



Psicothema

ISSN: 0214-9915

psicothema@cop.es

Universidad de Oviedo

España

Balluerka, Nekane; Gorostiaga, Arantxa; Alonso-Arbiol, Itziar; Haranburu, Mikel  
La adaptación de instrumentos de medida de unas culturas a otras: una perspectiva práctica  
Psicothema, vol. 19, núm. 1, 2007, pp. 124-133  
Universidad de Oviedo  
Oviedo, España

Disponible en: <http://www.redalyc.org/articulo.oa?id=72719119>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica  
Red de Revistas Científicas de América Latina, el Caribe, España y Portugal  
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

## La adaptación de instrumentos de medida de unas culturas a otras: una perspectiva práctica

Nekane Balluerka, Arantxa Gorostiaga, Itziar Alonso-Arbiol y Mikel Haranburu  
Universidad del País Vasco

El objetivo del presente estudio consiste en mostrar, desde una perspectiva práctica, las pautas que se deben seguir para adaptar adecuadamente un instrumento de medida de una cultura a otra. Para ello, se describe exhaustivamente el proceso de adaptación del NEO PI-R del castellano al euskera, siguiendo una serie de pasos que pretenden evitar las fuentes de error a las que se exponen los investigadores inmersos en un proceso de adaptación. En la fase de validación empírica del instrumento se describen cuatro estudios que examinan su dimensionalidad, su validez convergente y diferencial, la consistencia interna de cada dimensión y el posible funcionamiento diferencial de los ítems respecto al género y a la etapa vital (joven o adulto). A partir de los resultados obtenidos en esta fase, cabe concluir que la versión en euskera del NEO PI-R ofrece garantías suficientes para una evaluación adecuada de la personalidad desde la perspectiva de los «Cinco Grandes». Tales estudios muestran la forma en la que pueden obtenerse evidencias sobre el grado de validez de la versión adaptada del test, lo que constituye la piedra angular de cualquier proceso de adaptación.

*Test adaptation to other cultures: A practical approach.* The aim of this study was to show the practical steps to be followed in the adaptation of instruments for measuring psychological constructs across cultures. For this purpose, the adaptation process of the NEO PI-R from Spanish into Basque is thoroughly described. In order to avoid the sources of error that researchers could encounter during the adaptation process, established guidelines are followed. In the stage which focuses on the empirical validation of the instrument to be used, four studies are described that examine the following aspects: dimensionality, convergent and differential validity, internal consistency of each dimension, and the possible differential item functioning with regard to gender and life stage (youth or adulthood). In the light of the results obtained, we conclude that the Basque version of NEO PI-R guarantees an adequate assessment of personality from the Big Five viewpoint. These studies illustrate how to obtain evidence about the degree of validity of the adapted version of a test, which is the key feature of any adaptation process.

La adaptación de instrumentos de medida de unas culturas a otras es un proceso complejo que se halla expuesto a una gran cantidad de fuentes de error (Hambleton, 1996, 2005; Muñiz, 2000; Muñiz y Hambleton, 1996; Sireci, Patsula y Hambleton, 2005; van de Vijver y Hambleton, 1996; van de Vijver y Poortinga, 2005). Partiendo de una revisión exhaustiva de la literatura sobre el tema, Hambleton y Patsula (1999) propusieron una serie de pasos para evitar tales fuentes de sesgo y mejorar la calidad en la adaptación de los tests. Dichos pasos implican, fundamentalmente, la selección de traductores profesionales cualificados, la elección de un adecuado diseño de traducción, la realización de las correcciones necesarias en la primera adaptación lingüística del instrumento, la realización de un estudio piloto con el test adaptado, la aplicación empírica del test a una muestra amplia y representativa de la po-

blación objetivo y, si se pretenden llevar a cabo comparaciones interculturales, la comprobación de la equivalencia de los ítems en las versiones original y objetivo del instrumento.

Partiendo de tales recomendaciones en el presente trabajo se pretende mostrar, desde una perspectiva práctica, las pautas que se deben seguir para adaptar adecuadamente un instrumento de medida de una cultura a otra. Para ello, se describen exhaustivamente los pasos que se están siguiendo en el proceso de adaptación del NEO PI-R (Costa y McCrae, 1992) del castellano al euskera, así como los resultados obtenidos en dicho proceso.

Antes de comenzar con la descripción, cabe señalar que el NEO PI-R es uno de los instrumentos más utilizados para evaluar la personalidad desde el modelo de las «Cinco Grandes» dimensiones que han emergido en distintos estudios transculturales (para una revisión del modelo véanse Digman, 1990; John, 1990; McCrae y Costa, 1987; Wiggins, 1996). De hecho, aunque se trata de un inventario relativamente nuevo en el campo de la evaluación de la personalidad, ya se han llevado a cabo adaptaciones del mismo a 36 culturas diferentes (McCrae, 2002). Como señalan Costa, McCrae y Johnson (2002), la adaptación del instrumento al euskera añadiría otra familia de lenguas al conjunto de familias a las que

ya ha sido adaptado el NEO PI-R. Además, respondería a una demanda proveniente de una gran parte de la sociedad vasca que desea ser evaluada en su lengua materna y dentro de su propio grupo de referencia.

#### Fases del proceso de adaptación

##### 1. Fase de traducción de los ítems

Los tres primeros pasos propuestos por Hambleton y Patsula (1999) para mejorar la calidad en la adaptación de los tests, a saber, la selección de traductores profesionales cualificados, la elección de un adecuado diseño de traducción y la realización de las correcciones necesarias en la primera adaptación lingüística del instrumento, pueden integrarse dentro de una fase más amplia del proceso de adaptación como es la referida a la traducción de los ítems. En esta sección se describe la forma en la que se ejecutaron tales pasos en la adaptación del NEO PI-R del castellano al euskera. Es importante señalar que la traducción debe realizarse con sumo cuidado, puesto que los pasos arriba enumerados constituyen los cimientos sobre los que se apoyarán las fases posteriores del proceso.

La traducción de los ítems del NEO PI-R se llevó a cabo utilizando un diseño de traducción inversa, dado que es uno de los procedimientos más adecuados para detectar algunos de los problemas asociados con traducciones o adaptaciones deficientes y, por ende, para evaluar la calidad de la traducción (Hambleton, 1996, 2005). Cada uno de los ítems de la versión en castellano fue traducido al euskera de forma independiente por dos traductores con amplio dominio del castellano y del euskera, familiarizados con las dos culturas e instruidos en los aspectos psicométricos básicos relacionados con la construcción de ítems. Además, uno de ellos era experto en el ámbito de la evaluación de la personalidad. Las dos traducciones fueron comparadas y sometidas a discusión hasta obtener una versión consensuada de cada uno de los ítems. Partiendo de esta versión, otros dos traductores de características similares a los que realizaron la traducción directa, tradujeron, de forma independiente, todos los ítems de la versión en euskera al castellano y llegaron a una versión consensuada del test. Finalmente, los cuatro integrantes del equipo de traductores compararon cada uno de los ítems de las versiones original e inversamente adaptada del instrumento, examinando las posibles no-equivalencias en cuanto al significado. Partiendo de este análisis se llevaron a cabo las modificaciones oportunas en la versión en euskera del NEO PI-R. Posteriormente, esta versión del instrumento fue traducida al inglés por un bilingüe con amplio conocimiento de las lenguas inglesa y vasca, y ajeno al proceso de traducción anterior. Su traducción fue revisada por el profesor Robert R. McCrae, coautor de la versión original del instrumento, quien propuso pequeñas modificaciones en algunos de los ítems. Tras la incorporación de estas modificaciones, la versión en euskera del instrumento ha sido considerada como una «traducción autorizada» por los autores de la versión original del NEO PI-R.

