

eman ta zabal zazu



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

**ADAPTACIÓN Y VALIDACIÓN DE
INSTRUMENTOS PARA EVALUAR LA
PERSONALIDAD Y EL DESEMPEÑO
LABORAL**

Tesis doctoral presentada por:

Alexander Rodríguez López

Octubre 2020

**UNIVERSIDAD DEL PAÍS VASCO
/ EUSKAL HERRIKO UNIBERTSITATEA**

Facultad de Psicología

**Departamento de Psicología Clínica y de la Salud y
Metodología de Investigación**

**Programa de doctorado:
Psicología**

**Adaptación y validación de instrumentos para
evaluar la personalidad y el desempeño laboral**

Tesis doctoral presentada por:
Alexander Rodríguez López

Dirigida por:
Nekane Balluerka Lasa
Arantxa Gorostiaga Manterola

Laudio, Octubre 2020

A mi madre, por tu esfuerzo para darme las oportunidades que
nunca tuviste.

A Patricia, por estar siempre ahí y por tu paciencia.

A Mikel, por tu ejemplo.

Arantxari, bidean gidatu izanagatik.

*«Avoir une autre langue, c'est posséder une deuxième âme» –
Carlomagno*

*«Enjoying success requires the ability to adapt. Only by being
open to change will you have a true opportunity to get the most
from your talent» – Nolan Ryan*

«Põe quanto és no mínimo que fazes» – Fernando Pessoa

Agradecimientos

A las Dras. Nekane Balluerka y Arantxa Gorostiaga, por su rigor y su exigencia, que me han hecho aprender tantísimo.

A la Dirección del IVAP, por hacer posible este proyecto.

A los y las responsables de las instituciones públicas que han aceptado participar, así como a las personas voluntarias que han participado en la cumplimentación de los cuestionarios, por su apoyo inestimable.

Gracias a todos/as.

LISTA DE ACRÓNIMOS EMPLEADOS

AFC: Análisis Factorial Confirmatorio

AFE: Análisis Factorial Exploratorio

APO: Ajuste Persona-Organización

CAPV: Comunidad Autónoma del País Vasco

CCI: Curva Característica del Ítem

DA: Desempeño Adaptativo

ESEM: Exploratory Structural Equation Model

FDI: Funcionamiento Diferencial del Ítem

FFM: Five-Factor Model

FFT: Five-Factor Theory

ITC: International Test Commission

IVAP: Instituto Vasco de Administración Pública

OCB: Organizational Citizenship Behaviors

OCB-I: OCBs de tipo Interpersonal

OCB-O: OCBs de tipo Organizacional

TRI: Teoría de Respuesta al Ítem

LENGUAJE DE GÉNERO

En la redacción del texto de la tesis se ha intentado reflejar la voluntad de no discriminación y de uso de un lenguaje igualitario.

A tal fin, en general, se ha acudido a las siguientes alternativas, valorando siempre el contexto:

a) uso de términos colectivos, de genéricos, de construcciones impersonales o de aquellas que eviten formas marcadas solo como masculinas o solo como femeninas.

b) desdoblamientos, preferiblemente coordinados con y o con o.

En cuanto a las referencias hechas en plural a colectivos que pueden ser constituidos tanto por hombres como por mujeres, se entiende que el masculino tiene valor genérico, tal como hemos dicho más arriba, salvo que se especifique expresamente lo contrario (o se deduzca de forma inequívoca por el contexto). En estos casos, por tanto, habitualmente no se utilizan las formas dobles, porque, por un lado, dificultan la lectura del texto y lo cargan y, por otro lado, pueden llegar a crear contrasentidos si no se utilizan con coherencia.

De cualquier forma, y aunque se han seguido las recomendaciones de la Guía para el uso no sexista de la lengua castellana y de imágenes en la Universidad del País Vasco-Euskal Herriko Unibertsitatea (UPV/EHU), pedimos disculpas de antemano si alguna persona se siente incómoda o injustamente tratada por su condición femenina o no binaria.

ÍNDICE DE CONTENIDOS

Capítulo 1. Introducción: justificación y estructura del estudio. La adaptación de instrumentos de evaluación.....	25
1.1. Justificación y estructura del estudio.....	25
1.2. La adaptación de instrumentos de evaluación.....	29
1.2.1. La adaptación es mucho más que la traducción.....	30
1.2.2. Aspectos a tener en cuenta en la traducción.....	31
1.2.3. Métodos para establecer equivalencias de traducción.....	34
1.2.4. Equivalencia métrica a nivel de ítem y a nivel de test o escala.....	36
1.2.5. Sesgos más frecuentes y estrategias recomendadas.....	38
1.2.6. Directrices para la adaptación de instrumentos psicológicos.....	41
1.2.7. La adaptación de instrumentos en el ámbito organizacional.....	46
1.2.8. Particularidades del euskera para la adaptación de instrumentos...	49
Capítulo 2. Adaptación al castellano y validación de la escala de desempeño adaptativo de Charbonnier-Voirin y Roussel.....	53
2.1. El desempeño laboral.....	54
2.1.1. Naturaleza y taxonomía “clásica” del desempeño laboral.....	54
2.1.2. La inclusión del desempeño adaptativo en el marco del desempeño laboral.....	58
2.1.3. Definiciones y taxonomías del constructo desempeño adaptativo...	60
2.1.4. Conceptualización del desempeño adaptativo.....	66
2.1.5. Medición del desempeño.....	67
2.2. Revisión de instrumentos que evalúan el desempeño adaptativo.....	69
2.3. Adaptación del Cuestionario de desempeño adaptativo.....	72
2.3.1. Proceso de traducción de los ítems, entrevistas cognitivas y fase piloto.....	72
2.3.2. Método.....	74
2.3.2.1. Participantes.....	74
2.3.2.2. Instrumentos.....	76
2.3.2.3. Procedimiento.....	76

2.3.2.4. Análisis de datos.....	77
2.3.3. Resultados.....	82
2.3.3.1. Análisis de los ítems del instrumento.....	82
2.3.3.2. Dimensionalidad y validez de constructo: modelo confirmatorio, ESEM, y estructura interna.....	89
2.3.3.3. Fiabilidad.....	92
2.3.3.4. Funcionamiento diferencial de los ítems según lengua materna y complejidad del puesto.....	97
2.3.4. Conclusiones.....	99
Capítulo 3. Adaptación al euskera y validación del cuestionario de personalidad OPERAS.....	103
3.1. Revisión teórica del concepto de personalidad.....	103
3.1.1. Taxonomía de la personalidad: el modelo de los Cinco Grandes....	104
3.1.2. Origen y generalización transcultural del modelo de los Cinco Grandes.....	107
3.1.3. La medición de la personalidad: estado del arte y principales problemas.....	110
3.1.4. Contextualización de los ítems de los cuestionarios de personalidad.....	112
3.1.5. Ventajas y desventajas de las escalas breves de personalidad.....	114
3.1.6. La distorsión de respuestas en los cuestionarios de personalidad....	116
3.2. Revisión de los cuestionarios de personalidad más empleados en España en el ámbito de los Recursos Humanos.....	119
3.3. Adaptación de OPERAS.....	127
3.3.1. Proceso de traducción de los ítems, entrevistas cognitivas y fase piloto.....	127
3.3.2. Método.....	128
3.3.2.1. Participantes.....	128
3.3.2.2. Instrumentos.....	129
3.3.2.3. Procedimiento.....	129
3.3.2.4. Análisis de datos.....	130
3.3.3. Resultados.....	133
3.3.3.1. Análisis de los ítems del instrumento.....	133

3.3.3.2. Dimensionalidad del instrumento y validez de constructo: AFE, AFC, ESEM.....	141
3.3.3.3. Validez convergente: asociación entre las dimensiones de OPERAS y del BFI.....	145
3.3.3.4. Fiabilidad.....	145
3.3.3.5. Funcionamiento diferencial de los ítems según la lengua materna de los y las participantes.....	150
3.3.4. Conclusiones.....	152
Capítulo 4. Relaciones entre los cinco grandes factores de personalidad, los constructos vinculados con el desempeño laboral, el ajuste persona-organización y las variables sociodemográficas y laborales.....	157
4.1. Introducción.....	157
4.2. Red de relaciones nomológicas de la escala de Desempeño adaptativo de Charbonnier-Voirin y Roussel en castellano.....	158
4.2.1. El modelo I-ADAPT de Ployhart y Bliese (2006) y la conceptualización del desempeño adaptativo.....	159
4.2.2. La relación entre desempeño adaptativo y conductas cívico-organizacionales.....	161
4.2.3. El ajuste persona-organización (APO) y el desempeño laboral.....	164
4.3. Red nomológica de OPERAS en euskera: la relación entre personalidad y desempeño laboral.....	167
4.3.1. La personalidad y el desempeño adaptativo.....	169
4.3.2. La personalidad y las conductas cívico organizacionales (OCBs)...	170
4.3.3. La personalidad y el ajuste persona-organización (APO).....	171
4.3.4. El modelo I-ADAPT y la relación entre personalidad y desempeño laboral. El papel del ajuste persona-organización y del desempeño adaptativo...	171
4.4. Las variables sociodemográficas y laborales.....	173
4.4.1. La relación del género con los constructos analizados.....	174
4.4.2. La relación de la edad con los constructos analizados.....	175
4.4.3. La relación de la experiencia laboral con los constructos analizados.....	177
4.4.4. La relación del desempeño de puestos de mando con los constructos analizados.....	178

4.4.5. La relación entre nivel educativo, nivel de complejidad del puesto y los constructos analizados.....	180
4.4.6. La relación entre la estabilidad laboral y los constructos analizados.....	181
4.5. Objetivos e hipótesis.....	182
4.5.1. Objetivos.....	182
4.5.2. Hipótesis.....	182
4.5.2.1. Relación entre las dimensiones de desempeño adaptativo y las OCBs.....	182
4.5.2.2. Relación entre las dimensiones de desempeño adaptativo y el APO.....	183
4.5.2.3. Relación entre factores de personalidad y dimensiones de desempeño adaptativo.....	183
4.5.2.4. Relación entre factores de personalidad y dimensiones de las conductas cívico-organizacionales (OCBs).....	183
4.5.2.5. Relación entre factores de personalidad y el ajuste persona-organización (APO).....	184
4.5.2.6. Relación entre la personalidad y el desempeño: el rol del APO y del desempeño adaptativo.....	184
4.5.2.7. Relación entre el género y los constructos analizados.....	185
4.5.2.8. Relación entre la edad y los constructos analizados.....	185
4.5.2.9. Relación entre la experiencia laboral y los constructos analizados.....	185
4.5.2.10. Relación entre el ejercicio de puestos de mando y los constructos analizados.....	186
4.5.2.11. Relación entre el nivel educativo y la complejidad del puesto y los constructos analizados.....	186
4.5.2.12. Relación entre la estabilidad laboral y los constructos analizados.....	186
4.6. Metodología.....	187
4.6.1. Participantes.....	187
4.6.2. Instrumentos.....	188
4.6.3. Procedimiento.....	190
4.6.4. Análisis de datos.....	190

4.7. Resultados.....	192
4.7.1. Varianza del método común.....	192
4.7.2. Estadísticos descriptivos y fiabilidad de los constructos analizados	192
4.7.3. Correlaciones entre los constructos analizados.....	193
4.7.4. Modelos de regresiones múltiples jerárquicas.....	195
4.7.5. Relación entre los constructos evaluados y las variables sociodemográficas y laborales.....	198
4.8. Conclusiones.....	207
Capítulo 5. Conclusiones generales.....	219
Referencias.....	231

ÍNDICE DE TABLAS Y FIGURAS

Tablas

Tabla 1.1. Estrategias ante los Tipos de Sesgos (adaptado de Van de Vijver y Tanzer, 2004).....	40
Tabla 1.2. Comparativa de Directrices sobre Adaptación de Instrumentos Psicométricos.....	44
Tabla 1.3. Recomendaciones para la Adaptación de Escalas (adaptado de Heggstad et al., 2019).....	47
Tabla 2.1. Taxonomías “Clásicas” del Desempeño de tarea (adaptado de Gorriti y López, 2009).....	56
Tabla 2.2. Taxonomías del Desempeño Laboral que Incluyen el Desempeño Adaptativo.....	59
Tabla 2.3. Comparación de enfoques de la Adaptabilidad (adaptado de Murphy, 2015).....	60
Tabla 2.4. Definiciones del Desempeño Adaptativo.....	62
Tabla 2.5. Dimensiones y Definiciones de la Taxonomía de Pulakos et al. (2000).....	63
Tabla 2.6. Dimensiones y Definiciones de la Taxonomía de Charbonnier-Voirin y Roussel (2012).....	65
Tabla 2.7. Información Sociodemográfica de las Personas Participantes.....	75
Tabla 2.8. Estadísticos Descriptivos de los Ítems de la Escala de Charbonnier-Voirin y Roussel e Índice de Homogeneidad.....	82
Tabla 2.9. Porcentajes de Elección de las Opciones de Respuesta en cada Ítem de la Escala de Charbonnier-Voirin y Roussel.....	83
Tabla 2.10. Parámetros Estimados según el Modelo de Respuesta Graduada de los ítems de la Escala de Charbonnier-Voirin y Roussel.....	85
Tabla 2.11. Saturaciones Factoriales y Porcentaje de Varianza de los Factores del Modelo ESEM de la Escala de Charbonnier-Voirin y Roussel.....	90
Tabla 2.12. Correlaciones entre los Factores de la Escala de Charbonnier-Voirin y Roussel.....	91
Tabla 2.13. Correlación entre los Ítems de la Escala de Charbonnier-Voirin y Roussel y su Dimensión Correspondiente.....	91
Tabla 2.14. Intercorrelaciones entre los Ítems de la Escala de Charbonnier-Voirin y Roussel.....	92
Tabla 2.15. Fiabilidad (Consistencia interna) de las dimensiones de la Escala de Charbonnier-Voirin y Roussel.....	93
Tabla 2.16. Fiabilidad Test-retest de las dimensiones de la Escala de Charbonnier-Voirin y Roussel.....	96

Tabla 2.17. Funcionamiento Diferencial de los Ítems de la Escala de Charbonnier-Voirin y Roussel en base a la Complejidad del Puesto Desempeñado.....	97
Tabla 2.18. Funcionamiento Diferencial de los Ítems de la Escala de Charbonnier-Voirin y Roussel en base a la Lengua Materna de las Personas Participantes.....	98
Tabla 3.1. Definición de los 5 grandes Factores de Personalidad.....	105
Tabla 3.2. Constructos, Tipo de Escala, Fiabilidad y Validez de los Cuestionarios de Personalidad.....	123
Tabla 3.3. Editorial, Ítems Contextualizados o no, Número de Ítems, Tiempo de Aplicación y Medidas contra la Deseabilidad Social de los Cuestionarios de Personalidad.....	126
Tabla 3.4. Estadísticos Descriptivos e Índice de Homogeneidad de los Ítems de OPERAS	134
Tabla 3.5. Porcentajes de Elección de las Opciones de Respuesta en cada Ítem de OPERAS.....	135
Tabla 3.6. Parámetros Estimados según el Modelo de Respuesta Graduada de los ítems de OPERAS.....	137
Tabla 3.7. Saturaciones de los Ítems Corregidas por DS y AC en los Factores de OPERAS.....	142
Tabla 3.8. Correlaciones entre los Factores de OPERAS.....	144
Tabla 3.9. Correlación entre los Ítems de OPERAS y su Factor Correspondiente.....	144
Tabla 3.10. Correlaciones entre las Dimensiones de OPERAS y BFI.....	145
Tabla 3.11. Fiabilidad (Consistencia Interna) de los factores de OPERAS.....	146
Tabla 3.12. Fiabilidad test-retest de los factores de OPERAS.....	149
Tabla 3.13. Funcionamiento Diferencial de los Ítems de OPERAS en base a la Lengua Materna de las Personas Participantes.....	151
Tabla 4.1. Estadísticos Descriptivos.....	193
Tabla 4.2. Correlaciones entre los Constructos Estudiados.....	195
Tabla 4.3. Regresión Múltiple Jerárquica con OCB-O como Variable Criterio.	196
Tabla 4.4. Regresión Múltiple Jerárquica con OCB-I como Variable Criterio.....	196
Tabla 4.5. Regresión Múltiple Jerárquica con Creatividad como Variable Criterio.....	197
Tabla 4.6. Regresión Múltiple Jerárquica con Gestión del Estrés como Variable Criterio.....	197
Tabla 4.7. Regresión Múltiple Jerárquica con Adaptabilidad Interpersonal como Variable Criterio.....	197
Tabla 4.8. Regresión Múltiple Jerárquica con Reactividad ante Emergencias como Variable Criterio.....	197
Tabla 4.9. Regresión Múltiple Jerárquica con Esfuerzo Formativo como Variable Criterio.....	198
Tabla 4.10. Puntuaciones Medias en función de las Variables Sociodemográficas y Laborales.....	199

Tabla 4.11. Puntuaciones Medias en los Constructos Analizados según los Grupos de Edad.....	200
Tabla 4.12. Resumen de las Diferencias de Medias Estadísticamente Significativas en las Puntuaciones según Variables Sociodemográficas y Laborales.....	202
Tabla 4.13. Resumen de las Hipótesis Formuladas y Contraste desde el Punto de Vista del Tamaño del Efecto.....	203

Figuras

Figura 2.1. Curvas Características del Ítem para los Ítems que Miden Gestión del estrés.....	86
Figura 2.2. Curvas Características del Ítem para los Ítems que Miden Esfuerzo Formativo.....	87
Figura 2.3. Curvas Características del Ítem para los Ítems que Miden Adaptabilidad.....	87
Figura 2.4. Curvas Características del Ítem para los Ítems que Miden Reactividad.....	88
Figura 2.5. Curvas Características del Ítem para los Ítems que Miden Creatividad....	88
Figura 2.6. Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para la Dimensión Gestión del Estrés.....	94
Figura 2.7. Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para la Dimensión Esfuerzo Formativo.....	94
Figura 2.8. Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para la Dimensión Adaptabilidad Interpersonal.....	95
Figura 2.9. Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para la Dimensión Reactividad ante Emergencias.....	95
Figura 2.10. Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para la Dimensión Creatividad.....	96
Figura 3.1. Curvas Características del Ítem para los Ítems que Miden Amabilidad.....	138
Figura 3.2. Curvas Características del Ítem para los Ítems que Miden Apertura a la experiencia.....	138
Figura 3.3. Curvas Características del Ítem para los Ítems que Miden Estabilidad Emocional.....	139
Figura 3.4. Curvas Características del Ítem para los Ítems que Miden Extroversión...	139
Figura 3.5. Curvas Características del Ítem para los Ítems que Miden Responsabilidad.....	140
Figura 3.6. Curvas Características del Ítem para los Ítems que Miden Deseabilidad Social.....	140
Figura 3.7. Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para el Factor Amabilidad.....	146
Figura 3.8. Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para el Factor Apertura a la experiencia.....	147
Figura 3.9. Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para el Factor Estabilidad Emocional.....	147
Figura 3.10. Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para el Factor Extroversión.....	148
Figura 3.11. Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para el Factor Responsabilidad.....	148
Figura 3.12. Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para el Factor Deseabilidad social.....	149
Figura 4.1. Teoría de la Adaptabilidad Individual (I-ADAPT). Adaptado de Ployhart y Bliese (2006).....	161

CAPÍTULO 1.

INTRODUCCIÓN: JUSTIFICACIÓN Y ESTRUCTURA DEL ESTUDIO. LA ADAPTACIÓN DE INSTRUMENTOS DE EVALUACIÓN.

1.1. Justificación y estructura del estudio

El presente estudio responde a la necesidad de disponer de instrumentos que permitan evaluar de forma rigurosa la personalidad y el desempeño laboral en el ámbito organizacional. En nuestro contexto bilingüe son escasos los instrumentos adaptados a las dos lenguas oficiales que permitan llevar a cabo, de forma adecuada, procesos de evaluación y selección de personal. De hecho, actualmente únicamente los cuestionarios de personalidad NEO-PI-R (Balluerka et al., 2007) y su versión breve NEO-FFI (Haranburu et al., 2007) están adaptados al euskera. Asimismo, es necesario considerar que el Instituto Vasco de Administración Pública (IVAP) lleva a cabo más de 30 procesos selectivos anuales en los que se emplean pruebas de personalidad. En dichos procesos selectivos un número cada vez mayor de personas demandan poder realizar este tipo de pruebas en euskera y la legislación insta a garantizar los derechos lingüísticos de la ciudadanía en las oposiciones. Por todo ello resulta imprescindible adaptar más pruebas al euskera.

Por otra parte, hay una gran escasez de instrumentos que miden el *desempeño adaptativo*, una dimensión del desempeño laboral, en castellano. Como resultado, aún no se han podido estudiar en dicha lengua y cultura las relaciones de ese constructo y las dimensiones que lo componen con otra serie de variables relevantes en el contexto laboral, tales como la personalidad, las *conductas cívico-organizacionales* (organizational citizenship behaviors, OCB, en inglés) o el *ajuste persona-organización* (en adelante, APO).

La relación entre la personalidad y variables relevantes en el ámbito organizacional, tales como el desempeño laboral o el APO constituye un tema de gran interés tanto a nivel teórico como a nivel práctico. La literatura muestra que los factores de personalidad son predictores de aspectos importantes para las organizaciones, como el trabajo en equipo, las conductas contraproductivas, la orientación al cliente, las OCBs, la satisfacción laboral, el compromiso con la organización y la seguridad laboral (Barrick et al., 2011; Barrick y Mount, 2014; Ones et al., 2005). Por otra parte, como

indican Ryan y Ployhart (2014), dada la importancia del desempeño adaptativo en el entorno laboral actual, resulta sorprendente la escasez de trabajos que estudien su relación con variables utilizadas en la selección de personal. Estos autores también han remarcado la necesidad de realizar trabajos en los que se profundice en el APO en el contexto de selección.

En lo que se refiere a la relación entre el modelo de cinco factores de personalidad (Five factor model o FFM, por sus siglas en inglés) y las OCBs, son dos los factores, a saber, Responsabilidad y Amabilidad, que han recibido mayor atención en cuanto a su vinculación con las OCBs (Bowling y Burns, 2010; Chiaburu et al. 2011). Así, diferentes estudios (Cooper et al. 2013; Ilies et al. 2009) coinciden en señalar que los mencionados factores se relacionan de forma positiva con dos dimensiones de los OCBs, comportamientos dirigidos a la organización (OCB-O) y comportamientos de tipo interpersonal (OCB-I), respectivamente. Por otro lado, los hallazgos relativos a los factores Neuroticismo y Extroversión no han sido consistentes, y la relación entre Apertura mental y las OCBs se ha mostrado débil (Hoffmann y Dilchert, 2012).

Menor atención ha recibido la relación entre los cinco factores de personalidad y el desempeño adaptativo, siendo el estudio llevado a cabo por Charbonnier-Voirin (2013), uno de los escasos trabajos que han explorado esta cuestión. Sus resultados sugieren que cada factor de personalidad presenta una correlación de magnitud diferente con las dimensiones de desempeño adaptativo y muestran la importancia de los factores Apertura mental y Responsabilidad. No obstante, como señalan Pulakos et al. (2012), se necesita profundizar en mayor medida en la relación entre los factores de personalidad y las diferentes dimensiones del desempeño adaptativo.

A su vez, también son escasos los trabajos que han centrado su atención en la relación entre los factores de personalidad y el APO. Gillespie et al. (2003), encontraron relaciones positivas entre Extroversión y Responsabilidad y APO. Tsai et al. (2012) llegaron a conclusiones similares, encontrando además una relación positiva entre Amabilidad y APO.

Por otro lado, la relación que pueden presentar variables sociodemográficas y laborales como la edad, el género, el nivel educativo o la experiencia laboral con variables como la personalidad, el APO o el desempeño laboral es un aspecto de

especial relevancia por sus implicaciones a nivel de selección, gestión de recursos humanos y evaluación del desempeño.

Una cuestión frecuentemente estudiada en selección de personal es el *impacto adverso*, que se produce cuando grupos definidos (culturales, de género, de edad, raciales) presentan puntuaciones medias diferentes en instrumentos selectivos. Dichas diferencias en las medias generan un riesgo potencial de ratios de contratación diferentes para colectivos protegidos y pueden implicar discriminación indirecta (Aramburu-Zabala, 2005; Guion, 2011; Hough et al., 2001). En cuanto a las pruebas de personalidad, existe un consenso generalizado sobre la ausencia de impacto adverso en función de género, edad o raza (Barrick y Mount, 2014), cuestión vital para garantizar el cumplimiento del principio de igualdad en el acceso al empleo público exigido por la legislación.

Este consenso no se da en las pruebas que miden las OCBs. Existen resultados discrepantes acerca de la relación entre variables como el género, la edad, los años de experiencia en el puesto y las puntuaciones en OCB. Así, Akinbode (2011) encontró que la edad, el género y la experiencia laboral no son predictores válidos de las puntuaciones en OCB, pero Chou y Pearson (2011) encontraron que la edad y la experiencia laboral sí presentaban correlaciones significativas con las OCBs. Allen y Rush (2001), por su parte, señalaron que era más probable que los varones realizaran OCBs con el fin de aumentar sus posibilidades de promoción mientras que Dávila et al. (2011) encontraron que las mujeres ponen en práctica con más frecuencia OCBs de tipo interpersonal. No obstante, el meta-análisis de Ng et al. (2016) no ha hallado diferencias de género en las puntuaciones en OCB.

En lo que se refiere al desempeño adaptativo, tal y como han señalado Jundt et al. (2015), se han realizado pocos estudios. Un ejemplo sería el de O'Connell et al. (2008), en el que se encontraron niveles diferentes de adaptabilidad según el género o el nivel educativo, pero no así en función de la edad.

Del mismo modo, se ha prestado escasa atención a la relación entre variables como el género, la edad o los años de experiencia laboral y el APO. Tsai et al. (2012), o Xu (2014) no han encontrado relación entre estas variables o el nivel educativo y el APO, si bien Su et al. (2015), en su revisión sobre el APO indican que un mayor ajuste está vinculado a un incremento en la permanencia en un mismo puesto.

Por lo tanto, teniendo en cuenta los interrogantes planteados por investigaciones previas y la necesidad de disponer de más pruebas adaptadas a nuestro contexto cultural, se plantea el siguiente objetivo general del estudio: proporcionar instrumentos que permitan una evaluación rigurosa de los factores de personalidad y el desempeño laboral en el ámbito organizacional. Con ello, se pretende contribuir a preservar los derechos lingüísticos de las personas en los procesos selectivos y a incrementar el bagaje de conocimientos existentes en el ámbito de la Psicología de las organizaciones. Todo ello con el fin de satisfacer los dos valores principales de la selección de personal: la eficacia y la igualdad. La eficacia supone seleccionar a las personas más adecuadas para desempeñar los puestos de trabajo, mientras que la igualdad implica que personas con las mismas características o niveles en una variable tengan las mismas probabilidades de ser seleccionadas (Gorriti, 2010, 2018).

Para responder a este objetivo general, se han establecido cuatro objetivos específicos:

1. Adaptar a la lengua y cultura castellana el cuestionario de desempeño adaptativo de Charbonnier-Voirin y Roussel.
2. Adaptar a la lengua y cultura vasca la prueba OPERAS, basada en el modelo Big Five de personalidad.
3. Analizar la relación de los cinco grandes factores de personalidad con constructos vinculados con el desempeño laboral (desempeño adaptativo y OCBs) y el APO. Para ello se examinarán las redes de relaciones nomológicas de las adaptaciones tanto del cuestionario de Charbonnier-Voirin y Roussel como de la prueba OPERAS.
4. Analizar la posible relación existente entre los constructos de interés en materia de psicología organizacional (personalidad, desempeño adaptativo, OCBs y APO) y un conjunto de variables sociodemográficas y laborales.

Con el propósito de responder a los mencionados objetivos, la presente tesis se estructura de la siguiente manera. En el capítulo que nos ocupa, se justifica la relevancia del estudio realizado, y además se resumen los aspectos más relevantes acerca de la metodología para la adaptación y validación de instrumentos psicométricos. Seguidamente, el segundo capítulo se centra en el objetivo específico 1, la adaptación y

validación de la escala de desempeño adaptativo de Charbonnier-Voirin y Roussel (2012). Para contextualizar este proceso de adaptación se aporta una revisión teórica del concepto de desempeño adaptativo y de los instrumentos que evalúan dicho constructo y se describen los principales resultados obtenidos en diversos estudios. A continuación, en el tercer capítulo se describe la adaptación y validación del cuestionario de personalidad OPERAS (Vigil-Colet et al., 2013), que corresponde al segundo objetivo específico. Adicionalmente, se revisa el constructo de la personalidad centrado en el modelo de los cinco grandes factores de personalidad, así como los instrumentos disponibles para su medición en castellano. El capítulo 4, por su parte, se centra en los objetivos específicos 3 y 4 de la tesis. En primer lugar y para cumplir con el tercer objetivo, se describen las redes de relaciones nomológicas de los instrumentos adaptados en el marco de la tesis. De esta manera, se analizan las relaciones de los cinco grandes factores de personalidad con constructos vinculados con el desempeño laboral (OCBs y desempeño adaptativo) y con el ajuste persona-organización en aras a contribuir a mejorar la eficacia en la selección de personal. En segundo lugar, en el cuarto capítulo se hace hincapié en cuestiones vinculadas a la igualdad en la selección. A tal fin se analizan las relaciones de variables sociodemográficas y laborales, como la edad, el género, la complejidad de los puestos, el nivel de titulación o el desempeño de puestos de mando con los constructos mencionados (personalidad, desempeño adaptativo, OCBs y APO). Por último, en el capítulo 5 se presentan las principales conclusiones del estudio.

1.2. La adaptación de instrumentos de evaluación

Los tests psicométricos se utilizan cada vez en mayor medida para evaluar las aptitudes cognitivas, las características de personalidad o los estilos de aprendizaje, lo que a su vez está provocando un gran auge de la adaptación de tests a diferentes idiomas o culturas (Valdivia, 2014). Esta tendencia se corresponde con los intereses de las organizaciones, que intentan mejorar continuamente la calidad de sus procedimientos de selección. Por ello, están optando cada vez más por la adaptación de tests previamente validados en otros idiomas y culturas (Hambleton y De Jong, 2003; Hambleton y Patsula, 1999; Valentine, 2013).

Asimismo, existen varias razones por las que las pruebas empleadas en la evaluación de personas en general, y en la selección de personal en las organizaciones en particular, deben ser ofrecidas en varios idiomas en sociedades plurilingües. En primer lugar, es una manera de mejorar la equidad permitiendo a los candidatos que se examinen en su idioma materno o preferido. Puede haber sesgos en las puntuaciones de los tests cuando las personas los realizan en su segunda o tercera lengua, provocados por diferencias en cómo se operacionaliza el constructo, cómo se administra la prueba y/o por cómo se traducen los elementos. Sin embargo, si la prueba es realizada en la lengua principal del examinando, cabe esperar que la validez predictiva de los resultados de las pruebas aumente en comparación con las pruebas realizadas en la segunda o tercera lengua. En segundo lugar, la adaptación de un test a diferentes idiomas facilita la investigación transcultural que identifica diferencias y similitudes entre los grupos lingüísticos y culturales. En tercer lugar, reduce el coste de las pruebas. La creación de una nueva prueba para una población diferente suele ser más costosa que la adaptación de una prueba de un idioma a otro, aunque este proceso también exige un esfuerzo considerable. Además, una prueba adaptada garantiza una adecuada evaluación, del mismo modo que un test arraigado y validado (Hambleton y Kanjee, 1993, 1995; Valentine, 2013).

En el caso del País Vasco, la legislación reconoce a todos los ciudadanos el derecho a usar tanto el euskera como el castellano en sus relaciones con la Administración Pública en el ámbito territorial de la Comunidad Autónoma, y a ser atendidos en la lengua oficial que elijan. Por ello, las Administraciones Públicas Vascas deben garantizar que las pruebas selectivas, y, entre ellas las pruebas de evaluación psicométrica, se puedan realizar en ambas lenguas (Parlamento Vasco, 1982, 1989). Lamentablemente, la mencionada escasez de pruebas de personalidad validadas en euskera conlleva que los derechos de la ciudadanía no puedan respetarse de manera eficaz, lo que hace necesario adaptar más pruebas a la lengua vasca.

1.2.1. La adaptación es mucho más que la traducción

La adaptación debe ir más allá de la mera traducción del test, ya que la validez de las pruebas psicológicas y su correcta aplicación dependen de sus propiedades psicométricas. Aunque un equipo de traducción intente asegurar la consistencia en el significado del contenido de dos versiones de un test, la mera traducción ha sido cuestionada por especialistas como Hambleton y Patsula (1999) o Van de Vijver y

Hambleton (1996), que argumentan que una buena traducción no es garantía suficiente para crear un test válido en otro idioma o para poder comparar los resultados de la misma prueba en dos versiones idiomáticas distintas.

Por ello, el método más empleado para reproducir un test elaborado en un idioma a otra lengua y cultura es la adaptación, que es un proceso riguroso en el que la traducción es solo uno de los pasos necesarios para lograr una transformación con garantías. Dicho proceso lingüístico y psicométrico estandarizado incluye otros pasos como la evaluación de la equivalencia de constructo, la selección de los miembros de un equipo de traducción, realizar los ajustes para que el test pueda ser utilizado en el segundo idioma o el análisis de la equivalencia global a fin de que la versión adaptada permita medir los rasgos de los hablantes de esa lengua y cultura sin ningún tipo de sesgo. Su objetivo último es la elaboración de baremos normativos que pueden ser tomados como referencia para medir sin discriminaciones las características de cualquier hablante de una población y para la comparación con los datos del conjunto de miembros de la población correspondiente. (Hambleton, 1996; Muñiz y Hambleton, 1996).

1.2.2. Aspectos a tener en cuenta en la traducción

La traducción de un test presenta una serie de dificultades, especialmente en el caso de pruebas de personalidad, ya que, como señalan Hendriks et al. (2003), algunas lenguas no poseen traducciones precisas de las palabras usadas para describir la personalidad en otras lenguas y culturas.

Consecuentemente, la mera familiaridad con los idiomas origen y objetivo no garantiza que el test sea traducido de manera adecuada. El proceso de adaptación debe tener en cuenta las diferencias lingüísticas y culturales entre las poblaciones origen y objetivo, por lo que hay que hacer énfasis no tanto en traducciones literales como en esmeradas adaptaciones de significados. De ahí que sea necesario que los miembros del equipo de traducción reciban formación para asegurar que se traduzca tanto el significado de las palabras como el de las frases, algo complicado en ciertas expresiones idiomáticas. Deben, además, conocer la materia objeto de estudio –el estudio de la personalidad o el desempeño laboral, en este caso- y sus tecnicismos, así como las reglas de redacción de ítems, de modo que no las violen en la traducción. No en vano, como señala Matthews-López (2003), el proceso de adaptación de un test es

prácticamente idéntico al de la construcción de un instrumento psicométrico. La diferencia estriba en que en la adaptación los constructos a evaluar ya han sido definidos y, al menos para una cultura o lengua, existen los instrumentos adecuados para su evaluación y estudio.

Los ítems son el elemento básico de un instrumento de evaluación, por lo que si se construyen o adaptan de manera deficiente se comprometerán las propiedades psicométricas del instrumento, así como la validez de las inferencias que se hagan a partir de las puntuaciones. Por ello, es fundamental que en la redacción de los ítems se cuide la claridad y la sencillez, evitando tecnicismos, negaciones, dobles negaciones o enunciados ambiguos. Asimismo, se debe asegurar que los ítems sean comprensibles para la población a la que se dirige el instrumento, empleando un lenguaje inclusivo y evitando términos ofensivos y/o discriminatorios. Ítems mal redactados o excesivamente vagos pueden incrementar la varianza explicada por factores irrelevantes, afectando negativamente a las evidencias de validez del instrumento o a la relación entre las puntuaciones del instrumento adaptado y constructos externos (Heggstad et al., 2019; Muñiz y Fonseca-Pedrero, 2019). En lo que se refiere a la redacción de ítems de personalidad, Ferrando y Demestre (2008) señalan que para aumentar su capacidad discriminativa debe prestarse atención a su indicatividad (el grado en que el ítem describe aspectos esenciales del rasgo medido) y se debe procurar que los enunciados sean cortos y directos.

Las recomendaciones en materia de redacción de ítems de Tornimbeni et al. (2004) por su parte, son elaborar ítems congruentes con el objetivo de medición y evitar los ítems demasiado largos (de más de 20 vocablos), las oraciones complejas con ambigüedades de sentido, el uso de expresiones extremas (nunca, siempre, todos) y utilizar el lenguaje más apropiado al nivel de maduración y nivel educativo de la población destino del test.

Brislin, autor de uno de los artículos más citados en materia de adaptación de instrumentos de evaluación, *The wording and translation of research instruments* (1986), estableció diversas directrices para la redacción de ítems aplicables a la traducción de un instrumento:

- Utilizar frases cortas y sencillas y evitar palabras innecesarias (a menos que la redundancia se busque deliberadamente).

- Emplear la forma activa en lugar de la voz pasiva porque es más fácil de comprender.
- Repetir los sustantivos en lugar de utilizar los pronombres para evitar referencias vagas.
- Evitar las metáforas, refranes y expresiones coloquiales. En muchos casos sus traducciones no son igual de concisas y familiares.
- Evitar los verbos y preposiciones que no tienen un significado preciso, como "pronto" y "a menudo".
- Evitar formas posesivas, cuando sea posible, ya que la propiedad puede ser difícil de determinar.
- Utilizar términos específicos en lugar de términos generales.

Otro aspecto relevante a considerar al realizar la traducción es, como señalaron Auchter y Stansfield (1997), el empleo del lenguaje natural en el ámbito de la lengua objetivo. Angoff y Cook (1988), por su parte, subrayaron la importancia de tener en cuenta el contexto de los ítems para mejorar la calidad de la traducción.

Autores como Solano-Flores et al. (2006), partiendo de la base de que es inevitable que se produzca algún error en la traducción, identificaron una serie de dimensiones de errores en la traducción que tienen que considerarse en la revisión de los ítems traducidos. Dichas dimensiones se fundamentan en la exactitud de la traducción o la corrección gramatical (*gramática, semántica*), las características editoriales y de producción de las pruebas traducidas (*estilo, formato, convenciones*), la correspondencia entre la traducción y el uso de las palabras empleadas en el contexto social por la población destinataria (*registro*), el contenido (*información, constructo*) y la presencia del ítem en la enseñanza de la cultura objetivo (*currículum*). Una última dimensión (*origen*) subraya que los errores del ítem en el lenguaje origen pueden transferirse a la traducción. Según los mencionados autores, los errores, cuando se producen, suelen ser multidimensionales.

Una recomendación generalizada es contar con un equipo de traducción que implemente un procedimiento iterativo de depuración de los ítems, porque ello redundará en un menor número de errores y en una mejora de la calidad de la traducción. La traducción por consenso, en la que un grupo diverso y representativo de traductores/as bilingües llegan a acuerdos sobre la traducción de los ítems, es un enfoque que goza de

gran aceptación en el área de la evaluación lingüística (Cattell et al., 1993; Hambleton y Kanjee, 1993; Woodcock y Muñoz-Sandoval, 1999).

Según Solano-Flores et al. (2006) el equipo de traducción debería ser multidisciplinar, y debería contar con especialistas en la materia objeto de adaptación, lingüistas, traductores/as profesionales y especialistas en medición. La revisión de la traducción de los ítems se debería llevar a cabo mediante discusiones de grupo coordinadas por quienes lideren el proyecto de investigación.

En la misma línea, Martínez Arias et al. (2006) señalaron que las personas expertas en el contenido de la prueba deben revisar los instrumentos para evitar que se produzca una infrarrepresentación del constructo o introducción de varianza irrelevante. Estas autoras, al igual que Muñiz y Fonseca-Pedrero (2019), aconsejan que los ítems sean revisados por personas pertenecientes a la población objetivo. De este modo se podrán evitar sesgos o características del lenguaje que puedan resultar ofensivas para colectivos concretos y se garantizará que las personas comprendan lo que se está preguntando en el test.

