used simply as a validity index for the individual test taker. The items are mostly exaggerated, true scored, and socially desirable, but unlikely to be completely true statements about anyone. Responses to this scale are not normally distributed, as most scores are around 0 or 1.

Procedure

To obtain this large sample, participants were contacted from very different sites (classroom, leisure associations, while waiting for a yearly health check, home, etc.), and the snowball technique was used as well. In all situations, the participants were provided written instructions and an introductory letter explaining the overall goal of the study ("the study you will collaborate in attempts to evaluate the functioning of an American question-

naire in our culture"). Most of the questionnaires were administered individually and others in a group situation. In the former situation, participants also received a prepaid envelope which had to be posted. Participants with 10% or more of blank or double answers were rejected. Around 90% of the respondents correctly completed the questionnaire. As this was the first time the psychometric properties of the Zuckerman-Kuhlman Personality Questionnaire were studied in a general population sample, no participant was excluded from scores on the Infrequency scale to avoid introducing *a priori* bias. Special attention was paid to having a representative number of participants across the full adult age-range; that elementary school, high school, and college education were equally represented; and that the distribution of both sexes was representative of the distribution in the general population. The questionnaire was administered anonymously, and only the demographic variables such as age, sex, and amount of education were obtained. Participants participated voluntarily and received no reward for their collaboration.

Statistical Analysis

Means, standard deviations, skewness, kurtosis, and internal consistency for the scales were obtained for the total sample. For women and men, means, standard deviations, internal consistencies, and *t* ratios were also obtained and Cohen *d* reported. The Type I error rate per comparison was set by default to .05. Pearson product-moment correlations were calculated for women and men separately and for the total sample as well. Finally, to investigate the internal structure of the Zuckerman-Kuhlman Personality Questionnaire, an exploratory factor analysis by principal components analysis with a varimax rotation for women and men as well as for the total sample was carried out.

RESULTS

Means, Standard Deviations, Coefficients Alpha, t Ratios for Comparisons of Sex, and Correlation Matrix

The distribution of age is shown in Fig. 1. The histogram illustrates how age is distributed homogeneously along the whole range of ages assessed.

Table 1 shows the means, standard deviations, skewness, and kurtosis of the total sample, and Cronbach coefficients alpha for men and women, as well as for the total sample, as indices of internal consistency. The skewness and kurtosis statistics showed a normal distribution of scales (values between -1 and +1). The magnitude of the coefficients alpha was adequate and similar to that found for the original American version. The mean alpha for the Zuckerman-Kuhlman Personality Questionnaire scales for the total sample