##### 2. Fase piloto

Siguiendo la propuesta de Hambleton y Patsula (1999), se llevó a cabo un estudio piloto a fin de obtener datos empíricos acerca del funcionamiento de la versión en euskera del NEO PI-R. A continuación se describe la forma en la que se realizó dicho estudio y los resultados obtenidos en el mismo.

#### Método

##### Participantes

En el estudio participaron 210 personas (54 varones y 156 mujeres) con edades comprendidas entre los 16 y los 59 años. Todas ellas eran residentes en la comunidad autónoma del País Vasco y presentaban un alto dominio de la lengua vasca. El 60% de la muestra tenía edades comprendidas entre los 16 y los 22 años (jóvenes) y el 40% restante oscilaba entre los 25 y los 59 años (adultos). La muestra de jóvenes estaba compuesta por estudiantes de Bachillerato y universitarios, mientras que la de adultos estaba compuesta por funcionarios de la Administración pública vasca. Todos los individuos participaron en el estudio bajo condiciones que preservaban su anonimato.

##### Instrumento de medida

Versión en euskera del inventario NEO PI-R. El inventario NEO PI-R consta de 240 ítems que se puntúan en una escala tipo Likert con 5 opciones de respuesta que van desde «Totalmente de acuerdo» («*Erabat ados*») hasta «En total desacuerdo» («*Erabat desados*»). La forma S del inventario NEO PI-R es una medida de autoinforme que evalúa la personalidad en base al modelo de los «Cinco Grandes» factores, a saber, el Neuroticismo (N), la Extraversión (E), la Apertura a la experiencia (O), la Amabilidad (A) y la Responsabilidad (C). Cada una de estas cinco dimensiones incluye seis facetas. El Neuroticismo es la dimensión que se refiere al nivel de inestabilidad emocional y está compuesta por las facetas de ansiedad, hostilidad, depresión, ansiedad social, impulsividad y vulnerabilidad. La Extraversión es la dimensión que muestra el nivel de sociabilidad y energía, y está integrada por las facetas de cordialidad, gregarismo, asertividad, actividad, búsqueda de sensaciones y emociones positivas. La Apertura a la experiencia refleja la sensibilidad y el interés por el mundo interior y exterior e incluye las facetas de fantasía, estética, sentimientos, acciones, ideas y valores. La Amabilidad se refiere a la naturaleza de las relaciones interpersonales, y las facetas que la componen son la confianza, la franqueza, el altruismo, la actitud conciliadora, la modestia y la sensibilidad a los demás. Por último, la Responsabilidad mide el nivel de autocontrol y la capacidad de planificación, a través de las facetas de competencia, orden, sentido del deber, necesidad de logro, autodisciplina y deliberación.

Cabe señalar que en la versión en castellano del inventario NEO PI-R se obtuvo una estructura factorial de cinco factores, en la que las 30 facetas saturaban de forma clara en sus correspondientes dimensiones, si bien algunas facetas también presentaban saturaciones altas en otros factores. En cuanto a la consistencia interna de las dimensiones, el índice  $\alpha$  de Cronbach adoptó valores de 0,90; 0,84; 0,82; 0,83 y 0,88 para el Neuroticismo, la Extraversión, la Apertura a la experiencia, la Amabilidad y la Responsabilidad, respectivamente. Asimismo, los análisis realizados con las variables, sexo, edad, inteligencia e intereses profesionales con el fin de examinar la validez diferencial del instrumento, corroboraron los resultados encontrados en estudios anteriores (Costa y McCrae, 1999).

##### Procedimiento

Los datos fueron recogidos por uno de los autores del presente artículo, en los centros de estudio o de trabajo de los participantes.

Además de aplicarles la versión en euskera del inventario NEO PI-R, se les pidió que realizaran las sugerencias que consideraran oportunas respecto a cualquier aspecto del test.

#### *Análisis cuantitativo y cualitativo de los ítems*

Con los datos recogidos en esta fase, se calcularon las puntuaciones medias, desviaciones típicas e índices de homogeneidad de cada uno de los ítems con respecto a sus correspondientes facetas. Aquellos ítems que presentaron un índice de homogeneidad menor que 0,30 fueron revisados por los cuatro integrantes del equipo de traductores, quienes, siguiendo las sugerencias de los expertos en la evaluación de la personalidad, llevaron a cabo algunas modificaciones a fin de que los ítems se aproximaran en mayor medida a sus correspondientes facetas. Por otra parte, se tomaron en consideración las sugerencias que realizaron los participantes en el estudio, las cuales se centraron básicamente en la complejidad o dificultad de comprensión de algunos términos incluidos en determinados ítems. En este caso, el equipo de traductores trató de reformular los ítems a fin de facilitar su comprensión. El instrumento derivado de este proceso de revisión es el que se está utilizando para realizar una validación empírica rigurosa de la versión del NEO PI-R adaptada al euskera.

#### *3. Fase de aplicación empírica del instrumento de medida*

Dado que la aplicación empírica del test a una muestra amplia de la población objetivo constituye el paso más importante en el proceso de adaptación, describiremos exhaustivamente una serie de estudios que se han llevado a cabo con el objetivo de buscar evidencias sobre el grado de validez de la versión adaptada al euskera del NEO PI-R. Es importante señalar que el proceso de adaptación no ha culminado todavía, por lo que los resultados derivados de tales estudios no pueden ser considerados como definitivos. No obstante, las pautas que se han seguido para la validación son extrapolables a cualquier proceso de adaptación.

En concreto, en esta fase se llevaron a cabo cuatro estudios. En el Estudio 1 se examinó la dimensionalidad del instrumento. El Estudio 2 se centró en el análisis de la fiabilidad del NEO PI-R en lo que se refiere a su consistencia interna. En el Estudio 3 se obtuvieron evidencias sobre la validez convergente y diferencial del instrumento. Finalmente, el último de los estudios examinó el funcionamiento diferencial de los ítems del NEO PI-R con respecto al sexo y a la etapa vital (joven o adulto).

#### ESTUDIO 1

El objetivo del Estudio 1 consistió en analizar la estructura dimensional de la versión en euskera del inventario NEO PI-R mediante un análisis factorial exploratorio y en comparar dicha estructura con la de la versión en castellano del test.

#### Método

##### *Participantes*

La versión en euskera del instrumento fue cumplimentada por 425 personas con amplio dominio de la lengua vasca y residentes en la comunidad autónoma del País Vasco. De ellas, 245 eran mujeres y 180 eran varones. El 70% de la muestra tenía edades com-

prendidas entre los 16 y los 20 años (jóvenes) y el 30% restante oscilaba entre los 25 y los 60 años (adultos). La muestra de jóvenes estaba compuesta por estudiantes de 1º y 2º de Bachillerato y la de adultos por funcionarios de la Administración pública vasca. Todos ellos participaron en el estudio bajo condiciones que preservaban su anonimato.