1.2.3. Métodos para establecer equivalencias de traducción

La equivalencia entre las versiones original y adaptada de un test se asegura mediante dos métodos: el método de criterio (analítico-razional), relacionado con la traducción, y el método estadístico o empírico (Cardoso-Ribeiro et al., 2010; Hambleton, 1993, 1994).

El método analítico-razional está basado en el acuerdo respecto a la equivalencia de traducción de cada ítem. Se lleva a cabo antes de la aplicación del instrumento y, por tanto, de que se pueda emplear el método estadístico. Cabe distinguir entre dos métodos: la traducción directa (forward-translation) y la traducción inversa (back-translation).

El primero consiste en que un/a traductor/a, o, preferentemente un grupo de traductores/as (Epstein et al., 2015), traduce el instrumento del idioma fuente al idioma objetivo, y después otro grupo juzga la equivalencia entre ambas versiones (original y traducida), pudiendo hacerse revisiones a partir de los problemas identificados. Sin embargo, este método no identifica la varianza no relacionada con el constructo

introducida por el grupo de traductores/as (Cardoso-Ribeiro et al., 2010; Hambleton y Kanjee, 1993; Matthews-López, 2003).

En la traducción inversa, por su parte, un equipo de traductores/as adapta un instrumento del idioma origen al idioma destino, otro grupo de traductores/as vuelve a traducir el instrumento adaptado al idioma origen, y por último se comparan las dos versiones (traducida y retro-traducida o inversamente adaptada) para analizar la equivalencia. De esta manera aumenta la seguridad de la equivalencia si las similitudes entre ambas versiones son elevadas (Brislin, 1970; Cardoso-Ribeiro et al., 2010; Valentine, 2013).

El método de traducción inversa es el más utilizado y recomendado puesto que personas que no dominen la lengua objetivo pueden examinar las versiones original y retro-traducida del instrumento (Brislin, 1970). No obstante, no está exento de limitaciones, ya que, como señala Valentine (2013), el equipo de traducción y las personas que responden a un test funcionan en ambientes distintos, por lo que no necesariamente tienen que llegar a las mismas conclusiones con respecto a la interpretación de un ítem. Por tanto, el análisis subjetivo realizado por traductores/as, aunque sean expertos/as, no es garantía suficiente de que las versiones original y adaptada de un instrumento sean equivalentes. Además, como indican Hambleton y Kanjee (1993), los profesionales de la traducción no suelen predecir adecuadamente qué ítems funcionarán de manera diferente en colectivos diversos.

Los problemas que conllevan los métodos basados en criterios racionales hacen que para evaluar adecuadamente la equivalencia entre distintas versiones de un test sea imprescindible llevar a cabo también un análisis empírico de las respuestas de las personas a dicho test, empleando diversos diseños y procedimientos estadísticos.

Así, en el método estadístico o empírico se emplean tres diseños de investigación con el objetivo de establecer equivalencias entre diferentes versiones idiomáticas de un instrumento (Hambleton y Patsula, 1999). Los diseños se distinguen según el tipo de participantes (monolingües en ambos idiomas o bilingües), la versión del instrumento administrada (versión original, traducida o retro-traducida), y la técnica estadística empleada. Estos diseños son: 1) aplicación de la versión original y adaptada del test a individuos bilingües; 2) aplicación de la versión original y retro-traducida a monolingües del idioma fuente; y 3) aplicación de la versión original a monolingües del

idioma origen y de la versión adaptada a monolingües del idioma objetivo (Cardoso-Ribeiro et al., 2010).

Los tres diseños son populares, aunque como señalaron Hambleton y Patsula (1999), el tercero sería el más fácil de aplicar. No obstante, en el País Vasco, donde apenas existen monolingües vasco parlantes, no es aplicable para la adaptación de pruebas al euskera, siendo lo deseable una combinación de los dos primeros diseños (Isasi et al., 2000).

1.2.4. Equivalencia métrica a nivel de ítem y a nivel de test o escala

Para que las puntuaciones obtenidas al utilizar los instrumentos originales y los adaptados tengan el mismo significado es necesario llevar a cabo la evaluación de su *equivalencia métrica*. Un test o una subescala tienen equivalencia métrica para distintos grupos o poblaciones si las personas con puntuaciones idénticas en el constructo subyacente/latente tienen la misma puntuación directa o puntuación verdadera esperada a nivel de ítem, a nivel de puntuación total en la subescala o test, o en ambos. Por tanto, es necesario evaluar que la relación entre las puntuaciones observadas y el constructo medido por un test o escala es idéntica en las poblaciones a las que se aplique o que los ítems son psicométricamente equivalentes, es decir, que el conjunto de respuestas posibles evoca la misma respuesta, con igual probabilidad, entre sujetos que tienen el mismo nivel de aptitud (Drasgow, 1984; Drasgow y Kanfer, 1985; Heggstad et al., 2019; Hulin, 1987; Raju et al., 2002). Todo ello contribuye a detectar y reducir la posibilidad de impacto adverso (Hough et al., 2001).

En la misma línea, Sireci y Allalouf (2003) indicaron que quienes adapten un test deben establecer que:

- 1) El constructo medido existe en todos los grupos de interés.
- 2) El constructo se mide de la misma manera en todos los grupos.
- 3) Los ítems que se presumen equivalentes en diferentes idiomas son lingüística y estadísticamente equivalentes; y
- 4) Puntuaciones similares en las diferentes versiones lingüísticas de la prueba reflejan niveles similares en el constructo medido.

Evidentemente, este proceso va más allá del estudio de la calidad de las traducciones del test a las distintas lenguas. Como señaló Zumbo (2003), la equivalencia (o invarianza) métrica de las versiones en diferentes lenguas de un instrumento se asegura mediante dos grandes grupos de técnicas estadísticas: análisis a nivel de escala y análisis a nivel de ítem.

En los *análisis a nivel de escala*, el conjunto de ítems que componen una escala o un test son analizados mediante el análisis factorial y mediante el coeficiente alfa para estimar la consistencia interna de los tests original y adaptado. El primero es el método principal para estudiar la equivalencia a nivel de escala, puesto que permite comprobar si un constructo o una serie de constructos pueden ser medidos en dos lenguas diferentes y si se miden de manera equivalente en las diferentes versiones (Valentine, 2013; Zumbo, 2003).

La equivalencia a nivel de constructo puede evaluarse antes y después del estudio de campo, aunque cuando se hace antes -juicio de un grupo de expertos/as en la materia que representen los diferentes idiomas y culturas (Hui y Triandis, 1985)-, dado que no hay datos disponibles, la información que se obtiene es diferente. Después del estudio de campo, los datos obtenidos mediante las respuestas de los participantes pueden ser sometidos a análisis factorial exploratorio o confirmatorio para examinar la equivalencia de constructo. Si hay diferencias significativas en rasgos o comportamientos, será necesario adecuar la medida al contexto cultural, por lo que algunos ítems serán diferentes en las dos lenguas. A continuación, la equivalencia de constructo debería abordarse de una manera más indirecta. Un procedimiento común es examinar las redes nomológicas, analizando la validez convergente (patrón de correlaciones con instrumentos que midan el mismo constructo) y discriminante (correlaciones con las medidas de otros constructos) de los constructos medidos en cada grupo cultural (Van de Vijver y Tanzer, 2004).

En el *análisis a nivel de ítem* se estudia el funcionamiento diferencial de cada uno de los ítems de la escala o test. El *funcionamiento diferencial del ítem* (FDI) ocurre cuando personas de distintos grupos que tienen el mismo nivel en un constructo no responden de una forma similar al ítem en cuestión. Para que se pueda afirmar que existe sesgo del ítem es necesario identificar alguna característica del mismo que esté perjudicando a uno o más grupos (por ejemplo, cuando el ítem aborda un concepto que es más familiar a un grupo que a otro y sin embargo ese concepto no es central en el

constructo que se está midiendo). Las técnicas estadísticas que se utilizan para identificar ítems sesgados buscan ítems que funcionen de manera diferente en personas pertenecientes a los distintos grupos pero que tienen el mismo nivel en el constructo que está siendo medido. Una vez que los ítems con FDI son identificados, mediante un análisis cualitativo se intentarán explicar las diferencias observadas. Si al analizar la explicación para estas diferencias se observa que no están relacionadas con el objetivo del test, dichos ítems se considerarán “sesgados”, lo que puede tener implicaciones a nivel de validez predictiva y puntos de corte (Gómez-Benito et al., 2010; Hidalgo-Montesinos et al. 2015).

Hay muchos procedimientos estadísticos utilizados para detectar el funcionamiento diferencial de los ítems, la mayoría basados en la Teoría de Respuesta al ítem (TRI). Los más populares son los siguientes: Mantel-Haensel, Regresión logística, SIBTEST, Chi-cuadrado de Lord, y Área de Raju (Sireci y Allalouf, 2003). En las últimas dos décadas, el interés creciente acerca de la detección de FDI en ítems politómicos –los habituales en instrumentos de evaluación de la personalidad o el desempeño- ha provocado que estos procedimientos estadísticos, originalmente desarrollados para trabajar con ítems dicotómicos, se hayan adecuado para trabajar también con los politómicos (Valentine, 2013).

1.2.5. Sesgos más frecuentes y estrategias recomendadas

Los factores que comprometen la validez de las comparaciones intergrupales se denominan sesgos. Un sesgo puede producirse por diferentes razones y sus efectos pueden limitarse a una parte o afectar a todo el instrumento que se desea adaptar. La ausencia de sesgos (o la equivalencia métrica), es un prerrequisito para las comparaciones transculturales.

Según Van de Vijver y Hambleton (1996) y Van de Vijver y Tanzer (2004) hay tres tipos de sesgos inherentes al complejo proceso de adaptación de pruebas psicológicas: sesgos de constructo, sesgos metodológicos y sesgos de ítem.

El *sesgo de constructo* implica que el constructo objeto de estudio difiere entre las versiones original y adaptada de un instrumento, bien en la conceptualización del mismo, o bien en cuanto a la importancia que cada cultura otorga a ciertas conductas. Este sesgo puede producirse cuando el solape del constructo en distintas culturas sea solo parcial, cuando las conductas asociadas a un constructo sean pertinentes para una

cultura, pero no para otra, cuando la representación del constructo en el instrumento sea insuficiente, o cuando el instrumento mida constructos distintos en grupos culturales diferentes (Van de Vijver y Hambleton, 1996; Van de Vijver y Tanzer, 2004).

Según Van de Vijver y Hambleton (1996), el *sesgo metodológico* ocurre cuando el procedimiento de aplicación del instrumento crea diferencias en las respuestas entre las versiones original y adaptada. Este sesgo puede surgir por tres razones. Una causa es que las muestras no sean comparables por diferencias relativas a nivel educativo, sociocultural o de motivación, o por la composición en cuanto a género y edad de las personas participantes. La segunda causa son los problemas en la administración del test (diferentes condiciones de aplicación de la prueba, problemas de comunicación entre evaluador/a y candidatos/as, manuales de administración inadecuados, instrucciones ambiguas). La tercera causa tiene que ver con las características del instrumento, la familiaridad con los estímulos presentados o el formato de respuesta y diferencias intergrupos en cuanto a estilos de respuesta (deseabilidad social, aquiescencia, puntuaciones extremas). Si el sesgo metodológico está presente, influye en todos los ítems y puede provocar diferencias en los resultados de diferentes grupos no atribuibles al constructo medido.

El tercer sesgo es el *sesgo de ítem*, que se produce cuando un ítem tiene un significado psicológicamente distinto para diferentes culturas y está vinculado a un concepto mencionado anteriormente, el funcionamiento diferencial de los ítems. Como posibles causas cabe mencionar una mala traducción del ítem, ambigüedades en el ítem original, baja familiaridad o pertinencia del contenido del ítem en ciertas culturas, cambios en la redacción, formato o contenido del ítem, o la influencia de especificidades culturales como factores perturbadores o connotaciones asociadas con la redacción del ítem (Balluerka et al., 2010; Balluerka et al., 2014; Gómez-Benito et al., 2018; Van de Vijver y Tanzer, 2004).

Van de Vijver y Tanzer (2004) presentaron las estrategias más destacadas para abordar cada uno de los tres tipos de sesgos mencionados, indicando que algunas técnicas son comunes para los sesgos de constructo y/o sesgos metodológicos. Las estrategias analizadas combinan aspectos relativos al diseño de investigación, métodos estadísticos y procedimientos basados en el juicio de expertos/as, algunos de los cuales ya han sido mencionados en el apartado anterior (Tabla 1.1).

Tabla 1.1

Estrategias ante los Tipos de Sesgos (adaptado de Van de Vijver y Tanzer, 2004)

Tipo de sesgo	Estrategias
Sesgo de constructo	<ul style="list-style-type: none"> ▪ Descentralización (desarrollar al mismo tiempo el mismo instrumento en varias culturas) ▪ Enfoque de convergencia (desarrollo intracultural independiente de instrumentos y subsiguiente administración intercultural de todos los instrumentos)
Sesgo metodológico	<ul style="list-style-type: none"> ▪ Amplia formación a quien administre el test (por ejemplo, para aumentar su sensibilidad cultural) ▪ Manual o protocolo detallado para la administración, puntuación, y la interpretación de los resultados ▪ Instrucciones detalladas (con número suficiente de ejemplos y/o de ejercicios) ▪ Incluir variables propias de los sujetos o del contexto en el diseño (nivel educativo) o análisis estadísticos que controlen estas “variables ruido”, por ejemplo, un análisis de la covarianza, un diseño de bloques aleatorios y otras técnicas estadísticas (análisis de regresión, correlación parcial, etc.) ▪ Uso de información colateral (actitudes o motivación ante un test prueba) ▪ Evaluación de estilos de respuesta ▪ El uso de test-retest, formación y/o estudios de intervención
Sesgo de constructo y/o sesgo metodológico (aspectos comunes)	<ul style="list-style-type: none"> ▪ Uso de informadores/as con experiencia en la cultura y el lenguaje de interés ▪ Uso de muestras de sujetos bilingües ▪ Uso de encuestas con muestras locales (por ejemplo, analizando el contenido de preguntas de respuesta abierta) ▪ Administración no estándar del instrumento (por ejemplo, prueba piloto en la que se anime a los y las participantes a indicar en voz alta cómo interpretan los estímulos y la razón de sus respuestas) ▪ Comparación intercultural de redes nomológicas (estudios de validez convergente/discriminante, estudios monorasgo-multimétodo, equivalencias en la connotación de las frases clave mediante asociación libre o diferenciales semánticos)
Sesgo de ítem	<ul style="list-style-type: none"> ▪ Métodos de juicio de detección de sesgo en los ítems (análisis lingüístico y psicológico) ▪ Métodos psicométricos de detección de sesgo en los ítems (análisis de FDI) ▪ Análisis de errores o distractores ▪ Documentación de "ítems de repuesto" en el manual de la prueba que sean tan representativos del constructo como los ítems finalmente utilizados en el instrumento.

1.2.6. Directrices para la adaptación de instrumentos psicológicos

Ante la importancia de estos sesgos potenciales, en 1985, la American Educational Research Association (AERA) junto con la American Psychological Association (APA) y el National Council on Measurement in Education (NCME) publicaron los *Standards for Educational and Psychological Testing*. En dichos estándares se recogían una serie de recomendaciones relativas a la adaptación de tests de un lenguaje o dialecto a otro, tales como la necesidad de garantizar la fiabilidad y validez de lo que se pretendía medir en los grupos lingüísticos a evaluar, o que si el objetivo es comparar dos versiones de un test que están en dos idiomas distintos, es necesario aportar evidencia de esa comparabilidad.

Posteriormente, la International Test Commission desarrolló, a partir de 1994, unas directrices que sirvieran para prevenir las fuentes de error y para proporcionar soluciones a las mismas. Dicho documento ha sido citado en centenares de ocasiones en publicaciones científicas y profesionales, lo que avala su repercusión en el ámbito de la adaptación de los tests. En los últimos años las directrices se han sometido a una revisión, y la nueva edición está compuesta por veinte directrices agrupadas en seis apartados: directrices previas, desarrollo del test, confirmación, aplicación, puntuación e interpretación y documentación (International Test Commission, 2017).

Además de estas directrices, y desde un punto de vista más aplicado, autores como Hambleton y Patsula (1999), Matthews-López (2003), Balluerka et al. (2007) o Epstein et al. (2015), han planteado una serie de pasos concretos a seguir en la adaptación de un test. Carretero-Dios y Pérez (2005), partiendo de la similitud entre los procesos de construcción y adaptación de un test, también plantearon una serie de fases concretas a tener en cuenta.

Hambleton y Patsula (1999) propusieron una serie de pasos para evitar las fuentes de sesgo mencionadas en el epígrafe anterior y mejorar la calidad en la adaptación de los tests. Dichos pasos implican, fundamentalmente, la selección de profesionales de la traducción cualificados/as, 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 población

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.

Matthews-López (2003) se basó en las indicaciones de la ITC y de Hambleton y Patsula para elaborar sus propias directrices, enfocadas a la adaptación al español de una prueba de aptitudes, GMAT verbal.

Balluerka et al. (2007), en la adaptación del castellano al euskera del NEO-PI-R se basaron igualmente en las recomendaciones de Hambleton y Patsula, distinguiendo entre las fases de traducción, aplicación piloto y aplicación empírica del instrumento de medida.

Epstein et al. (2015), por su parte, en su revisión sobre directrices de adaptación de instrumentos transculturales, incidieron en las cuestiones más relevantes señaladas por la literatura con respecto al proceso de adaptación previo a la validación empírica. Estos autores vincularon su enfoque al cumplimiento de las directrices del estudio COSMIN sobre estándares para la selección de instrumentos en el ámbito de la salud (Mokkink et al., 2010).

Carretero-Dios y Pérez (2005) propusieron una serie de fases de investigación para procesos de construcción o adaptación de instrumentos psicométricos poniendo el acento en los análisis estadísticos a llevar a cabo:

- análisis estadístico de los ítems (correlación ítem-total corregida; correlación media inter-ítem; medias y desviaciones típicas);
- estudio de la dimensionalidad del instrumento (estructura interna medida mediante análisis factorial prestando atención a las saturaciones de cada ítem con su dimensión correspondiente);
- estimación de la fiabilidad (alfa de Cronbach y correlación media inter-ítem);
- obtención de evidencias externas de validez (validez convergente y discriminante; correlación con otros constructos, regresión múltiple y ecuaciones estructurales)

Con respecto a las propiedades psicométricas de un instrumento, Muñiz y Fonseca-Pedrero (2019) señalan que en la estimación de la fiabilidad de las puntuaciones se debe ir más allá del clásico Alfa de Cronbach, utilizando el coeficiente

Omega. Asimismo, consideran que no cabe limitarse a la Teoría Clásica del Test, sino que debe tenerse en cuenta la función de información desde la perspectiva de la Teoría de Respuesta al Ítem (TRI). Clark y Watson (2019) también se decantan por la TRI, indicando que puede ser de especial utilidad de cara a la selección de ítems. Los métodos basados en la TRI permiten identificar ítems que garantizan una evaluación precisa, fiable y válida a lo largo de todo el continuo del rasgo medido. En este sentido, indican que se debe optar por mantener ítems con un amplio rango de distribuciones de respuesta, es decir, con diferentes medias o porcentajes de elección de las opciones de respuesta, ya que la mayor parte de constructos representan dimensiones distribuidas de manera continua, y las puntuaciones se pueden producir a lo largo de toda la dimensión medida. Es crucial contar con ítems que discriminen en diferentes puntos del continuo, de manera que se puedan definir incluso puntuaciones extremas del mismo. Asimismo, Clark y Watson (2019) consideran necesario prestar atención a la validez incremental, esto es, a la demostración de que una medida añade de manera significativa validez a la predicción de un criterio más allá de la predicción de otras fuentes de datos (Hunsley y Meyer, 2003).

Estas directrices y pautas coinciden al señalar la importancia de contar con diseños metodológicos, procedimientos racionales y análisis estadísticos adecuados, así como en la relevancia de la formación de un comité de expertos/as multidisciplinar y en la trascendencia de las aplicaciones piloto de test previos a un estudio de campo con una muestra amplia (Tabla 1.2).

El papel del comité de expertos/as es asesorar acerca de la interpretación lingüística y cultural de los ítems. Para ello, es recomendable que sus miembros reciban formación con respecto a la evaluación estandarizada, los constructos medidos, técnicas de construcción de pruebas psicométricas y conceptos psicométricos básicos como sesgos, dificultad y discriminación de los ítems (Matthews-López, 2003)

Tabla 1.2.

Comparativa de Directrices sobre Adaptación de Instrumentos Psicométricos

Hambleton y Patsula (1999)	Matthews-López (2003)	Balluerka et al. (2007)	Epstein et al. (2015)	Carretero-Dios y Pérez (2005)
<ol style="list-style-type: none"> 1. Estudiar la equivalencia en el constructo en la lengua y cultura de interés. 2. Decidir entre adaptar una prueba existente o desarrollar una nueva. 3. Seleccionar traductores/as especializados/as. 4. Traducir y adaptar la prueba. 5. Repasar y revisar la versión adaptada de la prueba. 6. Realizar una prueba piloto de la versión adaptada. 7. Realizar un estudio de campo más ambicioso. 8. Análisis estadístico para comparar las versiones original y adaptada. 9. Garantizar la equivalencia de las versiones de la prueba. 10. Realizar un procedimiento de validación según el caso. 11. Documentar el proceso y preparar un manual para usuarios/as. 12. Entrenar a usuarios/as de la prueba. 13. Realizar un seguimiento continuo de la prueba adaptada. 	<ol style="list-style-type: none"> 1. Reunir un equipo de adaptación multidisciplinar. 2. Determinar si el constructo de interés es relevante, medible con fiabilidad y adaptable. 3. Crear un comité directivo y formarlo. 4. Organizar formación sobre redacción de ítems. 5. Traducir y/o adaptar el material del test. 6. Realizar una prueba piloto. 7. Revisar los datos de la prueba piloto. 8. Construir la versión operativa del test. 9. Realizar una prueba a pequeña escala. 10. Revisar la información obtenida. 11. Realizar un estudio estadístico y psicométrico. 12. Evaluar el test adaptado documentando los análisis realizados. 	<ol style="list-style-type: none"> 1. Fase de traducción de los ítems mediante un diseño de traducción inversa. 2. Fase piloto, con un análisis cuantitativo y cualitativo de los ítems. 3. Fase de aplicación empírica del instrumento de medida, examinando la dimensionalidad del instrumento, su consistencia interna, obteniendo evidencias de su validez convergente y diferencial y examinando el funcionamiento diferencial de los ítems en base a variables sociodemográficas. 	<ol style="list-style-type: none"> 1. La retrotraducción no es obligatoria, pero es una herramienta de comunicación útil con el/la autor/a del instrumento original. 2. Se debe trabajar con un comité de expertos/as para asegurar la equivalencia entre los instrumentos original y adaptado. El comité debe recibir formación en redacción de ítems. 3. Recoger información sobre la población objetivo, mediante una aplicación piloto del cuestionario con entrevistas cognitivas o focus groups. 4. La validación del instrumento adaptado debe incluir las siguientes propiedades de medida: <ul style="list-style-type: none"> -Consistencia interna -Fiabilidad -Error de medida -Validez de contenido -Validez de constructo -Validez de criterio 	<ol style="list-style-type: none"> 1. Justificación del estudio. 2. Delimitación conceptual del constructo a evaluar. 3. Construcción y evaluación cualitativa de ítems. 4. Análisis estadístico de los ítems. 5. Estudio de la dimensionalidad del instrumento (estructura interna). 6. Estimación de la fiabilidad. 7. Obtención de evidencias externas de validez.

Como señalan Matthews-Lopez (2003) o Epstein et al. (2015), el proceso de adaptación depende en buena medida de las pruebas piloto de un instrumento. En ellas, los ítems son evaluados en cuanto a sus medias aritméticas, dificultad, su discriminación, cargas factoriales o su equidad (FDI), y dichos valores son comparados con los valores de la prueba original. En base a dicha evaluación, se pueden incorporar nuevos ítems o bien eliminar ítems con carga factorial baja ($< .30$) o que no se ajusten al modelo teórico, dejando constancia de qué ítems fueron modificados y por qué, explicando los criterios empleados; todo ello para elegir los ítems con mayores garantías de maximizar las propiedades finales del instrumento (Muñiz y Fonseca-Pedrero, 2019). También se evalúa la adecuación o la claridad de las instrucciones o la validez aparente y el equilibrio del test y se recoge información sobre posibles errores de contenido o formato que se pueden corregir antes de pasar a una fase operacional. Se recomienda la recogida de información cualitativa mediante cuestionarios (o entrevistas cognitivas) antes o durante la sesión piloto, de forma que los y las participantes puedan comentar aspectos del instrumento adaptado, tales como la calidad percibida de los ítems, contenidos controvertidos o polémicos, adecuación del tiempo de administración, claridad de las instrucciones o relevancia de los dialectos en la formulación de los ítems. En esta fase es importante la recogida de tiempos de realización como referencia para aplicaciones en contextos selectivos. El éxito de la prueba piloto dependerá de la cantidad y representatividad de las personas participantes, que deben ser de características similares a las de la población objetivo (Martínez Arias et al. 2006; Woodcock y Muñoz-Sandoval, 1999).

En muchos casos se emplean más ítems de los que se prevé utilizar en el test definitivo e incluso se introducen ítems “variantes” que no se utilizan en la puntuación, y que se pilotan para futuras aplicaciones. Con los datos de la prueba piloto se realizan algunos análisis estadísticos de los ítems para poder seleccionar los que se adapten mejor al objetivo del test. Es importante examinar el rendimiento de diferentes subgrupos para evaluar problemas potenciales de sesgo (Martínez Arias et al., 2006).

Como señalaron Martínez Arias et al. (2006), en los tests referidos a normas, la puntuación de cada sujeto debe interpretarse en relación a otros sujetos de la población a la que pertenece, lo que implica recoger datos procedentes de grupos de referencia (poblaciones a las que va dirigido el test) en un estudio de campo. Así, el test se debe aplicar a una muestra representativa de la población (grupo normativo), que sirve de

grupo de referencia para interpretar las respuestas de los sujetos. Para ello, se debe comenzar por seleccionar una muestra representativa de la población objetivo, estableciendo los estratos necesarios (género, profesiones, edades...) si se considera que puede haber diferencias entre subpoblaciones.

Una vez aplicado el test a las personas seleccionadas, se realizarán los análisis estadísticos (fiabilidad, errores de medida, validez para el uso al que esté destinado, estructura interna...). A partir de estos datos se obtendrán las escalas de puntuaciones en las que basar las normas o baremos (Martínez Arias et al., 2006).

Con estos datos se debe elaborar un manual, lo que facilitará que el test pueda ser utilizado por otros profesionales distintos al constructor. Según Martínez Arias et al. (2006), un manual debe incluir:

1. Fundamentación teórica del constructo o explicación del dominio que cubre el test
2. Usos a los que se puede destinar el test
3. Poblaciones a las que está dirigido
4. Instrucciones completas para su administración
5. Coeficientes de fiabilidad o datos de precisión de la medida
6. Evidencias de validez para los usos a los que está destinado
7. Sugerencias sobre posibles usos de las puntuaciones de los tests
8. Normas de interpretación de las puntuaciones

1.2.7. La adaptación de instrumentos en el ámbito organizacional

La adaptación de instrumentos es una práctica habitual en el ámbito organizacional, y es frecuente que en dichas adaptaciones se produzcan cambios en los instrumentos que pueden tener un efecto en la validez. Heggstad et al. (2019) han estudiado más de 2000 escalas y han elaborado un listado de recomendaciones (Tabla 1.3) para preservar la validez de los instrumentos utilizados en este ámbito ante los cambios más comunes en la adaptación de instrumentos.

Tabla 1.3.

Recomendaciones para la Adaptación de Escalas (adaptado de Heggstad et al., 2019)

Tipo de cambio realizado	Descripción del cambio	Recomendaciones
Traducir una escala	Traducción de una escala de una lengua a otra	-Procedimientos estandarizados de retro-traducción (diseño inverso) -Análisis de equivalencia o invarianza comparando la escala original con la traducida
Acortar una escala	Eliminar algunos ítems de la escala original antes de la administración de la escala adaptada	-Elegir ítems de la escala original mediante una combinación de evidencia empírica y comprensión conceptual del constructo para mantener la validez de contenido -Comparar las correlaciones ítem-total de la escala original y la adaptada -Administrar la escala completa en una 1ª aplicación. En una 2ª aplicación administrar la escala total a la mitad de la muestra y la acortada al resto
Eliminar ítems post hoc	Eliminar ítems de una escala tras la recogida de datos	-Práctica no recomendable -Una excepción se da cuando se hayan documentado los mismos problemas con los mismos ítems. Se debe informar al respecto con criterios conceptuales y empíricos
Cambiar una escala de respuesta	Modificar el número de puntos de una escala o cambiar los anclajes verbales	-Transparencia con respecto al cambio -Aportar evidencias de la validez de la escala adaptada
Cambiar la redacción de los ítems	Adaptar la redacción de los ítems pero no para cambiar el contexto, referente o marco temporal	-Práctica a evitar -Describir los cambios en las instrucciones -Cada cambio en la redacción se debe acompañar de evidencias de validez adicionales -Administrar la escala adaptada en una 1ª aplicación. En una 2ª aplicación administrar al azar la escala original a la mitad de la muestra y la adaptada al resto. Cabe añadir un tercer grupo en ambas aplicaciones y medir la equivalencia o invarianza. Asimismo, se puede aplicar una variable externa en la 2ª aplicación para comparar las correlaciones

Tabla 1.3 (continuación)

Recomendaciones para la Adaptación de Escalas (adaptado de Heggstad et al., 2019)

Tipo de cambio realizado	Descripción del cambio	Recomendaciones
Cambiar el referente de sí mismo a otra persona	Cambiar el referente del ítem o las instrucciones de la propia persona encuestada a una tercera persona	<ul style="list-style-type: none"> -Considerar si el tercero que evalúa puede aportar respuestas relevantes y válidas a los ítems -Incluir los ítems en un anexo -Análisis Factorial Confirmatorio (AFC) para comprobar si la estructura es similar a la obtenida mediante auto-informe -Comparar auto y hetero-evaluaciones mediante correlaciones y análisis de equivalencia o invarianza
Cambiar el referente del individuo al grupo	Cambiar el referente del ítem o las instrucciones de la propia persona encuestada a un constructo a nivel de grupo	<ul style="list-style-type: none"> -Aportar fundamentación teórica para el significado del constructo de nivel superior -Realizar un AFC multinivel
Cambiar el contexto	Adaptar los ítems del contexto original (p. ej. general) a un nuevo contexto (p. ej. específico del trabajo)	<ul style="list-style-type: none"> -Transparencia sobre los cambios -AFC para comparar las estimaciones de los parámetros de la escala original y adaptada -Si algún ítem no tiene sentido en el contexto adaptado se necesita una mayor labor de validación de contenido (panel de expertos/as)
Cambiar el marco temporal	Adaptar las instrucciones o los ítems de una escala para reflejar un marco temporal diferente	<ul style="list-style-type: none"> -Transparencia sobre los cambios -AFC para comparar el número de factores y las saturaciones de los factores de la escala original y la adaptada -Analizar correlaciones test-retest de las versiones de la escala asociadas a diferentes marcos temporales
Añadir o combinar ítems	Crear nuevos ítems para la escala o combinar ítems de diferentes escalas en una única escala	<ul style="list-style-type: none"> -Práctica a realizar en caso de una razón conceptual convincente, ya que añadir ítems implica crear una nueva escala -Transparencia sobre los cambios -Realizar un AFC de la escala con los ítems añadidos. Comprobar la estructura factorial con la escala original y la magnitud de las saturaciones de los ítems añadidos

Tabla 1.3 (continuación)

Recomendaciones para la Adaptación de Escalas (adaptado de Heggstad et al., 2019)

Tipo de cambio realizado	Descripción del cambio	Recomendaciones
Emplear una pequeña parte de una escala	Administrar subescalas de una medida excluyendo el resto de la medida completa	-Emplear una única escala de un instrumento con múltiples escalas previsiblemente no afectará a la validez de las puntuaciones de la escala

1.2.8. Particularidades del euskera para la adaptación de instrumentos

Según Isasi et al. (2000), en contextos en los que haya contacto de lenguas es necesario contemplar tres ámbitos fundamentales antes de iniciar el proceso de adaptación de un test: 1) lingüístico: distancia interlingüística, existencia de dialectos, estándar común, desarrollo lexical de la(s) lengua(s); 2) demolingüístico: número, proporción y distribución de los tipos de hablantes (monolingües de cada lengua, proporción de bilingües); y 3) sociolingüístico: características de las poblaciones monolingües competentes en la(s) misma(s) lengua(s), estatus de cada lengua, alfabetización, educación, medios de comunicación, etc.

La adaptación de instrumentos psicométricos conlleva ciertas peculiaridades en el caso del euskera, una de las lenguas vivas más antiguas de Europa, y una de las pocas lenguas continentales que no desciende del tronco lingüístico indo-europeo. Por todo esto, presenta importantes diferencias con respecto al castellano en cuanto a léxico y estructura gramatical (Trask, 1997). Además de las diferencias semánticas, el euskera también muestra importantes diferencias sintácticas en comparación con el español. Por ejemplo, la lengua vasca es una lengua aglutinante y postposicional, mientras que el castellano es no-aglutinante y preposicional; presenta un orden sintáctico sujeto-objeto-verbo en frases breves mientras que la estructura sintáctica en castellano sigue un orden sujeto-verbo-objeto; el euskera tiene un léxico más limitado que el español; en la lengua vasca no existen artículos de género mientras que en castellano son utilizados con frecuencia; y resulta más sencillo construir palabras compuestas en euskera que en español (Gorostiaga et al., 2011).

Dadas estas diferencias entre el euskera y el castellano, el grado esperado de correspondencia estructural debería ser menor que en el caso de las lenguas más similares, lo cual ha de tenerse en cuenta en el proceso de adaptación (Balluerka et al., 2009; Gorostiaga et al., 2011; Su-ju, 2006). De hecho, algunas investigaciones han revelado problemas de equivalencia psicométrica entre distintas versiones lingüísticas del programa PISA, dándose los mayores niveles de no-equivalencia en países donde se hablan lenguas no indo-europeas (Grisay y Monseur, 2007; Grisay et al., 2007).

En cuanto al uso del euskera, según la VI Encuesta Sociolingüística del Gobierno Vasco (2017) un 33.9 % (631.000 habitantes) de la población de la Comunidad Autónoma Vasca mayor de 16 años era vascoparlante bilingüe, un 19.1 % (356.000) vascoparlante bilingüe pasivo –saben hablarlo, pero no lo usan- y un 47% (877.000) era castellanoparlante exclusivo. En la actualidad, prácticamente no existen vascófonos monolingües, dado que los vascos que desconocen el castellano son menos del 1% de la población total. En Navarra es vascohablante uno de cada ocho habitantes (el 12.9 %) y en el País Vasco francés uno de cada cinco (20.5 %). En cuanto a la evolución del uso del euskera, en 2016 había 212.000 hablantes más que en 1991, fecha de la primera encuesta sociolingüística, siendo el porcentaje de nuevos vascohablantes cada vez mayor, especialmente entre los jóvenes.

Desde un punto de vista sociolingüístico, el euskera presenta una serie de dialectos: vizcaíno, guipuzcoano, labortano, suletino, altonavarro y bajonavarro (Zuazo, 1998). Sin embargo, el registro más utilizado en la administración, la educación o los medios de comunicación es el *euskera batua* (o euskera unificado), lengua estándar impulsada por Euskaltzaindia-Academia de la Lengua Vasca con el propósito de recuperar el uso de la lengua. Esta lengua estándar, durante las últimas cuatro décadas, ha contribuido a aumentar el número de vascoparlantes, fundamentalmente gracias a su presencia en el contexto educativo (en el que la mayor parte del alumnado estudia la mayoría de asignaturas impartidas en euskera), a su utilización en la radiotelevisión pública vasca (EITB) y a la exigencia de la acreditación de un nivel o perfil lingüístico mínimo en euskera para la mayoría de puestos en el sector público vasco. El euskera batua, además, ha permitido que sea posible la adaptación a la lengua vasca de instrumentos psicométricos originalmente creados en otras lenguas (Isasi et al., 2000).

La lengua vasca ha coexistido históricamente en un entorno bilingüe, con el castellano en la Comunidad Autónoma del País Vasco (CAPV) y en la Comunidad Autónoma de Navarra, y con el francés en el departamento de Pirineos Atlánticos en Francia. En la CAPV, el euskera se ha relegado históricamente a situaciones informales, mientras que el castellano, al menos hasta hace poco tiempo, ha sido considerado como lengua de cultura y reservado para situaciones más formales (Etxebarria, 1992).

Según Etxebarria (2003), la situación lingüística del País Vasco se ha definido en términos de bilingüismo y diglosia si bien la variación en las comunidades lingüísticas y los usos de las dos lenguas, que coexisten en la comunidad, es muy diferenciada. No obstante, en las últimas décadas se ha llevado a cabo una política lingüística que ha tenido como objetivo fundamental garantizar progresivamente la consecución de una sociedad cada vez más bilingüe e integrada que ha supuesto una notable mejoría en la situación del euskera, particularmente en el sistema educativo, en los medios de comunicación, en la Administración Pública y en la sociedad en general, articulándose multitud de programas de promoción del uso del euskera.

Habida cuenta de las peculiaridades del euskera y de su coexistencia con otras lenguas en un mismo ámbito geográfico, diferentes investigaciones se han interesado por el impacto de la lengua materna y del idioma de aplicación de tests en los procesos de evaluación psicológica (Balluerka et al., 2007; Balluerka y Gorostiaga, 2012; Gorostiaga et al., 2011).

CAPÍTULO 2.

ADAPTACIÓN AL CASTELLANO Y VALIDACIÓN DE LA ESCALA DE DESEMPEÑO ADAPTATIVO DE CHARBONNIER-VOIRIN Y ROUSSEL

Las transformaciones que se han producido en la sociedad en los últimos años, derivadas de fenómenos como los avances tecnológicos, la globalización, la crisis económica o la transición a una economía basada en el conocimiento han provocado que la capacidad adaptativa del personal sea un elemento esencial de las organizaciones del siglo XXI. Estos fenómenos han creado un entorno de trabajo imprevisible en el que empleados y empleadas deben afrontar constantemente situaciones nuevas y complejas que hacen que los puestos de trabajo sean más exigentes a nivel cognitivo y perceptivo (Howe, 2014; Kozlowski et al., 1999; Pulakos, et al., 2012).

Además, los puestos, se definen cada vez más en función de roles cambiantes, en vez de definirse en base a funciones y tareas estables (Murphy y Jackson, 1999). Esta cuestión está vinculada con la planificación de la gestión de la carrera del personal o la sucesión de empleados/as a punto de jubilarse o abandonar una organización, dado que actualmente al trabajador o trabajadora se le puede pedir que sea capaz de adaptarse a diferentes roles organizativos (Pulakos et al., 2012).

Todo ello hace que la adaptabilidad de las personas, y, por ende, su desempeño adaptativo sea una destreza clave en el mundo del trabajo actual y probablemente en el de las próximas décadas. En un contexto como el presente, identificar las diferencias individuales que pueden tener un impacto positivo en las organizaciones es esencial. Por esta razón, el desempeño adaptativo se ha convertido en un constructo de gran interés en la psicología aplicada en los últimos años (Bidshahri, 2017; Dorsey, et al., 2010; Gorriti, 2018).

En este capítulo se aborda la visión tradicional del desempeño laboral y la inclusión del desempeño adaptativo en dicho marco. A continuación, se define el constructo en base a los principales modelos teóricos, distinguiendo este tipo de desempeño de otros constructos relacionados, y, tras analizar diferentes maneras de proceder a su medición, se examinan los principales instrumentos desarrollados para medirlo. Todo ello con el objetivo de adaptar al castellano una prueba breve que mida de manera parsimoniosa, exhaustiva, fiable y válida el desempeño adaptativo y sea

aplicable a todo tipo de puestos de trabajo: la escala de Charbonnier-Voirin y Roussel (2012). Por último, se explicita el procedimiento de adaptación de la escala de Charbonnier-Voirin y Roussel (2012) a la lengua y cultura castellana y su proceso de validación en esa lengua.