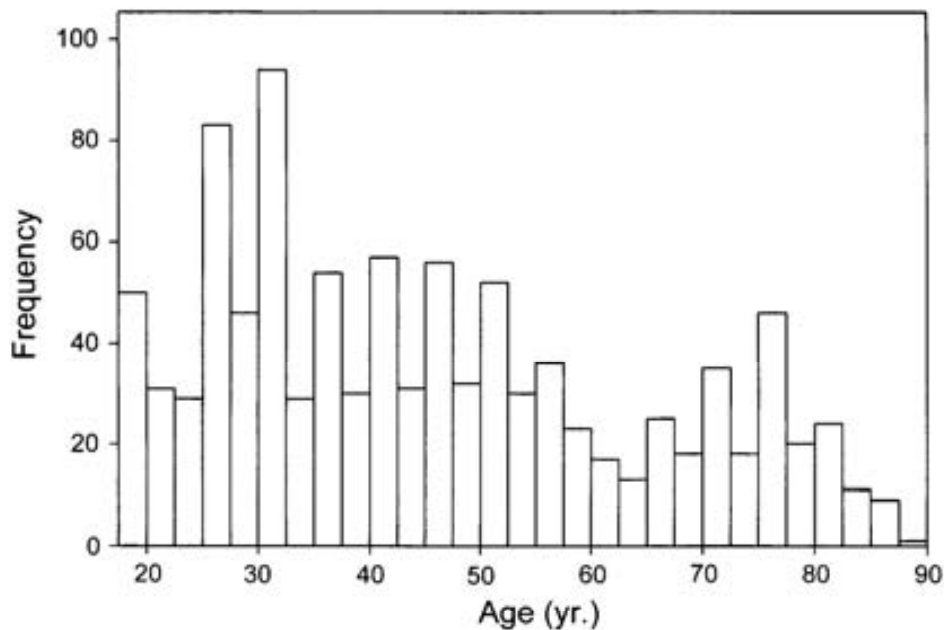


FIG. 1. Age distribution of sample.

was .77, with values ranging from .70 to .85. The Neuroticism-Anxiety scale had the highest reliability (.85) and the Aggression-Hostility scale had the lowest (.70). Activity, Sociability, and Impulsive Sensation Seeking scales had intermediate values (.72-.80), but all were satisfactory.

TABLE 1
MEANS, STANDARD DEVIATIONS, SKEWNESS, KURTOSIS, AND COEFFICIENTS ALPHA FOR
SCALES OF ZUCKERMAN-KUHLMAN PERSONALITY QUESTIONNAIRE (N = 1,000)

Scale	M	SD	Total Sample		Cronbach Alpha		
			Skewness	Kurtosis	Total (N = 1,000)	Men (n = 486)	Women (n = 514)
Neuroticism-Anxiety	8.14	4.72	.23	-.89	.85	.84	.84
Activity	8.53	3.47	-.09	-.59	.72	.73	.72
Sociability	7.31	3.49	.01	-.65	.76	.75	.76
Impulsive Sensation Seeking	7.76	4.19	.34	-.52	.80	.81	.80
Aggression-Hostility	6.82	3.19	.19	-.49	.70	.72	.69
Infrequency	2.29	1.98	.85	.28	.64	.62	.66

Means and standard deviations for men and women and differences between the sexes are shown in Table 2. The sex groups did not differ significantly on age, although they did differ significantly on the Neuroticism-Anxiety, Sociability, Impulsive Sensation Seeking, and Infrequency scales. Women scored significantly higher on the Neuroticism-Anxiety and Sociability

scales, and men scored higher on the Impulsive Sensation Seeking and Infrequency scales. There were no statistically significant sex differences on Activity and Aggression-Hostility. Cohen *d* showed a moderate effect for the Neuroticism-Anxiety scale, while for the rest of the scales, the effect is low.

TABLE 2
MEANS, STANDARD DEVIATIONS, AND *t* RATIOS FOR MEN AND WOMEN AND COHEN *d*

Variable	Men (<i>n</i> = 486)		Women (<i>n</i> = 514)		<i>t</i>	Cohen <i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Age, yr.	44.46	18.35	46.26	19.07	1.52	.09
Zuckerman-Kuhlman Personality Questionnaire						
Neuroticism-Anxiety	6.74	4.41	9.46	4.63	9.49†	.60
Activity	8.53	3.50	8.54	3.44	0.06	.00
Sociability	7.02	3.49	7.57	3.48	2.49*	.16
Impulsive Sensation Seeking	8.03	4.27	7.51	4.09	1.96*	.12
Aggression-Hostility	6.95	3.26	6.69	3.12	1.32	.08
Infrequency	2.43	1.97	2.17	1.99	2.10*	.14

**p* < .05. †*p* < .001.

Table 3 presents the Pearson correlation matrix for the total sample and for men and women separately. In the total sample, correlations among scales ranged from .03 to .28, with an absolute mean interscale correlation of *r* = .10. Although almost all of the correlations were significant given the large sample size, the absolute values were very low.