Las respuestas a la versión en castellano del inventario fueron recogidas en una muestra normativa de 2.000 personas que participaban en procesos de selección, cuya procedencia y características eran muy diversas. De ellos, 1.000 eran mujeres y 1.000 eran varones (Costa y McCrae, 1999).

#### *Instrumento de medida*

En el presente estudio se utilizaron la versión en euskera del inventario NEO PI-R que incorporaba las modificaciones derivadas del estudio piloto y la versión en castellano de dicho instrumento. Dado que la descripción general del NEO PI-R ya ha sido realizada en el epígrafe referido a la fase piloto, no redundaremos en ese aspecto.

#### Resultados

A fin de examinar la dimensionalidad del inventario, se llevó a cabo un análisis de componentes principales con rotación Varimax sobre sus 30 facetas. Se decidió aplicar dicha rotación en base a que los cinco «Grandes Factores» de la personalidad, y su operacionalización en el NEO PI-R, han sido conceptualizados como ortogonales (Costa y McCrae, 1995; Goldberg, 1993). De hecho, en otras versiones lingüísticas del test (McCrae, 2002; McCrae y Allik, 2002; McCrae y Costa, 1997) se utiliza el mismo procedimiento de rotación al analizar su estructura factorial. Antes de realizar el análisis, se calcularon la medida de adecuación muestral de Kaiser-Meyer-Olkin (KMO) y el test de esfericidad de Bartlett. El índice KMO mostró un valor de 0,86 y el test de Bartlett resultó estadísticamente significativo,  $\chi^2(435) = 5592,82; p < 0,001$ , lo que nos llevó a concluir que el análisis de componentes principales resultaba a priori pertinente.

Partiendo del criterio de extracción de factores de Kaiser se obtuvieron 5 factores con valores propios iguales o superiores a la unidad, que explicaban conjuntamente el 57,49% de la varianza total. En la tabla 1 se presentan los pesos factoriales rotados para cada faceta, así como los porcentajes de varianza explicada para cada factor. Considerando el valor de 0,35 como punto de corte de los elementos de la matriz factorial, todas las facetas, a excepción de la Asertividad (3ª faceta de la Extraversión), saturaban de forma clara en sus correspondientes dimensiones.

Con el objetivo de examinar si la estructura factorial de la versión en euskera del NEO PI-R es equivalente a la de la versión en castellano, se utilizó la técnica de rotación ortogonal Procrustes (McCrae, Zonderman, Costa, Bond, y Paunonen, 1996). En la tabla 2 se presenta la estructura factorial rotada Procrustes para la muestra vascoparlante. El coeficiente de congruencia total adoptó el valor de 0,95 y los de los factores oscilaron entre 0,91 y 0,97. Además, a excepción de la faceta Deliberación (C6) de la Responsabilidad, las restantes facetas mostraron una alta congruencia entre ambas versiones lingüísticas del test. Tales resultados ponen de manifiesto que la solución factorial de la versión en euskera del NEO PI-R es equivalente a la de la versión en castellano.

## ESTUDIO 2

El objetivo del Estudio 2 consistió en examinar la consistencia interna de las distintas dimensiones del NEO PI-R y en compararla con la de las dimensiones de la versión en castellano del test.

## Método

Las muestras y los instrumentos de medida utilizados fueron los mismos que los del Estudio 1.

## Resultados

La consistencia interna de cada una de las dimensiones del inventario fue estimada mediante el coeficiente  $\alpha$  de Cronbach. En

la versión en euskera, dicho índice adoptó valores de 0,92 ( $IC_{95\%}=0,91-0,93$ ); 0,86 ( $IC_{95\%}=0,84-0,88$ ); 0,83 ( $IC_{95\%}=0,80-0,85$ ); 0,85 ( $IC_{95\%}=0,83-0,87$ ) y 0,90 ( $IC_{95\%}=0,89-0,91$ ) para el Neuroticismo, la Extraversión, la Apertura a la experiencia, la Amabilidad y la Responsabilidad, respectivamente. Estos valores permiten concluir que la consistencia interna es elevada en todas las dimensiones, superando el punto de corte de 0,75 generalmente aceptado para instrumentos pertenecientes al ámbito de las ciencias de la salud (Streiner y Norman, 1989).

La comparación entre los índices de consistencia interna de ambas versiones lingüísticas del test se llevó a cabo mediante la *W* de Feldt. Los resultados obtenidos en tales análisis se presentan en la tabla 3. Como puede observarse en dicha tabla, exceptuando las dimensiones de Apertura a la experiencia y Amabilidad, en las que no se encontraron diferencias estadísticamente significativas entre

	1	2	3	4	5
Neuroticismo (N)					
N1	,825				
N2	,609			-,502	
N3	,842				
N4	,789				
N5	,380				-,478
N6	,769				
Extraversión (E)					
E1		,776			
E2		,734			
E3	-,434			-,382	
E4		,560			
E5		,583			
E6		,755			
Apertura a la experiencia (O)					
O1			,652		
O2			,682		
O3		,489	,538		
O4			,429		
O5			,675		
O6			,499		
Amabilidad (A)					
A1				,581	
A2				,758	
A3		,504		,558	
A4				,701	
A5	,413			,533	
A6				,579	
Responsabilidad (C)					
C1					,728
C2					,712
C3				,426	,633
C4					,736
C5					,789
C6					,675
Porcentaje de varianza explicada	19,76	10,24	6,47	7,78	13,23
Autovalores	5,93	3,07	1,94	2,34	3,97

Nota: Se ha utilizado una rotación Varimax. Los pesos factoriales menores que 0,35 han sido omitidos de la tabla

Facetas NEO PI-R	Factor					Congruencia entre facetas
	N	E	O	A	C	
Neuroticismo (N)						
N1	<b>,83</b>	-,01	,02	,03	,06	,99
N2	<b>,65</b>	,11	-,12	<b>-,45</b>	-,13	,97
N3	<b>,85</b>	-,12	-,02	-,01	-,21	,99
N4	<b>,78</b>	-,17	,04	,15	-,13	,97
N5	<b>,43</b>	,27	,14	-,27	<b>-,49</b>	,97
N6	<b>,77</b>	-,02	-,02	,00	-,26	,94
Extraversión (E)						
E1	-,08	<b>,79</b>	,10	,28	,00	,91
E2	-,05	<b>,72</b>	-,28	,17	-,12	,92
E3	<b>-,38</b>	,32	,26	<b>-,42</b>	,20	,98
E4	,03	<b>,57</b>	,19	-,13	,08	,94
E5	,09	<b>,56</b>	,02	-,20	-,23	,92
E6	-,02	<b>,75</b>	,18	,05	-,09	,96
Apertura a la experiencia (O)						
O1	,17	,16	<b>,64</b>	-,07	-,16	,98
O2	,21	,08	<b>,67</b>	,12	,16	,96
O3	,19	,51	<b>,51</b>	-,03	,13	,99
O4	-,09	,21	<b>,43</b>	-,03	-,07	,95
O5	-,19	-,16	<b>,67</b>	-,03	,31	,97
O6	-,16	,05	<b>,52</b>	,24	-,03	,97
Amabilidad (A)						
A1	<b>-,37</b>	,15	,13	<b>,55</b>	,13	,98
A2	-,05	,17	,02	<b>,76</b>	,17	,92
A3	-,05	<b>,52</b>	,06	<b>,55</b>	,15	,95
A4	-,16	-,18	,05	<b>,69</b>	,06	,94
A5	,37	-,07	-,02	<b>,56</b>	-,16	,96
A6	,09	,17	<b>,36</b>	<b>,57</b>	,11	,96
Responsabilidad (C)						
C1	-,29	,06	,21	,00	<b>,73</b>	,98
C2	-,10	,09	-,20	-,02	<b>,70</b>	,98
C3	-,07	,09	,13	<b>,41</b>	<b>,63</b>	,98
C4	-,01	,26	,06	-,19	<b>,73</b>	,96
C5	-,24	,01	-,01	,15	<b>,79</b>	,97
C6	-,14	-,29	-,05	,20	<b>,69</b>	,78
Congruencia factorial/total	,97	,91	,95	,97	,96	,95