2.1. El desempeño laboral

Las pruebas selectivas, al ser empleadas en el sector público, deben tener en cuenta lo mencionado por el artículo 61.2 del Estatuto Básico del Empleado Público (Congreso de los Diputados, 2007, 2015): “Los procedimientos de selección cuidarán especialmente la conexión entre el tipo de pruebas a superar y la adecuación al desempeño de las tareas de los puestos de trabajo convocados”. Por tanto, para comprobar que las pruebas de una oposición han cumplido su cometido, es preciso evidenciar que el desempeño de las personas contratadas ha sido eficaz.

Según dicha ley, la evaluación del desempeño es el procedimiento mediante el cual se mide y valora la conducta profesional y el rendimiento o el logro de resultados. Dicho enfoque coincide con el de Viswesvaran (2001), que apostilla que la conducta evaluada debe estar “bajo control del trabajador”. Asimismo, la medición o evaluación del desempeño laboral implica considerar comportamientos relevantes para la organización en el sentido de que una mala ejecución u omisión, tiene consecuencias trascendentes para ella (Gorriti, 2007; Gorriti y López, 2009; Viswesvaran, 2001). En definitiva, el desempeño es el criterio que refleja cuantitativamente la concepción del éxito para la organización. En el ámbito de la selección de personal, el éxito está vinculado a la eficacia de los procedimientos selectivos a la hora de seleccionar a las personas idóneas para los perfiles de los puestos a cubrir (Guion, 1997).

2.1.1. Naturaleza y taxonomía “clásica” del desempeño laboral

Por lo que se refiere a la naturaleza del desempeño, la visión “clásica” (Beck et al., 2014; Gorriti y López, 2014; Lievens, Conway, et al., 2008; Rotundo y Sackett, 2002) plantea que existen tres ámbitos o tipos de desempeño para cualquier organización y para cualquier trabajo: el *desempeño de tarea* -comportamientos técnico-productivos directamente exigidos para la ejecución de las tareas-, el *desempeño contextual* -comportamientos de naturaleza social o psicológica, que facilitan o

dificultan el desempeño de tarea y ayudan a conseguir los objetivos organizacionales- y las *conductas contra-productivas* -conductas contrarias a los legítimos intereses de la organización como robos, abusos de alcohol y drogas...-. Estas últimas están muy ligadas a aspectos disciplinarios, son difíciles de medir y, como indican Gorriti y López (2014) o Ramírez (2013), no se recomienda su uso como criterio de desempeño laboral en un estudio de validación vinculada a la selección de personal.

La investigación en materia de desempeño de tarea es amplia, y fruto de estudios factoriales se han identificado varias taxonomías de dimensiones relevantes para evaluarlo (Tabla 2.1). Como ejemplo cabría mencionar la de Bernardin y Beatty (1984), que identificó las dimensiones de Calidad, Cantidad, Tiempo, Costo-eficacia, Necesidad de supervisión e Impacto interpersonal, o la de Campbell (1990), que determinó que las dimensiones de desempeño de tarea eran: Eficacia en tareas específicas, Eficacia en tareas no específicas, Comunicación, Esfuerzo demostrado, Disciplina personal sostenida, Facilitación de ejecución, Supervisión y Gestión o administración (Gorriti y López, 2009).

Tal y como indican Gorriti y López (2009), algunas de las dimensiones mencionadas en la Tabla 2.1 podrían considerarse más bien antecedentes o condiciones del desempeño (conocimientos, destrezas de supervisión y liderazgo), mientras que otras (impacto interpersonal, comunicación) entran dentro del ámbito del desempeño contextual, y solamente algunas (cantidad o productividad, calidad, eficacia o gestión y administración) deberían considerarse aspectos de ejecución de tarea pura.

En este sentido, diferentes autores conceden especial relevancia a tres componentes del desempeño de tarea: Cantidad, Calidad y Precisión (Edwards et al., 2008; Tubré et al., 2006; Welbourne et al., 1998). Como indica Ramírez (2013), dichos componentes pueden encontrarse en la mayoría de las propuestas de los autores que han intentado delimitar el dominio general del desempeño individual (Bernardin y Beatty, 1984; Campbell, 1990; Jenkins et al., 1998; Murphy, 1989; Toops, 1944; Viswesvaran, 1993; Wherry, 1957).

Tabla 2.1.

Taxonomías “Clásicas” del Desempeño de tarea (adaptado de Gorriti y López, 2009)

Autor/es	Naturaleza de la medida	Dimensiones
Toops (1944)	Distinguiendo entre adecuación: calidad y ausencia de errores, y cantidad: volumen de producto	-Unidades de producción -Calidad del trabajo -Tiempo de permanencia -Destrezas de supervisión y liderazgo
Wherry (1957)		-Resultados -Calidad -Tiempo perdido en la ejecución -Cambio de trabajos -Tiempo de formación o promoción -Satisfacción
Bernardin & Beatty (1984)	Definen la ejecución como los registros de resultados producidos en un puesto durante un periodo de tiempo	-Calidad -Cantidad -Tiempo -Costo-eficacia -Necesidad de supervisión -Impacto interpersonal
Campbell (1990)	Describe 8 dimensiones e indica que la correlación entre las 8 es baja, lo que dificulta identificar esa estructura latente	-Eficacia en tareas específicas -Eficacia en tareas no específicas (para todos) -Comunicación oral y escrita. -Esfuerzo demostrado (en acabar una tarea) -Disciplina personal sostenida -Facilitación de la ejecución de pares y equipos -Supervisión -Gestión o administración
Murphy (1989)	Entiende la ejecución como el compendio de cuatro dimensiones	-Conductas que hacen perder el tiempo -Ejecución de las tareas o roles definidos para el puesto -Interpersonales: de ayuda y de equipo -Conductas destructivas: violencia, robos, etc.
Viswesvaran (1993)		-Desempeño laboral global o general -Rendimiento o productividad -Calidad -Competencia de comunicación -Competencia administrativa -Esfuerzo -Competencia interpersonal -Liderazgo -Conocimiento del puesto -Conformidad u obediencia a la autoridad

Nota. En negrita las dimensiones de tarea “puras”

En cuanto al desempeño contextual, la principal taxonomía es la de Borman y Motowidlo (1993), que distinguen tres dimensiones, cada una de las cuales tiene tres subdimensiones: (1) ayuda a otros (cooperación, cortesía, motivar a los demás); (2) vinculación organizacional (representación, lealtad, compromiso); y (3) desempeño más allá del puesto (persistencia, iniciativa, auto-desarrollo).

Esta taxonomía del desempeño contextual comparte similitudes con los patrones comportamentales del constructo *conductas cívicas organizacionales* (Organizational Citizenship Behaviors, OCB, en inglés). Según Organ y Ryan (1995), este constructo está vinculado con “contribuciones individuales en el lugar de trabajo que van más allá de los requerimientos de rol y de los logros laborales contractualmente recompensados”. Estas conductas incluyen dimensiones como altruismo, cortesía, deportividad, responsabilidad y virtud cívica (Ramírez, 2013).

Según Organ (1997), la diferencia radica en que el desempeño contextual tal y como es definido no requiere que la conducta sea extra-rol ni que exista ausencia de recompensa, es decir, puede tratarse de un conjunto de expectativas informales acompañadas por algún tipo de compensación. Sin embargo, investigaciones posteriores han vinculado el desempeño contextual al comportamiento ciudadano organizacional (Borman et al., 2001; Hanson y Borman, 2006; Van Dyne et al., 1995).

Gorriti y López (2014) recomiendan distinguir entre desempeño de tarea y desempeño contextual a la hora de evaluar el constructo pese a que la correlación entre ambos desempeños sea alta ($r = .74$; Salgado et al., 2007). Estos autores coinciden con Campbell et al. (1996), que señalan que una única medida de desempeño global sólo es interesante a nivel informativo, por lo que es recomendable centrarse en la investigación de facetas de desempeño concretas.

Por otra parte, Gorriti y López (2014), dentro de la disquisición entre *desempeño típico* -lo que una persona hace o hará en circunstancias típicas, medible mediante cuestionarios de evaluación del desempeño- y *desempeño máximo* -lo que una persona puede hacer en la mejor de sus circunstancias, medible mediante exámenes prácticos-, entienden que en el sector público se debe optar por el primero, salvo excepciones como el trabajo policial.

2.1.2. La inclusión del desempeño adaptativo en el marco del desempeño laboral

En la realidad laboral actual los puestos de trabajo se definen cada vez más en base a roles y contextos y se enfatiza en el cambio constante para satisfacer las demandas de los clientes o la naturaleza dinámica del trabajo. Esta tendencia dificulta captar todos los elementos propios de las tareas en las descripciones de los puestos, lo que a su vez ha provocado cambios en la conceptualización del desempeño. Recientes estudios han puesto de manifiesto que las dimensiones clásicas pueden no ser suficientes para representar completamente el dominio del desempeño. Las maneras clásicas de definir y evaluar el desempeño laboral pueden haber dejado de ser válidas debido a la creciente demanda de nuevas conductas como el aprendizaje continuo o la adaptabilidad, y, además, los requisitos del desempeño de tarea y contextual son susceptibles de evolucionar a lo largo del tiempo. Todo ello conlleva una reducción sistemática de la validez de los predictores más comunes del desempeño, especialmente cuando el tiempo entre la medición de predictores y criterios se incrementa (Bidshahri, 2017; Cascio, 1995; Henry y Hulin, 1987; Ilgen y Hollenbeck, 1991; Schmitt et al, 2003).

Esta reflexión ha provocado que diferentes investigadores hayan propuesto una dimensión adicional del desempeño centrada en la medición de la capacidad de empleados y empleadas para gestionar los cambios en el lugar de trabajo, dimensión clave en las organizaciones actuales (Carpini y Parker, 2017; Howe, 2014; Ramírez, 2013). Por ello, autores como Johnson (2003), Griffin et al. (2007) o Koopmans et al. (2011) han incluido el desempeño adaptativo como una faceta diferenciada dentro de sus modelos teóricos de desempeño laboral (Tabla 2.2), e incluso Campbell (1999) reconoció que su taxonomía original debería haber incluido este componente “genuino” del desempeño. Asimismo, Bartram (2005), desde un enfoque basado en competencias laborales concebidas como modelo de desempeño laboral, identificó una competencia denominada “Adaptación y afrontamiento”. No obstante, Borman y Motowidlo (1993) o Van Iddekinge y Ployhart (2008) señalan que no está suficientemente claro que esta dimensión del desempeño pueda distinguirse del desempeño de tarea o del desempeño contextual, algo en lo que coinciden con Johnson (2001, 2003), quien indicó que solo algunos elementos identificados como desempeño adaptativo pueden ser considerados como una dimensión independiente del desempeño laboral.

Tabla 2.2.

Taxonomías del Desempeño Laboral que Incluyen el Desempeño Adaptativo

Autor/es	Naturaleza de la medida	Dimensiones y subdimensiones
Johnson (2003)	Distingue tres grandes bloques: desempeño de tarea, desempeño de ciudadanía y desempeño adaptativo	<ul style="list-style-type: none"> -Desempeño de tarea <ul style="list-style-type: none"> ▪ Eficacia en tareas específicas ▪ Eficacia en tareas no específicas (para todos) ▪ Comunicación oral y escrita. ▪ Supervisión ▪ Gestión o administración ▪ Iniciativa consciente hacia las tareas -Desempeño de ciudadanía <ul style="list-style-type: none"> ▪ Iniciativa consciente (esfuerzo para conseguir objetivos) ▪ Apoyo personal ▪ Apoyo organizacional -Desempeño adaptativo <ul style="list-style-type: none"> ▪ Lidar con situaciones laborales inciertas
Griffin et al. (2007)	Dos grandes dimensiones que reflejan las variables de incertidumbre (nivel de dinamismo necesario: eficacia, adaptabilidad y proactividad) e interdependencia (el objetivo de la contribución: tarea individual, miembro del equipo, miembro de la organización) en el contexto laboral	<ul style="list-style-type: none"> -Eficacia <ul style="list-style-type: none"> ▪ Individual ▪ Equipo ▪ Organización -Adaptabilidad <ul style="list-style-type: none"> ▪ Individual ▪ Equipo ▪ Organización -Proactividad <ul style="list-style-type: none"> ▪ Individual ▪ Equipo ▪ Organización
Bartram (2005)	Modelo basado en competencias laborales	<ul style="list-style-type: none"> -Liderar y decidir -Apoyar y cooperar -Interactuar y presentar -Analizar e interpretar -Crear y conceptualizar -Organizar y ejecutar -Adaptar y afrontar -Emprender y realizar
Koopmans et al. (2011)	Describe 4 grandes dimensiones definidas mediante una serie de indicadores que pueden ser medidos directamente	<ul style="list-style-type: none"> -Desempeño de tarea -Desempeño contextual -Desempeño adaptativo -Conductas contraproductivas

2.1.3. Definiciones y taxonomías del constructo desempeño adaptativo

La literatura en materia de desempeño adaptativo ha evolucionado a partir de diferentes perspectivas, lo que ha producido una gran fragmentación a la hora de conceptualizar el constructo de desempeño adaptativo (Jundt et al., 2015). Por ello, la revisión teórica de Baard et al. (2014) ha supuesto un avance que ha venido a solucionar la necesidad de integrar y sintetizar la investigación en esta materia (Ployhart y Bliese, 2006).

Baard et al. (2014) distinguieron dos grandes ámbitos en los que situar los cuatro grandes enfoques del desempeño adaptativo: dominio general y dominio específico (Tabla 2.3). Este último dominio se centra en cambios y/o procesos clave para la adaptación y no es generalizable a todo tipo de circunstancias laborales. Está muy vinculado a la formación y al aprendizaje, ya que su supuesto fundamental es que los conocimientos y destrezas pueden ser aprendidos y aplicados a contextos específicos, mitigando declives en el desempeño derivados de cambios.

Tabla 2.3

Comparación de enfoques de la Adaptabilidad (adaptado de Murphy, 2015)

Enfoque	Dominio	Conceptualización	Medida
Constructo del desempeño (Pulakos et al., 2000, 2002)	General	El desempeño adaptativo describe situaciones en las que las personas modifican su comportamiento para satisfacer las demandas de una nueva situación o un cambio en el ambiente	JAI
Diferencia individual (Ployhart y Bliese, 2006)	General	La adaptabilidad individual no es solo una capacidad de respuesta ante un cambio del entorno, sino también un conjunto de aptitudes, destrezas y motivaciones que tiene una persona para reaccionar o ser proactivo ante los cambios en diferentes situaciones	I-ADAPT
Cambio del desempeño (Kozlowski et al., 2001)	Específico	El desempeño adaptativo es la capacidad de las personas para abordar la brecha entre aprender y transferir tareas que son novedosas o están poco estructuradas	Evaluaciones experimentales
Proceso (Bell y Kozlowski, 2002)	Específico	El desempeño adaptativo se evidencia en situaciones de cambio en las que el conocimiento y las destrezas aprendidas durante la formación deben adaptarse para el desempeño eficaz en situaciones nuevas o más complejas	Tareas o simulaciones realizadas mediante ordenador

El dominio general, por su parte, asume que la adaptabilidad es un conjunto estable de rasgos y paradigmas de desempeño generalizable a todo tipo de situaciones. Este enfoque se ha aplicado principalmente en selección y evaluación del desempeño y ha dado lugar a dos tipos de conceptualizaciones: por un lado, el desempeño adaptativo como diferencias individuales (una serie de rasgos estables) que contribuyen a la predicción del desempeño en circunstancias cambiantes; por otro, como constructo del desempeño (una serie de dimensiones que caracterizan el desempeño laboral de tipo adaptativo) con el mismo nivel de relevancia que el desempeño de tarea o el contextual.

En consecuencia, dentro del dominio general, existen, por un lado, investigaciones que conciben la adaptabilidad como un conjunto de diferencias individuales que predisponen a empleados y empleadas para adaptarse con éxito a los cambios en las tareas. Este enfoque sugiere que las dimensiones tradicionales del desempeño (tarea, contextual, conductas contraproductivas) pueden ser adaptativas. El trabajo de Ployhart y Bliese (2006) es el más destacado en esta área, y teoriza que los conocimientos, las destrezas, los valores, las aptitudes y la personalidad predicen el desempeño adaptativo, y que este, a su vez, predice el desempeño de tarea o el contextual. Este postulado ha sido parcialmente confirmado por diversos estudios que han analizado la relación entre APO y adaptabilidad o entre adaptabilidad y desempeño contextual (Almahamid et al., 2010; Chan y Schmitt, 2002; Wang et al., 2011). Sin embargo, los estudios sobre la conceptualización del predictor (la adaptabilidad) y el criterio (el desempeño) han carecido de consistencia, presentando una considerable variabilidad en cuanto a su definición y su medida, lo que dificulta extraer conclusiones significativas acerca de la utilidad del desempeño adaptativo concebido como diferencia individual (Baard et al., 2014).

Por otro lado, el enfoque que concibe el desempeño adaptativo como un constructo del desempeño asume que es una dimensión diferenciada del desempeño laboral que puede ser predicha por una serie de diferencias individuales y está limitada por el contexto (Pulakos et al., 2012). Según Pulakos et al. (2006) este enfoque es más operativo que el de las diferencias individuales porque no todos los puestos requieren adaptabilidad y es más eficaz medir el resultado de conductas adaptativas que predecir el resultado de conductas adaptativas que pueden no llegar a producirse.

Siguiendo a Baard et al. (2014), existen una serie de definiciones del enfoque del desempeño adaptativo como constructo del desempeño. Sin ánimo de ser exhaustivos, recogemos algunas de tales definiciones en la Tabla 2.4. Una de estas definiciones es la de Griffin et al. (2007), que distinguieron tres niveles de análisis del desempeño adaptativo: individual (cambios en equipamientos o procedimientos), grupal (incorporación de un nuevo miembro) y organizacional (cambios en operaciones). En la presente investigación el desempeño adaptativo se conceptualizará esencialmente como constructo de desempeño a nivel individual. No obstante, para extraer conclusiones sobre este tipo de desempeño es recomendable no segregar los hallazgos en diferentes dominios, teniendo en cuenta las conclusiones de estudios con métodos, contextos y objetivos diversos (Calarco, 2016; Jundt et al., 2015).

Tabla 2.4.

Definiciones del Desempeño Adaptativo

Autor/es	Resumen de la Definición
Allworth y Hesketh (1999)	La adaptabilidad consiste en conductas que demuestran la capacidad de la persona para hacer frente al cambio y aplicar estrategias de aprendizaje de una tarea a otra en función de los requerimientos del puesto
Griffin y Hesketh (2003)	El desempeño adaptativo se compone de los aspectos del desempeño referidos a las demandas cambiantes del puesto.
Griffin et al. (2007)	La adaptabilidad es un constructo multinivel que se refiere a la manera en que las personas dentro de los equipos y de las organizaciones responden a cambios en el entorno y los apoyan
Johnson (2001)	El desempeño adaptativo se refiere a la efectividad de la persona en cumplir con las demandas del entorno o de nuevas situaciones
Pulakos et al. (2000)	El desempeño adaptativo describe situaciones en las que las personas modifican su comportamiento para satisfacer las demandas de una nueva situación o evento o de un entorno cambiado
Koopmans et al. (2011)	Ejemplos de indicadores: generar ideas nuevas e innovadoras, ajustar objetivos y planes a la situación, aprender nuevas tareas e innovaciones, ser flexible y abierto de mente con otras personas, entender a otras personas y culturas, mostrar resiliencia, permanecer en calma, analizar rápidamente, actuar de manera adecuada
Shoss et al. (2012)	El desempeño adaptativo es una faceta de desempeño dirigida a respuestas efectivas a los cambios en el entorno

Como indican Baard et al. (2014) o Campbell y Wiernik (2015), el modelo taxonómico más riguroso para conceptualizar el desempeño adaptativo como constructo de criterio es el de Pulakos et al. (2000). Estos autores, mediante una revisión de la literatura y de más de mil incidentes críticos¹ extraídos de 21 puestos de los sectores público y privado, identificaron 8 dimensiones del desempeño adaptativo que mostraron un buen ajuste en el análisis factorial confirmatorio. Dichas dimensiones, definidas en la Tabla 2.5, son el fruto de trabajos desarrollados durante varias décadas y reflejan un consenso generalizado acerca de la naturaleza multidimensional del desempeño adaptativo (Campbell y Wiernik, 2015).

Tabla 2.5.

Dimensiones y Definiciones de la Taxonomía de Pulakos et al. (2000)

Dimensión	Definición
1. Gestionar emergencias y situaciones de crisis	Reaccionar de manera apropiada y decisiva ante situaciones peligrosas
2. Gestionar el estrés laboral	Permanecer en calma frente a la presión, manejar la frustración y ejercer una influencia calmante
3. Resolver problemas de manera creativa	Resolver problemas atípicos, mal definidos o complejos
4. Gestionar situaciones laborales inciertas e impredecibles	Adaptarse a situaciones impredecibles y lidiar con ellas, cambiar el foco de atención, emprender acciones decisivas
5. Aprender nuevas tareas, tecnologías y procedimientos del puesto	Anticiparse, prepararse y aprender destrezas necesarias para futuras exigencias del puesto
6. Demostrar adaptabilidad interpersonal	Adaptar el estilo interpersonal para alcanzar objetivos al trabajar con nuevos equipos, compañeros/as o clientes.
7. Demostrar adaptabilidad cultural	Rendimiento efectivo en diferentes culturas aprendiendo nuevos idiomas, valores, tradiciones y políticas
8. Demostrar adaptabilidad física	Adaptarse a diferentes condiciones físicas, como calor, ruido, climas incómodos y entornos difíciles

¹ Descripciones de actividades laborales suficientemente complejas como para permitir que un observador haga inferencias acerca de la habilidad de un trabajador (Gorriti y López, 2009)

Esta taxonomía fue diseñada para generar una lista exhaustiva de dimensiones del desempeño adaptativo. Sin embargo, los puestos de trabajo o roles organizacionales tienen perfiles diferentes en cuanto a sus requerimientos de adaptabilidad, por lo que en cada caso las dimensiones tendrán una importancia diferente, siendo en ocasiones mínimamente importantes o incluso no aplicables. Además, debido a las diferencias entre dimensiones, es previsible que tengan diferentes predictores, lo que tiene grandes implicaciones en materia de selección de personal (Dorsey et al., 2010; Howe, 2014; Pulakos et al., 2000; Pulakos et al., 2002; Pulakos et al., 2006).

La taxonomía de Pulakos et al. (2000) no está exenta de problemas y el modelo no ha recibido un apoyo conceptual y empírico generalizado. Pese a ello, autores que trabajan desde conceptualizaciones distintas, como Ployhart y Bliese (2006), en el marco del desempeño adaptativo conceptualizado como diferencias individuales, han encontrado un modelo con 8 dimensiones muy similares a las de dicha taxonomía. Así, Griffin y Hesketh (2003) propusieron un modelo más parsimonioso que pretendía reducir las dimensiones de Pulakos et al. (2000) a 2 dimensiones: conductas proactivas y conductas reactivas. No obstante, dada la elevada correlación entre ambas dimensiones también cabría ser considerado como modelo unidimensional (Samale, 2016), algo que coincide con lo apuntado por los propios Pulakos et al. (2002), quienes encontraron que las 8 dimensiones convergían en una única dimensión de desempeño adaptativo. Por otro lado, las investigaciones de Han y Williams (2008) y Griffin y Hesketh (2005) redujeron las 8 dimensiones a las 4 dimensiones (*Gestión de emergencias*, *Gestión del estrés*, *Resolver problemas de manera creativa*, y *Lidiar con situaciones laborales inciertas e impredecibles*) y las 5 dimensiones (*Reactividad ante emergencias*, *Producción creativa*, *Gestión del estrés*, *Esfuerzo de aprendizaje*, y *Adaptación interpersonal*) que, respectivamente, mejor se adaptaban a los contextos laborales objeto de sus estudios.

En vista de estas dificultades y de la falta de aplicabilidad de la dimensión Adaptabilidad física a muchos puestos de trabajo, algunos autores la eliminaron de sus instrumentos de desempeño adaptativo (Griffin y Hesketh, 2003; Marlow et al., 2015). A la misma conclusión llegaron Charbonnier-Voirin y Roussel (2012), que al tratar de desarrollar un instrumento que midiese el desempeño adaptativo a partir del modelo de

Pulakos et al. (2000), encontraron que las dimensiones *Adaptabilidad interpersonal* y *Adaptabilidad cultural* saturaban en un único factor, lo mismo que ocurría con las dimensiones *Gestionar emergencias o situaciones de crisis* y *Gestionar el estrés laboral*. En consecuencia, su modelo se compuso de cinco dimensiones confirmadas mediante análisis factoriales exploratorios y confirmatorios: *Gestión del estrés*; *Esfuerzo formativo y de aprendizaje (en adelante, Esfuerzo formativo)*; *Adaptabilidad interpersonal*; *Reactividad ante emergencias*; y *Creatividad* (Tabla 2.6).

Tabla 2.6.

Dimensiones y Definiciones de la Taxonomía de Charbonnier-Voirin y Roussel (2012)

Dimensión	Definición
1. Gestión del estrés	Capacidad para mantener la compostura y canalizar el estrés del grupo
2. Esfuerzo formativo	Tendencia a iniciar acciones para promover el desarrollo personal
3. Adaptabilidad interpersonal	Capacidad para adecuar el estilo interpersonal para trabajar de manera efectiva con otras personas de la propia organización o empresas aliadas
4. Reactividad ante emergencias o circunstancias inesperadas	Capacidad para gestionar prioridades y adaptarse a nuevas situaciones de trabajo
5. Creatividad	Capacidad del personal para encontrar soluciones o nuevos enfoques ante problemas complejos o previamente desconocidos

La escala de Charbonnier-Roussel y Roussel (2012) ha sido empleada en diferentes investigaciones vinculadas al desempeño adaptativo (Jin y Zhong, 2014; Kanten et al., 2015; Monteiro, 2015; Naami et al, 2014; Zhang et al., 2015) y ha sido apoyada por diferentes revisiones de la literatura en la materia (Pini y Mariani, 2014; Walker, 2015).

2.1.4. Conceptualización del desempeño adaptativo

Para poder conceptualizar mejor el desempeño adaptativo es preciso explorar su red nomológica tanto en lo que se refiere a constructos conceptualmente cercanos como en lo relativo a la diferencia entre esta dimensión de desempeño y el resto de dimensiones del desempeño laboral.

Así, si bien autores como Dorsey et al. (2010) entienden que las 8 dimensiones de Pulakos et al. (2000) subsumen constructos como innovación, creatividad o aprendizaje, existe un gran número de estudios que rebaten esta similitud conceptual (Howe, 2014). Siguiendo a LePine (2003) o Howe (2014), la innovación es inherentemente proactiva, frente a la naturaleza fundamentalmente reactiva del desempeño adaptativo; se trata de un concepto más amplio y que implica más procesos, y además no constituye una condición necesaria para el desempeño adaptativo.

La creatividad, por su parte, puede conceptualizarse como la generación de ideas novedosas (Perry-Smith y Shalley, 2003), de lo que se deduce que difiere del desempeño adaptativo en tres aspectos: la creatividad implica la creación de una idea, mientras que el desempeño adaptativo implica acción; la creatividad está relacionada con la novedad de una idea, algo que no es un requisito para el desempeño adaptativo; y, mientras que el desempeño adaptativo es esencialmente una respuesta al cambio, la creatividad puede ser tanto reactiva como proactiva (Howe, 2014). En ese sentido, en el modelo de Charbonnier-Voirin y Roussel (2012) se identifica una dimensión denominada Creatividad, si bien, como en el modelo de Pulakos et al. (2000), está relacionada con la capacidad de respuesta ante problemas complejos.

En cuanto al aprendizaje, se trata de un constructo eminentemente cognitivo vinculado a cambios en conocimientos o destrezas, que puede facilitar o ser el resultado del desempeño adaptativo, pero no necesariamente tiene que derivar de un cambio en los elementos sociales, ambientales o propios de las tareas del entorno laboral (Howe, 2014).

Más debate ha suscitado la distinción del desempeño adaptativo de dos dimensiones clásicas del desempeño laboral: el desempeño de tarea y el desempeño contextual. Así, Johnson (2001) vinculó algunas dimensiones del modelo de Pulakos et

al. (2000) con las dos dimensiones clásicas: Gestión de emergencias o crisis, Resolución creativa de problemas y Adaptabilidad física con desempeño de tarea, al estar relacionadas con el desempeño de las tareas de un puesto; y Adaptabilidad cultural, Adaptabilidad interpersonal y Gestión del estrés con desempeño contextual, al tratarse de conductas dirigidas al contexto social.

Sin embargo, investigaciones como las de Tucker et al. (2010), señalan que las grandes dimensiones del desempeño (adaptativo, contextual y de tarea), pese a estar correlacionadas, son distintas, lo que coincide con lo señalado anteriormente por Allworth y Hesketh (1999) o Griffin et al. (2007). Siguiendo a Pini y Mariani (2014), un tipo de desempeño puede ser adaptativo en función de las características de la situación, de modo que un desempeño de tarea adaptativo puede implicar cambiar una tecnología en la realización de una tarea, mientras que el desempeño contextual adaptativo puede incluir la ayuda a compañeros/as nuevos/as ante situaciones inesperadas. De hecho, el estudio de Shoss et al. (2012), encontró que la relación entre desempeño adaptativo y desempeño de tarea estaba mediada por variables individuales como el factor Conciencia de personalidad o la política organizacional. En consecuencia, las evidencias empíricas actuales muestran que las dimensiones de la taxonomía de Pulakos et al. (2000) no están subsumidas ni en el desempeño de tarea ni en el contextual y son diferentes conceptualmente. Por tanto, el desempeño adaptativo debe ser considerado al mismo nivel que dichas dimensiones “clásicas” del desempeño laboral (Mariani, 2011; Cortina y Luchman, 2012).

2.1.5. Medición del desempeño

Para medir el desempeño existen medidas “objetivas” como los registros organizacionales. En la Administración podrían tenerse en cuenta aspectos como el cumplimiento de plazos para emitir informes, el número de informes realizados, el número de días de absentismo laboral, etc. Estos registros organizacionales computan resultados medibles, observables, lo que disminuye los sesgos de la evaluación subjetiva; sin embargo, pueden estar afectados por contaminación de criterio y deficiencia de criterio. La contaminación se debe al hecho de que los resultados pueden deberse a factores que están fuera del control del individuo (carga excesiva de trabajo, varias bajas en el servicio al mismo tiempo); la deficiencia se debe a que los resultados evaluados pueden no tener en cuenta aspectos importantes del desempeño individual.

Este hecho hace que sea preferible optar por las medidas subjetivas (las obtenidas mediante escalas, cuestionarios cumplimentados por observadores, etc.), siendo las evaluaciones de los y las superiores la manera más habitual de medir el criterio en estudios de validación (Baard et al., 2014; Gorriti y López, 2010; Pulakos et al., 2012; Quijano, 1997).

Tradicionalmente se ha recomendado que sean los y las superiores quienes evalúen el desempeño de tarea y que compañeras y compañeros se encarguen de la evaluación contextual (con todas las reservas de halo y sesgos que tiene este tipo de evaluación). Según los datos meta-analíticos que aportan Viswesvaran et al. (1996), los y las superiores evalúan de forma más consistente, más fiable, que los compañeros de trabajo, encontrando que la fiabilidad inter-jueces en el caso de los superiores era de .52, mientras que la de los/las compañeros/as era de .42.

No obstante, en lo relativo al desempeño contextual y a las OCBs, las investigaciones de Demerouti et al. (2014) o Carpenter et al. (2014) indican que las auto-evaluaciones recogen de manera más precisa las OCBs realizadas y que compañeros/as o superiores no pueden observar todos los aspectos de las conductas ajenas. Asimismo, según Demerouti et al. (2014), las auto-evaluaciones explican más varianza del desempeño contextual que las evaluaciones de compañeros/as de trabajo, de lo que se deduce que las auto-evaluaciones deberían ser la manera preferible de medir las conductas vinculadas a este tipo de desempeño laboral.

En la misma línea y en lo que se refiere al desempeño adaptativo, Pulakos et al. (2012) señalaron que las auto-evaluaciones pueden ser importantes para percibir conductas que previsiblemente no se llevan a cabo delante de un superior, dado que, por ejemplo, las personas son reacias a mostrar que han trabajado duro ante cambios del entorno, prefiriendo dar la impresión de que tienen un talento natural (Dweck, 1986, 2002). Otra razón es que pueden ocultar sus conductas adaptativas ante la necesidad de aprender tareas nuevas o manejar el estrés laboral, lo que hace que los y las superiores no tengan una visión precisa sobre cómo manejan sus subordinados/as este tipo de situaciones. Consecuentemente, las auto-evaluaciones en estudios de validación pueden reflejar mejor que las evaluaciones de los y las superiores las conductas de desempeño adaptativo. Asimismo, este tipo de evaluaciones forman parte de las evaluaciones de

360° que autores como Baard et al. (2012) demandan para mejorar la medición del desempeño adaptativo y evaluar de manera más precisa su estructura factorial.

2.2. Revisión de instrumentos que evalúan el desempeño adaptativo

A lo largo de los últimos 20 años diferentes autores han desarrollado instrumentos para evaluar el desempeño adaptativo individual. Uno de los primeros intentos fue el Graphic Distributional Performance Rating (GDPR) de Hesketh y Feiler (1994), que Allworth y Hesketh (1999) utilizaron para medir desempeño de tarea, contextual y adaptativo. Cuatro de sus 16 ítems medían desempeño adaptativo, aunque no aportaron estadísticos de ajuste del modelo.

Poco tiempo después, Pulakos et al. (2000) crearon el Job Adaptability Inventory (JAI), de 132 ítems en su versión extensa y 68 en su versión breve, y que mide las 8 dimensiones del modelo de estos autores, centrándose en los requisitos de adaptabilidad de los puestos de trabajo. Desafortunadamente, los mencionados autores no han puesto a disposición pública los ítems del JAI. En una línea similar, Griffin y Hesketh (2003) desarrollaron un instrumento de 20 ítems a partir de las definiciones de 7 de las 8 dimensiones de Pulakos et al. (2000). La dimensión Adaptabilidad física no fue considerada relevante para las organizaciones en las que se realizó su investigación y la dimensión Adaptabilidad cultural se transformó en Adaptabilidad a la cultura transdepartamental o laboral. Su propósito era someter a comprobación un modelo de tres dimensiones: conductas adaptativas, reactivas y tolerantes, pero en base a sus resultados su modelo recibió un apoyo moderado, dado que el factor de conductas tolerantes no fue corroborado por los datos del estudio. Griffin y Hesketh (2004, 2005) emplearon instrumentos de 18 y 20 ítems centrados en la forma en la que los colaboradores ejecutan una serie de tareas que requieren adaptabilidad basándose una vez más en las dimensiones de Pulakos et al. (2000), aunque su contenido no es público y la información sobre sus características psicométricas es escasa.

Han y Williams (2008) desarrollaron un instrumento de 12 ítems que medía 4 de las 8 dimensiones de Pulakos et al. (2000): Gestión de emergencias, Resolución creativa de emergencias, Gestión del estrés laboral y Lidar con la incertidumbre. Sin embargo, como indican Marques-Quinteiro et al. (2015), no hicieron públicos sus ítems y no analizaron en profundidad las propiedades psicométricas de su instrumento.

Por otro lado, otros instrumentos han sido diseñados para ser empleados en contextos concretos, como los sectores comercial o militar. Entre otros, cabe mencionar la escala de 6 ítems elaborada por De Jong y De Ruyter (2004) para evaluar el ajuste de los empleados de banca a sus clientes, el instrumento de Chen et al. (2005) para medir el desempeño de pilotos en simulaciones de vuelo, o la escala de Joung et al. (2006) que, aplicando simulaciones de situaciones de guerra, evaluaba el número de problemas identificados y las respuestas viables generadas (Charbonnier-Voirin y Roussel, 2012).

Dentro del marco del desempeño adaptativo entendido como diferencia individual, destacan dos instrumentos con un mayor número de ítems. Por un lado, Ployhart y Bliese (2006) crearon, a partir del modelo de Pulakos et al. (2000), el I-ADAPT-M, de 55 ítems que están accesibles en el dominio público, si bien se echa en falta información detallada con respecto al desarrollo y la validación de la escala (Howe, 2014). Por otro lado, cabe mencionar el Measure of Adaptive Performance (MAP), desarrollado a partir del Adaptive Performance Scale (Frame et al., 2006). El MAP se basa en el modelo de Pulakos et al. (2000) y en el trabajo de Griffin y Hesketh (2003), y ha tenido diferentes versiones con un número de ítems que oscila entre 41 y 63 ítems y que, según el caso, ha identificado 4, 7 o 9 dimensiones de desempeño adaptativo (Lillard et al., 2012; Marlow et al., 2015; Watts et al., 2011).

D-ADAPT (Oprins et al., 2018), por su parte, es un instrumento holandés de 31 ítems creado a partir de JAI y I-ADAPT-M y que mide 6 dimensiones: Gestión de situaciones de crisis, Resolución de problemas difíciles, Resolución de situaciones culturalmente exigentes, Resolución de circunstancias físicamente exigentes, Gestión del estrés laboral e Interacciones interpersonales. Se diseñó inicialmente para el entorno militar, aunque puede ser empleado en otros ámbitos. Permite evaluar los requerimientos de adaptabilidad de un puesto, así como llevar a cabo la autoevaluación del personal con respecto a su capacidad de adaptación. Sus autores únicamente informan de la fiabilidad de las dimensiones.

Kröger y Staufenbiel (2012) desarrollaron un instrumento de 18 ítems en alemán basado en el cuestionario JAI y en el trabajo de Griffin y Hesketh (2003). El análisis factorial confirmó la estructura propuesta de dos dimensiones: desempeño adaptativo relacionado con la tarea y desempeño adaptativo social. Esta última dimensión se componía a su vez de dos de las dimensiones del modelo de Pulakos et al. (2000):

Adaptación interpersonal y Adaptación intercultural a costumbres, tradiciones y perspectivas extranjeras. Sin embargo, la Adaptación intercultural es una subdimensión difícilmente aplicable a un gran número de puestos de trabajo.

Más cerca de nuestras fronteras, Marques-Quinteiro et al. (2015) desarrollaron un instrumento para medir desempeño adaptativo tanto individual como de equipo con una muestra portuguesa. El desempeño individual se evaluó mediante 8 ítems que medían 4 de las dimensiones de Pulakos et al. (2000): Resolución creativa de problemas, Lidiar con situaciones laborales imprevisibles e inciertas, Aprender tareas, procedimientos y tecnologías laborales y Gestionar el estrés laboral.

Ramírez (2013), por su parte, desarrolló un instrumento (PRO-4) en castellano que mide cuatro dimensiones del desempeño laboral individual: desempeño de tarea, contextual, adaptativo y proactivo. PRO-4 está compuesto por 29 ítems, 7 de los cuales miden desempeño adaptativo.

De todo lo antedicho y siguiendo a Charbonnier-Voirin y Roussel (2012), cabe concluir que la mayoría de instrumentos son demasiado extensos o demasiado breves para evaluar adecuadamente el dominio del desempeño adaptativo, sus ítems no están disponibles, falta información sobre sus propiedades psicométricas, están enfocados a un contexto muy específico o no son generalizables a todo tipo de puestos. Por consiguiente, es necesario disponer de una escala adecuada para medir las diferentes facetas del constructo de desempeño adaptativo, con buenas propiedades psicométricas y accesible para quienes se dedican a la investigación. La escala multidimensional desarrollada por estas autoras para medir el desempeño adaptativo ha recibido un apoyo considerable debido a sus características psicométricas, a que mide 5 dimensiones basadas en el modelo de Pulakos et al. (2000) aplicables a todo tipo de puestos y a que es una escala breve, ya que está compuesta por 19 ítems (Marques-Quinteiro et al., 2015; Pini y Mariani, 2014; Walker, 2015).

Inicialmente fue desarrollada en francés por Charbonnier-Voirin y Roussel, que desarrollaron una versión en inglés (2012). La escala original fue aplicada en una muestra de 524 personas empleadas en empresas de telecomunicaciones, servicios a personas y sector aeronáutico. Los valores de Alfa de Cronbach de las 5 dimensiones oscilaron entre .70 y .88. Los índices Rho de Jöreskog oscilaron entre .77 y .85. Estos

valores de fiabilidad son aceptables, especialmente considerando que cada escala consta de un máximo de 4 ítems. Los índices de ajuste al modelo teórico obtenidos en la fase de análisis factorial confirmatorio mostraron un ajuste satisfactorio: $\chi^2/g.l.=2.87$, RMSEA=.06, GFI=.92, AGFI=.90, CFI=.97.