TABLE 3
CORRELATIONS AMONG SCALES OF ZUCKERMAN-KUHLMAN PERSONALITY QUESTIONNAIRE
FOR TOTAL SAMPLE (UPPER RIGHT), MEN, AND WOMEN

Scale	1		2		3		4		5		6	
	M	W	M	W	M	W	M	W	M	W	M	W
1			.04		-.06		.14*		.17*		.03	
2	.04	.03			.13*		.26*		.07*		.24*	
3	-.10*	-.07	.07	.20*			.24*		.09*		.05	
4	.27*	.07	.20*	.31*	.21*	.28*			.28*		.05	
5	.28*	.12*	.04	.10*	.04	.15*	.25*	.31*			-.18*	
6	.08	.03	.25*	.24*	.09	.04	.05	.04	-.11*	-.26*		

Note.—1 = Neuroticism-Anxiety, 2 = Activity, 3 = Sociability, 4 = Impulsive Sensation Seeking, 5 = Aggression-Hostility, 6 = Infrequency; M = Men, W = Women. **p* < .05.

Exploratory Factor Analysis

A factor analysis, with principal components analysis followed by normalized varimax rotation of the 89 items remaining after excluding the items of the Infrequency scale, was conducted. The Kaiser-Meyer-Olkin measure of sampling adequacy was .85, and the Bartlett Test of Sphericity yielded a chi square approximately equal to 20,928.37 (*df* = 3,916, *p* = .0005), indicating

the appropriateness of the factor analysis. Fig. 2 shows the scree plot of the first 40 eigenvalues used to evaluate more precisely the course of the eigenvalues. Eigenvalues began to level off after five factors, and factors beyond the sixth showed little change. The first seven eigenvalues were 7.35, 5.86, 3.90, 3.30, 2.95, 2.28, and 2.02.

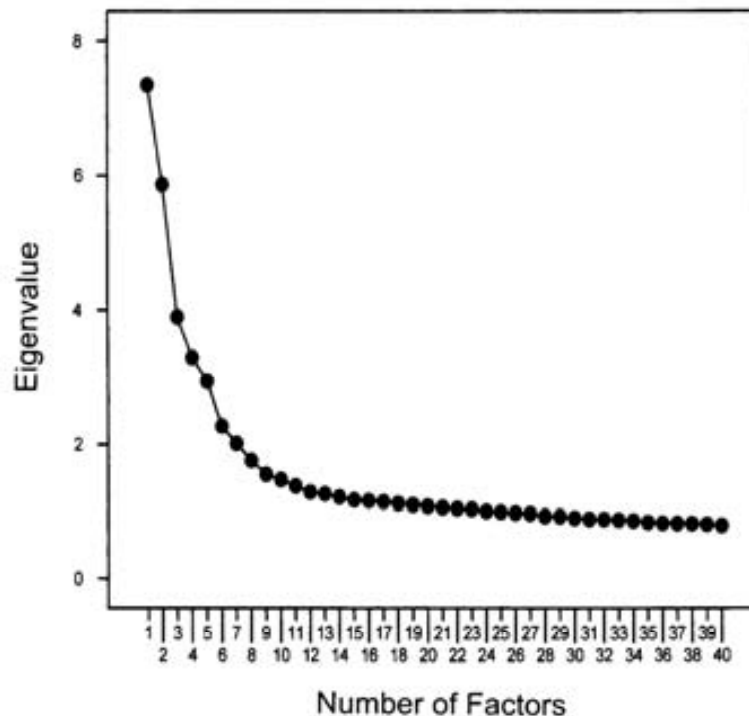


FIG. 2. Scree plot of the first 40 eigenvalues for factor analysis of the Zuckerman-Kuhlman Personality Questionnaire

Table 4 shows the factorial solution for the total sample with five factors extracted and rotated. Factor loadings of each item in the rotated components retained, as well as item communalities and percent of variance explained by each component are also reported. The first two factors were Neuroticism-Anxiety and Impulsive Sensation Seeking, which accounted for 7.24% and 5.95% of the explained variance, respectively. The next three factors were Activity, Sociability, and Aggression-Hostility, which accounted for 4.65%, 4.22%, and 4.19% of the explained variance, respectively. In all, the first five factors of the total sample explained only 26.3% of the variance, leaving about 74% unexplained. For the additional factor analyses conducted separately for men and women, five factors were also extracted and rotated, explaining 26.59% and 26.34% of the variance, respectively.