Nota: Se ha utilizado una rotación Procrustes. Los pesos factoriales superiores a 0,35 en valor absoluto se han destacado en negrita

los índices de consistencia interna de ambas versiones lingüísticas del instrumento, los  $\alpha$ -s obtenidos en las restantes dimensiones fueron más altos para la versión en euskera que para la versión en castellano.

### ESTUDIO 3

El objetivo del Estudio 3 consistió en examinar la validez convergente y diferencial del instrumento.

#### Método

##### Participantes

En el presente estudio se utilizó la muestra de vascoparlantes empleada en los Estudios 1 y 2.

##### Instrumentos de medida

Además de la versión en euskera del Inventario NEO PI-R descrita en la fase piloto, en este estudio se utilizó un cuestionario que recababa información acerca de distintas variables sociodemográficas y la versión en castellano del cuestionario «Big Five» (BFQ) (Caprara, Barbaranelli y Borgogni, 1995).

- Cuestionario de variables sociodemográficas. Este cuestionario fue elaborado *ad hoc* para la investigación e incluía preguntas acerca de variables que, a tenor de la revisión de la literatura sobre el modelo de los «Cinco Grandes», se hallan relacionadas con las distintas dimensiones que mide el NEO PI-R. Consta de 21 ítems con diferentes formatos de respuesta. Además de preguntarle cuál era su sexo, su fecha de nacimiento, su lugar de nacimiento y residencia y su estado civil, el participante debía responder a preguntas que abordaban aspectos tales como el orden de nacimiento, el nivel de estudios, la situación laboral, el nivel socioeconómico, las aficiones, las tendencias artísticas, la propensión a sufrir accidentes, enfermedades y acontecimientos negativos, el consumo de sustancias peligrosas, la orientación religiosa y la tendencia política.
- Versión en castellano del Cuestionario «Big Five» (BFQ).- El BFQ consta de 132 ítems con una escala de respuesta tipo Likert de cinco puntos que va desde «Completamente verdadero para mí» hasta «Completamente falso para mí». Se trata de un instrumento que, al igual que el NEO PI-R,

está basado en el modelo de los «Cinco Grandes» factores. En este cuestionario los factores se han denominado Energía, Afabilidad, Tesón, Estabilidad emocional y Apertura mental. En cada uno de los cinco factores se han identificado dos subdimensiones, que hacen referencia a distintos aspectos de la propia dimensión y que constan de 12 elementos. Asimismo, el BFQ dispone de una escala de Distorsión que tiene por finalidad proporcionar una medida de la tendencia del individuo a ofrecer un perfil falseado de sí mismo al responder al cuestionario.

Cabe señalar que el análisis de la estructura factorial de la versión en castellano del BFQ confirmó la presencia de cinco factores que subyacen a las diez subdimensiones y que explicaban conjuntamente el 57% de la varianza total. Cada una de las subdimensiones presentó saturaciones elevadas en su factor correspondiente y saturaciones bajas o nulas en el resto de los factores. En cuanto a la consistencia interna de tales factores, el índice  $\alpha$  de Cronbach adoptó valores de 0,75; 0,73; 0,79; 0,87 y 0,76 para las dimensiones de Energía, Afabilidad, Tesón, Estabilidad emocional y Apertura mental, respectivamente. Asimismo, cabe destacar que las correlaciones existentes entre las dimensiones del BFQ y los constructos evaluados mediante otros instrumentos (JASSE-H, EPQ, STAI-R...) mostraron una adecuada validez convergente y discriminante (Caprara et al., 1995).

#### Resultados

##### Asociación entre las dimensiones del NEO PI-R y del BFQ

A fin de obtener evidencia convergente del instrumento, se calcularon los coeficientes de correlación de Pearson entre las puntuaciones de las dimensiones del NEO PI-R y del BFQ que miden el mismo constructo, así como los índices de correlación corregidos por atenuación. Tales coeficientes se presentan en la tabla 4. Como cabía esperar, las puntuaciones en Neuroticismo del NEO PI-R mostraron una asociación inversa estadísticamente significativa y de gran magnitud con las puntuaciones en Estabilidad emocional del BFQ. A su vez, la asociación entre las puntuaciones de Responsabilidad del NEO PI-R y de Tesón del BFQ también fue alta, aunque de signo positivo. Las puntuaciones en Extraversión, Apertura a la experiencia y Amabilidad del NEO PI-R mostraron correlaciones positivas de una intensidad entre moderada y alta con las de las dimensiones Energía, Apertura mental y Afabilidad del BFQ, respectivamente.

Dimensión	$\alpha$ de Cronbach		W de Feldt
	Versión en castellano	Versión en euskera	
Neuroticismo	0,90	0,92	1,25*
Extraversión	0,84	0,86	1,14*
Apertura a la experiencia	0,82	0,83	1,06
Amabilidad	0,83	0,85	1,13
Responsabilidad	0,88	0,90	1,20*

\*  $p < 0,05$

Dimensiones NEO PI-R	Dimensiones BFQ	R	R corregida por atenuación
Neuroticismo	Estabilidad emocional	-0,707**	-0,804
Extraversión	Energía	0,547**	0,705
Apertura a la experiencia	Apertura mental	0,537**	0,667
Amabilidad	Afabilidad	0,581**	0,709
Responsabilidad	Tesón	0,710**	0,826

\*\*  $p < 0,0001$



*Relación entre las dimensiones del NEO PI-R y otras variables*

Con el objetivo de obtener evidencia adicional sobre la relación entre los constructos medidos por el NEO PI-R y otras variables se utilizaron procedimientos de validación diferencial basados en pruebas paramétricas (*t* de Student), siempre que se cumplieran los supuestos para la aplicación de tales pruebas, y no paramétricas (*U* de Mann-Whitney), cuando no se cumplía alguno de tales supuestos. Atendiendo a la naturaleza de las variables que trataban de predecir las posibles diferencias existentes entre los individuos en las distintas dimensiones del NEO PI-R, se plantearon hipótesis que hacían referencia a dos tipos de aspectos. Por una parte, se examinaron dos variables de carácter biológico, tales como la edad y el sexo, y, por otra, se analizaron dos variables de naturaleza más sociocultural, a saber, la tendencia política y la orientación religiosa. Los análisis realizados en el presente estudio trataron de replicar los resultados obtenidos en investigaciones previas en las que se han encontrado diferencias en algunas de las dimensiones globales o de las facetas del NEO PI-R en función de dichas variables.