En consecuencia, dada la necesidad de evaluar de manera fiable y válida el desempeño adaptativo y la escasez de instrumentos que permiten evaluarlo en castellano, resulta preciso adaptar a dicha lengua y cultura un instrumento con las características que presenta la herramienta de Charbonnier-Voirin y Roussel (2012), lo que constituye el Objetivo 1 de la presente tesis. En este capítulo se detalla el proceso de adaptación y validación del mismo.

2.3. Adaptación del Cuestionario de desempeño adaptativo

2.3.1. *Proceso de traducción de los ítems, entrevistas cognitivas y fase piloto*

La adaptación a la lengua y cultura castellanas de la escala de Charbonnier-Voirin y Roussel se llevó a cabo teniendo en cuenta las recomendaciones de Hambleton y Patsula (1999), que propusieron una serie de pasos para evitar posibles fuentes de sesgo y mejorar la calidad en la adaptación de los tests. Dichos pasos son: seleccionar un equipo de traducción especializado, elegir un diseño de traducción adecuado, realizar las correcciones precisas en la adaptación lingüística inicial del instrumento, llevar a cabo un estudio piloto con el test adaptado, y aplicar el test a una muestra amplia y representativa de la población objetivo (Balluerka et al., 2007).

De esta manera, la traducción de los ítems de la escala de Charbonnier-Voirin y Roussel se realizó utilizando un diseño de traducción inversa basada en la versión en inglés, si bien la versión original en francés se tomó como referencia en todo momento. En la adaptación participaron una traductora, cuatro psicólogos/as expertos/as en pruebas de personalidad y dos metodólogas. Tras llegar a una versión consensuada traducida al castellano y, a partir de ésta, a una versión retro-traducida consensuada, las personas integrantes del equipo de traducción compararon cada uno de los ítems de las versiones original e inversamente adaptadas del instrumento, con el objetivo de detectar posibles no-equivalencias en cuanto al significado. Partiendo de este análisis se llevaron a cabo las modificaciones oportunas en la versión en castellano de la escala.

Siguiendo la propuesta de Hambleton y Patsula (1999), tras la elaboración de la versión adaptada consensuada por el equipo multidisciplinar de traducción y antes de proceder a la aplicación del instrumento a una muestra amplia de participantes con fines de tipificación, se realizaron entrevistas cognitivas a una serie de personas y se llevó a cabo un estudio piloto a fin de recabar datos empíricos acerca del funcionamiento de la versión en castellano de la escala de Charbonnier-Voirin y Roussel.

Se verificó la adecuación de los ítems del cuestionario a la población objetivo mediante la realización de entrevistas cognitivas ($N = 10$) a 5 varones y 5 mujeres (Media de edad = 48.6 años; $SD = 9.70$). En dichas entrevistas se exploró, de manera individual y empleando un guion de entrevista, el grado de comprensión del instrumento, de sus ítems, de las instrucciones y de la tarea a realizar. Asimismo, se recogieron sus interpretaciones sobre los ítems del cuestionario, profundizando en aquellos que generaron más debate en el proceso de traducción.

Analizada esta información se elaboraron los cuestionarios para la aplicación piloto. En el estudio piloto participaron 80 personas (31 varones y 49 mujeres) con edades comprendidas entre los 23 y los 63 años ($M = 46.01$ años; $SD = 7.72$). Tanto en las entrevistas como en la aplicación piloto la muestra estaba compuesta por empleados públicos de la Administración Pública Vasca y los datos fueron recogidos en los centros de trabajo de las y los participantes bajo condiciones que preservaban su anonimato.

Finalmente, se procedió al análisis cuantitativo y cualitativo de los ítems y a la reformulación de los mismos tras dicha aplicación piloto. Así, con los datos recogidos en esta fase, se calcularon las puntuaciones medias y las desviaciones típicas de cada uno de los ítems, así como los índices de homogeneidad corregidos con respecto a sus correspondientes facetas. Aquellos ítems que presentaron un índice de homogeneidad menor que .30 (Ítem 10) fueron revisados por los integrantes del equipo de traducción, quienes, siguiendo las sugerencias de dos expertas en herramientas de evaluación psicológica, llevaron a cabo algunas modificaciones a fin de que se aproximaran en mayor medida a su correspondiente dimensión teórica. El ítem 10 pertenecía a la dimensión *Gestionar el estrés*, que presentó un Alfa de Cronbach inferior a .70, con un valor de .60.

Dado que en la versión original del instrumento dicha dimensión estaba compuesta por solo 3 ítems y que un ítem problemático podía comprometer las propiedades psicométricas de la misma, se consideró oportuno añadir 3 ítems más a dicha dimensión para la versión a emplear en la muestra de validación, los ítems 20, 21 y 22. Otra razón para añadir estos 3 ítems es que, a diferencia del contenido de los ítems 20, 21 y 22, el contenido de los ítems originales 10 y 15 podría vincularse a la dimensión Adaptabilidad Interpersonal. Estos 3 ítems adicionales fueron adaptados a partir de ítems empleados en el desarrollo de la versión inicial de la escala de Charbonnier-Voirin y Roussel (2012). Dos de esos tres ítems añadidos a la escala en castellano presentaban la particularidad de estar redactados en sentido inverso al resto de ítems, es decir, un alto grado de acuerdo con los mismos implicaba poca capacidad de gestión del estrés, lo que podría reducir la tendencia a la aquiescencia en las respuestas.

Por otra parte, se tomaron en consideración las sugerencias que realizaron los y las participantes en las entrevistas cognitivas y la aplicación piloto. Dichas sugerencias se centraron básicamente en la complejidad o dificultad de comprensión de algunos términos incluidos en determinados ítems. En este caso, el equipo de traductores/as reformuló los elementos a fin de facilitar su comprensión o añadió sinónimos a los términos empleados en la aplicación piloto (Ítems 13, 16, 18). El instrumento derivado de este proceso de revisión fue el que se utilizó para realizar la validación empírica rigurosa y la tipificación de la versión de la escala de Charbonnier-Voirin y Roussel adaptada al castellano

2.3.2. Método

2.3.2.1. Participantes

El análisis de las propiedades psicométricas de la versión en castellano de la escala de Charbonnier-Voirin y Roussel se realizó con una muestra incidental de 678 empleados/as públicos/as (444 mujeres y 234 varones) pertenecientes a diferentes entidades públicas ubicadas en las tres provincias de la Comunidad Autónoma del País Vasco (CAPV). Las características de la muestra se resumen en la Tabla 2.7.

Tabla 2.7.
Información Sociodemográfica de las Personas Participantes

Variables	N	%
Género		
Mujeres	444	65.49
Hombres	234	34.51
Lengua materna		
Euskera	318	49.23
Castellano	328	50.77
Desempeño puestos de mando		
Puesto con funciones de mando	213	31.84
Subordinado	456	68.16
Nivel de complejidad del puesto		
Alta (requisito de acceso: titulación universitaria)	384	56.64
Baja (requisito de acceso: educación primaria o secundaria)	294	43.36
Nivel académico		
Nivel alto (Estudios universitarios)	540	79.65
Nivel bajo (Educación primaria o secundaria)	138	20.35
Grado de estabilidad del puesto desempeñado		
Fijo	245	36.51
Temporal	426	63.49
Grupos de Edad		
Grupo1 (21 a 40 años)	207	30.58
Grupo 2 (41 a 50 años)	304	44.90
Grupo 3 (51 a 65 años)	166	24.52
Experiencia en la organización		
Alta (>= 2 años)	129	19.03
Baja (< 2 años)	549	80.97
Experiencia en el puesto		
Alta (>= 2 años)	198	29.20
Baja (< 2 años)	480	70.80
	Media (SD)	Rango
Edad	44.63 (7.66)	21-63 años
Años en la organización	11.46 (9.71)	.08-43 años
Años en el puesto	7.71 (7.82)	.08-39 años

Para reclutar a las personas participantes, el Instituto Vasco de Administración Pública (IVAP) solicitó su colaboración a diferentes entidades públicas vascas, aceptando participar en el estudio un total de 32 ayuntamientos, Diputaciones Forales o entidades públicas vascas, además de 16 departamentos o entidades pertenecientes al Gobierno Vasco. Cada entidad designó a una persona encargada de reclutar personas voluntarias con las que se acordaron la fecha y lugar de las sesiones de recogida de datos, que tuvieron lugar en las correspondientes sedes de las entidades participantes. Este procedimiento ayudó a que la tasa de respuesta fuera elevada: participaron 678 personas (96%) de las 703 apuntadas inicialmente.

2.3.2.2. Instrumentos

Se han empleado, además de la adaptación al castellano del cuestionario de desempeño adaptativo de Charbonnier-Voirin y Roussel (2012), un cuestionario de variables sociodemográficas y laborales (género, edad, lengua materna, desempeño de puestos de mando, nivel de complejidad del puesto, grado de estabilidad en el puesto, titulación académica, años de experiencia en el puesto y la organización) y una serie de instrumentos para evaluar su red de relaciones nomológicas: el cuestionario de personalidad OPERAS (Vigil-Colet et al., 2013) en su versión en euskera, adaptada en el marco de esta tesis; la escala para evaluar OCBs de Lee y Allen (2002); y la adaptación al castellano del cuestionario desarrollado por Piasentin (2007) para medir ajuste persona-organización (APO). Dado que en este capítulo no se describen las relaciones entre el cuestionario adaptado al castellano y los instrumentos destinados a evaluar su red de relaciones nomológicas, cuestiones que se abordan en el capítulo 4, nos remitimos a dicho capítulo para una descripción detallada de los mencionados instrumentos.

2.3.2.3. Procedimiento

Se administró el instrumento, junto con las pruebas necesarias mencionadas anteriormente para el proceso de validación y para evaluar la red nomológica de la prueba. El orden de aplicación fue el siguiente: Cuestionario de variables sociodemográficas y laborales, cuestionario de personalidad OPERAS de Vigil-Colet et al. (2013), Escala de Lee y Allen para evaluar OCBs (2002), Escala para medir APO de Piasentin (2007), y Escala de Charbonnier-Voirin y Roussel (2012).

Se realizó una segunda aplicación al cabo de un mes para evaluar la fiabilidad test-retest de la escala de Charbonnier-Roussel y Voirin en castellano. Para ello se utilizó una submuestra de 151 personas.

Se contó con el informe favorable del Comité de Ética para las Investigaciones relacionadas con Seres Humanos (CEISH/GIEB) de la UPV/EHU, por lo que previamente se recogieron los consentimientos informados de las personas participantes. Las y los participantes cumplimentaron los cuestionarios en sus respectivos centros de trabajo de manera grupal. En cada sesión de aplicación se explicó el propósito de la investigación y en qué consistía el trabajo a realizar, respondiendo a las dudas planteadas. Para garantizar el anonimato, las personas participantes tuvieron que elaborar, al comienzo de cada sesión, una clave de identificación de 5 cifras, las tres primeras correspondientes a un número de teléfono (móvil, laboral...) y las dos últimas referidas al día de nacimiento.

2.3.2.4. Análisis de datos

Una vez codificados los datos, se realizaron los análisis de las propiedades psicométricas del cuestionario mediante los siguientes procedimientos:

a) Análisis de los ítems del instrumento: estadísticos descriptivos (media, desviación típica, asimetría y curtosis), índice de homogeneidad (correlación total de elementos corregida, es decir, la correlación entre un determinado elemento y el factor al que éste pertenece excluyendo dicho elemento), porcentajes de elección de las opciones de respuesta.

En lo que se refiere a la media, si esta supera significativamente el promedio teórico de 4 de la escala Likert de 7 puntos empleada en este instrumento (siendo 1 “*muy en desacuerdo*” y 7 “*muy de acuerdo*”), el elemento puede ser etiquetado como atractivo en la muestra, deseable socialmente. Cabe indicar que, para facilitar la comprensión de las medias de los ítems, se proporcionan los datos de las mismas a partir de la transformación de los ítems inversos (21 y 22), de forma que en todos los casos las mayores puntuaciones en el ítem se relacionen con un mayor desempeño adaptativo.

De manera complementaria, y siguiendo las recomendaciones de Clark y Watson (2019), se examinó el comportamiento de los ítems del test desde la óptica de la Teoría de Respuesta al Ítem (TRI). A tal fin se analizó la utilidad de los ítems para el propósito de la evaluación y en qué medida contribuyen a cumplir con los objetivos de evaluación definidos en el propio test empleando el programa IRTPRO, publicado por ssicentral.com. Una de las principales características de la TRI es que proporciona estimaciones invariantes tanto de las propiedades psicométricas de los ítems (su nivel de discriminación o dificultad) como de las características de los sujetos. En consecuencia, los parámetros que caracterizan al ítem y al test son menos dependientes de la muestra particular de participantes utilizada y los parámetros que caracterizan al sujeto no dependen de la muestra particular de ítems utilizada (Hidalgo-Montesinos y French, 2016).

Un modelo de TRI que se puede aplicar a escalas que utilizan varias opciones de respuesta es el Modelo de Respuesta Graduada (MRG). Este modelo implica que las opciones en un ítem (por ejemplo, 1, 2, 3, 4, 5, 6 y 7 si se emplea una escala Likert de 7 puntos, como en este caso) están ordenadas, y que una puntuación o respuesta más alta de un sujeto indica un nivel más alto en la dimensión evaluada. En el MRG, la capacidad del ítem para discriminar entre los niveles del rasgo latente (niveles de θ) se mantiene constante mientras que la dificultad del ítem se establece en cada “paso del ítem”, el umbral entre cada punto de la escala Likert (Hidalgo-Montesinos y French, 2016; Penfield, 2014).

Un paso inicial en la adaptación de un instrumento es analizar los ítems con el objetivo de identificar qué ítems funcionan de forma apropiada en cada escala y qué ítems pueden necesitar una revisión o incluso ser sustituidos por otros. Un ítem que funciona bien debe discriminar a los sujetos evaluados en todos los niveles de θ , lo que implica que debe presentar un valor del parámetro a (nivel de discriminación) relativamente alto, además de utilizar todas las opciones de respuesta (cada curva de respuesta debe distribuirse en distintos valores de θ ; en este caso, entre -3 y 3 desviaciones típicas, siendo 0 la media poblacional).

Otra característica esencial de la TRI que se analizó es la Curva Característica del Ítem (CCI), que aporta información gráfica acerca del funcionamiento del ítem en el continuo de la característica medida y para cada una de sus opciones de respuesta.

b) Análisis de la dimensionalidad de la prueba y de su validez de constructo. Para analizar cómo se adecúan los datos empíricos al modelo teórico de partida, las cinco dimensiones del modelo de Charbonnier-Voirin y Roussel basadas en el modelo de Pulakos et al. (2000), se empleó el análisis factorial confirmatorio (AFC). El AFC permite estudiar el ajuste de lo empírico a lo modelizado dentro de unos parámetros numéricos interpretables.

No obstante, en medidas empleadas en Psicología del Trabajo es habitual que los datos no presenten un ajuste adecuado a la teoría en análisis factoriales confirmatorios porque habitualmente los ítems no son indicadores puros de un sólo factor y se dan correlaciones elevadas entre dimensiones (Marsh et al., 2014; Vassend y Skrandal, 2011). Asimismo, Marsh et al. (2005) sugieren que los valores de corte tradicionalmente aceptados pueden ser demasiado estrictos para un AFC a nivel de ítem relativo a un instrumento multidimensional.

Por esta razón, se complementó el AFC con modelos de ecuaciones estructurales exploratorios (Exploratory Structural Equation Models – ESEM en inglés), que son menos restrictivos a la hora de medir la estructura factorial e integran los mejores aspectos de los AFCs, los modelos de ecuaciones estructurales y el Análisis Factorial Exploratorio, proporcionando tests confirmatorios de estructuras factoriales a priori (Booth y Hugues, 2014; Marsh et al., 2010; Marsh et al., 2014).

Por otro lado, para evaluar la dimensionalidad del instrumento se analizaron las saturaciones de los ítems en las dimensiones, así como las correlaciones entre las dimensiones y entre los ítems del instrumento. En el análisis de las correlaciones entre los ítems Briggs y Cheek (1986) señalaron que la media de las correlaciones inter-ítem de cada dimensión teórica aporta información acerca de la unidimensionalidad de la misma. Según estos autores, el nivel óptimo de homogeneidad se da cuando la media está entre .2 y .4, siendo recomendable que los valores de las correlaciones se encuentren entre .15 y .50. Un valor inferior a .1 indica que es improbable que el resultado total en la escala pueda representar adecuadamente la complejidad de los

ítems. Una media superior a .5, por su parte, indicaría que los ítems de una escala son muy redundantes y el constructo medido muy específico.

c) Evaluación de la fiabilidad. Siguiendo el ejemplo de la escala original (Charbonnier-Voirin y Roussel, 2012), se tuvieron en cuenta el estadístico Rho de Jöreskog, y el Alfa de Cronbach para evaluar la fiabilidad basada en la consistencia interna. No obstante, dado el escaso número de ítems de escalas breves como la de Charbonnier-Voirin y Roussel, las correlaciones inter-ítems no pueden ser altas, lo que a su vez conlleva una baja consistencia interna si se mide mediante Alfa de Cronbach. Como alternativa complementaria al alfa de Cronbach, se ha propuesto el Omega de McDonald (Schipolowski et al., 2014; Viladrich et al., 2017), que se ha calculado junto con el índice de fiabilidad marginal basado en la Teoría de Respuesta al Ítem (TRI) ofrecida por el programa IRTPRO.

Adicionalmente, y siguiendo la recomendación de Muñiz y Fonseca-Pedrero (2019), se calculó la función de información desde la perspectiva de la TRI. En la TRI la fiabilidad está vinculada a la función de información, es decir, cuanta más información proporciona un test, una dimensión o un ítem, mayor es la precisión en la estimación del rasgo o característica del individuo. Asimismo, en el marco de la TRI la precisión puede ser evaluada en cada nivel del rasgo psicológico que está siendo medido y, por lo tanto, puede no ser constante.

Fiabilidad test-retest: para evaluar la estabilidad temporal de los factores y de las dimensiones se utilizó la técnica test-retest, empleando el coeficiente de correlación de Pearson (Leppink y Pérez-Fuster, 2017; Rammstedt y Beierlein, 2014).

d) Realización de estudios de funcionamiento diferencial de ítems con distintos subgrupos de participantes en función de su lengua materna (euskera o castellano) y de la complejidad del puesto que desempeñaban. El propósito de estos análisis fue examinar si el lenguaje mediante el cual estaban expresados los ítems podía favorecer que los individuos educados en una determinada lengua o que desempeñaban un puesto de mayor o menor complejidad obtuviesen puntuaciones de mayor o menor valor en alguno de los ítems de cada dimensión.

El análisis del funcionamiento diferencial de los ítems (FDI) fue llevado a cabo mediante el procedimiento de regresión logística (Swaminathan y Rogers, 1990) dentro de cada una de las dimensiones del instrumento de Charbonnier-Voirin y Roussel (2012). Se empleó la sintaxis específica para la detección del FDI mediante SPSS creada por Slocum et al. (2004).

Para detectar el FDI mediante este método se realizaron análisis empleando una regresión logística en tres pasos, en la que se calcula la cantidad de varianza de las respuestas de los participantes que puede explicarse por las variables incluidas en el modelo en cada paso y en los pasos anteriores. En el primer paso se calcula la varianza explicada por la puntuación total en la dimensión a la que pertenece el ítem. En el segundo paso del modelo se calcula la cantidad de varianza explicada tanto por la puntuación total en la dimensión como por la pertenencia al grupo de comparación (lengua materna o nivel de complejidad del puesto desempeñado). En el tercer paso del modelo se calcula la cantidad de varianza explicada tanto por las variables del paso 2 como por la interacción entre la puntuación total en la dimensión correspondiente y la pertenencia al grupo. La cantidad de varianza explicada en cada paso se representa por un valor de R^2 de Nagelkerke que se puede comparar en los diferentes pasos para detectar la presencia de FDI. Para determinar la presencia de FDI no uniforme, se comparan los valores de R^2 de Nagelkerke del paso 3 y del paso 1. Una segunda comparación de los valores de R^2 de Nagelkerke en los pasos 2 y 1 permite determinar la presencia de DIF uniforme. Según los criterios de Jodoin y Gierl (2001), para que se considere que un ítem presenta FDI, el tamaño del efecto debe tener un valor de R^2 de, al menos, .035 (Whiteman, 2011).

Los análisis se realizaron mediante los programas informáticos Factor, SPSS, MPlus e IRTPRO. En el siguiente epígrafe se describen los resultados obtenidos en los análisis realizados, a excepción de la asociación del constructo evaluado por el instrumento con otras variables, cuestión que se describe en el capítulo 4 de la presente tesis.

2.3.3. Resultados

2.3.3.1. Análisis de los ítems del instrumento

Se realizaron varios análisis para estudiar las características psicométricas de los elementos de la adaptación. En la tabla 2.8 se proporcionan los estadísticos descriptivos de los elementos de la escala de Charbonnier-Voirin y Roussel. Las medias de los ítems 21 y 22, sin revertir, fueron de 3.54 y 3.5, respectivamente, mientras que los ítems transformados (21R y 22R en la Tabla 2.8) presentaron una media de 4.45 y 4.5. Todos los ítems presentaron valores inferiores a los límites que indican problemas de normalidad, 2 para asimetría y 7 para curtosis (Curran et al., 1996). Los valores de las desviaciones típicas superaron el valor deseable de 1, y los pocos que no lo hicieron (ítems 6, 11, 12, 14, 15, 18) se aproximaron mucho a dicho valor.

Asimismo, en la tabla 2.8 se proporcionan los índices de homogeneidad corregidos (correlación total de elementos corregida) referidos a cada una de las 5 dimensiones de la escala. El índice de homogeneidad corregido se calculó para los 22 elementos de la escala y solo fueron problemáticos los ítems 13 y 14 ($< .30$).

Tabla 2.8.

Estadísticos Descriptivos de los Ítems de la Escala de Charbonnier-Voirin y Roussel e Índice de Homogeneidad

Ítem	Media	Desviación típica	Asimetría	Curtosis	Correlación total de elementos corregida
1	5.50	1.21	-1.17	1.45	0.44
2	5.73	1.21	-1.25	2.81	0.50
3	6.17	1.15	-1.20	2.17	0.53
4	5.72	1.27	-1.28	1.78	0.58
5	5.61	1.00	-1.09	1.52	0.53
6	5.41	0.95	-0.95	0.88	0.30
7	5.24	1.24	-0.74	0.53	0.43
8	6.14	1.29	-1.41	4.21	0.63
9	5.56	1.60	-1.07	1.24	0.59
10	5.71	1.71	-1.05	1.88	0.43
11	5.44	0.86	-0.77	0.63	0.31
12	5.40	0.80	-0.75	0.58	0.36
13	5.49	2.55	-0.96	1.49	0.19
14	5.66	0.83	-0.90	0.90	0.26
15	4.68	0.92	-0.38	0.02	0.49
16	4.81	1.12	-0.52	0.16	0.69
17	5.65	1.06	-0.88	1.31	0.66
18	5.83	0.94	-0.86	1.72	0.51
19	5.61	1.20	-1.04	1.65	0.63
20	5.07	1.17	-0.63	-0.29	0.64
21R	4.45	1.12	-0.18	-1.09	0.71
22R	4.50	1.04	-0.28	-1.11	0.52

A continuación, en la tabla 2.9 se proporciona, de cada una de las opciones de respuesta puntuables, el porcentaje de atracción de cada alternativa, es decir, el porcentaje de casos que la seleccionaron en cada ítem. Se observa que las opciones más seleccionadas fueron los valores entre 4 y 7 de la escala Likert.

Tabla 2.9.

Porcentajes de Elección de las Opciones de Respuesta en cada Ítem de la Escala de Charbonnier-Voirin y Roussel

Ítem	1	2	3	4	5	6	7
1	0.59	2.67	4.59	8.00	24.00	43.11	17.04
2	0.15	0.74	1.63	6.37	20.74	55.56	14.81
3	0.00	0.44	0.15	4.15	12.15	43.41	39.70
4	0.44	2.37	3.56	6.22	20.00	41.04	26.37
5	0.00	1.04	4.00	6.22	24.59	50.07	14.07
6	0.59	2.52	3.11	15.70	20.59	42.52	14.96
7	0.30	1.48	6.37	13.19	32.44	36.89	9.33
8	0.00	0.44	1.04	1.19	11.70	52.74	32.89
9	0.44	1.78	4.15	9.63	22.07	42.96	18.96
10	0.15	0.44	2.52	6.37	23.26	50.52	16.74
11	0.30	1.78	3.11	14.22	26.52	37.33	16.74
12	0.15	0.89	4.30	12.74	28.59	41.93	11.41
13	0.44	0.74	2.96	11.26	26.96	44.74	12.89
14	0.30	0.89	3.26	9.93	23.26	38.81	23.56
15	1.04	4.74	7.26	32.30	26.07	23.85	4.74
16	1.33	3.70	8.15	24.74	29.78	25.63	6.67
17	0.00	0.59	2.52	6.22	27.85	47.85	14.96
18	0.00	0.30	1.19	3.85	23.41	52.89	18.37
19	0.30	1.04	2.37	9.48	23.70	46.22	16.89
20	0.15	3.56	12.15	10.96	29.33	33.78	10.07
21R	1.93	10.67	21.93	13.33	18.07	25.78	8.30
22R	3.85	10.22	21.19	10.67	14.96	29.19	9.93

En la Tabla 2.10 se presentan los parámetros estimados junto a sus errores típicos para los ítems que componen cada escala. Hay que tener en cuenta que había 6 estimaciones para el parámetro b , una para cada umbral, dado que el test cuenta con 7 opciones de respuesta, es decir, $k-1$ umbrales. La segunda columna de la tabla contiene el parámetro a (discriminación del ítem), y se estima que .65 es el umbral mínimo para que un ítem tenga un funcionamiento aceptable, en base a lo establecido por Baker (2001). Las columnas 5, 7, 9, 11, 13 y 15 contienen los parámetros umbrales referidos al nivel mínimo para pasar de una respuesta de 1 a 2, de 2 a 3, de 3 a 4, de 4 a 5, de 5 a 6 y de 6 a 7 respectivamente (b_1 , b_2 , b_3 , b_4 , b_5 y b_6). Las tres últimas columnas de la tabla informan de un valor de chi-cuadrado, sus grados de libertad y su probabilidad asociada (p), que hacen referencia a una prueba de ajuste estadístico del ítem al MRG (Hidalgo-Montesinos y French, 2016; Orlando y Thissen, 2000, 2003). Al igual que en las tablas anteriores, los ítems 21 y 22 se transformaron en ítems positivos para facilitar la comprensión de los datos.

Los resultados muestran que todos los ítems superaron el umbral de .65 para la discriminación, y, en términos de Baker (2001) se puede considerar que tienen un nivel de discriminación alto (valores del parámetro $a > 1.7$). Los únicos ítems que presentaron un mal ajuste al modelo MRG ($p < .01$) fueron los ítems 10, 13, 17 y 20, que sin embargo presentaron índices adecuados en materia de discriminación.

En las Figuras 2.1 a 2.5 se muestran las CCI de todos los ítems de la adaptación al castellano del cuestionario de Charbonnier-Voirin y Roussel (2012). En general, se observa un buen uso de todas las alternativas de respuesta y dado el apuntamiento de las curvas, cabe afirmar que proporcionan un nivel de información adecuado. En cuanto a los ítems más problemáticos señalados en los párrafos anteriores, destacaron por enfocarse en el centro de la distribución y por un uso razonable de las alternativas de respuesta. Asimismo, los ítems 14 y 20 presentaron unos de los valores de a más elevados, 2.97 y 2.55, respectivamente.

Tabla 2.10.

Parámetros Estimados según el Modelo de Respuesta Graduada de los ítems de la Escala de Charbonnier-Voirin y Rousset

Ítem	Dimensión	<i>a</i>	<i>e.t.</i>	<i>b1</i>	<i>e.t.</i>	<i>b2</i>	<i>e.t.</i>	<i>b3</i>	<i>e.t.</i>	<i>b4</i>	<i>e.t.</i>	<i>b5</i>	<i>e.t.</i>	<i>b6</i>	<i>e.t.</i>	X ²	<i>g.l.</i>	Prob.
5	GE	1.69	0.14	-3.44	0.30	-2.33	0.17	-1.69	0.12	-0.51	0.07	1.53	0.12			70.17	64	.28
10	GE	0.75	0.09	-9.02	1.70	-7.17	1.06	-4.89	0.62	-3.27	0.40	-1.09	0.16	2.36	0.29	114.41	70	.00
15	GE	1.08	0.10	-4.71	0.52	-3.01	0.28	-2.10	0.19	-0.24	0.09	0.99	0.12	3.21	0.30	98.76	82	.10
20	GE	2.55	0.22	-3.69	0.45	-2.16	0.13	-1.22	0.08	-0.76	0.06	0.17	0.06	1.56	0.10	102.68	63	.00
21R	GE	1.97	0.16	-2.78	0.20	-1.52	0.10	-0.58	0.06	-0.14	0.06	0.49	0.07	1.89	0.13	73.95	74	.48
22R	GE	1.20	0.11	-3.18	0.28	-1.89	0.16	-0.69	0.09	-0.21	0.08	0.44	0.09	2.28	0.19	106.10	92	.15
4	EF	2.27	0.19	-3.51	0.33	-2.48	0.16	-1.94	0.12	-1.43	0.09	-0.55	0.06	0.78	0.07	57.42	45	.10
9	EF	2.30	0.19	-3.45	0.32	-2.60	0.17	-1.94	0.12	-1.26	0.08	-0.38	0.06	1.09	0.08	60.51	40	.02
14	EF	2.97	0.29	-3.35	0.31	-2.69	0.18	-1.97	0.11	-1.20	0.07	-0.36	0.05	0.82	0.07	56.13	38	.03
19	EF	1.51	0.12	-4.64	0.56	-3.56	0.31	-2.76	0.21	-1.67	0.12	-0.50	0.07	1.44	0.12	59.57	46	.09
3	AI	1.66	0.17	-4.03	0.46	-3.85	0.41	-2.45	0.20	-1.39	0.11	0.36	0.07			33.34	26	.15
8	AI	1.93	0.21	-3.69	0.40	-2.99	0.26	-2.64	0.21	-1.42	0.11	0.58	0.07			32.00	26	.19
13	AI	1.25	0.13	-4.91	0.61	-4.09	0.43	-3.00	0.27	-1.72	0.15	-0.37	0.08	1.89	0.17	92.80	34	.00
18	AI	1.60	0.16	-4.37	0.55	-3.34	0.31	-2.43	0.19	-0.84	0.08	1.32	0.11			41.15	28	.05
2	REA	1.56	0.13	-4.86	0.70	-3.70	0.34	-2.97	0.24	-2.01	0.15	-0.80	0.08	1.56	0.12	63.88	39	.01
7	REA	3.40	0.34	-3.12	0.28	-2.34	0.15	-1.55	0.09	-0.90	0.06	0.09	0.05	1.50	0.09	41.66	30	.08
12	REA	2.88	0.25	-3.55	0.41	-2.72	0.19	-1.87	0.11	-1.06	0.07	-0.11	0.05	1.42	0.09	49.69	29	.01
17	REA	1.44	0.12	-4.16	0.43	-2.94	0.23	-2.03	0.15	-0.51	0.08	1.63	0.13			69.28	38	.00
1	CRE	1.24	0.12	-4.78	0.54	-3.28	0.29	-2.43	0.21	-1.69	0.15	-0.44	0.08	1.61	0.15	53.41	56	.57
6	CRE	1.56	0.14	-4.10	0.42	-2.88	0.23	-2.32	0.18	-1.15	0.10	-0.30	0.07	1.54	0.12	56.39	48	.19
11	CRE	1.75	0.16	-4.31	0.50	-2.98	0.23	-2.31	0.16	-1.17	0.09	-0.12	0.06	1.34	0.11	72.45	47	.01
16	CRE	2.19	0.21	-2.98	0.23	-2.11	0.14	-1.40	0.10	-0.38	0.06	0.56	0.06	1.93	0.13	69.42	47	.02

Nota. GE = Gestión del estrés; EF = Esfuerzo formativo; AI = Adaptabilidad interpersonal; REA = Reactividad ante emergencias; CRE= Creatividad; *a* = Discriminación; *b1*, *b2*, etc. = parámetros umbrales; *e.t.*= error típico; *g.l.* = grados de libertad; Prob. = probabilidad

Figura 2.1.

Curvas Características del Ítem para los Ítems que Miden Gestión del estrés

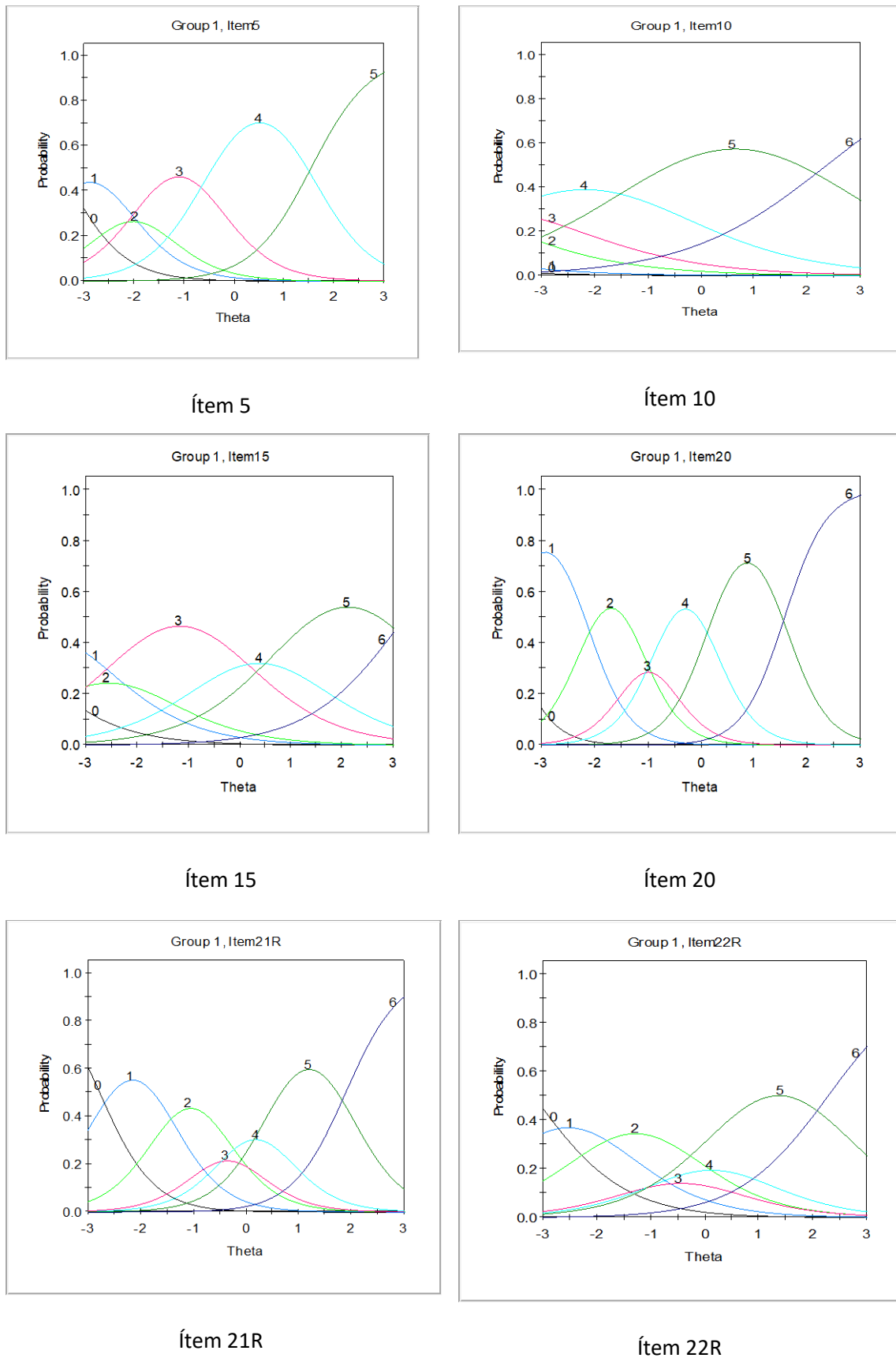


Figura 2.2.

Curvas Características del Ítem para los Ítems que Miden Esfuerzo Formativo

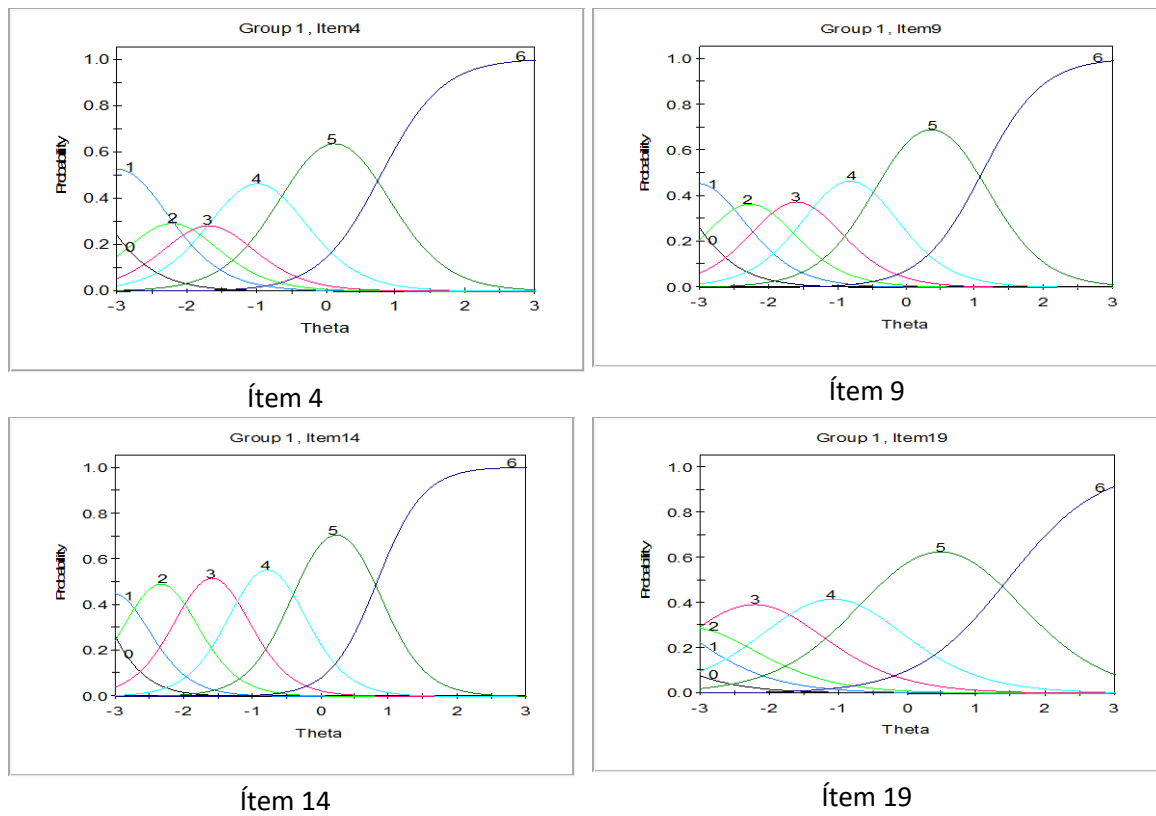


Figura 2.3.