TABLE 4
 PRINCIPAL COMPONENTS ANALYSIS AND VARIMAX ROTATION OF FIVE PRINCIPAL COMPONENTS
 OF ZUCKERMAN-KUHLMAN PERSONALITY QUESTIONNAIRE FOR TOTAL SAMPLE, *A Priori*
 FACTOR ASSIGNMENT, ITEM COMMUNALITIES, AND PERCENT VARIANCE EXPLAINED

Item	<i>A Priori</i>	Factor ^a					<i>b</i> ²
		I	II	III	IV	V	
61	+Neuroticism-Anxiety	.68	.00	-.05	-.02	-.04	.46
25	+Neuroticism-Anxiety	.63	.08	.01	-.08	.16	.44
15	+Neuroticism-Anxiety	.63	.07	.06	-.13	.14	.44
30	+Neuroticism-Anxiety	.58	.11	.01	.01	.13	.37
7	+Neuroticism-Anxiety	.55	-.02	-.11	-.02	-.13	.33
46	+Neuroticism-Anxiety	.55	-.02	-.02	.07	.04	.31
80	+Neuroticism-Anxiety	.55	.07	-.01	.04	.05	.31
90	+Neuroticism-Anxiety	.54	.14	-.03	-.18	.17	.37
76	+Neuroticism-Anxiety	.53	.03	.02	-.05	-.03	.29
41	+Neuroticism-Anxiety	.53	-.11	-.04	.13	-.03	.31
35	+Neuroticism-Anxiety	.52	-.02	.18	-.05	.10	.31
96	+Neuroticism-Anxiety	.51	.02	.08	.01	.02	.27
56	+Neuroticism-Anxiety	.51	-.02	.05	.06	.01	.26
51	+Neuroticism-Anxiety	.50	.10	.11	.00	-.13	.29
66	+Neuroticism-Anxiety	.48	.02	.15	-.12	.13	.29
23	-Activity	.46	-.01	-.20	-.07	-.18	.29
71	+Neuroticism-Anxiety	.45	.03	-.01	.06	.02	.21
99	+Activity	.43	.04	.18	.03	.34	.34
20	+Neuroticism-Anxiety	.39	-.04	.04	-.10	.23	.22
88	-Activity	.16	-.13	-.08	-.07	-.09	.06
39	+Impulsive Sensation Seeking	-.08	.63	.08	.06	.04	.42
24	+Impulsive Sensation Seeking	-.12	.60	.05	.09	.12	.39
55	+Impulsive Sensation Seeking	.02	.56	.08	.09	.11	.34
95	+Impulsive Sensation Seeking	-.04	.54	-.12	.24	.10	.38
65	-Impulsive Sensation Seeking	.02	.54	.13	.02	.13	.32
84	-Impulsive Sensation Seeking	.10	.52	.20	-.01	.17	.35
70	+Impulsive Sensation Seeking	.05	.51	-.07	.18	.17	.32
34	+Impulsive Sensation Seeking	-.10	.49	-.09	-.15	.03	.28
75	+Impulsive Sensation Seeking	-.20	.46	.08	-.12	.01	.27
79	+Impulsive Sensation Seeking	.08	.43	.10	.14	-.01	.23
45	+Impulsive Sensation Seeking	.03	.43	.21	.11	-.05	.24
1	+Impulsive Sensation Seeking	.21	.42	-.17	.03	-.07	.25
14	+Impulsive Sensation Seeking	.26	.39	.08	.02	.14	.25
60	+Impulsive Sensation Seeking	.11	.38	.31	.02	-.04	.25
28	+Activity	-.28	.37	.17	.00	.13	.27
29	-Impulsive Sensation Seeking	-.14	-.37	.31	-.03	.06	.26
6	-Impulsive Sensation Seeking	-.12	-.34	.24	-.12	-.03	.20
27	+Sociability	.03	.31	.14	.28	.07	.20
50	+Impulsive Sensation Seeking	.26	.28	.02	.10	.12	.17
64	+Activity	.05	.26	.21	.00	-.02	.12
2	-Neuroticism-Anxiety	-.14	.21	.01	-.01	-.21	.11
36	+Aggression-Hostility	.16	.18	-.03	-.03	.07	.06
16	-Aggression-Hostility	-.07	-.18	.16	-.16	-.13	.10
19	+Impulsive Sensation Seeking	.09	.17	-.08	-.07	-.14	.07
3	+Aggression-Hostility	.13	.16	-.07	.07	-.01	.05