En primer lugar, se presentan las diferencias encontradas en las puntuaciones de las dimensiones del NEO PI-R en función de la edad. Así, partiendo de los resultados obtenidos por diversos autores (Costa, Terracciano y McCrae, 2001; McCrae y John, 1992; McCrae et al., 1999; Yang, McCrae y Costa, 1998), se predijo que los jóvenes (entre 16 y 20 años) presentarían un mayor nivel de Neuroticismo, Extraversión y Apertura a la experiencia que los adultos (entre 25 y 60 años). A su vez, se hipotetizó que la Amabilidad y la Responsabilidad caracterizarían en mayor medida a los adultos que a los jóvenes. El contraste *t* de diferencia de medias entre los niveles de Neuroticismo de los jóvenes y de los adultos resultó estadísticamente significativo ( $t(423)=6,06$ ;  $p=0,0001$ ;  $d$  de Cohen=0,65), corroborando la hipótesis planteada respecto a dicha dimensión. De la misma forma, los jóvenes también mostraron un mayor nivel de Extraversión ( $t(423)=6,04$ ;  $p=0,0001$ ;  $d$  de Cohen=0,63) que los adultos. Por otra parte, aunque la comparación entre los promedios de los rangos de las puntuaciones presentadas por los adultos y por los jóvenes en la Apertura a la experiencia, llevada a cabo mediante la *U* de Mann-Whitney ( $U$  de Mann-Whitney=15450;  $p=0,003$ ), puso de manifiesto que desde el punto de vista estadístico los adultos obtenían significativamente puntuaciones más altas que los jóvenes, el efecto fue de una escasa magnitud ( $\eta^2=0,02$ ). Sin embargo, como veremos en el Estudio 4, esta dimensión fue la que mostró una mayor cantidad de ítems con DIF con respecto a la etapa vital, lo que nos obliga a mantener cierta cautela ante el presente resultado. En el caso de la Amabilidad, el contraste *t* de diferencia de medias ( $t(423)=-4,54$ ;  $p=0,0001$ ;  $d$  de Cohen=0,48) corroboró nuestra hipótesis, mostrando los adultos un mayor nivel de amabilidad que los jóvenes. Por último, también se confirmó que la Responsabilidad es un rasgo que caracteriza en mayor medida a los adultos que a los jóvenes ( $t(423)=-6,73$ ;  $p=0,0001$ ;  $d$  de Cohen=0,71).

En segundo lugar, se presentan las diferencias encontradas en las puntuaciones de las dimensiones del NEO PI-R en función del sexo. Así, partiendo de las investigaciones que han mostrado que las mujeres presentan puntuaciones consistentemente más altas que los hombres en Neuroticismo (Costa et al., 2001; De Miguel, 2005; Lynn y Martin, 1997; Nolen-Hoeksema, 1987) y en Amabilidad (Costa et al., 2001; Feingold, 1994), se hipotetizó que la puntuación de las mujeres de la muestra superaría a la de los varones

en ambas dimensiones. Confirmando tal predicción, el contraste *t* de diferencia de medias entre los niveles de Neuroticismo de las mujeres y de los hombres resultó estadísticamente significativo ( $t(423)=5,96$ ;  $p=0,0001$ ;  $d$  de Cohen=0,58). De la misma forma, las mujeres mostraron un mayor nivel de Amabilidad que los hombres ( $t(423)=5,02$ ;  $p=0,0001$ ;  $d$  de Cohen=0,49). La diferencia entre las mujeres y los hombres resultó estadísticamente significativa en todas las facetas del Neuroticismo y en cuatro de las facetas de la Amabilidad. No obstante, las diferencias entre los individuos de ambos sexos no fueron estadísticamente significativas ni en Confianza (A1) ni en Actitud conciliadora (A4). Aquí también hemos de señalar la necesidad de contemplar los resultados asociados al Neuroticismo con cierta cautela, dado que, en el Estudio 4, ésta fue la dimensión que mostró una mayor cantidad de DIF con respecto al sexo.

Con respecto a la Extraversión, los resultados de nuestro estudio concuerdan con los obtenidos por Costa et al. (2001). Así, las mujeres puntuaron significativamente más alto que los hombres en Cordialidad (E1) ( $t(423)=4,25$ ;  $p=0,0001$ ;  $d$  de Cohen=0,42), Gregarismo (E2) ( $t(423)=4,76$ ;  $p=0,0001$ ;  $d$  de Cohen=0,47) y Emociones positivas (E6) ( $t(423)=4,22$ ;  $p=0,0001$ ;  $d$  de Cohen=0,41), mientras que en Asertividad (E3) y Búsqueda de emociones (E5) no se observaron diferencias estadísticamente significativas en función del sexo.

En cuanto a la dimensión de Apertura a la experiencia, nuestro estudio también confirmó los resultados encontrados por Costa et al. (2001) y por De Miguel (2005). De hecho, las mujeres obtuvieron puntuaciones significativamente más altas en Apertura a la estética (O2) ( $t(423)=5,99$ ;  $p=0,0001$ ;  $d$  de Cohen=0,59), Apertura a los sentimientos (O3) ( $t(423)=5,11$ ;  $p=0,0001$ ;  $d$  de Cohen=0,50), y Apertura a la acción (O4) ( $t(423)=2,85$ ;  $p=0,005$ ;  $d$  de Cohen=0,28), si bien, con respecto a esta última faceta, la magnitud del efecto fue moderada. Por otra parte, los hombres se mostraron más abiertos a las ideas (O5) que las mujeres ( $t(423)=-2,52$ ;  $p=0,012$ ;  $d$  de Cohen=0,25), aunque también en este caso el tamaño del efecto fue moderado.

En tercer lugar, se presentan las diferencias encontradas en las puntuaciones de las dimensiones del NEO PI-R en función de la tendencia política. En este caso, diversas investigaciones han mostrado que las personas cuya tendencia política es de centro-derecha presentan mayores puntuaciones en Responsabilidad y Extraversión que las que se manifiestan de izquierdas, mientras que estas últimas se caracterizan por ser más amables y abiertas a la experiencia que las de centro-derecha (Altemeyer, 1996; Caprara, Barbaranelli y Zimbardo, 1999; Lippa y Arrad, 1999; Van Hiel y Mervielde, 2002). En contra de lo esperado, el contraste *t* de diferencia de medias entre los niveles de Extraversión de los individuos de centro-derecha y los individuos de izquierdas puso de manifiesto que estos últimos eran significativamente más extravertidos que los primeros ( $t(381)=2,21$ ;  $p=0,028$ ;  $d$  de Cohen=0,24), aunque con una magnitud del efecto moderada. En dos de las dimensiones restantes se cumplieron nuestras predicciones. Así, los contrastes *t* de diferencia de medias mostraron un mayor nivel de Amabilidad ( $t(381)=2,33$ ;  $p=0,021$ ;  $d$  de Cohen=0,26) y un menor nivel de Responsabilidad ( $t(381)=-3,42$ ;  $p=0,001$ ;  $d$  de Cohen=0,34) en los votantes de izquierdas que en los de centro-derecha, siendo los efectos observados de una magnitud moderada. Sin embargo, a pesar de que la comparación entre los promedios de los rangos de las puntuaciones presentadas por las personas de estas dos orientaciones políticas, llevada a cabo me-

dante la  $U$  de Mann-Whitney, puso de manifiesto que desde el punto de vista estadístico los individuos de izquierdas obtenían puntuaciones significativamente más altas en Apertura a la experiencia que los de derechas ( $U$  de Mann-Whitney= 13275,50;  $p=0,022$ ), los efectos fueron de una escasa magnitud ( $\eta^2=0,02$ ).