Curvas Características del Ítem para los Ítems que Miden Adaptabilidad

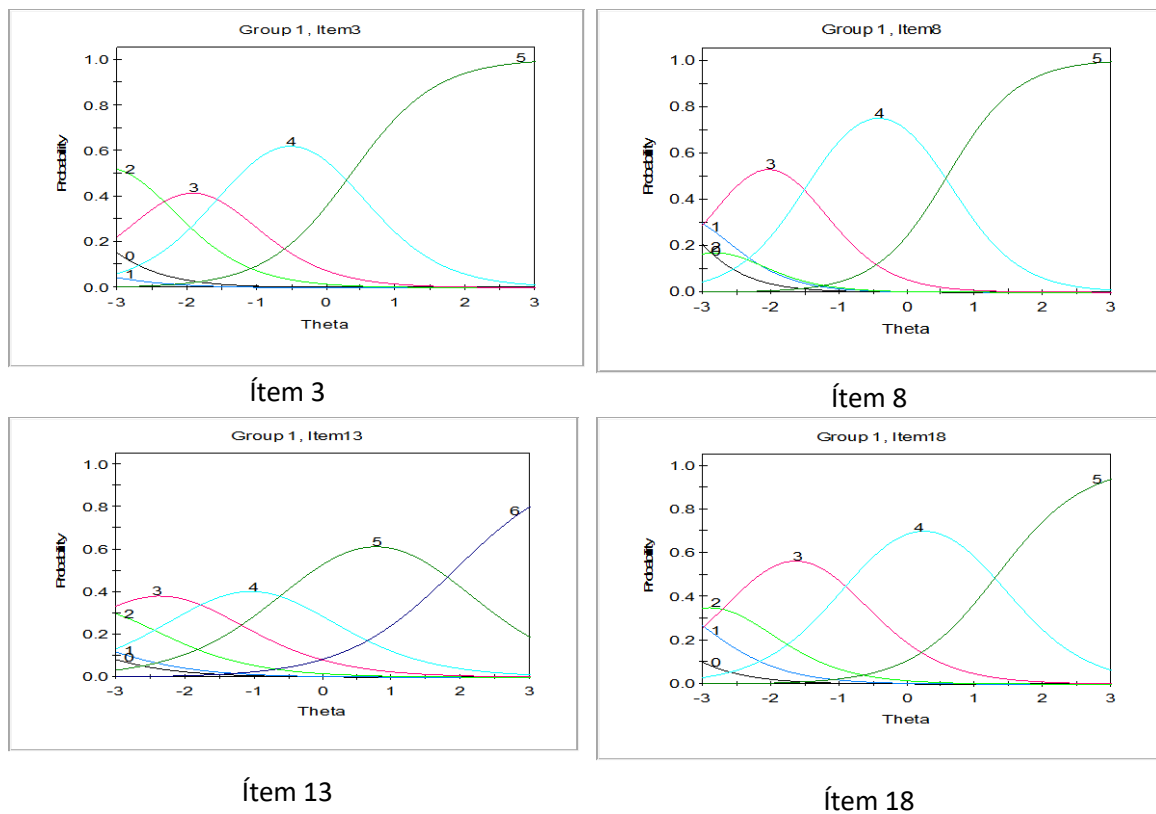


Figura 2.4.

Curvas Características del Ítem para los Ítems que Miden Reactividad

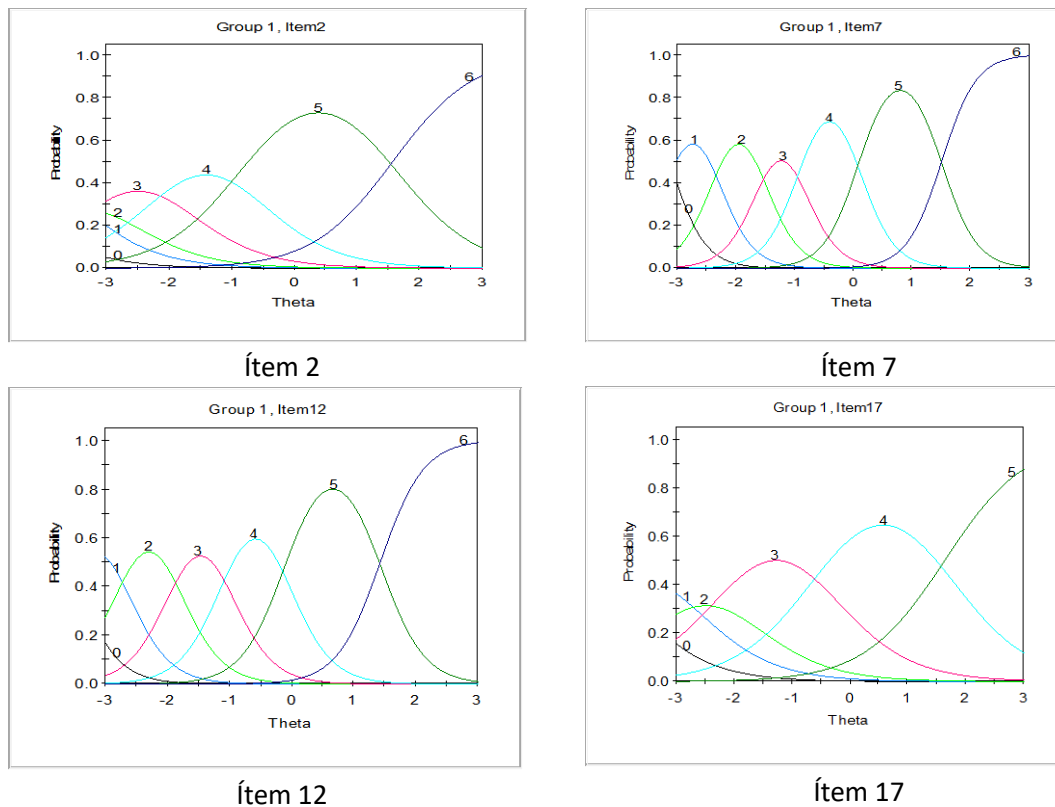
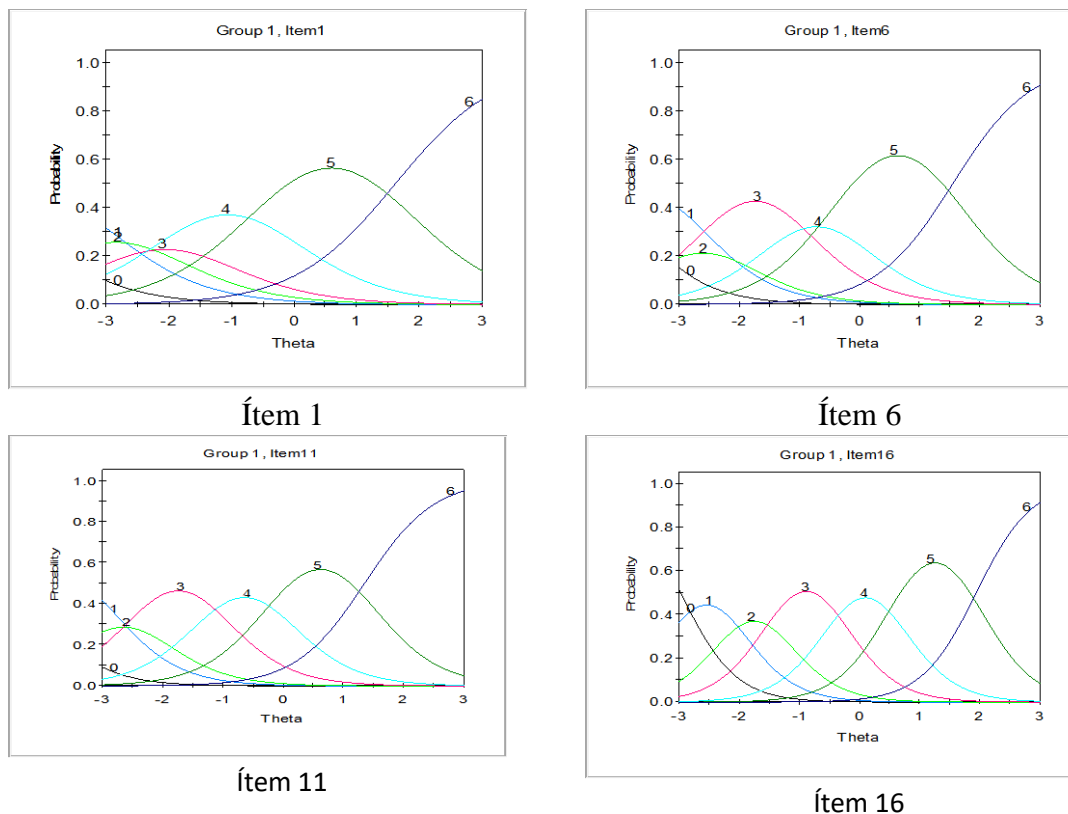


Figura 2.5.

Curvas Características del Ítem para los Ítems que Miden Creatividad



2.3.3.2. Dimensionalidad y validez de constructo: modelo confirmatorio, ESEM, y estructura interna

El AFC, mediante el método de estimación de Máxima Verosimilitud, mostró los siguientes datos de ajuste al modelo: $\chi^2(199) = 1167.40$, $p=.01$; RMSEA = .08; CFI = .81; TLI = .78. Se comprueba un ajuste moderado de los datos a la teoría, con un resultado aceptable en cuanto a RMSEA, el estadístico más relevante en enfoques confirmatorios (Rigdon, 1996), pero relativamente bajo en lo referido a CFI y TLI (Hu y Bentler, 1999).

Por su parte, el análisis ESEM mostró los siguientes resultados de ajuste al modelo: $\chi^2(131) = 525.10$, $p=.01$; RMSEA = .07; CFI = .97; TLI = .94. Se utilizó el método de estimación de Máxima Verosimilitud, método más adecuado que otros como ULS o DWLS para estudiar la adecuación de los datos (Xia y Yang, 2019). En este caso se comprueba un buen ajuste de los datos al modelo (Hu y Bentler, 1999).

En la Tabla 2.11 se muestran las saturaciones, señalando en negrita y fondo oscuro las cargas factoriales primarias, es decir, los datos correspondientes a los ítems concretos que corresponden a cada dimensión en base a la teoría y el diseño de Charbonnier-Voirin y Roussel (2012). Las cargas factoriales primarias estuvieron en un rango entre .13 y 1.07, con una mediana de .57, mientras que las cargas secundarias (saturaciones en dimensiones distintas de las teorizadas) estuvieron entre .00 y .54, con una mediana de .08. Las saturaciones de los ítems en la dimensión correspondiente fueron adecuadas (DiStefano et al, 2009; Floyd y Widaman, 1995; Swisher et al., 2004) a excepción de los ítems 10 (Gestión del estrés) y 13 (Adaptabilidad interpersonal), que presentaron saturaciones inferiores a .25. Dichos ítems, así como los ítems 11 (Creatividad) y 15 (Gestión del estrés) presentaron saturaciones más elevadas con dimensiones distintas a las teorizadas: Esfuerzo formativo (Ítems 11 y 13), Adaptabilidad interpersonal (Ítem 10), y Creatividad (Ítem 15).

El ítem 7 (Reactividad), por su parte, presentó una saturación superior a 1. No obstante, si los factores están correlacionados, las saturaciones de los factores, al ser coeficientes de regresión y no correlaciones, pueden ser mayores que uno en magnitud, lo que puede ser indicio de multicolinealidad en los datos (Jöreskog, 1999). Se observa que los ítems añadidos a la dimensión Gestión del estrés (Ítems 20, 21 y 22) presentaron

saturaciones de intensidad elevada y superiores a las del resto de ítems de dicha dimensión, lo que es un argumento a favor de su inclusión en la adaptación al castellano.

Tabla 2.11.

Saturaciones Factoriales y Porcentaje de Varianza de los Factores del Modelo ESEM de la Escala de Charbonnier-Voirin y Roussel

Ítem	GE	EF	AI	REA	CRE
5. Conservo la calma en las situaciones en las que debo tomar muchas decisiones	.50	-.14	.13	.14	.17
10. Busco soluciones discutiendo con tranquilidad con mis colegas	.13	-.01	.46	-.08	.33
15. Cuando las situaciones son difíciles, normalmente mis colegas me piden consejo debido a mi auto-control	.28	-.03	.03	.15	.50
20. Me siento tranquilo/a incluso si mis tareas cambian o se presentan muy rápidamente	.71	.08	.04	.09	-.06
21. Tener que asumir un trabajo extra de manera inesperada me pone nervioso/a	.92	-.02	-.14	-.10	-.04
22. El estrés ocasionado por mi trabajo repercute en la calidad de lo que hago	.69	.01	-.09	-.03	-.19
4. Me formo con regularidad dentro o fuera del trabajo para mantener al día mis competencias	.06	.71	-.08	-.01	.13
9. Estoy atento/a a las últimas innovaciones en mi profesión para mejorar mi forma de trabajar	.07	.67	-.11	.02	.28
14. Busco todas las oportunidades que me permiten mejorar mi desempeño (formación, grupos de trabajo, intercambios con mis colegas, etc.)	-.07	.80	.19	.02	.00
19. Me preparo para el cambio participando en cada proyecto o tarea que me permita hacerlo	.15	.46	.37	.07	-.07
3. Desarrollar buenas relaciones con todos mis interlocutores/as es un elemento importante de mi eficacia	.03	-.02	.59	.09	.01
8. Intento entender los puntos de vista de mis interlocutores/as para poder interactuar mejor con ellos/as	.05	.01	.59	.05	.13
13. Aprendo nuevas maneras de hacer mi trabajo para colaborar con otras personas	-.02	.54	.22	.17	.07
18. Adecúo mi comportamiento cuando lo necesito para trabajar bien con otros	.10	.10	.55	.15	-.07
2. Consigo centrar toda mi atención en la situación para actuar rápidamente	.15	-.07	.17	.46	.04
7. Decido rápidamente las acciones a llevar a cabo para resolver problemas	-.08	-.14	-.09	1.07	.01
12. Analizo rápidamente las posibles soluciones y sus implicaciones para elegir la más apropiada	-.05	.04	.01	.79	.02
17. Reorganizo fácilmente mi trabajo para adaptarme a las nuevas circunstancias	.29	.12	.16	.39	-.08
1. No dudo en cuestionar ideas establecidas para proponer una solución innovadora	.03	.06	.05	.11	.40
6. En mi departamento la gente cuenta conmigo para proponer nuevas soluciones	.10	-.06	.10	.04	.65
11. Utilizo fuentes de información muy variadas para encontrar una solución innovadora	.10	.40	.02	.09	.36
16. Desarrollo nuevas herramientas y métodos para resolver problemas nuevos	.08	.23	-.04	.21	.46
Porcentaje de varianza explicada	39.13	9.23	6.53	5.67	4.56

Nota. GE = Gestión del estrés; EF = Esfuerzo formativo; AI = Adaptabilidad interpersonal; REA = Reactividad ante emergencias; CRE = Creatividad. En negrita, las saturaciones de los ítems en su dimensión teorizada.

La matriz de correlaciones entre los factores de la escala de Charbonnier-Voirin y Roussel (2012) muestra los resultados que figuran en la Tabla 2.12, con correlaciones de tamaño alto entre la mayoría de factores.

Tabla 2.12.

Correlaciones entre los Factores de la Escala de Charbonnier-Voirin y Roussel

	GE	EF	AI	REA	CRE
GE	1				
EF	.46**	1			
AI	.56**	.66**	1		
REA	.61**	.50**	.63**	1	
CRE	.52**	.62**	.59**	.59**	1

Nota. **. La correlación es significativa en el nivel 0.01 (bilateral). GE = Gestión del estrés; EF = Esfuerzo formativo; AI = Adaptabilidad interpersonal; REA = Reactividad ante emergencias; CRE = Creatividad

Como se puede comprobar en la Tabla 2.13, en 4 de las dimensiones la media de las correlaciones entre los ítems y su correspondiente dimensión se encontró en el rango recomendado por Briggs y Cheek (1986), magnitudes entre .25 y .50. La excepción fue Esfuerzo formativo con un valor ligeramente superior, .51.

Tabla 2.13.

Correlación entre los Ítems de la Escala de Charbonnier-Voirin y Roussel y su Dimensión Correspondiente

Dimensión	Mínimo	Máximo	Media
Gestión del estrés	.10	.54	.33
Esfuerzo formativo	.36	.61	.51
Adaptabilidad interpersonal	.10	.40	.25
Reactividad	.37	.69	.46
Creatividad	.33	.49	.40

Las correlaciones entre los ítems se muestran en la siguiente página, en la Tabla 2.14. Las casillas oscurecidas son las que indican correlaciones entre ítems que miden la misma dimensión.

Tabla 2.14.

Intercorrelaciones entre los Ítems de la Escala de Charbonnier-Voirin y Roussel

Ítem	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22		
1	1																							
2	.21	1																						
3	.14	.31	1																					
4	.18	.19	.17	1																				
5	.15	.34	.24	.16	1																			
6	.33	.29	.24	.22	.27	1																		
7	.26	.47	.24	.18	.35	.31	1																	
8	.15	.27	.40	.26	.27	.25	.27	1																
9	.22	.26	.18	.61	.21	.31	.26	.30	1															
10	.33	.22	.26	.18	.25	.33	.20	.40	.27	1														
11	.35	.22	.21	.43	.26	.38	.31	.33	.58	.36	1													
12	.25	.37	.22	.26	.34	.26	.69	.31	.30	.27	.40	1												
13	.25	.23	.27	.43	.19	.28	.29	.36	.47	.33	.45	.38	1											
14	.22	.18	.27	.58	.16	.24	.21	.26	.56	.27	.45	.28	.61	1										
15	.28	.32	.25	.28	.43	.50	.41	.28	.36	.36	.35	.41	.36	.31	1									
16	.38	.26	.19	.29	.28	.45	.41	.27	.45	.28	.49	.39	.50	.43	.50	1								
17	.23	.37	.24	.20	.34	.28	.43	.25	.29	.25	.32	.45	.37	.30	.32	.38	1							
18	.16	.26	.34	.18	.28	.20	.27	.34	.21	.31	.28	.33	.32	.30	.30	.22	.38	1						
19	.23	.24	.27	.36	.19	.26	.27	.31	.40	.30	.41	.27	.53	.57	.32	.38	.43	.41	1					
20	.16	.26	.24	.20	.50	.23	.35	.22	.25	.25	.29	.37	.32	.25	.39	.30	.49	.31	.39	1				
21	.11	.25	.17	.16	.38	.13	.24	.19	.16	.17	.21	.20	.14	.10	.27	.20	.33	.21	.19	.54	1			
22	.00	.17	.07	.08	.25	.05	.14	.09	.09	.10	.13	.16	.13	.07	.16	.08	.26	.14	.15	.38	.50	1		

2.3.3.3. Fiabilidad

En la tabla 2.15 se presentan los índices de consistencia interna de los factores de la escala de Charbonnier-Voirin y Roussel: Rho de Jöreskog, Alfa de Cronbach, Omega de McDonald e Índice de Fiabilidad marginal.

Los resultados obtenidos en la adaptación al castellano fueron muy similares a los de la versión original y los coeficientes Rho oscilaron entre .60 y .83, lo que permite concluir que todos ellos presentaron una adecuada consistencia interna (Nunnally, 1978). Únicamente estuvo por debajo de los estándares aceptables la magnitud de Alfa de Cronbach de la dimensión Adaptabilidad Interpersonal, que sin embargo presentó valores adecuados en el resto de índices de evaluación de la fiabilidad. Las medianas de los índices Rho, Alfa de Cronbach y Omega fueron, respectivamente, .77, .73 y .75, adecuadas, mientras que la mediana del índice de fiabilidad marginal fue de .82, buena según Muñiz et al. (2011).

Tabla 2.15.

Fiabilidad (Consistencia interna) de las dimensiones de la Escala de Charbonnier-Voirin y Roussel

	Rho de Jöreskog	Alfa de Cronbach	Omega	Fiabilidad marginal
Gestión del estrés	.77	.74	.75	.82
Esfuerzo formativo	.83	.81	.80	.83
Adaptabilidad Interpersonal	.60	.37	.56	.71
Reactividad	.78	.78	.76	.85
Creatividad	.73	.72	.69	.77

En las Figuras 2.6 a 2.10 se aportan las funciones de información del test para cada dimensión, así como el correspondiente error típico de medida. Se observa que el nivel de precisión fue mayor en niveles bajos, medios y medio-altos (niveles theta de -3 a 1.5) en las dimensiones medidas, por lo que cabe interpretar que la adaptación al castellano del instrumento de Charbonnier-Voirin y Roussel (2012) es útil como herramienta de descarte de candidatos no idóneos, ya que permite detectar de manera fiable a individuos con niveles bajos en las dimensiones analizadas.

Figura 2.6.

Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para la Dimensión Gestión del Estrés

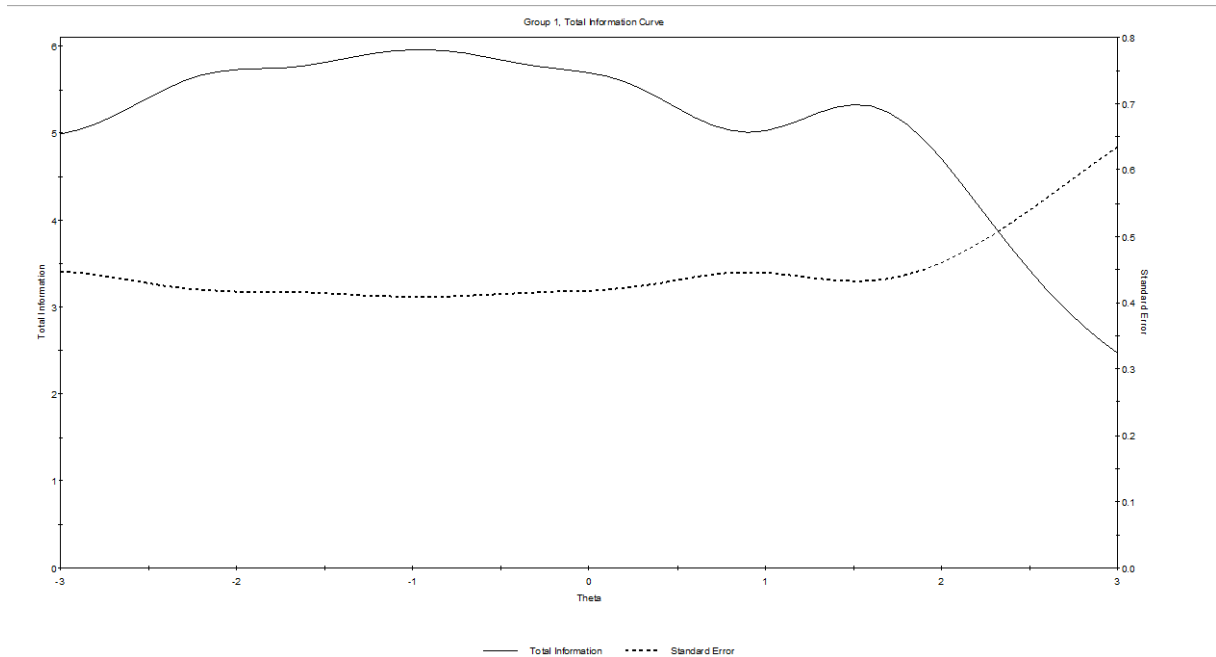


Figura 2.7.

Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para la Dimensión Esfuerzo Formativo

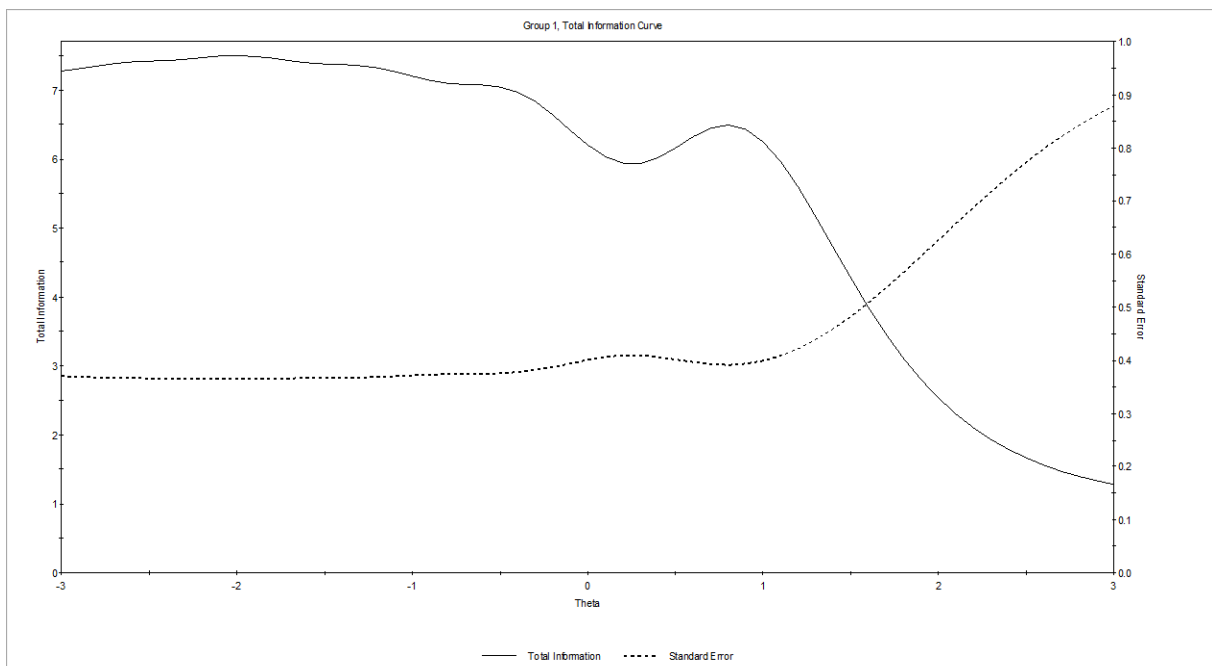


Figura 2.8.

Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para la Dimensión Adaptabilidad Interpersonal

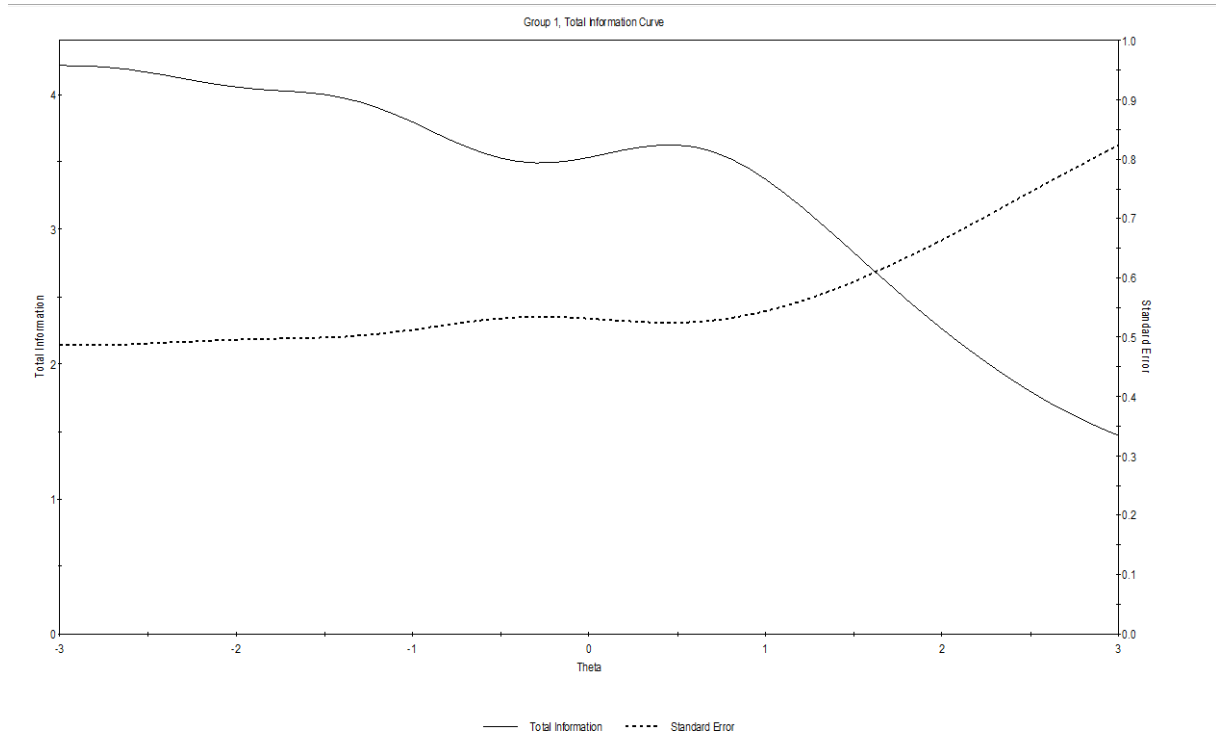


Figura 2.9.

Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para la Dimensión Reactividad ante Emergencias

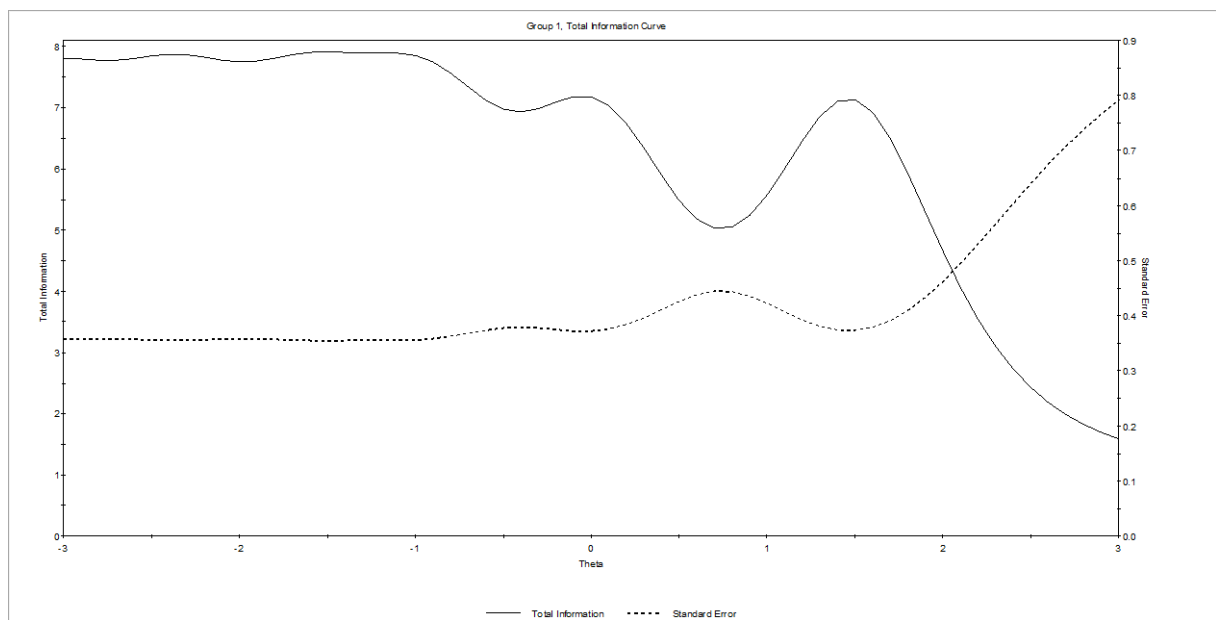
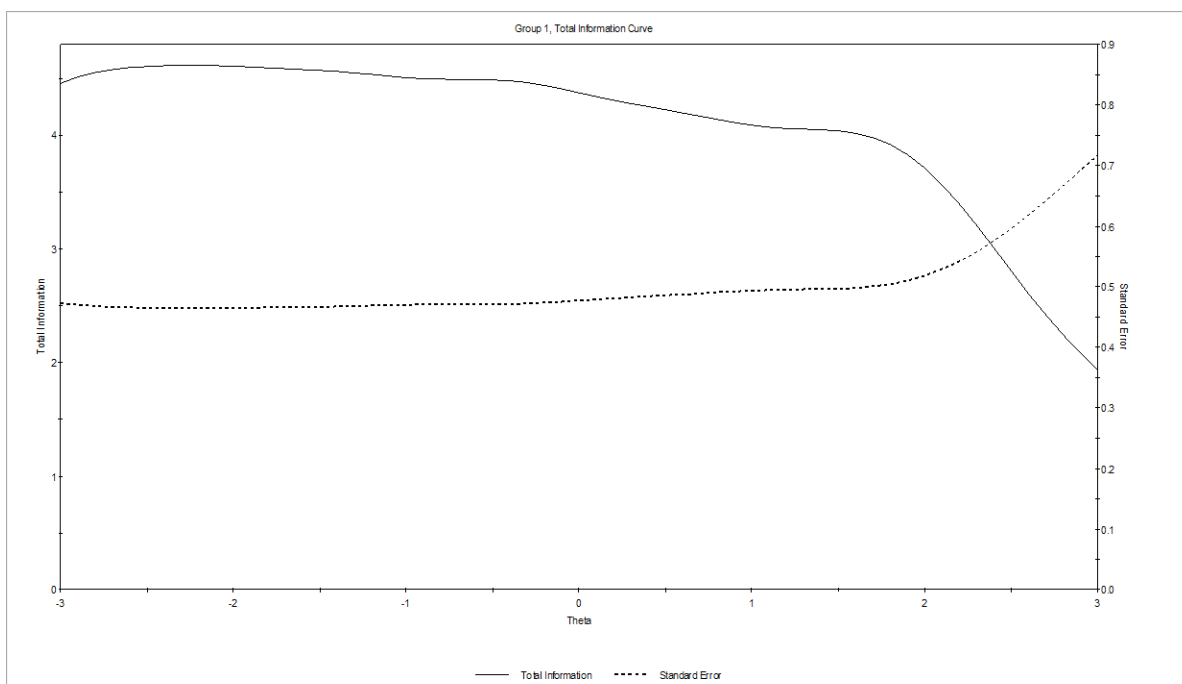


Figura 2.10.

Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para la Dimensión Creatividad



En la tabla 2.16 se presentan los índices de correlación obtenidos entre las puntuaciones de las aplicaciones test-retest. La mediana de los índices de fiabilidad test-retest obtenidos un mes después de la primera aplicación fue de .75, que cabe calificar como buena (Muñiz et al., 2011).

Tabla 2.16.

Fiabilidad Test-retest de las dimensiones de la Escala de Charbonnier-Voirin y Roussel

	<i>r</i> (N = 151)
Gestión estrés	.75
Esfuerzo formativo	.82
Adaptabilidad Interpersonal	.70
Reactividad	.75
Creatividad	.66

2.3.3.4. Funcionamiento diferencial de los ítems según lengua materna y complejidad del puesto

En las Tablas 2.17 y 2.18 se pueden observar los resultados obtenidos en los análisis realizados, y se comprueba que ningún ítem presentó FDI en base a las variables analizadas, puesto que los valores de R^2 de Nagelkerke fueron inferiores a .035 en todos los casos.

Tabla 2.17.

Funcionamiento Diferencial de los Ítems de la Escala de Charbonnier-Voirin y Roussel en base a la Complejidad del Puesto Desempeñado

	FDI Total	FDI Uniforme	FDI No Uniforme
	ΔR^2	ΔR^2	ΔR^2
	Nagelkerke	Nagelkerke	Nagelkerke
Ítem1	.00	.00	.00
Ítem2	.00	.00	.00
Ítem3	.00	.00	.00
Ítem4	.00	.00	.00
Ítem5	.00	.00	.00
Ítem6	.00	.00	.00
Ítem7	.00	.00	.00
Ítem8	.00	.00	.00
Ítem9	.00	.00	.00
Ítem10	.01	.01	.00
Ítem11	.00	.00	.00
Ítem12	.00	.00	.00
Ítem13	.00	.02	.02
Ítem14	.00	.00	.00
Ítem15	.01	.01	.00
Ítem16	.00	.00	.00
Ítem17	.00	.00	.00
Ítem18	.01	.00	.00
Ítem19	.00	.00	.00
Ítem20	.00	.00	.00
Ítem21R	.01	.00	.01
Ítem22R	.02	.02	.00

Tabla 2.18.

Funcionamiento Diferencial de los Ítems de la Escala de Charbonnier-Voirin y Roussel en base a la Lengua Materna de las Personas Participantes

	FDI Total	FDI Uniforme	FDI No Uniforme
	ΔR^2	ΔR^2	ΔR^2
	Nagelkerke	Nagelkerke	Nagelkerke
Ítem1	.00	.00	.00
Ítem2	.01	.00	.01
Ítem3	.01	.01	.00
Ítem4	.01	.01	.00
Ítem5	.01	.00	.00
Ítem6	.00	.00	.00
Ítem7	.01	.01	.00
Ítem8	.00	.00	.00
Ítem9	.01	.01	.00
Ítem10	.01	.01	.00
Ítem11	.01	.00	.01
Ítem12	.01	.00	.00
Ítem13	.01	.00	.01
Ítem14	.01	.00	.00
Ítem15	.01	.01	.01
Ítem16	.00	.00	.00
Ítem17	.01	.00	.00
Ítem18	.02	.01	.01
Ítem19	.02	.02	.00
Ítem20	.00	.00	.00
Ítem21R	.00	.00	.00
Ítem22R	.00	.00	.00

2.3.4. Conclusiones

En el presente capítulo se pretendía adaptar al castellano un instrumento válido y fiable para la medida del desempeño adaptativo. A tal fin se emplearon métodos cualitativos y cuantitativos. Un equipo multidisciplinar se encargó de la traducción de los ítems al castellano mediante la metodología de traducción inversa. A continuación, se emplearon análisis cuantitativos para evaluar las propiedades de los ítems, examinar la dimensionalidad del instrumento, así como para evaluar su consistencia interna y su estabilidad temporal y para obtener evidencias de su validez de constructo, diferencial y externa, cuestión esta última que se aborda en el Capítulo 4.

Los ítems que componen el cuestionario de desempeño adaptativo en castellano presentan propiedades psicométricas aceptables en lo que respecta a su capacidad de discriminación, variabilidad (desviaciones típicas cercanas o superiores a 1), índice de homogeneidad, asimetría, curtosis o el grado en el que saturan en las dimensiones teorizadas. Los ítems más problemáticos desde el punto de vista de la saturación en las dimensiones del modelo (10, 11, 13 y 15) presentan aspectos deseables desde la óptica de la TRI (niveles satisfactorios de discriminación y buen uso de las alternativas de respuesta). Asimismo, son elementos de utilidad para que el cuestionario pueda evaluar de manera óptima el constructo porque contribuyen a la amplitud conceptual del mismo (Clark y Watson, 1995).

El análisis de la dimensionalidad empleando ESEM apoya la estructura interna propuesta de 5 dimensiones: Gestión del estrés, Esfuerzo Formativo, Adaptabilidad interpersonal, Reactividad ante emergencias y Creatividad. Este análisis confirma la estructura del cuestionario original y avala la inclusión en la adaptación al castellano de 3 nuevos ítems cuyas saturaciones en la dimensión Gestión del estrés han sido elevadas y cuyo contenido reafirma el núcleo conceptual de la dimensión que miden.

Las correlaciones entre las dimensiones son ligeramente superiores a las observadas en el estudio original, aunque similares a las halladas por Charbonnier-Voirin et al. (2010) en una versión previa de la presente escala, o por Hamtiaux et al. (2013), que emplearon el instrumento I-ADAPT de Ployhart y Bliese (2006), basado en el modelo de Pulakos et al. (2000).

Como señalan Hamtiaux et al. (2013), cabe esperar cierto solapamiento entre las dimensiones debido al contenido de los ítems ya que, por ejemplo, Reactividad y Gestión del estrés tienen en común la necesidad de afrontar situaciones inestables. Asimismo, Esfuerzo formativo y Creatividad están basadas en el aprendizaje y la resolución de problemas.

El contenido de los ítems 10 y 13 puede haber afectado a la correlación de tales ítems con dimensiones distintas a las teorizadas, afectando a las saturaciones bajas con las dimensiones que pretendían medir. Así, el ítem 10 hace referencia a discutir con colegas, lo que podría explicar que presente algunas de sus correlaciones más elevadas con los ítems 8 y 13, que miden Adaptabilidad interpersonal. Las puntuaciones de las y los participantes en los ítems 11 y 13 se relacionan con los 4 ítems que miden Esfuerzo Formativo, probablemente por su alusión al aprendizaje de nuevos métodos y al uso de fuentes de información variadas. La vinculación entre el ítem 15 y la Creatividad es menos clara conceptualmente, si bien el que otras personas recurran a uno mismo en situaciones estresantes puede estar relacionado con la mencionada dimensión.