^aI = Neuroticism-Anxiety, II = Impulsive Sensation Seeking, III = Activity, IV = Sociability, V = Aggression-Hostility.

TABLE 4 (CONT'D)
 PRINCIPAL COMPONENTS ANALYSIS AND VARIMAX ROTATION OF FIVE PRINCIPAL COMPONENTS
 OF ZUCKERMAN-KUHLMAN PERSONALITY QUESTIONNAIRE FOR TOTAL SAMPLE, *A Priori*
 FACTOR ASSIGNMENT, ITEM COMMUNALITIES, AND PERCENT VARIANCE EXPLAINED

Item	<i>A Priori</i>	Factor ^a					<i>b</i> ²
		I	II	III	IV	V	
33	+Activity	.04	.03	.66	.07	-.03	.45
83	+Activity	.08	.09	.66	-.01	-.01	.45
5	+Activity	.10	.03	.57	-.01	-.04	.33
74	+Activity	.07	.01	.56	-.04	-.09	.33
44	-Activity	.02	-.09	-.51	-.07	-.04	.27
38	-Activity	-.01	.05	-.50	-.09	.06	.27
94	+Activity	-.12	.01	.49	.08	.16	.28
59	+Activity	-.11	.25	.37	.15	-.07	.23
13	+Activity	.03	.08	.34	-.03	.05	.12
18	+Activity	-.16	.16	.30	-.15	.03	.16
54	+Activity	.23	.06	.28	-.05	.21	.18
49	-Activity	-.03	.09	-.20	.02	.05	.05
58	+Sociability	-.01	.10	.05	.56	-.04	.33
17	-Sociability	-.06	.30	-.04	-.55	.02	.39
63	-Sociability	-.05	.32	.00	-.55	.05	.40
87	-Sociability	.19	.03	.17	-.54	-.05	.36
68	-Sociability	.24	.05	.17	-.53	-.04	.37
12	-Sociability	.04	-.02	.17	-.53	-.05	.31
22	+Sociability	.02	.14	.14	.48	-.05	.28
43	-Sociability	.17	-.15	.13	-.48	-.01	.30
37	-Sociability	.01	-.07	-.04	-.46	-.03	.22
78	+Sociability	-.02	.09	.30	.42	-.02	.27
48	+Sociability	.01	.33	.11	.41	.04	.29
9	+Sociability	-.07	.21	.15	.37	.13	.23
82	+Sociability	.10	.16	.12	.35	.11	.18
98	+Sociability	.11	.19	.26	.26	-.12	.20
92	+Sociability	.15	.21	.07	.25	.00	.13
53	+Sociability	.01	.21	-.04	.22	.02	.09
72	+Aggression-Hostility	.08	.03	.11	.04	.52	.29
77	+Aggression-Hostility	.03	.26	.03	.02	.50	.32
97	+Aggression-Hostility	-.02	.09	.06	.00	.50	.26
86	-Aggression-Hostility	.07	.09	.19	-.01	-.50	.30
11	+Aggression-Hostility	-.04	.15	-.22	.05	.47	.30
57	-Aggression-Hostility	.08	-.12	.22	-.03	-.47	.29
67	+Aggression-Hostility	.08	.05	.14	.06	.47	.25
42	+Aggression-Hostility	-.09	.07	.20	.10	.43	.24
89	+Impulsive Sensation Seeking	.25	.28	.19	.05	.42	.35
62	-Aggression-Hostility	.07	.14	.08	.05	-.40	.19
47	+Aggression-Hostility	-.01	.03	.05	.02	.40	.17
8	+Aggression-Hostility	.25	.09	-.05	-.07	.40	.23
85	-Neuroticism-Anxiety	-.28	.11	.22	-.01	-.39	.30
31	-Aggression-Hostility	-.09	.00	.12	-.08	-.36	.16
91	+Aggression-Hostility	.21	.13	.04	-.00	.35	.19
21	-Aggression-Hostility	-.05	-.07	.02	.05	-.16	.04
% variance		7.24	5.95	4.65	4.22	4.19	