Finalmente, se presentan las diferencias encontradas en las puntuaciones de las dimensiones del NEO PI-R en función de la orientación religiosa. Así, partiendo de los resultados obtenidos en diversos estudios (Saroglou y Fiasse, 2003; Saucier y Goldberg, 1998; Strejffeler y McNally, 1998), se hipotetizó que los individuos creyentes, fuesen o no practicantes, mostrarían mayores niveles de Responsabilidad y Amabilidad que los no creyentes. A su vez, se predijo que estos últimos obtendrían puntuaciones significativamente superiores a las de los primeros en Impulsividad (N5), Búsqueda de emociones (E5) y Apertura a los valores (O6). De acuerdo con la primera de tales predicciones, el contraste  $t$  de diferencia de medias puso de manifiesto que la puntuación de los creyentes en Responsabilidad era significativamente más alta que la de los no creyentes ( $t(415)=-3,20$ ;  $p=0,001$ ;  $d$  de Cohen= 0,33), aunque en contra de lo esperado, estos últimos presentaron una puntuación similar a la de los primeros en Amabilidad. Por otra parte, los resultados obtenidos confirmaron que, como cabía esperar, los no creyentes presentaban puntuaciones significativamente superiores a las de los creyentes en Impulsividad (N5) ( $t(415)=1,98$ ;  $p=0,049$ ;  $d$  de Cohen= 0,20), aunque la magnitud del efecto fue moderada. Con respecto a las facetas Búsqueda de emociones (E5) ( $U$  de Mann-Whitney= 17116,50;  $p=0,023$ ;  $\eta^2=0,02$ ) y Apertura a los valores (O6) ( $U$  de Mann-Whitney= 16462,50;  $p=0,004$ ;  $\eta^2=0,02$ ), los resultados pusieron de manifiesto que, aunque desde el punto de vista de la significación estadística existían diferencias en la dirección esperada, los efectos presentaban una magnitud demasiado pequeña como para considerar tales diferencias como reales.

#### ESTUDIO 4

El objetivo del Estudio 4 consistió en estudiar el funcionamiento diferencial de los ítems del NEO PI-R con respecto al sexo y a la etapa vital (joven o adulto). Con ello se pretendía examinar si el lenguaje mediante el cual estaban expresados los ítems podía favorecer que los individuos de un determinado sexo o de una determinada edad obtuviesen puntuaciones de mayor o menor valor en alguno de los ítems de cada dimensión.

#### Método

En el estudio se utilizó la misma muestra de vascoparlantes empleada en los Estudios 1, 2 y 3.

#### Resultados

El análisis del funcionamiento diferencial de los ítems (DIF) fue llevado a cabo mediante el procedimiento de regresión logística (Swaminathan y Rogers, 1990) dentro de cada una de las dimensiones del NEO PI-R, utilizando un software específico para detección del DIF (Gómez, Hidalgo, Padilla y González, 2005; Hidalgo, Gómez y Padilla, 2005). Previamente, los ítems fueron dicotomizados, recodificándose las categorías 0 (Total desacuerdo), 1 (Desacuerdo) y 2 (Neutral) en la categoría 0 y las categorías 3 (Acuerdo) y 4 (Total acuerdo) en la 1.

El modelo de regresión logística adopta la siguiente expresión:

$$y = \tau_0 + \tau_1 D + \tau_2 G + \tau_3 DG$$

donde  $y$  hace referencia a la respuesta obtenida en el ítem,  $D$  al nivel alcanzado en una determinada dimensión,  $G$  al grupo de pertenencia y  $DG$  a la interacción entre ambas variables. El parámetro  $\tau_1$  representa las diferencias existentes en la ejecución del ítem entre los individuos que presentan distintos niveles en una determinada dimensión,  $\tau_2$  corresponde a las diferencias existentes entre los grupos de comparación en la ejecución del ítem y el parámetro  $\tau_3$  se refiere a la interacción entre el nivel alcanzado en la dimensión y la pertenencia grupal. Siguiendo el modelo de regresión logística, un ítem muestra DIF uniforme, si  $\tau_2$  es distinto de cero y  $\tau_3$  es igual a cero; y DIF no uniforme, si  $\tau_3$  es distinto de cero, independientemente del valor que adopte  $\tau_2$ . En el presente estudio, se utilizó el estadístico de Wald para examinar la hipótesis referida a la existencia de DIF en los ítems de cada dimensión del NEO PI-R. Dicho estadístico compara el parámetro estimado con su error estándar. En la tabla 5 se presentan los valores de los estadísticos de Wald y sus niveles de significación para los parámetros  $\tau_2$  y  $\tau_3$  en cada una de las dimensiones y facetas del NEO PI-R tomando como grupos de comparación las mujeres y los hombres de la muestra.

Tabla 5  
Estadísticos de Wald y niveles de significación de los ítems pertenecientes al inventario NEO PI-R en función de la variable sexo

Dimensión	Faceta	Nº de ítem	$\tau_2$	$p$	$\tau_3$	$p$
Neuroticismo	Ansiedad	31	5,00	0,025		
		151	4,03	0,045		
	Hostilidad	36	5,19	0,023		
		131			4,07	0,044
	Depresión	191			8,18	0,004
		16			4,95	0,026
	Ansiedad social	21			7,75	0,005
		81			8,60	0,003
Extraversión	Impulsividad	111			4,86	0,028
		201			6,03	0,014
	Vulnerabilidad	116			4,15	0,042
		132			4,59	0,032
	Asertividad	197	4,48	0,034		
		117			5,18	0,023
	Actividad	68	5,22	0,022		
		233			4,12	0,042
Apertura a la experiencia	Estética	68	5,22	0,022		
	Ideas	233			4,12	0,042
Amabilidad	Confianza	154	4,82	0,028		
	Actitud conciliadora	199	4,53	0,033		
Responsabilidad	Competencia	155	4,05	0,044		
		130			4,46	0,035
	Orden	80			3,93	0,047
		60	4,27	0,039		
	Deliberación	90			6,86	0,009

Nota: Se incluye el valor de los estadísticos para los parámetros significativos



Los valores obtenidos en el estadístico de Wald pusieron de manifiesto que en la dimensión Neuroticismo, 11 de los ítems presentaban DIF. Además, en 8 de ellos el DIF fue no uniforme. Con respecto a la Extraversión, la cantidad de ítems con DIF disminuyó hasta 3, siendo en 2 de ellos no uniforme. Tanto en la dimensión Apertura a la experiencia como Amabilidad, únicamente se observaron 2 ítems con funcionamiento diferencial. En el primer caso, uno de ellos mostró DIF uniforme y el otro presentó DIF no uniforme, mientras que en la segunda dimensión ninguno de ellos mostró DIF no uniforme. Por último, en la dimensión Responsabilidad, 5 ítems presentaron un funcionamiento diferencial para mujeres y hombres, siendo el DIF no uniforme en 3 de ellos.