La media de las correlaciones inter-ítem de cada dimensión presenta valores que se corresponden con las recomendaciones de Briggs y Cheek (1986), lo que constituye un argumento a favor de la unidimensionalidad de las dimensiones de la escala de Charbonnier-Voirin y Roussel (2012).

Por otro lado, los análisis acerca de la fiabilidad del instrumento muestran índices aceptables de consistencia interna (evaluada desde múltiples enfoques), presentando medianas adecuadas en todos los casos. La mediana de los índices de Rho de Jöreskog es de .77, muy similar a la obtenida en el estudio original de Charbonnier-Voirin y Roussel (2012), .80, registrando valores superiores a los del estudio original para las dimensiones de la presente investigación Esfuerzo Formativo y Gestión del estrés. A diferencia del estudio original, en este caso se ha calculado la fiabilidad test-retest para evaluar la estabilidad de la medida, con un resultado adecuado. Asimismo, las funciones de información de las dimensiones de esta escala indican que estamos ante un instrumento capaz de descartar a candidatos con niveles bajos en las variables medidas, aspecto relevante en selección de personal.

Por lo que respecta a la validez diferencial, ningún ítem presenta funcionamiento diferencial en base a la lengua materna o en función del nivel de complejidad del puesto de trabajo desempeñado, según los análisis realizados. Por ello, cabe concluir que el test no presenta sesgo en dichas variables sociodemográficas y laborales, cuestión de especial importancia en el ámbito de la selección en el sector público.

Como limitación debe señalarse que la muestra empleada es muy homogénea en cuanto al sector laboral, ya que únicamente han participado empleados públicos. Una mayor heterogeneidad en la muestra siempre es deseable porque puede contribuir a aumentar la variabilidad de las respuestas y la fiabilidad (Streiner, 2003). Dado que la validación de un instrumento es un proceso continuo en el que se van incorporando evidencias que avalen su idoneidad y relevancia como instrumento de medida, sería interesante extender esta investigación a participantes del sector privado y también a estudiantes o a aspirantes en procesos selectivos.

Con todo, en base a los datos obtenidos en el proceso de adaptación al castellano del cuestionario de Charbonnier-Voirin y Roussel (2012), cabe afirmar que este instrumento ofrece suficientes garantías para medir adecuadamente el desempeño adaptativo. Como se ha mencionado en la introducción, muchos de los instrumentos empleados con propósitos similares muestran deficiencias metodológicas, mientras que el presente instrumento presenta propiedades psicométricas adecuadas, lo que posibilita evaluar el desempeño adaptativo del personal de manera fiable y válida. Asimismo, el presente instrumento contribuye a paliar la escasez de herramientas en castellano que evalúan el desempeño adaptativo y supone una aportación de interés en la definición del constructo de desempeño adaptativo.

CAPÍTULO 3.

ADAPTACIÓN AL EUSKERA Y VALIDACIÓN DEL CUESTIONARIO DE PERSONALIDAD *OPERAS*

3.1. Revisión teórica del concepto de personalidad

La personalidad puede definirse como el conjunto de características (la manera habitual de pensar, sentir y actuar) que definen a una persona y determinan su manera de interactuar con el entorno (Barrick et al., 2011; Hugues y Batey, 2017). Esas características se combinan de manera distinta en cada persona, inducen sus conductas y son relativamente estables o consistentes, es decir, permiten predecir patrones de conducta en diferentes situaciones o momentos porque las personas tienden a actuar de manera acorde a su auto-concepto (Feist y Barron, 2003; Robinson y Sedikides, 2018). En la edad adulta las características de la personalidad se muestran estables pese a posibles cambios en roles, relaciones o estado de salud y a las experiencias personales traumáticas o exitosas a lo largo de la vida (Hugues y Batey, 2017; McCrae y Costa, 2003).

Esta capacidad predictiva de la personalidad fue el eje de uno de los artículos más influyentes en la historia de la Psicología del Trabajo y las organizaciones, el meta-análisis de Barrick y Mount (1991), que subrayó la influencia crítica de la personalidad, el principal descriptor de la conducta humana, sobre el desempeño laboral. En consecuencia, la personalidad ha sido objeto de especial interés en materia de selección de personal, ámbito en el que la personalidad se refiere a las variables que moderan toda ejecución, que predicen que un trabajador o trabajadora que posee unas determinadas aptitudes, destrezas y conocimientos los usará correctamente o estará influenciado/a por aspectos emocionales (estabilidad emocional) o predisposiciones personales (extroversión, necesidad de logro, etc.). No en vano, la selección es el proceso de elegir a aquellos aspirantes que se comportarán con mayor probabilidad de una manera que alcanzará o sobrepasará los estándares de éxito definidos por las organizaciones (Barrick et al., 2011; Hugues y Batey, 2017; Smith y Smith, 2005).

El uso de datos de personalidad implica clasificar y medir a las personas en base a una serie de características de personalidad conocidas como rasgos. Un rasgo es una estructura disposicional subyacente a la persona que explica la consistencia en el comportamiento de dicho sujeto ante una serie de situaciones. Cabe definirlo como una dimensión continua en la que se pueden medir las diferencias individuales en base a la cantidad de la característica exhibida por el individuo. Dicho concepto se fundamenta en la observación del comportamiento humano y parte de la base de que las personas presentan conductas diferentes ante las mismas situaciones. Los rasgos son, por tanto, descriptores utilizados para definir la personalidad de los individuos, sus características diferenciales de funcionamiento, su manera relativamente estable de pensar, sentir y actuar (Barrick et al., 2011; Mischel, 2003; Solé, 2006).

En este capítulo se revisa el constructo de la personalidad, particularmente en lo que concierne al modelo predominante, el modelo de los cinco grandes factores de personalidad. Seguidamente, se analizan las cuestiones más relevantes a tener en cuenta en el análisis de instrumentos que evalúan la personalidad. A continuación, se realiza una breve revisión de los instrumentos más relevantes disponibles en el contexto español en el ámbito de selección de personal. Finalmente, se explicita el procedimiento de adaptación del cuestionario de personalidad OPERAS a la lengua y cultura vasca y su proceso de validación en esa lengua.

3.1.1. Taxonomía de la personalidad: el modelo de los Cinco Grandes

En el ámbito de la psicología de la personalidad existe un consenso generalizado acerca de una taxonomía que ha servido para agrupar miles de rasgos en un número manejable de factores o dimensiones. Dicha taxonomía recibe tres denominaciones distintas que se corresponden con enfoques diferentes, si bien en la práctica se emplean prácticamente como sinónimos: los *Cinco Grandes* (Big Five – BF), desde el enfoque que estudia los rasgos de personalidad expresados en el lenguaje natural (enfoque psicoléxico); el *Modelo de los Cinco Factores* (Five-Factor Model – FFM), desde el enfoque de los estudios de rasgos que parten de cuestionarios de personalidad como el NEO-PI (Costa y McCrae, 1985); y la *Teoría de los Cinco Factores* (Five-Factor Theory - FFT) formulada por McCrae y Costa (1996, 1999) y que incluye una serie de proposiciones acerca de la naturaleza, origen y desarrollo de los rasgos de personalidad (Srivastava, 2019).

Los tres enfoques coinciden al señalar la presencia de cinco factores (*Amabilidad o Afabilidad, Apertura a la experiencia, Estabilidad Emocional, Extroversión y Responsabilidad o Conciencia*) que son considerados como los rasgos esenciales que influyen en la conducta, si bien difieren en las denominaciones otorgadas a dichos factores (Tabla 3.1). Numerosas investigaciones han probado la robustez de esta estructura de la personalidad, además de demostrar que es generalizable a diferentes muestras de población y a distintas lenguas y culturas y que es consistente y estable a través del tiempo. Asimismo, el modelo FFM explica buena parte de la varianza de las estructuras factoriales de cuestionarios de personalidad diseñados desde diversas orientaciones teóricas, incluso las ajenas al FFM (Allik y Realo, 2017; Barrick et al., 2001; Barrick y Mount, 2014; Carlo et al., 2014; Hugues y Batey, 2017; Kajonius y Mac Giolla, 2017; O'Connor, 2002; Salgado, 2003; Soto y John, 2009).

Tabla 3.1.
Definición de los 5 grandes Factores de Personalidad

Factor	Nombre alternativo	Definición
Amabilidad	Afabilidad, Agradabilidad, Amigabilidad, Cordialidad	Rasgo propio de las personas que confían en los demás, son generosas, simpáticas, serviciales, compasivas, cariñosas y cooperativas, consideran sus necesidades e intereses al mismo nivel que los de las otras personas.
Apertura a la experiencia	Apertura Mental, Apertura al cambio, Cultura, Intelecto, Imaginación, Cultura	Describe a las personas que desean considerar nuevas ideas, pero no buscan la novedad por sí misma, y desean considerar nuevos modos de hacer las cosas. Hace referencia a la tendencia a interesarse por cuestiones conceptuales o intelectuales. Implica apreciación del arte, autorreflexión, preferencia por la diversidad y búsqueda intelectual, aunque también cuestionamiento de la autoridad.
Estabilidad emocional	Neuroticismo (Inverso a Estabilidad Emocional)	Facilidad para controlar sus emociones y para resistir a las situaciones de tensión o presión. Caracteriza a las personas que son habitualmente tranquilas, pacientes y sosegadas. Afrontan las situaciones de emergencia con un buen control emocional. En algunos cuestionarios se emplea el término Neuroticismo para referirse a este factor, siendo en tal caso el factor inverso a Estabilidad emocional.
Extroversión	Extraversión	Facilidad para la sociabilidad, el contacto con las personas y los grupos grandes. Las personas con puntuaciones altas en Extroversión son activas y habladoras, tienden a ser optimistas y llenas de energía. Les gusta la excitación y tienden a ser asertivas y dominantes.
Responsabilidad	Conciencia, Tesón	Es un rasgo propio de personas con propósitos claros y determinados, y que también suelen mostrar un alto deseo de logro, así como rigurosidad, orden y responsabilidad. Suelen ser personas sistemáticas, eficientes, constantes, prudentes en sus actos y minuciosas en sus quehaceres.

Pese a que hay investigadores/as que señalan que algunos rasgos esenciales de la personalidad se encuentran fuera del modelo Big Five (Ashton et al., 2000) o son partidarios/as de un modelo alternativo al de FFM (Van Hemert et al., 2002; Zuckerman et al., 1991), el modelo Big Five es el más reconocido y el que se utiliza con mayor frecuencia en la investigación en personalidad (Allik y Realo, 2017; Costa y McCrae, 2006; Hugues y Batey, 2017). Costa y McCrae (1992a) resumieron las evidencias de la validez del FFM en cuatro aspectos: (1) estudios longitudinales y transversales han encontrado que los cinco factores de personalidad son disposiciones comportamentales duraderas; (2) los rasgos asociados con los cinco factores emergen de diferentes sistemas de personalidad y de estudios del lenguaje natural; (3) los cinco factores se han encontrado en diferentes edades, géneros, razas y grupos lingüísticos; y (4) estudios de genética conductual han demostrado una base biológica para cada uno de los cinco factores.

Una fuente adicional de apoyo a la taxonomía de los cinco grandes proviene de los estudios sobre las concepciones teóricas populares de la personalidad. Diversos estudios han hallado que las conceptualizaciones sobre la personalidad y las observaciones de las manifestaciones cotidianas de la conducta por parte de personas no expertas en materia de personalidad, coinciden con la visión de los expertos sobre las dimensiones del FFM (Hall et al., 2019; Mehl et al., 2006).

No obstante, pese a su popularidad, autores como Vassend y Skrandal (2011), han cuestionado la adecuación del modelo, su validez de constructo, o la noción fundamental de que se pueden identificar cinco factores ortogonales. La razón para ello es que el FFM no presenta un ajuste adecuado en análisis factoriales confirmatorios (AFCs), pese a que existe una gran cantidad de investigaciones basadas en análisis factoriales exploratorios que lo apoyan (McCrae y Costa, 1997) y que mencionamos en el epígrafe siguiente. Esto llevó inicialmente a que algunos autores (McCrae et al., 1996) cuestionasen la utilidad de los AFCs en la investigación sobre el FFM. Sin embargo, eventualmente Marsh et al. (2010) demostraron que los modelos de ecuaciones estructurales exploratorios (Exploratory Structural Equation Models – ESEM en inglés), que son menos restrictivos a la hora de medir la estructura factorial, eran una alternativa válida a los AFCs y suponen una prueba adicional de la robustez del modelo factorial del FFM, algo que ha sido replicado en investigaciones posteriores

(Booth y Hugues, 2014; Furnham et al., 2013; Herrmann y Pfister, 2013; Marsh et al., 2014; Perera et al., 2015).

3.1.2. Origen y generalización transcultural del modelo de los Cinco Grandes

El modelo de los Cinco Grandes surgió del enfoque psicoléxico, cuyo supuesto de partida es que todos los comportamientos o diferencias individuales relevantes para la sociedad están incluidos en el lenguaje empleado por sus miembros. Además, según este enfoque el grado de representación de un atributo en el lenguaje se corresponde con la importancia de dicho atributo. De acuerdo con esta hipótesis, una manera de definir el dominio de un rasgo es analizar la versión completa de un diccionario en busca de términos relativos a la personalidad, ya que los rasgos existen en el lenguaje natural por su utilidad en las interacciones sociales. Asimismo, los términos vinculados a los cinco grandes se mantienen estables en el tiempo y entre generaciones (Choragwicka, 2010; Costa y McCrae, 2006; De Raad, 2000; Hendriks, 2002; Roivainen, 2020).

Galton fue uno de los primeros autores en aplicar el enfoque psicoléxico a la personalidad, observando que el diccionario contenía más de 1000 palabras relativas al carácter, e investigadores como Thurstone (1934) hallaron que cinco factores eran suficientes para describir a las personas. No obstante, el punto de partida de los estudios léxicos fue el de Allport y Odbert (1936), que en base al trabajo de Baumgarten (1933) en alemán, encontraron cerca de 18.000 términos en inglés que podían “distinguir la conducta de un ser humano con respecto a la de los demás”. De estos términos, 4.504 palabras fueron identificadas como rasgos personales estables, y suponen el origen del modelo de los cinco grandes (Goldberg, 1993; John y Srivastava, 1999).

La lista de palabras de Allport y Odbert fue la base de los trabajos de Cattell (1943, 1945 y 1947) que dieron lugar a su modelo de 16 factores (16 PF), el cual, pese a su popularidad no ha sido replicado por investigaciones posteriores (Digman y Takemoto-Chock, 1981; Tupes y Christal, 1961). No obstante, los factores de segundo orden del 16PF muestran cierta correspondencia con los Cinco Grandes factores de personalidad (John y Srivastava, 1999).

A partir de los datos de Cattell, Fiske (1949) encontró cinco factores muy similares a los del modelo actual. Estos estudios tuvieron continuación en los trabajos de Tupes y Christal (1961) y Norman (1963). Tupes y Christal son considerados como “padres” del modelo, ya que fueron los primeros en hipotetizarlo y en proponer el nombre inicial a las cinco dimensiones (Extroversión o Surgencia; Agradabilidad; Conciencia, Estabilidad emocional y Cultura). Norman, por su parte, encontró que se podían extraer cinco factores similares mediante evaluaciones de la personalidad realizadas por terceras personas. Autores como Borgatta (1964) o Smith (1967) confirmaron la estructura de cinco factores antes de que el número de investigaciones en materia de personalidad se redujese en los años 70 y principios de los 80 debido a críticas sobre su limitada capacidad para predecir la conducta.

La siguiente figura de referencia en los estudios psicoléxicos fue Goldberg (1990, 1992 y 1993), que partiendo del trabajo de Norman llegó a una estructura de cinco factores, a los que denominó Cinco Grandes, que se mantenía invariante incluso cuando más de cinco factores eran rotados, lo que aumentó el apoyo para el nuevo modelo en la comunidad científica. A lo largo de los años, investigadores e investigadoras de diferentes países han llevado a cabo análisis léxicos cuyos datos han convergido en el modelo de cinco factores o dimensiones (De Raad, 1992; Digman, 1990; John, 1990; Ostendorf, 1990). Este consenso creciente provocó un notable incremento en el número de investigaciones sobre los Cinco Grandes a partir de la segunda mitad de la década de los 90 (John et al., 2008).

La investigación transcultural y la replicación de los Cinco Grandes en culturas que difieren de forma notable en cuanto su lengua, estructuras económicas, religiones o prácticas educativas es la prueba más contundente de la precisión del modelo (Costa y McCrae, 2006). A partir de los años 70 los estudios psicoléxicos se extendieron a idiomas próximos al inglés, como el holandés (Brokken, 1978, De Raad et al., 1992) o el alemán (Angleitner et al., 1990; Borkenau y Ostendorf, 1990; Ostendorf, 1990), y también a otras lenguas, como el húngaro (Szirmák y De Raad, 1994), ruso (Shmelyov y Pokhil'ko, 1993), polaco (Szarota, 1996), checo (Hrebickova, 1995), italiano (Caprara y Perugini, 1994), español (García et al., 2004; Quevedo-Aguado et al., 1996; Silva et al., 1994), francés (Boies et al., 2001), portugués (Hutz et al., 1998), chino (Yang y

Bond, 1990; Trull y Geary, 1997), japonés (Aoki, 1971; Isaka, 1990), hindi (Lodhi et al., 2002) o turco (Somer y Goldberg, 1999), entre otras (Choragwicka, 2010).

En definitiva, el modelo de los Cinco Grandes ha sido replicado en un gran número de países y culturas, incluso no indoeuropeas, lo que avala su validez transcultural. Como indica Choragwicka (2010), la mayoría de estudios psicológicos internacionales concluyen que, si bien resulta prácticamente imposible observar una estructura de rasgos idéntica en todos los idiomas, las estructuras encontradas son muy semejantes. Las mayores similitudes se han hallado entre las lenguas germánicas y en un contexto más amplio, entre las lenguas europeas. Aunque algunos estudios en países en vías en desarrollo y lenguas no occidentales hayan postulado que un número variable de factores (2, 3 o 6) define mejor el dominio de la personalidad (Laajaj et al. 2019; Smaldino et al., 2019; Thalmayer et al., 2019), la mayoría de investigaciones han demostrado que la semejanza entre las estructuras encontradas no permite concluir que cada idioma o cultura refleje un modelo de personalidad alternativo al modelo de los Cinco Grandes. Si bien la calidad de la replicación varía ligeramente en diferentes culturas, esta variación se ha debido fundamentalmente a la calidad de los datos recopilados, particularmente en culturas no-occidentales, en las que los participantes estaban menos familiarizados con los cuestionarios, así como a tamaños muestrales reducidos y deficiencias metodológicas. Por lo tanto, aunque la universalidad lingüística de los cinco grandes factores derivados de los estudios psicológicos presente algunas excepciones, no hay una hipótesis alternativa con más reconocimiento dentro del enfoque de rasgos (Allik y McCrae, 2004; Allik y Realo, 2017; John et al., 2008).

Desde un enfoque alternativo al psicológico, en el enfoque basado en el estudio de cuestionarios de personalidad, también conocido como disposicional o psicométrico, se ha llegado a conclusiones similares y se ha acuñado el término Modelo de los Cinco Factores (FFM). Este modelo es descriptivo, puesto que propone un resumen empírico de la covariación de rasgos de personalidad y establece, a un nivel elevado de abstracción, los aspectos comunes en los sistemas de personalidad mediante el análisis factorial y el análisis y comparación de los cuestionarios de personalidad (McCrae y Costa, 1985, 1987; McCrae, Terracciano, et al., 2005). Este enfoque busca identificar conjuntos de variables que covarían entre sí, y no están relacionadas con otros conjuntos de variables. De esta forma, ha servido para reducir grandes cantidades de datos,

llevando de manera consistente a la identificación de los mismos cinco factores empleando diferentes instrumentos como NEO-PI o BFI (John et al., 1991), en diferentes culturas y partiendo de datos obtenidos mediante diferentes perspectivas (comparación de auto-evaluaciones y evaluaciones de terceras personas), en los que las diferencias halladas entre las puntuaciones medias de dichas perspectivas han sido mínimas (Allik y Realo, 2017; Costa y McCrae, 2006; Gatewood et al., 2015; McCrae, Terracciano et al., 2005; Schmitt et al., 2007). Los principales exponentes de este enfoque fueron McCrae y Costa, cuya prolífica obra en los años 80 constituyó un gran paso para el consenso actual sobre la estructura de la personalidad y contribuyó a que antiguos detractores del modelo, como Goldberg, se sumaran al mismo (Goldberg, 1993).

Los propios McCrae y Costa fueron los impulsores de la Teoría de los Cinco Factores (FFT), cuyo objetivo es explicar las propiedades universales o prácticamente universales de la personalidad humana. Según esta proposición teórica, los rasgos de personalidad son tendencias básicas con una base biológica. Estas tendencias básicas, en interacción con influencias externas, causan adaptaciones características (hábitos, destrezas, creencias). Las adaptaciones características en interacción con influencias externas son la causa de las conductas y las reacciones de las personas (Allik y Realo, 2017, 2019; McCrae y Sutin, 2018). La FFT se basa en la universalidad, la estabilidad y la heredabilidad de los rasgos (McCrae y Sutin, 2018), y se propuso con el objetivo de explicar los resultados de estudios longitudinales que mostraban que el aprendizaje o la experiencia ejercen una influencia moderada en los Cinco Grandes (Srivastava, 2019): investigaciones sobre la extraordinaria estabilidad de los rasgos de personalidad a lo largo de la vida (McCrae y Costa, 2003), la similitud en los patrones de covariación en países con diferentes tradiciones culturales, niveles de prosperidad económica o trayectorias históricas (McCrae, Terracciano, et al. 2005; Schmitt et al., 2007), la gran magnitud de los efectos de la herencia en comparación con el pequeño efecto del ambiente compartido (Yamagata et al., 2006) o la escasa influencia de los eventos vitales a nivel de rasgos de personalidad (Costa et al., 2000).

3.1.3. La medición de la personalidad: estado del arte y principales problemas

Los tres grandes actores en el ámbito de la personalidad aplicada a la selección de personal son las organizaciones, las editoriales que publican tests, y los/as

académicos/as. Estos tres actores tienen un objetivo común: poder disponer de evaluaciones fiables y válidas de la personalidad. Sin embargo, mientras el principal foco de interés para el ámbito académico es el estudio teórico sobre la naturaleza y estructura de la personalidad, las organizaciones desean instrumentos capaces de predecir conductas productivas o contraproductivas o el desempeño al menor coste, y las editoriales se centran en lo que se puede comercializar, lo que consideran que puede ser utilizado por las organizaciones. Todo ello produce una cierta fractura entre los ámbitos teórico y práctico, entre la base teórica apreciada por los académicos y los productos que acaban adquiriendo las organizaciones, a menudo más guiadas por la moda que por prácticas basadas en la evidencia (Hugues y Batey, 2017). Por ello, resulta crucial abordar cuestiones vinculadas con la utilidad de los instrumentos de personalidad y la medición de la misma.

Los instrumentos de evaluación de la personalidad por excelencia son los cuestionarios que recogen las respuestas expresadas por una persona para determinar la personalidad. Un tipo de cuestionario muy utilizado es el de auto-informe, que consiste en una serie de ítems o frases que solicitan a las personas evaluadas que, mediante un formato de respuesta de opción múltiple, indiquen su grado de acuerdo en torno a pensamientos, emociones y experiencias pasadas. Estos cuestionarios asumen que existe una correspondencia entre lo que las personas dicen de sí mismas y lo que realmente es cierto.

A la hora de evaluar la calidad y la utilidad de los cuestionarios, en la actualidad existen una serie de cuestiones relevantes que generan debate en los ámbitos académico, organizacional o editorial (Barrick et al., 2011; Hugues y Batey, 2017; Muñiz y Fernández-Hermida, 2010) y que desarrollamos en los puntos siguientes:

-Frente a los tests de personalidad basados en el modelo Big Five con ítems genéricos aplicables a todos los ámbitos de la vida, existen tests que incorporan ítems contextualizados en situaciones laborales, no necesariamente basados en el modelo FFM. Aunque algunos estudios indican que las medidas que no emplean el modelo FFM presentan peores datos de validez predictiva que los tests basados en dicho modelo (Salgado, 2003), autores como Shaffer y Postlethwaite (2012) han hallado que los tests dirigidos al trabajo presentan una mayor validez.

-Muchos de los cuestionarios creados en los años 70, 80 o 90 del pasado siglo son extensos. Un ejemplo es el NEO-Personality Inventory (Costa y McCrae, 1992b), adaptado a más de 40 lenguas (McCrae y Allik, 2002), entre ellas el euskera (Balluerka et al., 2007; Gorostiaga et al., 2011), y compuesto por 240 ítems -48 ítems miden cada factor-, siendo su tiempo de aplicación superior a los 40 minutos. Sin embargo, en las últimas décadas diferentes investigaciones han abogado por cuestionarios más breves para la evaluación de la personalidad.

-Una cuestión recurrente en la investigación de las pruebas de personalidad es el estudio de medidas para contrarrestar la falta de sinceridad de las y los aspirantes en procesos selectivos. Existen diferentes conceptualizaciones de este problema, así como diferentes metodologías para afrontarlo (Barrick et al., 2011; Barrick y Mount, 2014; Hugues y Batey, 2017)

3.1.4. Contextualización de los ítems de los cuestionarios de personalidad

En los últimos años, un gran número de tests han tratado de solventar las críticas a la escasa magnitud de la validez de las pruebas de personalidad (Morgeson et al., 2007), creando pruebas compuestas con ítems contextualizados a contextos específicos como el laboral y, de esta manera, elevar los coeficientes de validez de estas pruebas. Es lo que se conoce como el efecto del marco de referencia (Frame of reference, FOR, por sus siglas en inglés), que afecta al modo de responder de las personas y que parte de la idea de que la consistencia en la variabilidad de un rasgo puede incrementarse en situaciones similares (Mlinarić y Podlesek, 2013; Holtrop et al., 2014).

Este efecto implica reducir la variabilidad entre-sujetos y la inconsistencia intra-sujetos, lo que produce un aumento en la correlación de los factores de personalidad con el desempeño laboral (Lievens, De Corte, et al., 2008). Diferentes estudios, como el meta-análisis de Shaffer y Postlethwaite (2012), han encontrado que contextualizar ítems refiriéndolos al ámbito específico del trabajo, añadiendo al texto de un ítem de ámbito general, por ejemplo, la coletilla "en el trabajo", mejora la validez, pasando de un valor medio de .11 a un valor medio de .24.

Sin embargo, hay que tener en cuenta que emplear pruebas vinculadas a contextos laborales puede ser un inconveniente si la muestra de aspirantes no ha tenido

experiencia laboral (Arribas, 2014; Mlinarić y Podlesek, 2013). Esto es algo habitual en la selección en el sector público, al que se presentan muchas personas jóvenes en busca de su primer empleo que podrían alegar que emplear un test con ítems contextualizados al mundo del trabajo supone discriminación frente a quienes tienen experiencia laboral. En consecuencia, su utilización en procesos selectivos sería potencialmente objeto de litigio (Shen et al., 2017).

El énfasis en la vinculación con el mundo del trabajo ha hecho que en los últimos años se hayan comercializado diferentes pruebas, como las pruebas de competencias no necesariamente basadas en el modelo FFM, que evalúan la personalidad laboral y que emplean ítems contextualizados a situaciones laborales (Pamos, 2016).

No obstante, las pruebas que miden FFM son mejores predictores que las pruebas que no se basan en este modelo, con diferencias estadísticamente significativas en la validez de los factores Conciencia y Estabilidad emocional, los principales predictores del desempeño laboral del modelo de los cinco grandes. Cuando estos factores son medidos por pruebas basadas en el FFM presentan magnitudes superiores a las que registran cuando son medidos por pruebas basadas en constructos alternativos (Salgado, 2003).

Además, los modelos de competencias han recibido críticas por su menor rigor científico y metodológico en comparación con las pruebas de personalidad, la incoherencia en las definiciones de las competencias laborales, una excesiva simplificación de las conductas vinculadas a un desempeño eficaz o la falta de acuerdo sobre lo que es una competencia o cómo debe medirse (Bolden y Gosling, 2006; Kurz y Bartram, 2002; Leutner y Chamorro-Premuzic, 2018). Por ello, autores como Delamare Le Deist y Winterton (2005) señalan que las competencias pueden ser más útiles a la hora de organizar los recursos internos de una organización vinculando roles de puestos a características del personal, que como predictores del desempeño laboral.

La creación de pruebas contextualizadas o de competencias está vinculada a la búsqueda de una maximización de la validez externa de los instrumentos que miden personalidad. Dado que el objetivo último en selección de personal es maximizar la correlación entre personalidad y desempeño, en las últimas décadas se ha suscitado un

debate acerca de la amplitud de la medida y su fidelidad al contenido del desempeño. No en vano, Cronbach y Glesser (1957) señalaron que cabía esperar que las medidas más amplias tuviesen una capacidad predictiva moderada de medidas amplias de criterio, mientras que para lograr una validez máxima era necesario lograr un alto grado de fidelidad entre predictor y criterio, algo alcanzable cuando criterios específicos son predichos por dimensiones específicas.

En el ámbito de la personalidad, autores como Ones y Viswesvaran (1996) indicaron que la medición de los Cinco Grandes es suficiente cuando la variable criterio es en sí misma amplia o compleja (como en el caso del desempeño laboral), puesto que la predicción óptima se alcanza cuando predictor y criterio tienen el mismo grado de amplitud.

Por otro lado, Paunonen et al. (2003) o Judge et al. (2013) indicaron que las subdimensiones o facetas de los Cinco Grandes (Costa y McCrae, 1995; DeYoung et al., 2007) presentan una mayor capacidad predictiva independientemente de la complejidad del criterio.

No obstante, investigaciones recientes realizadas desde perspectivas metodológicas innovadoras (Salgado, Moscoso, et al., 2015; Soto y John, 2019) han contribuido a clarificar el debate y han puesto en valor la capacidad predictiva de los Cinco Grandes frente a la de sus facetas, ya que los ítems de una escala amplia tienden a ser menos homogéneos y explican más varianza de la personalidad, lo que aumenta su validez externa, siendo capaces de predecir tanto criterios amplios como específicos.

3.1.5. Ventajas y desventajas de las escalas breves de personalidad

En los últimos años, la demanda de escalas breves (de menos de 10 ítems por constructo) que midan fenómenos psicológicos relevantes es cada vez mayor, debido a razones como la escasez de recursos, o la falta de tiempo o de voluntad de los participantes para responder a cuestionarios extensos. Como señalan Ployhart y Bliese (2006), no es frecuente que en la recogida de datos en el “mundo real” se puedan administrar cuestionarios extensos ya que las organizaciones o los empleados rara vez se prestan a cumplimentar instrumentos de más de una página o que tengan más de 60

elementos. Cuestionarios como el NEO-PI, compuesto por 240 ítems y que requiere un tiempo superior a media hora para ser respondido, resultan excesivamente largos y, por tanto, difícilmente aplicables en investigaciones que pretendan evaluar múltiples constructos (Ziegler et al. 2014; Olaru et al. 2016) o en oposiciones en las que se pretenda aplicar varias pruebas que evalúan constructos distintos en poco tiempo (Ramió, 2011).

Así, se han elaborado varios cuestionarios breves en el marco del modelo de los Cinco Grandes. Como ejemplos se pueden mencionar el NEO-FFI (Costa y McCrae, 1989), versión breve del NEO-PI (Costa y McCrae, 1985), el BFI (John et al., 1991), el BFI-2 (Soto y John, 2017), el Big Five Personality Trait Short Questionnaire (BFPTSQ; Morizot, 2014) o la escala de diez ítems elaborada por Rammstedt y John (2007) a partir del cuestionario de 44 ítems, breve per se, BFI.

Otro ejemplo sería OPERAS (Overall Personality Assessment Scale). Este cuestionario fue desarrollado por Vigil-Colet et al. (2013) y es un test de 40 ítems basado en el modelo de los cinco factores de personalidad, que se caracteriza por proporcionar puntuaciones libres de los efectos de aquiescencia y deseabilidad social. OPERAS presenta una buena validez convergente con otras pruebas ampliamente utilizadas dentro del modelo de los cinco grandes. Asimismo, la fiabilidad factorial de las escalas medida mediante el estadístico ORION (Ferrando y Lorenzo-Seva, 2016) es de ($r_{\theta\theta} = .86$) para Extroversión, ($r_{\theta\theta} = .77$) para Responsabilidad, ($r_{\theta\theta} = .86$) para Estabilidad emocional, ($r_{\theta\theta} = .71$) para Amabilidad y ($r_{\theta\theta} = .81$) para Apertura a la experiencia.

No obstante, las escalas breves han sido criticadas por su limitada utilidad para la toma de decisiones individuales (Emons et al., 2007; Krueger, 2012; Krueger et al., 2012, 2013). Esta perspectiva ha sido discutida por autores que afirman que las escalas breves sí pueden ser empleadas en la toma de decisiones sobre personas siempre que se especifiquen los propósitos de medición de la escala (e.g. selección, investigación) y se aporten evidencias de su validez (Schipolowski et al., 2014; Ziegler et al., 2014).

Otra posible limitación de los cuestionarios breves podría ser su consistencia interna. A menudo el Alfa de Cronbach es el único estadístico de fiabilidad que aportan los cuestionarios más empleados. Por definición, el número de ítems de las escalas breves hace que las correlaciones inter-ítems no sean altas, lo que a su vez conlleva una

baja consistencia interna. Como alternativas al Alfa de Cronbach, se han propuesto el Omega de McDonald (Schipolowski et al., 2014), el estudio de la fiabilidad a partir de las puntuaciones factoriales –ORION–, o la fiabilidad marginal desde el ámbito de la Teoría de Respuesta al Ítem (Green et al., 1984; Sireci et al., 1991).

Varios autores consideran que los cuestionarios breves son peores en cuanto a su correlación test-criterio en comparación con los cuestionarios más extensos (Credé et al. 2012; Smith et al., 2000). Sin embargo, la validez de criterio de las pruebas breves ha mostrado resultados similares a los presentados por pruebas con un mayor número de ítems (Thalmayer et al., 2011), algo corroborado por Heene et al. (2014) cuando se emplean datos poblacionales y se asumen residuos no correlacionados. Asimismo, Rammstedt y John (2007) informaron de niveles satisfactorios de fiabilidad y de validez factorial, discriminante, congruente y de criterio de su escala Big Five de diez ítems.

El uso de escalas breves podría incluso presentar más ventajas en relación a su validez. Así, podría mejorar la validez aparente para las personas participantes, que no tendrían la impresión de que los ítems están repetidos, así como centrar la evaluación en los elementos nucleares de la variable de interés. Es más, un número elevado de ítems puede tener un impacto negativo en los procesos cognitivos y motivacionales de las personas que responden a un cuestionario (Tourangeau et al., 2000). Los cuestionarios extensos causan fatiga y aburrimiento a un gran número de participantes, lo que puede disminuir la motivación para cumplimentarlos totalmente, incrementar la cantidad de datos perdidos e influir negativamente en la calidad de las respuestas de quienes participan en una investigación, que pueden acabar leyendo de manera superficial los ítems o incluso respondiendo al azar. En consecuencia, los cuestionarios breves son una alternativa eficiente para incrementar la probabilidad de obtener respuestas generadas de manera óptima por parte de los candidatos y candidatas (Burisch, 1984; Krosnick, 1999; Torreblanca Murillo, 2017; Wanous et al., 1997).

3.1.6. La distorsión de respuestas en los cuestionarios de personalidad

Una crítica frecuente a las pruebas de personalidad es que no están libres de sesgos en las respuestas. Las personas tienden presentarse a sí mismas de una manera favorable, adaptando su imagen a factores que piensan que pueden proporcionarles

ventajas o beneficios en un proceso de evaluación, y algunas personas utilizan más el manejo de impresiones que otras (Barrick y Mount, 2014; Salgado, 2016).

Pese a esta visión predominante, algunos autores argumentan que las respuestas a un auto-informe de personalidad captan la impresión que el individuo desea causar, y que esta refleja lo que la mayor parte de la gente hace en sus interacciones con los demás, por lo que la distorsión en las respuestas no debería ser una preocupación primordial (Hogan, 1991; Hogan et al., 2007). Autores como Bangerter et al. (2012) o Seisdodos (1993) enfatizan que distorsionar las respuestas es una adaptación inteligente de las personas a las demandas del entorno evaluativo, adaptándose subjetivamente a un perfil que les favorece desde un punto de vista social o laboral, así como a las normas que se supone que deben respetarse. Asimismo, Bourdeau y Lock (2005) consideran que la distorsión representa una varianza válida e interpretable.

Además, diferentes investigaciones han encontrado que la distorsión en las respuestas no altera de manera significativa los coeficientes de validez de las pruebas de personalidad. Las diferencias en las medias y en el orden de los candidatos son mínimas al comparar las respuestas de quienes responden en contextos de selección con las de quienes responden en un contexto de promoción profesional y ya están contratados. Ello es así pese a que los primeros tienen una tendencia mayor que los segundos a distorsionar sus respuestas. Consecuentemente, muchos autores consideran que estas pruebas pueden ser utilizadas con garantías en contextos selectivos (Ellingson et al., 2007; Hogan et al., 2007; Ones et al., 1996; Schmitt y Oswald, 2006).

No obstante, aunque en un cuestionario bien diseñado el determinante principal de las respuestas sea el contenido de la dimensión que se pretenda medir, los sesgos en las respuestas influyen, en cierta medida, en los resultados de los participantes y no se deben ignorar (Ferrando et al., 2009). Los dos sesgos de respuesta principales según la literatura son la aquiescencia y la deseabilidad social (Paulhus, 1991), y ambos actúan de manera simultánea en mayor o menor medida (Hofstee et al., 1998; Neill y Jackson, 1970). La aquiescencia es la tendencia a responder afirmativamente o mostrarse de acuerdo con los ítems sin tener en cuenta su contenido (Rammstedt et al., 2013). En cuanto a la deseabilidad social, ha sido definida como “tendencia a contestar a los ítems de modo que se responde a las presiones sociales o normativas en lugar de proporcionar un auto-informe verídico” (Ellingson et al., 2001, p. 122).

Las propiedades psicométricas de los cuestionarios y su estructura factorial están afectadas por factores como la edad, la complejidad cognitiva o el nivel educativo, los cuales a su vez se ven afectados por efectos diferenciales de la aquiescencia (Rammstedt et al., 2010; Soto et al., 2008, 2011). Asimismo, hay evidencia que señala que algunos candidatos honestos son perjudicados por aquellos que distorsionan sus respuestas y que este efecto es máximo cuando sólo se contrata a una pequeña proporción de las personas que aspiran a un puesto (Barrick y Mount, 2014).

Esto es algo que coincide con lo apuntado por el meta-análisis de Salgado (2016), que ha hallado que la distorsión en las respuestas conlleva que estas sean más homogéneas, haciendo que las personas se muestren más similares que lo que realmente son y provocando que los candidatos que falsean sus respuestas obtengan mejores resultados. Este fenómeno hace que las medias de las variables aumenten o disminuyan (según lo que se presuponga conveniente para obtener un mejor resultado en cada caso) y produce una reducción en la desviación típica de las dimensiones. Todo ello afecta negativamente a los coeficientes de cuatro estadísticos: covarianza, fiabilidad (con un valor medio de .65 en el mencionado meta-análisis), error estándar y correlación (por ejemplo, en el caso de la validez). Estos efectos, obtenidos con muestras de aspirantes a puestos reales, no se habían observado en estudios similares anteriores ya que la mayoría de las investigaciones se habían basado en muestras de personas ya contratadas y, por tanto, no motivadas a falsear sus respuestas.

Los métodos utilizados históricamente para reducir la distorsión en las respuestas han sido emplear ítems de elección forzosa (Heggestad et al., 2006), descartar aquellos candidatos con puntuaciones altas en escalas que midan deseabilidad social o corregir las puntuaciones globales en función de la información de una escala de deseabilidad social (Barrick y Mount, 2014; Goffin y Christiansen, 2003), imponer límites de tiempo para responder (Komar et al., 2010), informar de la existencia de una escala que mide distorsión y crear baremos a partir de muestras de aspirantes a puestos de trabajo (Salgado, 2005, 2016).