^aI = Neuroticism-Anxiety, II = Impulsive Sensation Seeking, III = Activity, IV = Sociability, V = Aggression-Hostility.

DISCUSSION

This study assessed the psychometric properties of the Zuckerman-Kuhlman Personality Questionnaire in a general population sample in Spain. The obtained skewness and kurtosis statistics showed a normal distribution of frequencies, except for the scores on the Infrequency scale, which were highly skewed, as expected. Magnitudes of coefficients alpha found in the general population were adequate and higher than those reported for college student samples (Zuckerman, 2002; Gomà-i-Freixanet, Valero, Puntí, & Zuckerman, 2004). In the present sample, the Neuroticism-Anxiety scale was the most reliable and the Aggression-Hostility scale was the least reliable, being consistent with previously reported data for university students (Gomà-i-Freixanet, *et al.*, 2004). This same general trend for the Aggression-Hostility scale, i.e., having the lowest α reliability, was found for translations of the Zuckerman-Kuhlman Personality Questionnaire into different languages (Wu, *et al.*, 2000; De Pascalis & Russo, 2003), thereby probably suggesting a cultural difference from the USA samples.

The means of the scales in the general population were slightly lower than those obtained in student samples, except for the Infrequency scale, for which the mean increased. These results suggest the necessity of taking age into account in the normative data of personality questionnaires (Gomà-i-Freixanet & Valero, 2008). Regarding sex differences, women scored significantly higher on the Neuroticism-Anxiety and Sociability scales, while men scored significantly higher on the Impulsive Sensation Seeking and Infrequency scales. These differences follow the general trend found in the literature on sex differences (Costa, Terracciano, & McCrae, 2001). But unlike the results obtained with this questionnaire in the original American study with students, there were no significant sex differences on the Activity and on Aggression-Hostility scales. However, the results obtained for these two scales were similar to those of previous studies based on European student samples including Spain (Gutierrez-Zotes, Ramos, & Saiz, 2001; De Pascalis & Russo, 2003; Gomà-i-Freixanet, *et al.*, 2004). In the Spanish normative data, scores for these two scales seemed to vary with the age of participants (Gomà-i-Freixanet & Valero, 2008). While scores on the Activity scale increased evenly across age groups for both sexes, scores on the Aggression-Hostility scale decreased at a different rate for the two sexes. From 18 to 55 years, the decrease in groups for both sexes was similar in magnitude, but from 56 to 93 years, the decrease was more pronounced in groups of women than men. Thus, while there was a decrease on Aggression-Hostility for both sexes, for women the decrease was more abrupt than for men. This might explain the absence of significant sex differences in early middle ages and the presence of sex differences at later ages.

It is important to mention the statistically significant difference on the Infrequency scale. In the original American student sample, sex differences

on this scale were not reported, nor were they in most papers published on the Zuckerman-Kuhlman Personality Questionnaire. However, in the literature which refers to these data (Gomà-i-Freixanet, *et al.*, 2004, 2005; Gomà-i-Freixanet & Valero, 2008), a statistically significant difference was found, with men scoring higher than women. It is desirable that scores on this scale be included in the normative data of the Zuckerman-Kuhlman Personality Questionnaire as a measure of the reliability of the obtained scores.