En la tabla 6 se presentan los valores de los estadísticos de Wald y sus niveles de significación para los parámetros  $\tau_2$  y  $\tau_3$  en cada una de las dimensiones y facetas del NEO PI-R tomando como grupos de comparación los jóvenes (de entre 16 y 20 años) y los adultos (de entre 25 y 60 años).

Los valores obtenidos en el estadístico de Wald pusieron de manifiesto que en la dimensión Neuroticismo, 4 de los ítems pre-

sentaban DIF, siendo en todos los casos no uniforme. Con respecto a la Extraversión, la cantidad de ítems con DIF ascendió a 5, en 4 de los cuales el DIF era no uniforme. En la dimensión Apertura a la experiencia se observaron 9 ítems con funcionamiento diferencial, mostrando 2 de ellos un DIF no uniforme. En la dimensión Amabilidad la cantidad de ítems con DIF disminuyó a 4, siendo en todos los casos no uniforme. Por último, en la dimensión Responsabilidad 6 ítems mostraron un funcionamiento diferencial para jóvenes y adultos, presentando dos de ellos un DIF no uniforme.

### Discusión

El objetivo de este trabajo ha consistido en mostrar, desde una perspectiva práctica, las pautas que se deben seguir para adaptar adecuadamente un instrumento de medida de una cultura a otra, tomando como referencia el proceso de adaptación del NEO PI-R del castellano al euskera. Así, se ha tratado de explicar la forma en la que pueden llevarse a la práctica los pasos propuestos por Hambleton y Patsula (1999) para evitar las fuentes de error a las que se exponen los investigadores inmersos en un proceso de adaptación.

Centrándonos en los principales resultados obtenidos en la fase de validación empírica del instrumento objeto de adaptación, cabe señalar que las propiedades psicométricas observadas en la versión en euskera del NEO PI-R son muy similares a las que presenta la versión en castellano. Así, con respecto a su dimensionalidad, el análisis de componentes principales puso de manifiesto que, a excepción de la Asertividad, todas las facetas saturaban de forma clara en sus correspondientes dimensiones. A su vez, el análisis de la congruencia entre las soluciones factoriales de ambas versiones lingüísticas del test corroboró que, en la muestra de personas vasco-parlantes, se mantiene la estructura de cinco factores que subyace a la versión en castellano del inventario. Además, en todas las dimensiones se obtuvieron altos índices de consistencia interna, constatándose que el Neuroticismo, la Extraversión y la Responsabilidad presentan una mayor consistencia interna en la versión en euskera del instrumento que en la versión en castellano.

Por otra parte, el patrón de correlaciones observado entre las puntuaciones de las dimensiones del NEO PI-R y del BFQ proporcionó evidencias sobre la validez convergente de la versión del instrumento adaptada al euskera. Además, los procedimientos de validación diferencial utilizados corroboraron en gran medida los resultados obtenidos por otros autores. De hecho, las relaciones observadas entre las distintas dimensiones del NEO PI-R y el sexo, la edad, la tendencia política y la orientación religiosa fueron de naturaleza similar a las encontradas en estudios recientes. No obstante, cabe señalar que la magnitud de algunos de los efectos encontrados fue pequeña, concretamente en la dimensión Apertura a la experiencia, en el caso de la variable tendencia política, y en las facetas Búsqueda de emociones y Apertura a los valores, en el caso de la variable orientación religiosa.

En cuanto al análisis del funcionamiento diferencial de la versión adaptada al euskera en mujeres y hombres, cabe destacar que, a excepción de los que configuran la dimensión Neuroticismo, los ítems del inventario que presentaron tal funcionamiento fueron muy pocos. Por otra parte, cuando los grupos de comparación se establecieron en función de la variable edad, la dimensión que mostró una mayor cantidad de ítems con DIF fue la de Apertura a la experiencia. Aunque en el resto de dimensiones se produjo un ligero aumento de los ítems que presentaban DIF con respecto al caso en el que los grupos de comparación se establecieron en fun-

Dimensión	Faceta	Nº de ítem	$\tau_2$	p	$\tau_3$	p
Neuroticismo	Ansiedad social	76			8,97	0,003
	Impulsividad	231			4,22	0,040
	Vulnerabilidad	116			10,02	0,002
		146			4,39	0,036
Extraversión	Cordialidad	62			4,29	0,038
	Gregarismo	37	4,30	0,038		
	Asertividad	72			4,33	0,037
		162			8,00	0,004
	Búsqueda de emociones					
		202			4,39	0,036
Apertura a la experiencia	Fantasía	123	7,87	0,005		
	Estética	8			4,13	0,042
	Sentimientos	43	5,26	0,021		
		163			4,54	0,033
	Acciones	168	6,71	0,010		
	Ideas	23	5,46	0,019		
		203	5,75	0,016		
	Valores	58	4,03	0,044		
Amabilidad		148	7,83	0,005		
	Confianza	184			5,86	0,015
	Franqueza	39			3,90	0,048
	Altruismo	164			9,58	0,002
Responsabilidad	Actitud conciliadora	169			5,87	0,015
	Competencia	95	4,48	0,034		
		110			6,51	0,010
	Necesidad de logro	170	5,52	0,019		
		175			4,07	0,044
	Autodisciplina	235	6,19	0,013		
	Deliberación	210	4,50	0,033		

Nota: Se incluye el valor de los estadísticos para los parámetros significativos

ción del sexo, los ítems con DIF no alcanzaron el 15% del total de ítems en ninguna de las dimensiones.

A partir de los datos obtenidos en la fase de validación empírica del test, cabe concluir que la versión en euskera del NEO PI-R parece ofrecer garantías suficientes para una evaluación adecuada de la personalidad desde la perspectiva de los «Cinco Grandes». Sin embargo, es importante señalar que la muestra de individuos con la que se ha trabajado no es lo suficientemente amplia como para extraer conclusiones definitivas. Además, dado que la descripción del proceso de adaptación tenía como principal objetivo mostrar, desde una perspectiva práctica, las pautas que se deben seguir para adaptar adecuadamente un instrumento de medida de una cultura a otra,

huelga decir que estos resultados deben considerarse con cautela hasta que el proceso de adaptación llegue a culminarse.

#### Agradecimientos

Los resultados presentados en este trabajo forman parte de una investigación financiada por la Exma. Diputación Foral de Gipuzkoa y por la Unión Europea (fondos FEDER) en el marco del Programa «Red Guipuzcoana de Ciencia, Tecnología e Innovación» (OF863/2004), así como por el Vicerrectorado de Investigación de la Universidad del País Vasco (Euskal Herriko Unibertsitatea) (1) UPV 00109.231-H-15861/2004).