Sin embargo, la mayor parte de estas opciones presentan una serie de problemas (Hugues y Batey, 2017). Así, los ítems de elección forzosa crean escalas parcialmente ipsativas, lo que altera las propiedades psicométricas y plantea dudas sobre los constructos medidos, además de no permitir comparaciones normativas entre individuos (Heggestad et al., 2006), si bien en los últimos años se han propuesto metodologías que

pueden ayudar a paliar dichos problemas (Brown y Maydeu-Olivares, 2011; Salgado, Anderson, et al., 2015; Salgado y Lado, 2018). Asimismo, las opciones que implican descartar o corregir las puntuaciones de las personas con puntuaciones altas en deseabilidad social presentan problemas de tipo teórico, práctico y legal que desaconsejan su uso (Barrick y Mount, 2014; Salgado, 2005, 2016; Schmitt y Oswald, 2006). Reducir el tiempo para responder, por su parte, parece reducir la deseabilidad en personas con niveles de aptitud cognitiva más baja (Komar et al., 2010).

Las dos opciones más recomendables –informar a los y las aspirantes y emplear baremos específicos del contexto de selección-, se han mostrado de utilidad según diversas investigaciones (Barrick y Mount, 2014; Salgado, 2005, 2016). Ambas son compatibles con la propuesta de Ferrando et al. (2009), que defienden un método en el que la aquiescencia y la deseabilidad social sean consideradas como factores adicionales que puedan ser distinguidos el uno del otro y a su vez del factor que se pretenda medir en cada caso. La aplicación de este procedimiento a nivel de calibración de los ítems proporciona tres medidas de la saturación de cada elemento: una saturación sobre el factor de contenido que la prueba pretende medir, y dos saturaciones sobre dos factores ortogonales identificados como Deseabilidad Social y Aquiescencia, lo que permite obtener puntuaciones individuales libres de efectos de sesgo de respuesta.

3.2. Revisión de los cuestionarios de personalidad más empleados en España en el ámbito de los Recursos Humanos

A la espera de la próxima adaptación al castellano de cuestionarios como BFQ-2 (Caprara et al., 2008) o NEO-PI-3 (McCrae, Costa et al., 2005) a partir de 2021 por parte de la editorial TEA, en las siguientes páginas se analizan los cuestionarios de personalidad en castellano y validados con muestras españolas disponibles para los profesionales de recursos humanos. El análisis comprende tanto instrumentos que están a la venta por parte de editoriales especializadas como instrumentos empleados en investigaciones aplicadas. Se han tenido en cuenta los criterios de evaluación de Muñiz et al. (2011) y las cuestiones planteadas en los epígrafes precedentes.

Se han analizado los tests de personalidad más empleados según el análisis de Elosua (2012) en el ámbito de la Psicología del Trabajo (BFQ, NEO PI, PAPI y TPT) y el de Muñiz et al. (2020): 16PF, CompeTEA, BIP o PAPI. La excepción ha sido IPV (ECPA, 2016), al tratarse de un cuestionario que se ciñe exclusivamente a la personalidad de vendedores. Tampoco se han considerado cuestionarios que solo miden aspectos concretos de la personalidad, como, por ejemplo, la Escala de Habilidades Sociales de Gismero (2000) o herramientas que incluyen escalas de evaluación clínica, como PAI, MMPI o MCMI. Asimismo, se han descartado instrumentos cuestionados por sentencias judiciales (STS de 7 de mayo de 2008), como el CPI (Gough y Bradley, 1996). Además, se han examinado una serie de tests localizados en base a búsquedas realizadas en Google Académico, Psycinfo y EBSCOhost y al listado de tests en español de Schlueter et al. (2018), de los que se han podido conseguir datos de sus adaptaciones o validaciones con muestras españolas. Una excepción ha sido la prueba OPQ 32, instrumento que en su muestra de validación original en inglés presenta buenos resultados de fiabilidad, pero una validez de constructo cuestionable (Pamos, 2016; SHL, 2006)

La mayor parte de los cuestionarios se basan en el modelo FFM. Concretamente, hay que mencionar las pruebas BFCP (Aguado et al., 2008), BFI (Benet-Martínez y John, 1998), BFQ (Caprara et al., 2007), BFPTSQ (Ortet et al., 2017), CBP (Torreblanca Murillo, 2017), FFPI (Rodríguez-Fornells et al., 2001), IP/5F (Salgado, 1998), NEO-PI y su versión reducida -NEO-FFI- (Costa y McCrae, 2008; Solé, 2006), OPERAS (Vigil-Colet et al., 2013), QI5F (Salgado, 2014) y la Versión reducida de los Marcadores Bipolares de Goldberg, VRMBG (García et al., 2004).

Asimismo, existen cuestionarios que miden una serie de escalas englobadas en factores de segundo orden cercanos conceptualmente al modelo FFM: 16PF5 (Cattell et al., 1993; Aluja y Blanch, 2002), cuyos factores de segundo orden son Extraversión, Ansiedad, Dureza, Independencia y Auto-control; CPS (Fernández-Seara et al., 2008), cuyos factores superiores son Ajuste, Liderazgo, Independencia, Consenso y Extraversión; o TPT (Corral et al., 2002; Arribas-Águila, 2011), que tiene tres factores de segundo orden: Estabilidad Emocional, Apertura Mental y Responsabilidad. El cuestionario ZKA (Aluja et al., 2010), por su parte, está basado en el modelo alternativo al FFM, el Alternative Five Model (AFM), que tiene un enfoque psicobiológico y cuyos

factores se corresponden, en parte, con los de los Cinco Grandes: Neuroticismo, Búsqueda de sensaciones (que se corresponde parcialmente con baja Responsabilidad), Agresividad (constructo próximo a baja Amabilidad), Extroversión y Actividad, que sustituye al factor Apertura Mental del FFM (Zuckerman, 1991; Zuckerman, et al., 1991).

Por otro lado, existen cuestionarios que parten de modelos diferentes. Se trata, fundamentalmente de cuestionarios de competencias laborales como AXIS (Pamos, 2016), BIP (Hossiep y Paschen, 2008), COMPETEA (Arribas y Pereña, 2009), PRISMA (Psicólogos Empresariales, 2008), SOSIA (Gordon et al., 1997) o PAPI (Kostick, 1976; Cubiks, 2014; Sanz et al., 2006). En este último caso cabe distinguir entre la versión normativa (PAPI-N), la ipsativa (PAPI-I) y la versión más reciente (PAPI3). En este análisis nos hemos centrado en la versión que ha sido analizada en un mayor número de estudios, PAPI-N.

Otros instrumentos que parten de concepciones teóricas diferentes al FFM son los cuestionarios APP (Rodríguez González, 2004), CEP (Pinillos, 1964), Cuestionario de personalidad de Eysenck, EPQ-R (Aguilar et al., 1990; Eysenck y Eysenck, 2001; Ibañez, 1997), Inventario Millon de Estilos de Personalidad, MIPS (Millon, 2001; Aparicio y Sánchez Lopez, 1999), o PPG-IPG (Gordon, 2001).

La mayor parte de cuestionarios aportan información sobre la fiabilidad entendida como consistencia interna y medida mayoritariamente mediante el estadístico Alfa de Cronbach. En la Tabla 3.2 se presentan los valores de las medianas de los factores medidos por cada uno de los cuestionarios. En todos los casos tienen valores que cabe considerar adecuados, superiores a .60 (Muñiz et al., 2011), siendo los valores más altos, superiores a .80, los que presentan BFCP, BFPTSQ, BIP, CEP, FFPI, IP/5F, NEO, OPERAS, VRMBG y ZKA.

Habida cuenta de los problemas psicométricos asociados al Alfa de Cronbach, que solo estima adecuadamente la fiabilidad cuando se cumple el requisito de la tau-equivalencia (Reuterberg y Gustafsson, 1992; Viladrich et al., 2017), algunas pruebas aportan datos de coeficientes como Omega de McDonald (Pamos, 2016), Alfa Ordinal, en el caso de TPT, con una mediana de .74 (Arribas-Aguila, 2011), u ORION, con una mediana de .81 en el caso de OPERAS (Vigil-Colet et al., 2013).

Son escasos los cuestionarios que aportan datos de fiabilidad test-retest con muestras españolas. Concretamente, se trata de los siguientes: APP, AXIS, BFPTSQ, CPS, EPQ-R, IP/5F, OPERAS y QI5F.

En cuanto a las evidencias de validez aportadas, hay diferencias de unos cuestionarios a otros. La mayor parte ofrecen datos sobre correlaciones con otros tests (validez convergente). En muchos casos se proporcionan datos de saturaciones de los ítems en los factores correspondientes en análisis factoriales exploratorios. En menos casos se ofrecen datos de análisis factoriales confirmatorios y en ninguno de ellos se proporciona información sobre el sesgo de los ítems. Únicamente IP/5F ofrece datos de validez predictiva empleando como criterio el desempeño laboral, si bien con muestras muy reducidas.

En términos generales, los tests basados en el modelo FFM presentan un mejor ajuste factorial a la estructura teórica propuesta que los tests basados en modelos alternativos, especialmente por lo que se refiere a tests de competencias. A modo de ejemplo, entre las pruebas basadas en FFM, BFI presentó un índice CFI de .92, mientras que OPERAS obtuvo los siguientes índices: CFI = .97, GFI = .98, y RMSEA = .03. BPPTSQ, por su parte, tras obtener índices poco satisfactorios mediante AFC, presentó los siguientes estadísticos mediante ESEM: CFI = .90, TLI = .87, y RMSEA = .04.

Por lo que se refiere a las pruebas de competencias o de personalidad no basadas en el modelo de los 5 factores, TPT presentó los siguientes índices: CFI = .92, AGFI = .89, y RMSEA = .08. Los valores de ajuste al modelo en el caso de BIP fueron: GFI = .79, AGFI = .69, y RMSEA = .15. AXIS, por su parte, obtuvo los siguientes valores: CFI = .92, y RMSEA = .07. En cuanto a ZKA, presentó los mejores resultados al tener en cuenta los términos de error de las correlaciones: CFI = .93, TLI = .90, y RMSEA = .08.

Tabla 3.2.

Constructos, Tipo de Escala, Fiabilidad y Validez de los Cuestionarios de Personalidad

Test	Constructo	Escala / Puntos	Mediana fiabilidad	Evidencias validez
16PF 5	16 factores y 5 factores 2º orden	V-?-F	.76	2, 3
APP	4 factores	Ipsativa	.76	1,5
AXIS	21 competencias y 3 factores 2º orden	Likert / 7	.65	1, 2, 3
BFCP	FFM	Likert / 6	.88	1,2
BFI	FFM	Likert / 5	.79	1, 2, 3
BFPTSQ	FFM	Likert / 5	.83	1, 2, 3
BFQ	FFM	Likert / 5	.76	1, 2
BIP	14 dimensiones y 4 factores 2º orden	Likert / 6	.80	1, 2, 3
CBP	FFM	Likert / 5	.76	1, 2
CEP	3 factores	Sí-No-?	.81	2
COMPETEA	20 competencias y 5 factores 2º orden	Likert / 4	.73	1, 3
CPS	15 Escalas y 5 factores 2º orden	V/F	.73	1, 2
EPQ-R	3 factores	Sí-No	.79	2
FFPI	FFM	Likert / 5	.84	1, 2
IP/5F	FFM	Likert / 3	.86	1, 2, 3, 5
MIPS	24 escalas/ 3 factores	V-F	.72	1,2
NEO FFI	FFM	Likert / 5	.81	1, 2
NEO PI	FFM	Likert / 5	.84	1, 2
OPERAS	FFM	Likert / 5	.81	1, 2, 3
PAPI	20 escalas	Likert / 7	>.41	1, 2
PPG-IPG	9 factores	Ipsativa	.74	1, 2
PRISMA	32 escalas	Ipsativa	.70	1, 2, 3
QI5F	FFM	Cuasi-ipsativa	.73	1, 2, 3, 5
SOSIA	21 competencias/ 4 factores	Ipsativa	.69	2
TPT	15 facetas y 3 factores (EE, AM, RP)	Likert / 4	.62	1, 3
VRMPG	FFM	Likert / 9	.82	1, 2, 3
ZKA	AFM	Likert / 4	.88	1, 2, 3

Nota. 1. Validez convergente; 2. Análisis Factorial Exploratorio; 3. Análisis Factorial Confirmatorio; 4. Funcionamiento diferencial de ítems; 5. Validez predictiva

La mayor parte de los tests empleados en España son fruto de adaptaciones de otras lenguas, y en muchos casos no se informa sobre el procedimiento de adaptación seguido en sus respectivos manuales (Muñiz y Fernández-Hermida, 2010). En el caso de los cuestionarios de personalidad analizados, son adaptaciones de otras lenguas, principalmente del inglés, en 11 casos (16PF5, APP, BFI, BFPTSQ, EPQ-R, MIPS, NEO PI, NEO FFI, PAPI, PPG-IPG, VRMBG), pero también hay adaptaciones del alemán, en el caso de BIP, del francés, en el caso de SOSIA, o del italiano, en el caso de BFQ. FFPI se elaboró simultáneamente en inglés, holandés y alemán. ZKA, por su parte, se elaboró paralelamente en castellano e inglés. Los cuestionarios creados originalmente en castellano son AXIS, BFCP, CBP, CEP, COMPETEA, CPS, IP/5F, OPERAS, PRISMA, QI5F y TPT.

Por lo que respecta a los cuestionarios de personalidad adaptados a otras lenguas del Estado, hay versiones en catalán de diversos instrumentos: 16PF5, BFQ, BIP, COMPETEA, EPQ-R, NEO-PI, PAPI, SOSIA y TPT. En euskera, únicamente existen en la actualidad versiones de NEO-PI y NEO-FFI (Balluerka et al., 2007; Gorostiaga et al., 2011; Haranburu et al., 2007), con características psicométricas similares a las de las adaptaciones al castellano que se describen en este epígrafe.

Como se puede comprobar en la Tabla 3.3, únicamente presentan ítems contextualizados en el ámbito laboral las pruebas de competencias, es decir, AXIS, BIP, COMPETEA, PAPI y SOSIA. La excepción es PRISMA, que pese a medir competencias laborales, no emplea ítems referidos al trabajo.

Este último es uno de los escasos cuestionarios de tipo ipsativo, junto a APP, PPG-IPG, SOSIA o QI5F, este último, cuasi-ipsativo. El resto de tests son de tipo normativo y, en su mayoría emplean escalas Likert de 4 o 5 puntos (ver Tabla 3.2).

En cuanto a la longitud de los cuestionarios, la mayoría han optado por un número elevado de ítems, siendo los más breves (inferiores a 60 ítems) APP, BFI, BFPTSQ, CBP, NEO FFI, OPERAS y PPG-IPG, lo que los convierte en las pruebas más susceptibles de ser empleadas de manera conjunta con otro tipo de instrumentos de selección (assessment centers, tests de aptitudes, exámenes de conocimientos, pruebas prácticas...) en procesos masivos que se quieran realizar en breve espacio de tiempo.

Cabe señalar que los tiempos estimados de aplicación mencionados en los manuales técnicos están pensados para personas que respondan de manera sincera, algo que no sucede en pruebas competitivas, donde las personas que distorsionan sus respuestas tienden a emplear más tiempo para responder (Komar et al., 2010).

En lo que se refiere a las medidas empleadas para combatir la distorsión de respuestas por falta de sinceridad de los/as candidatos/as, la mayoría de los cuestionarios optan por aplicar una escala de deseabilidad social, o incluso dos o más escalas, en los cuestionarios 16PF5, CPS o MIPS. En COMPETEA se señala que hay que prestar atención tanto a puntuaciones elevadas como a puntuaciones muy bajas en dicha escala, ya que las puntuaciones extremas están vinculadas a Neuroticismo, Psicoticismo, o Dependencia (Arribas, 2014). Tests como BFCP, BFI, BFPTSQ, CBP, FFPI, NEO-FFI, VRMG o ZKA no emplean ninguna medida, mientras que OPERAS opta por el método de Ferrando et al. (2009) para corregir los sesgos de deseabilidad social y de aquiescencia.

En consecuencia, de entre los tests analizados cabe resaltar las propiedades de OPERAS. Este test destaca en cuanto a la calidad de los índices psicométricos que presenta, con una mediana de fiabilidad superior a .80, el modelo teórico del que parte (FFM) y su excelente ajuste al mismo, su brevedad, ya que tiene 40 ítems, o la novedosa técnica que emplea para combatir los sesgos de respuesta. Además, las características de OPERAS coinciden con las recomendaciones sobre la construcción de instrumentos para incrementar la validez externa de Soto y John (2019): escalas de entre 6 y 9 ítems, escalas que miden factores amplios como los Cinco Grandes (frente a posibles subescalas del modelo FFM), y equilibrio entre ítems de contenido positivo y negativo.

Asimismo, si bien Shaffer y Postlethwaite (2012) señalan que los instrumentos con ítems contextualizados al entorno laboral presentan una validez predictiva superior a la de los instrumentos con ítems genéricos por el efecto “marco de referencia”, estos últimos presentan menos problemas de ajuste a la legalidad en el ámbito de selección de personal en el sector público. Hay que subrayar que la jurisprudencia indica que se debe garantizar que todos los candidatos y las candidatas estén en condiciones de igualdad en el acceso, y un instrumento con ítems de contenido laboral podría interpretarse como perjudicial para candidatos sin experiencia de trabajo. Todo ello hace de OPERAS un

instrumento óptimo para su adaptación al euskera, lo que constituye el objetivo 2 de esta tesis.

Tabla 3.3.

Editorial, Ítems Contextualizados o no, Número de Ítems, Tiempo de Aplicación y Medidas contra la Deseabilidad Social de los Cuestionarios de Personalidad

Test	Editorial	Contextualizado (C) o no (NC)	Nº ítems/Duración	Medidas contra la deseabilidad social (DS)
16PF 5	TEA	NC	185 / 45 min	Escala DS, Infrecuencia y Aquiencia
APP	Thomas	NC	24	
AXIS	-	C	147	Escala DS
BFCP	IIC	NC	207	
BFI	-	NC	44	-
BFPTSQ	-	NC	50	-
BFQ	TEA	NC	132 / 30 min.	Escala DS
BIP	TEA	C	210 / 45 min.	Escala DS
CBP	-	NC	20	-
CEP	TEA	NC	145 / 25 min.	Escala sinceridad
COMPETEA	TEA	C	170 / 30 min	Escala DS
CPS	TEA	NC	233 / 35-50 min.	2 Escalas DS
EPQ-R	TEA	NC	83 / 25 min.	
FFPI	-	NC	100 / 15 min.	-
IP 5F	Metis	NC	200 / 20-30 min.	Escala DS
MIPS	-	NC	180 / 40 min.	1 escala de consistencia y 2 de impresiones
NEO FFI	TEA	NC	60 / 15 min.	-
NEO PI	TEA	NC	240 / 40 min.	Escala DS
OPERAS	-	NC	40 / 10 min	Ferrando et al. (2009)
PAPI	Grupo FactHum	C	126 N(137) I (126)	Escala DS
PPG-IPG	TEA	NC	38 / 25 min.	
PRISMA	SHL-PE	NC	96 elementos ipsativos / 30 min	Congruencia
Q15F	Metis	NC	140	
SOSIA	TEA	C	98 elementos ipsativos / 45 min	-
TPT	TEA	NC	160 / 30 min.	Escala DS
VRMPG	-	NC	25	-
ZKA	-	NC	200	-

Nota. NC = No contextualizado; C = Contextualizado. DS = Deseabilidad Social

3.3. Adaptación de OPERAS

3.3.1. *Proceso de traducción de los ítems, entrevistas cognitivas y fase piloto*

La traducción de los ítems del cuestionario OPERAS se realizó en base a las recomendaciones de Hambleton y Patsula (1999). Se utilizó un diseño de traducción inversa basada en la versión original en castellano. En la adaptación participaron dos traductoras del Servicio Oficial de Traducción del IVAP, cuatro psicólogos/as expertos/as en pruebas de personalidad y dos metodólogas. Tras llegar a una versión consensuada traducida al euskera y, a partir de ésta, a una versión consensuada retro-traducida al castellano, los/as integrantes del equipo de traductores/as compararon cada uno de los ítems de las versiones original e inversamente adaptadas del instrumento, examinando las posibles no-equivalencias en cuanto al significado. A partir de este análisis se llevaron a cabo las modificaciones oportunas en la versión en euskera de OPERAS.

Con posterioridad, se verificó la adecuación de los ítems del cuestionario a la población objetivo mediante la realización de entrevistas cognitivas y la aplicación de una versión piloto de la prueba. En las entrevistas cognitivas participaron 10 personas (5 mujeres y 5 varones) con edades comprendidas entre los 41 y los 63 años ($M = 48.6$ años; $SD = 9.70$), mientras que en el estudio piloto participaron 80 personas (31 varones y 49 mujeres) con edades comprendidas entre los 23 y los 63 años ($M = 46.01$ años; $SD = 7.72$). Todas ellas residían en la Comunidad Autónoma del País Vasco y presentaban un dominio entre medio y alto de la lengua vasca (Perfiles B2 a C2 de conocimiento del Marco Común Europeo de Referencia para las Lenguas). La muestra estuvo compuesta por empleados públicos de la Administración Pública Vasca y los datos fueron recogidos en sus centros de trabajo preservando su anonimato.

Las entrevistas cognitivas son una herramienta para detectar y corregir posibles problemas vinculados a la manera de responder a un cuestionario, así como para evaluar la adecuación y el grado de comprensión del lenguaje empleado en los ítems del mismo. Por tanto, el objetivo de las entrevistas cognitivas fue tratar de conocer por qué las personas participantes respondían de una manera determinada, proporcionando evidencias de validez basadas en el proceso de respuesta (Caicedo y Zalazar-Jaime, 2018). Posteriormente, en la aplicación piloto se realizó un primer análisis estadístico de

los ítems centrado fundamentalmente en la vinculación de cada elemento con su dimensión teórica. Con la información obtenida en las entrevistas cognitivas y la prueba piloto se procedió al análisis cuantitativo y cualitativo de los elementos del test y a la reformulación de algunos ítems de manera consensuada en el equipo de traducción.

Así, se tomaron en consideración las sugerencias que realizaron los y las participantes en las entrevistas cognitivas y en la aplicación piloto, 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/as reformuló los elementos a fin de facilitar su comprensión, teniendo en cuenta las sugerencias de las personas participantes, o añadió sinónimos a los términos empleados en la fase de traducción (Ítems 27, 30, 34). Por otro lado, con los datos recogidos en la fase piloto, se calcularon las puntuaciones medias y las desviaciones típicas de cada uno de los ítems, así como los índices de homogeneidad con respecto a sus correspondientes factores. El único ítem con un índice de homogeneidad menor que .30 (Ítem 13) fue revisado por los integrantes del equipo de traducción, quienes, siguiendo las sugerencias de los expertos en la evaluación de la personalidad, llevaron a cabo algunas modificaciones a fin de que se aproximara en mayor medida a su correspondiente factor de personalidad teórico. El instrumento derivado de este proceso de revisión fue el que se utilizó para realizar la validación empírica rigurosa y la tipificación de la versión de OPERAS adaptada al euskera.

3.3.2. Método

3.3.2.1. Participantes

La muestra de personas que participaron en la adaptación de OPERAS fue la misma que se describió en el capítulo anterior: 678 empleados/as públicos/as (444 mujeres y 234 varones), con una edad media de 44.63 años ($SD = 7.66$).

En base al tipo de titulación requerida para el acceso, cabe distinguir entre personas que desempeñaban puestos de alta complejidad, aquellos que requerían estudios universitarios ($N = 384$), frente a personas que desempeñaban puestos de baja complejidad ($N = 294$) o que no exigían estudios de educación superior.

En cuanto a la titulación académica que poseían, la mayoría señaló estar en posesión de una Licenciatura ($N = 420$), seguidos de 120 Diplomados/as, 113 personas con Educación Secundaria, 20 con Educación Primaria y 5 con Certificado de Escolaridad. En consecuencia, había 540 personas con estudios universitarios (nivel de educación alto) frente a 138 que no los poseían (nivel de educación bajo).

3.3.2.2. Instrumentos

Además de la adaptación al euskera de OPERAS, en el presente estudio se utilizaron el cuestionario de personalidad basado en los Cinco Grandes Big Five Inventory (BFI; John et al., 1991), en su versión española, adaptada por Benet-Martínez y John (1998), un cuestionario de variables sociodemográficas y laborales, ya descrito en el capítulo precedente, y una serie de instrumentos para analizar su red de relaciones nomológicas: escala para evaluar OCBs de Lee y Allen (2002); el cuestionario de desempeño adaptativo de Charbonnier-Voirin y Roussel (2012), cuya adaptación al castellano ha sido descrita en el capítulo precedente; y la adaptación al castellano del cuestionario desarrollado por Piasentin (2007) para medir ajuste persona-organización. Dado que en el presente capítulo no se evalúa la red de relaciones nomológicas de OPERAS en euskera, los instrumentos empleados para evaluar la relación de OPERAS en euskera con otras variables se describen de forma exhaustiva en el capítulo 4.

3.3.2.3. Procedimiento

Para la fase de validación experimental de OPERAS, se contó con la colaboración voluntaria de la muestra descrita en el capítulo 2. La versión adaptada al euskera de OPERAS fue administrada junto con los instrumentos destinados a evaluar su red nomológica citados en el epígrafe anterior. El orden de aplicación de dichos instrumentos fue: Cuestionario de variables sociodemográficas y laborales, cuestionario de personalidad OPERAS (2013), Escala de Lee y Allen para evaluar OCBs (2002), Escala para medir APO de Piasentin (2007), y Escala de Charbonnier-Voirin y Roussel (2012).

El Comité de Ética para las Investigaciones relacionadas con Seres Humanos (CEISH/GIEB) de la UPV/EHU informó favorablemente con respecto a la presente investigación. Las sesiones se desarrollaron de manera grupal en los respectivos centros de trabajo, donde las y los participantes cumplimentaron los cuestionarios. Al inicio de cada sesión se informó a quienes participaron en la investigación acerca de los

pormenores de la misma y del trabajo a realizar. Tras dicha explicación cumplimentaron un documento para declarar su consentimiento informado antes de responder a los instrumentos necesarios para la validación y adaptación al castellano de OPERAS. Para preservar su anonimato, cada participante tuvo que elaborar, al comienzo de cada sesión, una clave de identificación de 5 cifras. Para facilitar su memorización las tres primeras cifras debían extraerlas de un número de teléfono (fijo, móvil personal, teléfono de trabajo...) y las dos últimas del día de nacimiento.

3.3.2.4. Análisis de datos

Con el fin de analizar las propiedades psicométricas del instrumento, se llevaron a cabo los siguientes procedimientos de análisis de datos:

a) Análisis de los ítems del instrumento: estadísticos descriptivos (media, desviación típica, asimetría, curtosis). En los ítems positivos (en los que puntuaciones altas indican mayor nivel del factor de personalidad medido), cuando la media superaba el promedio teórico de 3 de la escala de medida tipo Likert de 5 puntos, el elemento podía ser etiquetado como atractivo en la muestra, mientras que, si la media estaba por debajo del citado promedio teórico, el elemento podía ser etiquetado como poco deseable socialmente en la muestra. En los ítems negativos sucedía a la inversa.

Se calculó el índice de homogeneidad corregido, a saber, la correlación entre cada elemento y el factor o la faceta a la que éste pertenecía excluyendo dicho elemento. Asimismo, se estimó, de cada una de las opciones de respuesta puntuables, el porcentaje de atracción de cada alternativa, es decir, el porcentaje de ocasiones en las que los participantes la seleccionaron en cada ítem (capacidad de discriminación de las opciones de respuesta).

Por otro lado, se evaluaron los ítems desde el enfoque de la Teoría de Respuesta al Ítem (TRI) empleando IRTPRO. Dadas las características de OPERAS, en el Modelo de Respuesta Graduada (MRG), se consideraron 4 umbrales de respuesta (ya que se empleaba una escala Likert de 5 puntos). En consecuencia, se aplicaron cuatro estimaciones para el parámetro b , una para cada umbral, es decir, $k-1$ umbrales (Hidalgo-Montesinos y French, 2016; Penfield, 2014). Asimismo, se calcularon las CCI para cada ítem.

b) Análisis de la dimensionalidad de la prueba y de su validez de constructo. A fin de examinar la dimensionalidad del inventario, se siguió el procedimiento empleado en la validación original de OPERAS. Así, se llevó a cabo un análisis factorial con una estimación basada en Unweighted Least Squares (ULS), mediante rotación ortogonal Varimax semi-especificada empleando el programa Psychological Test Toolbox de la Universitat Rovira i Virgili que permite realizar análisis factoriales controlando deseabilidad social y aquiescencia. (El programa está disponible en esta dirección electrónica: <http://psico.fcep.urv.cat/utilitats/PsychologicalTestToolbox/index.html>)

Asimismo, se llevaron a cabo análisis AFC y ESEM. En pruebas de personalidad basadas en los cinco grandes factores se ha llegado a cuestionar la noción fundamental relativa a poder identificar cinco factores ortogonales ya que el FFM no presenta un ajuste adecuado en análisis factoriales confirmatorios porque los ítems no son indicadores puros de un sólo factor y se dan correlaciones elevadas entre dimensiones (Vassend y Skrandal, 2011). Asimismo, y dado que la investigación sobre el FFM se ha basado en análisis factoriales exploratorios (AFE), se emplean con frecuencia modelos de ecuaciones estructurales exploratorios (ESEM). Este procedimiento es menos restrictivo que el AFC a la hora de medir la estructura factorial, integra las características más adecuadas de los AFCs, los modelos de ecuaciones estructurales y el AFE, y genera tests confirmatorios de estructuras teorizadas a priori (Booth y Hugues, 2014; Marsh et al., 2010; Marsh et al., 2014).

Por otro lado, se analizaron las correlaciones entre los factores de personalidad de OPERAS y la variable Deseabilidad Social. Asimismo, se analizaron las correlaciones inter-ítems para evaluar la unidimensionalidad de los mismos. Como se ha indicado en el Capítulo 2, el nivel óptimo de homogeneidad se da cuando la media está entre .2 y .4, y para el caso de escalas conceptualmente amplias como los factores de personalidad, valores entre .15 y .20 pueden ser aceptables (Briggs y Cheek, 1986; Clark y Watson, 2019).

c) A fin de obtener evidencia de validez convergente de la adaptación al euskera de OPERAS, se calcularon los coeficientes de correlación de Pearson entre las puntuaciones de las dimensiones de OPERAS y las del BFI (John et al., 1991) que miden los mismos constructos.

d) Evaluación de la fiabilidad. Al igual que en la validación del instrumento original, la fiabilidad entendida como consistencia interna se midió mediante el estadístico ORION (acrónimo de Overall Reliability of fully-Informative prior Oblique N-EAP scores), que emplea puntuaciones factoriales obtenidas a partir de las puntuaciones de la matriz inter-factorial (Ferrando y Lorenzo-Seva, 2016).

Como las escalas breves como OPERAS no tienen muchos ítems, las correlaciones inter-ítems no suelen ser altas, lo que a su vez conlleva una baja consistencia interna si se mide mediante Alfa de Cronbach. Dadas las características del Alfa de Cronbach, al igual que en el capítulo precedente, se empleó el Omega de McDonald (Schipolowski et al., 2014; Viladrich et al., 2017) y la fiabilidad marginal desde la perspectiva de la Teoría de Respuesta al Ítem. Además, se calcularon mediante IRTPRO las funciones de información para cada factor.

Asimismo, se evaluó la fiabilidad test-retest, utilizando un subgrupo de personas de la muestra. Para evaluar la estabilidad temporal de los factores de personalidad se utilizó la técnica test-retest, calculando el coeficiente de correlación de Pearson (Leppink y Pérez-Fuster, 2017; Rammstedt y Beierlein, 2014). A tal fin, se llevó a cabo una segunda aplicación del instrumento a una submuestra de 151 participantes que realizaron la segunda aplicación de OPERAS un mes después de la primera.

e) Realización de estudios de funcionamiento diferencial de ítems con distintos subgrupos de participantes. El objetivo del análisis del funcionamiento diferencial de los ítems respecto a la lengua materna (euskera o castellano) era examinar si el lenguaje mediante el cual estaban expresados los ítems podía favorecer que los individuos con diferente lengua materna obtuviesen puntuaciones de mayor o menor valor en alguno de los ítems de OPERAS.

El análisis del funcionamiento diferencial de los ítems (FDI) fue llevado a cabo mediante el procedimiento de regresión logística (Swaminathan y Rogers, 1990) dentro de cada una de las dimensiones de OPERAS, utilizando la sintaxis específica para la detección del FDI mediante SPSS creada por Slocum et al. (2004). Un ítem presenta FDI cuando el tamaño del efecto presenta un valor de ΔR^2 de Nagelkerke de al menos .035, según los criterios de Jodoin y Gierl (2001).

En el siguiente epígrafe se describen los resultados obtenidos en los análisis realizados, que se llevaron a cabo mediante los programas informáticos, Psychological Test Toolbox, Factor, SPSS, MPlus e IRTPRO.

3.3.3. Resultados

3.3.3.1. Análisis de los ítems del instrumento

Se llevaron a cabo diversos análisis para estudiar las características psicométricas de los ítems de la versión adaptada al euskera. En la tabla 3.4 se proporcionan los estadísticos descriptivos (medias, desviaciones típicas, asimetría, curtosis) de los elementos de OPERAS. Asimismo, esta tabla ofrece los índices de homogeneidad corregidos (correlación total de elementos corregida) referidos a cada uno de los 5 factores de OPERAS.

Los ítems pueden considerarse atractivos (medias >3 en la mayoría de ítems que reflejaban la dimensión medida de manera positiva, y <3 en los ítems que reflejaban la dimensión medida de manera negativa). Asimismo, presentaron puntuaciones medias en torno al punto medio de la escala y desviaciones típicas cercanas a 1 en la mayor parte de los casos, lo que se corresponde con la recomendación de Ware y Gandek (1998). Las medias obtenidas fueron muy similares a las del estudio original en castellano, con diferencias entre $-.5$ y $.5$ en todos los casos, salvo en los ítems 2, 22, 31 y 39. Las diferencias en cuanto a desviación típica también fueron mínimas, entre $-.2$ y $.2$ en todos los casos, salvo en los ítems 22 y 33.

En torno al 80% de los ítems presentó índices de homogeneidad superiores a $.30$. Los elementos con índices de homogeneidad inferiores al mencionado valor fueron los ítems 4, 5, 13, 18, 21, 28, 30, 35 y 38. Todos los ítems, con la excepción de los ítems 10 y 13, presentaron valores inferiores a los límites que indican problemas de normalidad, 2 para asimetría y 7 para curtosis (Curran et al., 1996).

Por otro lado, en la tabla 3.4 se muestra el porcentaje de atracción de cada alternativa de respuesta de los ítems de OPERAS en euskera. Para aclarar si el ítem era positivo o negativo, se indica en cada caso con un símbolo + o -. El ítem 1 era un ítem dummy, utilizado como entrenamiento para la versión online del test y solo se presentan sus resultados en la Tabla 3.5.

Tabla 3.4.

Estadísticos Descriptivos e Índices de Homogeneidad de los Ítems de OPERAS

Item	Media	Desviación Típica	Asimetría	Curtosis	Índice de homogeneidad
2.EX+	2.31	.93	0.29	-0.42	.50
3.ES+	4.04	.65	-0.76	2.22	.47
4.RE+	3.82	.75	-0.72	1.44	.19
5.DS-	4.29	.62	-0.60	1.60	.19
6.AM+	3.63	.67	-0.59	0.75	.40
7.AP-	2.00	.99	0.74	-0.17	.50
8.EX+	3.98	.73	-0.89	2.13	.50
9.ES-	1.88	.92	1.12	1.05	.56
10.RE-	1.54	.87	2.23	5.73	.37
11.DS+	2.62	1.40	0.33	-1.27	.38
12.AM+	4.54	.62	-1.73	6.21	.38
13.AP+	4.59	.71	-2.57	9.24	.24
14.EX	2.59	1.17	0.31	-0.86	.54
15.ES-	1.84	.85	1.08	1.27	.57
16.RE-	1.85	.87	1.18	1.56	.44
17.AM+	3.77	.76	-0.72	1.01	.39
18.AP-	2.43	1.10	0.47	-0.56	.24
19.DS+	3.50	1.16	-0.52	-0.67	.38
20.EX+	3.69	.94	-0.63	0.04	.48
21.ES+	2.50	1.03	0.34	-0.62	.20
22.RE-	1.97	.97	1.06	0.87	.40
23.AM-	2.50	.96	0.24	-0.59	.40
24.AP+	3.72	1.02	-0.58	-0.25	.51
25.EX-	3.46	.97	-0.34	-0.13	.48
26.DS+	2.25	1.11	0.70	-0.38	.45
27.ES-	1.93	.95	0.93	0.26	.40
28.RE+	3.71	.98	-0.76	0.31	.23
29.AM-	1.73	.85	1.26	1.68	.40
30.AP+	4.60	.65	-1.75	3.35	.27
31.EX	3.54	1.07	-0.54	-0.17	.44
32.ES-	2.15	.96	0.76	0.11	.47
33.RE-	2.03	.92	0.81	0.33	.43
34.AM+	3.94	.79	-0.84	1.39	.45
35.AP+	4.06	.88	-1.24	1.97	.22
36.EX+	3.48	.78	-0.53	0.40	.41
37.ES-	1.51	.79	1.89	4.47	.43
38.RE+	3.93	.78	-1.27	3.17	.28
39.AM-	2.06	.93	0.71	0.19	.40
40.AP-	2.08	1.12	0.84	-0.07	.38

Nota. AM = Amabilidad; AP = Apertura a la experiencia; ES = Estabilidad emocional; EX = Extroversión; RE = Responsabilidad; DS = Deseabilidad Social

Tabla 3.5.