The Zuckerman-Kuhlman Personality Questionnaire also has good interval discriminant validity, shown by the low intercorrelations among scales. This empirically replicated the relative independence of scales derived from factor analysis. This pattern was also found when analysing the data for the sexes separately. It is worth commenting on the pattern of correlations for the Infrequency scale, as this scale neither correlates with the more desirable traits, i.e., Activity and Sociability, nor with the less desirable ones, i.e., Neuroticism-Anxiety, Impulsive Sensation Seeking, and Aggression-Hostility, thus seeming to detect careless responding rather than social desirability (Zuckerman & Kuhlman, 1993).

The results of the principal components analysis, using orthogonal varimax rotation, supported the replicability of the original five-component structure of the alternative five-factor model of personality. The percentage of explained variance in the present study (26.3%) was not different from those reported for other samples, which ranged from 20.8% (Wu, *et al.*, 2000) to 28.0% (Shiomi, *et al.*, 1996).

The results obtained for the Zuckerman-Kuhlman Personality Questionnaire in a general population sample showed good internal consistency of scales, sex differences in the expected directions, independence of the scales derived from factor analysis, and replication of the original five-component structure. This is the first time the psychometric properties of the Zuckerman-Kuhlman Personality Questionnaire in a general community sample with a wide representation of adult ages, a representative distribution of both sexes, and a balanced distribution of amount of education within the sample were reported in any culture.

As a first step, it may be adequate to test the psychometric properties for university students, as this is often more convenient, but before embarking on higher scientific inference using results obtained from a questionnaire, the psychometric properties are necessary for a general population sample with an adequate representation of the adult ages, a representative distribution of both sexes, and an equal distribution of education.

REFERENCES

- BALL, S. A. (1995) The validity of an alternative five-factor measure of personality in cocaine abusers. *Psychological Assessment*, 7, 148-154.
- COSTA, P. T., TERRACCIANO, A., & MCCRAE, R. R. (2001) Gender differences in personality traits across cultures: robust and surprising findings. *Journal of Personality and Social Psychology*, 81, 322-331.