#### Referencias

- Altemeyer, B. (1996). *The authoritarian specter*. Cambridge, MA: Harvard University Press.
- Caprara, G.V., Barbaranelli, C., y Borgogni, L. (1995). *El Cuestionario «Big Five» (BFQ)*. Madrid: TEA.
- Caprara, G.V., Barbaranelli, C., y Zimbardo, P.G. (1999). Personality profiles and political parties. *Political Psychology*, 20, 175-197.
- Costa, P.T., Jr., y McCrae, R.R. (1992). *NEO PI-R, Profesional Manual. Revised NEO Personality Inventory (NEO PI-R) and NEO Five Factor Inventory (NEO-FFI)*. Odessa, FL.: Psychological Assessment Resources.
- Costa, P.T., Jr., y McCrae, R.R. (1995). Solid ground in the wetlands of personality: A reply to Block. *Psychological Bulletin*, 117, 216-220.
- Costa, P.T., Jr., y McCrae, R.R. (1999). *Inventario de personalidad NEO revisado (NEO PI-R) e Inventario NEO reducido de Cinco Factores (NEO-FFI). Manual profesional*. Madrid: TEA Ediciones.
- Costa, P.T., Jr., McCrae, R.R., y Johnson, F.H. (2002). Validity and utility of the Revised NEO Personality Inventory: Examples from Europe. En B. De Raad y M. Perugini (eds.): *Big Five Assessment* (pp. 61-77). Toronto: Hogrefe & Huber Publishers.
- Costa P.T., Jr., Terracciano, A., y 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 Miguel, A. (2005). Diferencias de edad y género en el NEO-PI-R en dos muestras con distinto nivel académico. *International Journal of Psychology and Psychological Therapy*, 1, 13-31.
- Digman, J.M. (1990). Personality structure: Emergence of the Five-factor model. *Annual Review of Psychology*, 41, 417-440.
- Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin*, 116, 429-456.
- Goldberg, L.R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48, 24-34.
- Gómez, J., Hidalgo, M.D., Padilla, J.L., y González, A. (2005, septiembre). *Desarrollo informático para la utilización de la Regresión Logística como técnica de detección de DIF*. Demostración informática presentada en el IX Congreso de Metodología de las Ciencias Sociales y de la Salud, Granada.
- Hambleton, R.K. (1996). Adaptación de tests para su uso en diferentes idiomas y culturas: fuentes de error, posibles soluciones y directrices prácticas. En J. Muñiz (coord.): *Psicometría*. Madrid: Universitat.
- Hambleton, R.K. (2005). Issues, designs and technical guidelines for adapting tests into multiple languages and cultures. En R.K. Hambleton, P.F. Merenda y S.D. Spielberger (eds.): *Adapting educational and psychological tests for cross-cultural assessment* (pp. 3-38). New Jersey: Lawrence Erlbaum Associates.
- Hambleton, R.K. y Patsula, L. (1999). Increasing the validity of Adapted Tests: Myths to be avoided and guidelines for improving test adaptation practices. *Journal of Applied Testing Technology*, 1. Accessible en <http://www.testpublishers.org/journal1.htm>.
- Hidalgo, M.D., Gómez, J., y Padilla, J.L. (2005). Regresión logística: alternativas de análisis en la detección del funcionamiento diferencial del ítem. *Psicothema*, 17, 509-515.
- John, O.P. (1990). The «Big Five» factor taxonomy: Dimensions of personality in the natural language and in questionnaires. En L.A. Pervin (ed.): *Handbook of Personality. Theory and research*. London: The Guilford Press.
- Lippa, R., y Arrad, S. (1999). Gender, personality, and prejudice: The display of authoritarianism and social dominance in interviews with college men and women. *Journal of Research in Personality*, 33, 463-493.
- Lynn, R., y Martin, T. (1997). Gender differences in Extraversion, Neuroticism and Psychoticism in 37 nations. *Journal of Social Psychology*, 137, 369-373.
- McCrae, R.R. (2002). NEO PI-R data from 36 cultures: Further intercultural comparisons. En R.R. McCrae y J. Allik (eds.): *The Five Factor model of personality across cultures. International and cultural psychology series* (pp. 105-125). New York, NY: Kluwer Academic/Plenum Publishers.
- McCrae, R.R., y Allik, J. (eds.). (2002). *The Five-Factor model of personality across cultures*. New York: Kluwer Academic/Plenum Publishers.
- McCrae, R.R., y Costa, P.T., Jr. (1987). Validation of a Five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52, 81-90.
- McCrae, R.R., y Costa, P.T. (1997). Personality Trait Structure as a Human Universal. *American Psychologist*, 52, 509-516.
- McCrae, R.R., y John, O.P. (1992). An introduction to the five-factor model and its applications. Special Issue: The five-factor model: Issues and applications. *Journal of Personality*, 60, 175-215.
- McCrae, R.R., Costa, P.T., Jr., Lima, M.P., Simões, A., Ostendorf, F., Angleitner, A., Marusic, I., Bratko, D., Caprara, G.V., Barbaranelli, C., Chae, J.H., y Piedmont, R.L. (1999). Age differences in personality across the adult life span: Parallels in five cultures. *Developmental Psychology*, 35, 466-477.
- McCrae, R.R., Zonderman, A.B., Costa, P.T., Jr., Bond, M.H., y Paunonen, S.V. (1996). Evaluating replicability of factors in the revised NEO personality inventory: Confirmatory Factor Analysis versus Procrustes rotation. *Journal of Personality and Social Psychology*, 70, 552-566.
- Muñiz, J. (2000). Adaptación de los tests de unas culturas a otras. *Metodología de las Ciencias del Comportamiento*, 2, 129-149.
- Muñiz, J., y Hambleton, R.K. (1996). Directrices para la traducción y adaptación de los tests. *Papeles del Psicólogo*, 66, 63-70.
- Nolen-Hoeksema, S. (1987). Sex differences in unipolar depression: Evidence and theory. *Psychological Bulletin*, 101, 259-282.
- Saroglou, V., y Fiasse, L. (2003). Birth order, personality and religion: A study among young adults from a three-sibling family. *Personality and Individual Differences*, 35, 19-29.
- Saucier, G., y Goldberg, L.R. (1998). What is beyond the Big Five? *Journal of Personality*, 66, 495-524.
- Sireci, S.G., Patsula, L., y Hambleton, R.K. (2005). Statistical methods for identifying flaws in the test adaptation process. En R.K. Hambleton, P.F. Merenda y S.D. Spielberger (eds.): *Adapting educational and psychological tests for cross-cultural assessment* (pp. 93-115). New Jersey: Lawrence Erlbaum Associates.

- Streiner, D.L., y Norman, G.R. (1989). *Health measurement scales. A practical guide to their development and use*. New York, NY: Oxford University Press.
- Streyffeler, L.L., y McNally, R.J. (1998). Fundamentalists and liberals: Personality characteristics of Protestant Christians. *Personality and Individual Differences*, 24, 579-580.
- Swaminathan, H., y Rogers, H.J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement*, 27, 361-370.
- van de Vijver, F.J.R., y Hambleton, R.K. (1996). Translating tests: Some practical guidelines. *European Psychologist*, 1, 89-99.
- van de Vijver, F.J.R., y Poortinga, Y.H. (2005). Conceptual and methodological issues in adapting tests. En R.K. Hambleton, P.F. Merenda y S.D. Spielberger (eds.): *Adapting educational and psychological tests for cross-cultural assessment* (pp. 39-63). New Jersey: Lawrence Erlbaum Associates.
- Van Hiel, A., y Mervielde, I. (2002). Explaining conservative beliefs and political preferences: a company of social dominance orientation and authoritarianism. *Journal of Applied Social Psychology*, 32, 965-976.
- Wiggins, J.S. (1996). *The Five-Factor Model of Personality. Theoretical Perspectives*. New York: The Guilford Press.
- Yang, J., McCrae, R.R., y Costa P.T., Jr. (1998). Adult age differences in personality traits in the United States and the People's Republic of China. *Journal of Gerontology: Psychological Sciences*, 53B, 375-383.