- DE PASCALIS, V., & RUSSO, P. (2003) Zuckerman-Kuhlman Personality Questionnaire: preliminary results of the Italian version. *Psychological Reports*, 92, 965-974.
- EYSENCK, H. J. (1991) Dimensions of personality: 16, 5 or 3? Criteria for a taxonomic paradigm. *Personality and Individual Differences*, 12, 773-790.
- EYSENCK, H. J. (1992) Four ways five factors are not basic. *Personality and Individual Differences*, 13, 667-673.
- GOMÀ-I-FREIXANET, M., SOLER, J., VALERO, S., PASCUAL, J. C., & PÉREZ SOLA, V. (2008) Discriminant validity of the ZKPQ in a sample meeting BPD diagnosis vs normal-range controls. *Journal of Personality Disorders*, 22, 178-190.
- GOMÀ-I-FREIXANET, M., & VALERO, S. (2008) Spanish normative data of the Zuckerman-Kuhlman Personality Questionnaire (ZKPQ) in a general population sample. *Psicothema*, 20, 318-324.
- GOMÀ-I-FREIXANET, M., VALERO, S., PUNTÍ, J., & ZUCKERMAN, M. (2004) Psychometric properties of the Zuckerman-Kuhlman Personality Questionnaire in a Spanish sample. *European Journal of Psychological Assessment*, 20, 134-146.
- GOMÀ-I-FREIXANET, M., WISMEIJER, A. J., & VALERO, S. (2005) Consensual validity parameters of the Zuckerman-Kuhlman Personality Questionnaire: evidence from self-reports and spouse reports. *Journal of Personality Assessment*, 84, 279-286.
- GRAY, J. A. (1982) *The neuropsychology of anxiety: an enquiry into the functions of the septo-hippocampal system*. New York: Oxford Univer. Press.
- GUTIÉRREZ-ZOTES, J. A., RAMOS, J. A., & SAIZ, J. (2001) Desarrollo de la versión española del cuestionario de personalidad Zuckerman-Kuhlman (ZKPQ-III) y propiedades psicométricas [Development of the Spanish version of the ZKPQ-III and psychometric properties]. *Psiquis*, 22, 239-250.
- JOIREMAN, J., & KUHLMAN, D. M. (2004) The Zuckerman-Kuhlman Personality Questionnaire: origin, development, and validity of a measure to assess an alternative five-factor model of personality. In R. M. Stelmack (Ed.), *On the psychobiology of personality*. Amsterdam: Elsevier. Pp. 49-64.
- OSTENDORF, F., & ANGLEITNER, A. (1994) A comparison of different instruments proposed to measure the Big Five. *European Review of Applied Psychology*, 44, 45-53.
- O'SULLIVAN, D., ZUCKERMAN, M., & KRAFT, M. (1996) The personalities of prostitutes. *Personality and Individual Differences*, 21, 445-448.
- O'SULLIVAN, D., ZUCKERMAN, M., & KRAFT, M. (1998) Personality characteristics of male and female participants in team sports. *Personality and Individual Differences*, 25, 119-128.
- SHIOMI, K., KUHLMAN, D. M., ZUCKERMAN, M., JOIREMAN, J., SATO, M., & YATA, S. (1996) Examining the validity and reliability of a Japanese version of the Zuckerman-Kuhlman Personality Questionnaire (ZKPQ). *Hyogo University of Teacher Education Journal*, 2, 1-13.
- THORNQUIST, M. H., & ZUCKERMAN, M. (1995) Psychopathy, passive-avoidance learning and basic dimensions of personality. *Personality and Individual Differences*, 19, 525-534.
- WU, Y., WANG, W., DU, W., LI, J., JIANG, X., & WANG, Y. (2000) Development of a Chinese version of the Zuckerman-Kuhlman Personality Questionnaire: reliabilities and gender/age effects. *Social Behavior and Personality*, 28, 241-250.
- ZUCKERMAN, M. (1992) What is a basic factor and which factors are basic? Turtles all the way down. *Personality and Individual Differences*, 13, 675-681.
- ZUCKERMAN, M. (2002) Zuckerman-Kuhlman Personality Questionnaire (ZKPQ): an alternative five-factorial model. In B. De Raad & M. Perugini (Eds.), *Big Five assessment*. Seattle, WA: Hogrefe & Huber. Pp. 377-396.
- ZUCKERMAN, M., & KUHLMAN, D. M. (1993) Norms for the Zuckerman-Kuhlman Personality Questionnaire (ZKPQ). (Unpublished manuscript, Univer. of Delaware)
- ZUCKERMAN, M., & KUHLMAN, D. M. (2000) Personality and risk-taking: common biosocial factors. *Journal of Personality*, 68, 999-1029.
- ZUCKERMAN, M., KUHLMAN, D. M., & CAMAC, C. (1988) What lies beyond E and N? Factor analyses of scales believed to measure basic dimensions of personality. *Journal of Personality and Social Psychology*, 54, 96-107.
- ZUCKERMAN, M., KUHLMAN, D. M., JOIREMAN, J., TETA, P., & KRAFT, M. (1993) A comparison of three structural models for personality: the big three, the big five, and the alternative five. *Journal of Personality and Social Psychology*, 65, 757-768.
- ZUCKERMAN, M., KUHLMAN, D. M., THORNQUIST, M., & KIERS, H. (1991) Five (or three): robust questionnaire scale factors of personality without culture. *Personality and Individual Differences*, 12, 929-941.

Accepted December 12, 2008.

ZKPO

El Modelo Alternativo de los Cinco Grandes