Porcentajes de Elección de las Opciones de Respuesta en cada Ítem de OPERAS

		1	2	3	4	5
1.Dummy	Gauza berriak probatzea gustatzen zait.	.44	2.77	8.60	57.00	29.45
2.EX+	Festako erregea/erregina izaten naiz.	20.41	36.88	31.63	8.31	1.02
3.ES+	Eroso sentitzen naiz nire buruarekin.	.29	2.19	10.79	65.60	19.39
4.RE+	Beti nago prest erantzukizunak nire gain hartzeko.	.29	4.37	21.72	57.87	13.85
5.DS-	Beti betetzen dut emandako hitza.	.29	.73	4.52	57.43	35.28
6.AM+	Ondo hitz egiten dut besteez.	.44	4.23	32.07	56.41	5.10
7.AP-	Artea aspergarria iruditzen zait.	1.17	7.00	19.83	32.80	37.46
8.EX+	Ondo moldatzen naiz gizarte-harremanetan.	.15	3.79	13.41	61.08	19.68
9.ES-	Maiz egoten naiz lur jota.	1.31	6.12	10.50	42.42	37.90
10.RE-	Ihes egiten diet nire betebeharrei.	2.92	1.75	2.62	32.36	58.60
11.DS+	Noizbait, nirea ez zen gauzaren bat hartu izan dut.	11.95	20.85	11.66	25.36	28.43
12.AM+	Besteak errespetatzen ditut.	.73	.44	1.02	39.21	56.85
13.AP+	Uste dut garrantzitsua dela gure burua kulturaz janzea.	1.60	.29	2.33	28.43	65.60
14.EX-	Hitz gutxikoa naiz.	5.54	19.10	22.16	32.07	19.39
15.ES-	Maiz sentitzen naiz triste.	.87	4.23	10.79	44.61	37.76
16.RE-	Gauzak erdizka uzten ditut.	1.17	5.10	8.45	47.23	36.30
17.AM+	Besteek asmo onez jokatzten dutela pentsatzen dut.	.73	5.10	22.45	57.87	12.10
18.AP-	Ihes egiten diet eztabaida filosofikoei.	3.79	14.29	23.32	35.42	21.43
19.DS+	Noizbait, aritu izan naiz gaizki-esaka norbaiti buruz.	19.68	38.63	17.49	16.76	5.69
20.EX+	Erraz egiten ditut lagunak.	1.75	10.35	21.43	47.67	17.06
21.ES+	Zaila da nik buruhausterik edukitzea.	15.74	38.78	24.20	17.06	2.33
22.RE-	Nire gauzak hankaz gora uzten ditut.	2.19	6.41	12.39	42.86	34.40
23.AM-	Oso kritikoa naiz besteekin.	1.17	15.31	29.45	38.48	13.85
24.AP+	Museoak bisitatzea gustatzen zait.	2.33	10.93	22.16	39.65	23.18
25.EX-	Nahiago dut jendearen arretea besteengan egotea.	13.12	36.73	32.65	13.56	1.90
26.DS+	Noizbait besteez aprobe txatu izan naiz.	3.21	13.99	14.58	39.36	27.11
27.ES-	Izuak (panikoak) erraz menderatzen nau.	.87	7.87	12.54	39.36	37.61
28.RE+	Perfekzionista naiz.	2.19	10.50	19.24	47.52	18.66
29.AM-	Askotan izaten naiz desatsegina besteekin.	1.02	3.50	9.48	38.63	45.63
30.AP+	Leku berriak ezagutzea gustatzen zait.	.15	1.17	4.81	25.36	66.76
31.EX	Erosoago egoten naiz nire burua nabarmendu gabe.	18.37	37.46	26.09	12.10	3.94
32.ES-	Maiz izaten ditut umore-aldaketak.	1.31	10.06	15.45	46.65	24.78
33.RE-	Denbora alferrik galtzen dut.	1.02	7.00	15.89	44.46	29.88
34.AM+	Pertsonak diren bezalakoak onartzen ditut.	.44	4.96	15.45	56.41	20.85
35.AP+	Nire inguruak jakin-mina pizten dit.	2.04	4.37	9.91	51.17	30.76
36.EX+	Badakit jendea neureganatzen.	1.31	8.16	36.01	47.08	5.69
37.ES-	Desatsegina egiten zait nire burua.	1.31	1.60	5.39	30.17	59.62
38.RE+	Planak egiten ditudanean, bete egiten ditut.	1.31	3.94	12.10	63.12	17.64
39.AM-	Norbait nirekin gaizki portatzen denean, beste horrenbeste egiten diot nik.	1.17	6.41	19.39	41.98	29.15
40.AP-	Antzerkia gutxi interesatzen zait.	3.21	10.50	14.43	34.26	35.28

Nota. AM = Amabilidad; AP = Apertura a la experiencia; ES = Estabilidad emocional; EX = Extroversión; RE = Responsabilidad; DS = Deseabilidad Social

En la Tabla 3.6 se presentan los parámetros estimados desde el enfoque de la TRI junto a sus errores típicos para los ítems que componen cada escala. La segunda columna de la tabla contiene el parámetro a (discriminación). Las columnas 5, 7, 9 y 11 contienen los parámetros umbrales (parámetro b) referidos al nivel mínimo para pasar de una respuesta de 1 a 2, de 2 a 3, de 3 a 4 y de 4 a 5, respectivamente. Las tres últimas columnas de la tabla informan del valor de chi-cuadrado, de sus grados de libertad y de su probabilidad asociada (p), que hacen referencia a la prueba de ajuste estadístico del ítem al MRG (Orlando y Thissen, 2000, 2003). Los resultados muestran que la gran mayoría de ítems superó el umbral de .65 para la discriminación (parámetro a). Las excepciones fueron 7, los ítems 18, 35, 21, 36, 4, 28 y 5. No obstante, todos estos ítems se ajustaron al modelo MRG. Los únicos ítems que presentaron un mal ajuste al mencionado modelo, $p < .01$, fueron 4, los ítems 13, 27, 10 y 38, que sin embargo presentaron índices adecuados respecto a la discriminación.

En las Figuras 3.1 a 3.6 se muestran las CCI de todos los ítems de OPERAS. En general, se observa un buen uso de todas las alternativas de respuesta y dado el apuntamiento de las curvas, cabe entender que proporcionan un nivel de información adecuado. En cuanto a algunos de los ítems más problemáticos señalados en los párrafos anteriores (ítems 13, 35 y 38), destacaron por el apuntamiento de las curvas referidas a los niveles más altos de los factores medidos (paso 4, de las opciones 4 a 5 en la escala Likert), lo que supone un complemento interesante a las CCI del resto de ítems y constituye un elemento a favor de su utilidad para el test. Los ítems 13 y 38 presentaron valores de discriminación aceptables.

Tabla 3.6.

Parámetros Estimados según el Modelo de Respuesta Graduada de los ítems de OPERAS

Ítem	Factor	<i>a</i>	<i>e.t.</i>	<i>b1</i>	<i>e.t.</i>	<i>b2</i>	<i>e.t.</i>	<i>b3</i>	<i>e.t.</i>	<i>b4</i>	<i>e.t.</i>	<i>X</i> ²	<i>g.l.</i>	Prob.
6	Amabilidad	1.32	.14	-4.79	.59	-2.83	.25	-.53	.08	2.75	.25	3.39	35	.69
12	Amabilidad	1.36	.16	-4.29	.50	-3.9	.42	-3.37	.34	-.32	.07	28.32	21	.13
17	Amabilidad	1.09	.12	-5.03	.62	-2.99	.29	-1.03	.12	2.16	.21	31.27	38	.77
23	Amabilidad	1.09	.11	-4.6	.52	-1.81	.18	-.18	.08	1.99	.20	48.79	38	.11
29	Amabilidad	1.24	.13	-4.29	.47	-2.97	.28	-1.83	.17	.14	.08	27.96	37	.86
34	Amabilidad	1.52	.15	-4.3	.49	-2.45	.20	-1.21	.10	1.19	.11	44.23	33	.09
39	Amabilidad	1.16	.12	-4.37	.48	-2.57	.24	-1.06	.11	.93	.11	27.55	38	.89
7	Apertura	2.22	.22	-2.91	.23	-1.75	.11	-.74	.06	.36	.06	35.10	38	.60
13	Apertura	.94	.12	-4.8	.64	-4.62	.60	-3.72	.46	-.9	.13	58.20	32	.00
18	Apertura	.57	.09	-5.86	.91	-2.76	.42	-.61	.16	2.37	.37	68.65	49	.03
24	Apertura	2.29	.22	-2.51	.18	-1.4	.09	-.47	.06	.90	.08	52.50	40	.09
30	Apertura	.79	.12	-8.61	1.73	-5.81	.88	-3.74	.52	-1.07	.17	3.72	31	.48
35	Apertura	.62	.09	-6.55	1.04	-4.61	.69	-2.86	.43	1.35	.23	7.72	47	.01
40	Apertura	1.25	.13	-3.08	.27	-1.76	.14	-.92	.09	.54	.09	58.95	46	.09
3	Estabilidad	1.45	.14	-4.73	.61	-3.17	.28	-1.71	.14	1.33	.12	49.18	36	.07
9	Estabilidad	2.30	.20	-2.85	.21	-1.84	.11	-1.16	.08	-.35	.06	5.61	37	.01
15	Estabilidad	2.51	.23	-2.97	.23	-2.03	.12	-1.21	.08	-.35	.06	59.37	36	.01
21	Estabilidad	.49	.08	-3.35	.61	-.47	.18	3.00	.53	7.86	1.40	66.91	52	.01
27	Estabilidad	1.26	.12	-4.34	.46	-2.34	.19	-1.37	.12	.45	.09	87.72	44	.00
32	Estabilidad	1.38	.12	-3.76	.35	-1.93	.15	-.97	.09	1.05	.11	48.38	42	.23
37	Estabilidad	1.52	.15	-3.53	.33	-2.96	.25	-2.11	.17	-.42	.07	57.60	41	.04
2	Extroversión	1.28	.12	-1.36	.13	.32	.08	2.20	.19	4.21	.43	39.49	46	.74
8	Extroversión	1.87	.18	-4.42	.61	-2.42	.18	-1.25	.10	1.14	.09	55.11	34	.01
14	Extroversión	1.69	.14	-2.31	.17	-.94	.08	-.07	.07	1.22	.10	57.47	47	.14
20	Extroversión	1.71	.15	-3.1	.26	-1.64	.12	-.58	.07	1.33	.11	54.27	45	.16
25	Extroversión	1.10	.11	-2.05	.19	.04	.09	1.83	.18	4.08	.42	57.50	51	.24
31	Extroversión	1.01	.11	-1.73	.18	.34	.09	1.94	.20	3.61	.37	66.05	56	.17
36	Extroversión	.12	.12	-3.92	.38	-2.16	.18	-.15	.08	2.65	.22	56.73	47	.16
4	Responsabilidad	.63	.10	-9.56	1.82	-5.05	.78	-1.75	.28	3.07	.47	48.94	37	.09
10	Responsabilidad	1.64	.18	-2.79	.25	-2.47	.21	-2.13	.17	-.38	.07	66.11	37	.00
16	Responsabilidad	1.92	.19	-3.11	.28	-2.08	.15	-1.43	.10	.41	.07	56.54	36	.02
22	Responsabilidad	1.44	.14	-3.22	.29	-2.13	.17	-1.28	.11	.55	.08	44.29	40	.29
28	Responsabilidad	.55	.09	-7.17	1.24	-3.68	.61	-1.43	.25	2.81	.47	57.81	48	.16
33	Responsabilidad	1.61	.15	-3.54	.34	-2.07	.16	-1.04	.09	.73	.08	47.01	36	.10
38	Responsabilidad	.99	.12	-4.81	.58	-3.32	.36	-1.84	.19	1.80	.20	79.37	40	.00
5	Deseabilidad	.64	.11	-9.45	1.93	-7.46	1.36	-4.7	.79	.99	.47	38.77	25	.04
11	Deseabilidad	1.15	.15	-2.11	.23	-.76	.11	-.21	.09	.98	.07	41.74	33	.14
19	Deseabilidad	1.19	.15	-1.46	.16	.39	.09	1.27	.14	2.82	.07	53.02	30	.01
26	Deseabilidad	1.91	.31	-2.55	.24	-1.23	.12	-.59	.08	.80	.08	48.32	28	.01

Nota. *a* = discriminación; *b1*, *b2*, etc. = parámetros umbrales; *e.t.*= error típico; *g.l.*= grados de libertad

Figura 3.1.

Curvas Características del Ítem para los Ítems que Miden Amabilidad

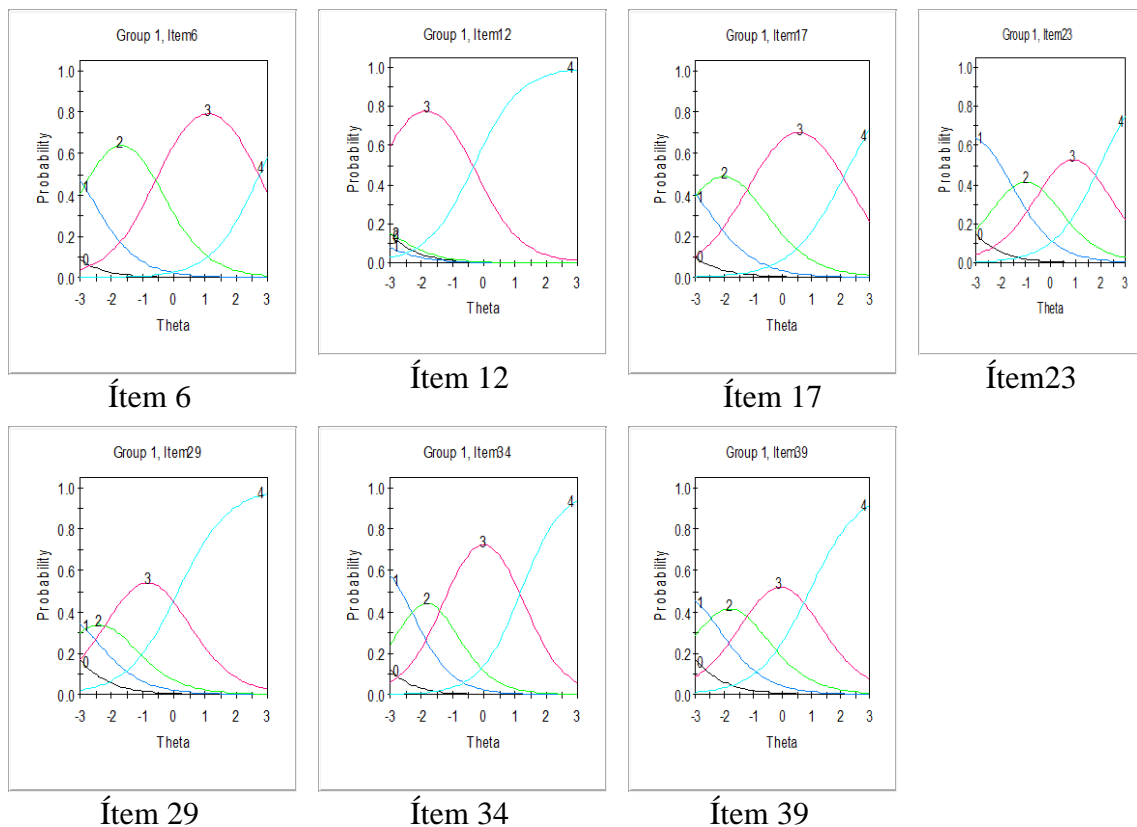


Figura 3.2.

Curvas Características del Ítem para los Ítems que Miden Apertura a la experiencia

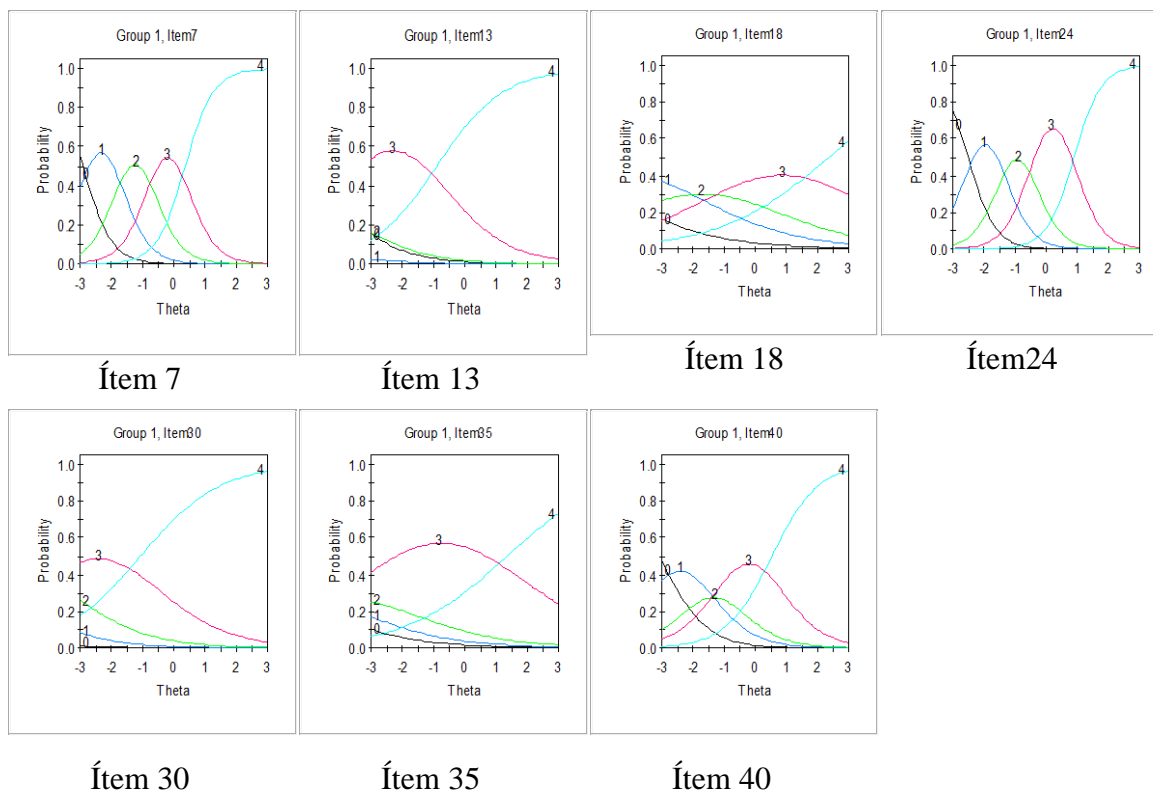


Figura 3.3.

Curvas Características del Ítem para los Ítems que Miden Estabilidad Emocional

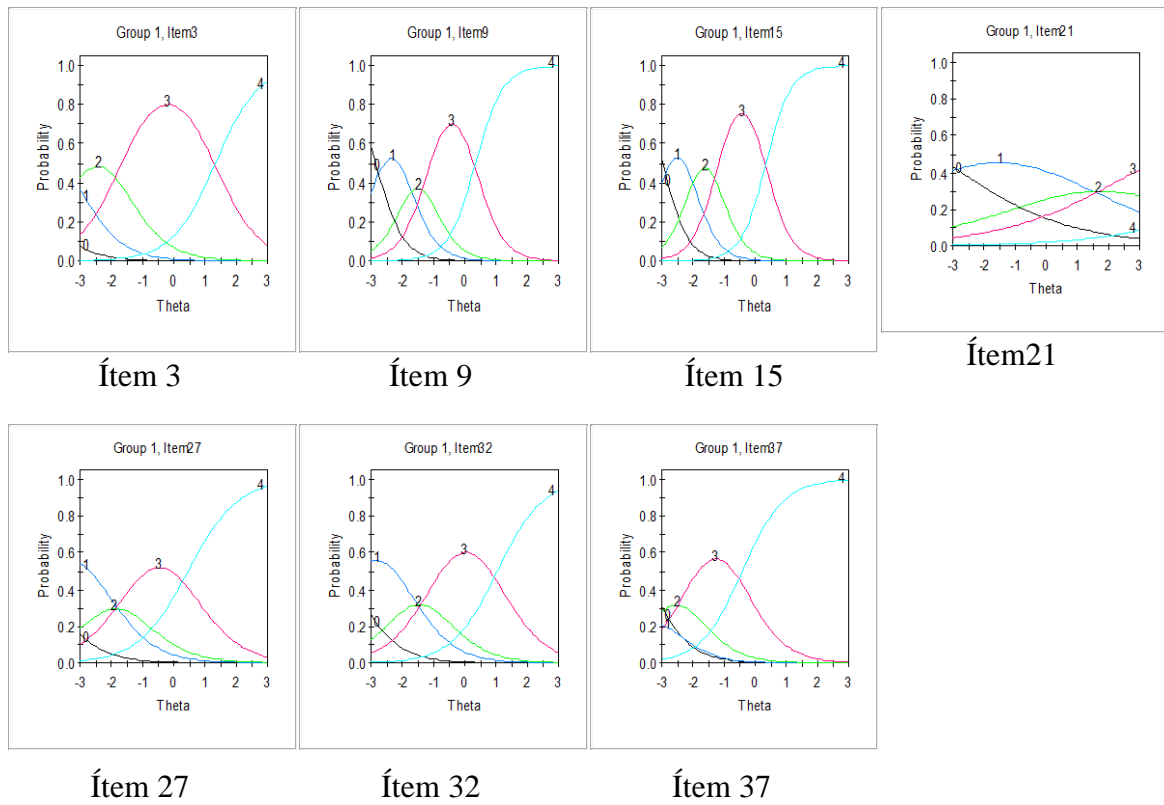


Figura 3.4.

Curvas Características del Ítem para los Ítems que Miden Extroversión

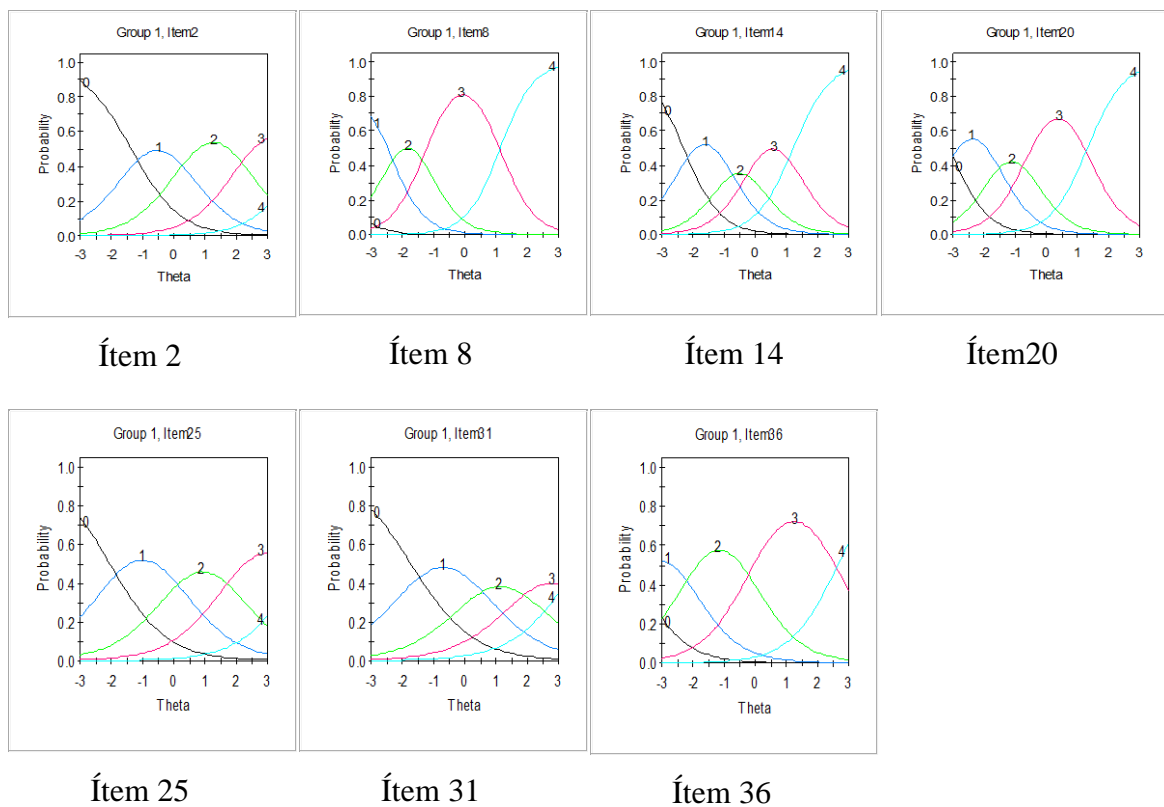


Figura 3.5.

Curvas Características del Ítem para los Ítems que Miden Responsabilidad

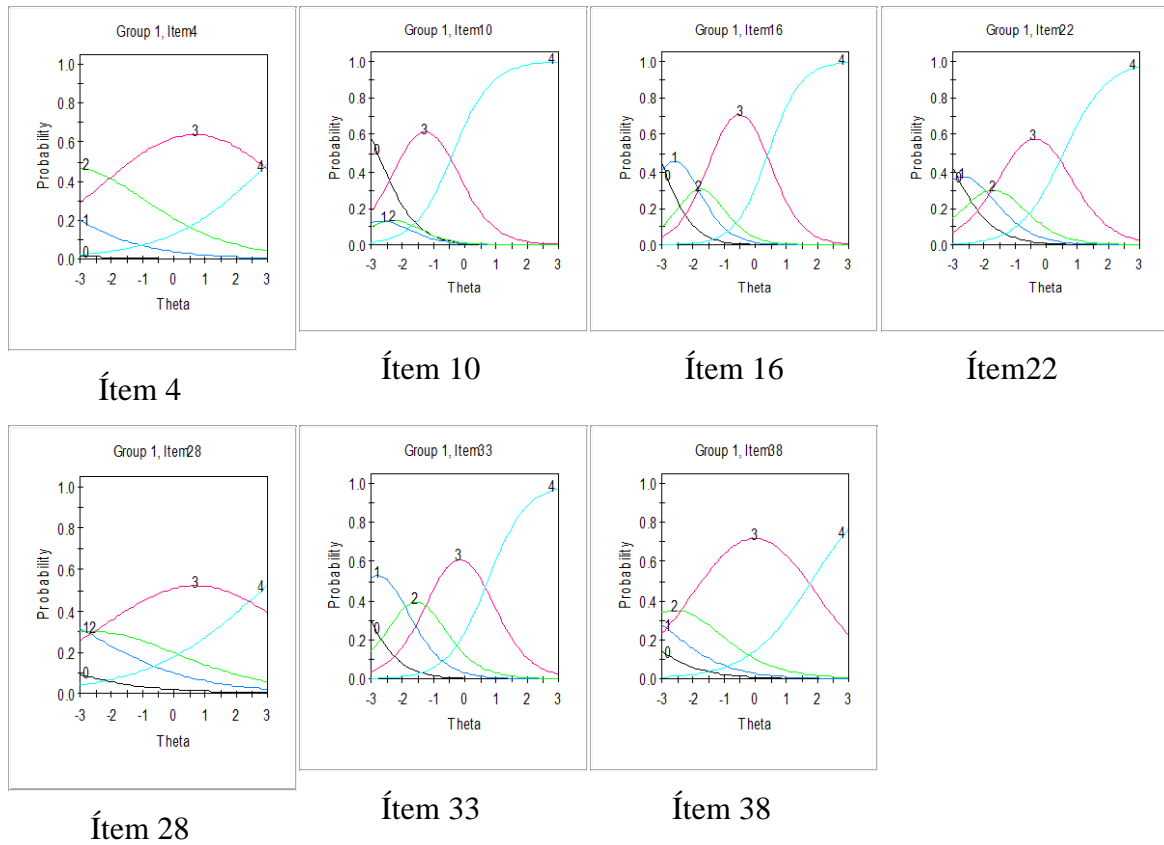
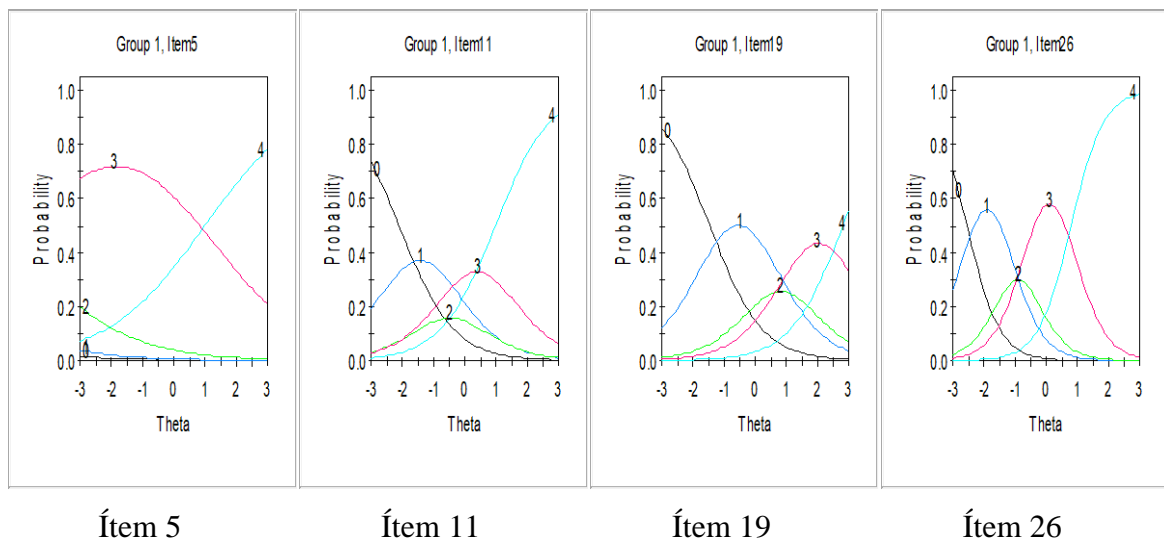


Figura 3.6.

Curvas Características del Ítem para los Ítems que Miden Deseabilidad Social



3.3.3.2. Dimensionalidad del instrumento y validez de constructo: AFE, AFC, ESEM

Antes de realizar el análisis factorial exploratorio controlando deseabilidad social y aquiescencia, 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 igual a .80 (bueno) y el test de Bartlett resultó estadísticamente significativo, $\chi^2(780) = 5500.5$; $p < .01$, lo que llevó a concluir que el análisis factorial resultaba pertinente a priori.

El análisis mediante Análisis Paralelo reflejó una estructura de 5 factores de contenido que explicaban el 49% de la varianza total: el primer factor, Estabilidad emocional, con sus 7 ítems, explicó el 22.04% de la varianza. El segundo, Amabilidad, explicó el 9.43% de la varianza y cuenta con 7 ítems. El tercero, Responsabilidad, explicó el 7.18% de la varianza integrando 7 ítems. El cuarto, Apertura a la experiencia, con 7 ítems, explicó un 5.39% de la varianza. Finalmente, el quinto, Extroversión, con 7 ítems, explicó un 5.18% de la varianza. En cuanto a los factores vinculados a estilos de respuesta, la Deseabilidad social, con sus 4 ítems empleados a modo de marcador, supuso un 4.18% de la varianza, mientras que la escala de control Aquiescencia explicó un 4.13% de la varianza.

En la tabla 3.7 se presentan las saturaciones factoriales rotadas de cada ítem en los factores del cuestionario corregidos por las escalas de control, Deseabilidad Social (SD) y Aquiescencia (AC). Se han oscurecido las casillas que reflejan las saturaciones de cada ítem en el factor al que corresponden teóricamente (cargas factoriales primarias), para facilitar la lectura. Al igual que en tablas anteriores, para aclarar si el ítem es positivo o negativo, se indica en cada caso con un símbolo + o -. El rango de las cargas factoriales primarias osciló entre .18 y .78, con una mediana de .48, mientras que las cargas secundarias (las saturaciones de los ítems en los factores con los que no están vinculados teóricamente) oscilaron entre .00 y .30, con una mediana de .08. Considerando el intervalo entre .25 y .30 como punto de corte de los elementos de la matriz factorial, todos los ítems saturaron de forma clara en sus correspondientes dimensiones a excepción de los ítems 4 y 35, con saturaciones inferiores a .25. Asimismo, los ítems 13, 21 y 38 mostraron saturaciones por debajo de .30 (Floyd y Widaman, 1995; DiStefano et al., 2009; Swisher et al., 2004).

Tabla 3.7.

Saturaciones de los Ítems Corregidas por DS y AC en los Factores de OPERAS

Ítem	DS	AC	Amabilidad	Apertura	Estabilidad	Extroversión	Responsabilidad
5.DS-	-.27	.00	.00	.00	.00	.00	.00
11.DS+	.52	.00	.00	.00	.00	.00	.00
19.DS+	.52	.00	.00	.00	.00	.00	.00
26.DS+	.67	.00	.00	.00	.00	.00	.00
6.AM+	-.33	.22	.40	.02	.10	.01	-.01
12.AM+	-.25	.13	.42	.11	.10	.00	.19
17.AM+	-.09	.20	.50	-.01	.09	.09	-.09
23.AM-	.16	.09	-.53	.00	-.19	-.02	.22
29.AM-	.20	.06	-.50	-.01	-.15	-.15	-.27
34.AM+	-.21	.20	.42	.07	.25	.02	.02
39.AM-	.15	.08	-.51	-.07	-.01	-.09	-.17
7.AP-	.14	.07	.03	-.71	-.08	-.03	.04
13.AP+	.00	.06	.05	.28	.09	-.02	.15
18.AP-	.07	.10	-.03	-.31	.01	-.09	-.10
24.AP+	-.14	.06	-.04	.78	.09	.06	-.10
30.AP+	-.10	.23	.21	.31	.06	.14	.18
35.AP+	-.08	.25	.06	.22	.08	.03	.23
40.AP-	.25	.06	-.15	-.48	.12	-.12	.00
3.ES+	-.17	.13	.09	.04	.54	.11	.04
9.ES-	.08	.09	-.08	.00	-.71	-.09	-.01
15.ES-	.06	.02	-.10	.01	-.71	-.14	.05
21.ES+	-.16	.03	-.08	-.01	.29	.14	-.29
27.ES-	.14	.09	-.17	-.02	-.39	-.19	-.05
32.ES-	.19	.20	-.30	-.11	-.47	-.08	-.06
37.ES-	.10	.09	-.19	-.10	-.49	-.12	-.18
2.EX+	.05	.04	-.13	.10	-.03	.60	-.07
8.EX+	-.11	.15	.25	.05	.08	.61	.08
14.EX-	.02	.09	.05	-.06	.03	-.65	-.05
20.EX+	-.12	.17	.25	.08	.01	.61	.15
25.EX-	.06	.24	.15	-.03	.02	-.59	.21
31.EX-	.11	.24	.14	-.05	.00	-.54	.27
36.EX+	-.04	.28	.07	.07	.11	.56	.27
4.RE+	-.08	.19	.05	.17	.16	.13	.18
10.RE-	.17	.06	-.17	-.05	-.25	.01	-.35
16.RE-	.15	.13	-.11	.07	-.28	.00	-.45
22.RE-	.12	.17	-.06	-.04	-.30	-.01	-.37
28.RE+	-.13	.15	-.11	.11	-.01	-.07	.41
33.RE-	.25	.10	-.10	.00	-.29	-.13	-.41
38.RE+	-.14	.29	.06	.00	.22	.16	.27
%	4.18	4.13	9.43	5.49	22.04	5.18	7.18
varianza							

Nota. AM = Amabilidad; AP = Apertura a la experiencia; ES = Estabilidad emocional; EX = Extroversión; RE = Responsabilidad; DS = Deseabilidad Social; AC = Aquiescencia

En la tabla 3.7 se reflejan las saturaciones de cada ítem en cada escala de control, de modo que se puede comprobar en qué medida están afectados los factores e ítems de personalidad por la Aquiescencia y la Deseabilidad social. Los ítems con saturaciones más altas en Deseabilidad Social fueron el 6 (Amabilidad), con una saturación de -.33, y los ítems 12 (Amabilidad), 33 (Responsabilidad) y 40 (Apertura a la experiencia) con saturaciones de -.25 o .25. Amabilidad, Estabilidad y Responsabilidad presentaron unas medianas de saturaciones en Deseabilidad Social (.20, .14 y .14, respectivamente) superiores a las de Apertura a la experiencia y Extroversión (.10 y .06, respectivamente).

En lo que se refiere a la Aquiescencia, los ítems con saturaciones más elevadas fueron el 36 (Extroversión), el 35 (Apertura a la experiencia) y el 38 (Responsabilidad). Los ítems 35 y 38, junto con el ítem 4 (Responsabilidad), presentaron saturaciones más elevadas en Aquiescencia que en sus factores teóricos. Los factores de personalidad con medianas más elevadas en esta escala de control fueron Extroversión y Responsabilidad, con magnitudes de .17 y .15 respectivamente.

Además del análisis factorial que acabamos de presentar, se llevó a cabo un AFC para estudiar el ajuste de los datos empíricos al modelo teórico de partida (Cinco grandes factores de personalidad) empleando el programa MPlus. El AFC presentó los siguientes datos de ajuste al modelo: $\chi^2(726) = 2281.99$, $p = .01$; RMSEA = .06; CFI = .71; TLI = .68. Se comprueba un ajuste moderado de los datos a la teoría (ajuste aceptable en el estadístico RMSEA; no así para TLI o CFI).

Con respecto al análisis ESEM, y dado que existía un problema de normalidad en algunos ítems, se utilizó el estimador WLSMV, con los siguientes resultados: $\chi^2(522) = 1131.08$, $p = .01$; RMSEA = .04; CFI = .93; TLI = .91. Se comprueba un ajuste aceptable de los datos al modelo (Hu y Bentler, 1999). Asimismo, las saturaciones de los ítems en sus respectivos factores fueron superiores a .30, salvo en el caso del ítem 4.

En lo que se refiere a la relación entre los factores de OPERAS, la matriz de correlaciones muestra los resultados que se presentan en la Tabla 3.8, con correlaciones de tamaño entre reducido y medio entre la mayoría de factores ($Mdn = .23$). Las correlaciones de los 5 factores presentaron tamaños similares siendo las más altas las que presentan Estabilidad con Amabilidad, .40, y Responsabilidad con Amabilidad y Estabilidad, .35 y .37 respectivamente. Deseabilidad social, por su parte, presentó

correlaciones de tamaño moderado con los cinco factores de personalidad ($Mdn = .19$), siendo las más elevadas las que presentó con Amabilidad, .28, y Responsabilidad, .26. El resto de correlaciones fueron más bajas, y la correlación de Deseabilidad Social con Extroversión no fue estadísticamente significativa.

Tabla 3.8.

Correlaciones entre los Factores de OPERAS

	Amabilidad	Apertura a la experiencia	Estabilidad	Extroversión	Responsabilidad	Deseabilidad social
Amabilidad	1					
Apertura a la experiencia	.24**	1				
Estabilidad	.40**	.18**	1			
Extroversión	.14**	.20**	.22**	1		
Responsabilidad	.35**	.21**	.37**	.11**	1	
Deseabilidad social	.28**	.17**	.19**	.07	.26**	1

Por otro lado, la gran mayoría de las correlaciones entre los ítems y su factor correspondiente se encontraron entre .15 y .50, lo que se corresponde con las recomendaciones de Briggs y Cheek (1986), que indicaron que la media de las correlaciones entre los ítems de cada factor es un estadístico de utilidad para evaluar su unidimensionalidad. Como se puede comprobar en la Tabla 3.9, la media de las correlaciones entre los ítems de cada factor se encontró en el rango recomendado por Briggs y Cheek (1986) en todos los casos.

Tabla 3.9.

Correlación entre los Ítems de OPERAS y su Factor Correspondiente

Dimensión	Mínimo	Máximo	Media
Amabilidad	.14	.30	.25
Apertura a la experiencia	.03	.55	.19
Estabilidad	.07	.59	.28
Extroversión	.11	.55	.31
Responsabilidad	.06	.36	.19
Deseabilidad social	.10	.34	.24

3.3.3.3. Validez convergente: asociación entre las dimensiones de OPERAS y del BFI

Los coeficientes de correlación de Pearson entre las puntuaciones de las dimensiones de OPERAS y las del BFI se presentan en la tabla 3.10. Como cabía esperar, las puntuaciones en Neuroticismo de BFI mostraron una asociación inversa, estadísticamente significativa y de gran magnitud, con las puntuaciones en Estabilidad Emocional de OPERAS. A su vez, la asociación entre las puntuaciones de Responsabilidad de OPERAS y de Conciencia del BFI también fue alta, y de signo positivo. El resto de dimensiones mostraron correlaciones positivas de una intensidad alta con las dimensiones equivalentes. La mediana de las correlaciones fue de .68, que cabe considerar excelente (Muñiz et al., 2011).

Tabla 3.10.

Correlaciones entre las Dimensiones de OPERAS y BFI

Dimensiones OPERAS	Dimensiones BFI	r
Amabilidad	Amabilidad	.68
Apertura a la experiencia	Apertura Mental	.62
Estabilidad emocional	Neuroticismo	-.71
Extroversión	Extroversión	.77
Responsabilidad	Conciencia	.68

3.3.3.4. Fiabilidad

En la tabla 3.11 se presentan los índices de consistencia interna de los factores de OPERAS en euskera. Los resultados obtenidos en la adaptación al euskera fueron similares a los de la adaptación española y los coeficientes ORION de los factores oscilaron entre .70 y .81, lo que permite concluir que todos ellos presentan una adecuada consistencia interna (Nunnally, 1978). Las medianas de los índices ORION, Alfa de Cronbach, Omega y Fiabilidad Marginal para los Cinco Grandes fueron, respectivamente, .75, .70, .68 y .76, adecuadas según Muñiz et al. (2011).

Como complemento a los índices anteriores, en las Figuras 3.7 a 3.12 se aportan las funciones de información del test para cada factor, así como el error típico de medida. Se observa que todos los factores fueron más precisos en niveles bajos o medios en las variables medidas, por lo que cabe interpretar que estamos ante un instrumento de gran utilidad para la selección entendida como descarte de aspirantes no idóneos, ya que permite detectar de manera fiable a individuos con niveles bajos en las variables analizadas.

Tabla 3.11.

Fiabilidad (Consistencia Interna) de los factores de OPERAS

	ORION	Alfa de Cronbach	Omega	Fiabilidad marginal
Amabilidad	.70	.69	.68	.72
Apertura a la experiencia	.77	.62	.63	.76
Estabilidad	.75	.72	.73	.80
Extroversión	.81	.76	.77	.79
Responsabilidad	.70	.62	.60	.74
Deseabilidad social	.67	.59	.59	.65

Figura 3.7.

Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para el Factor Amabilidad

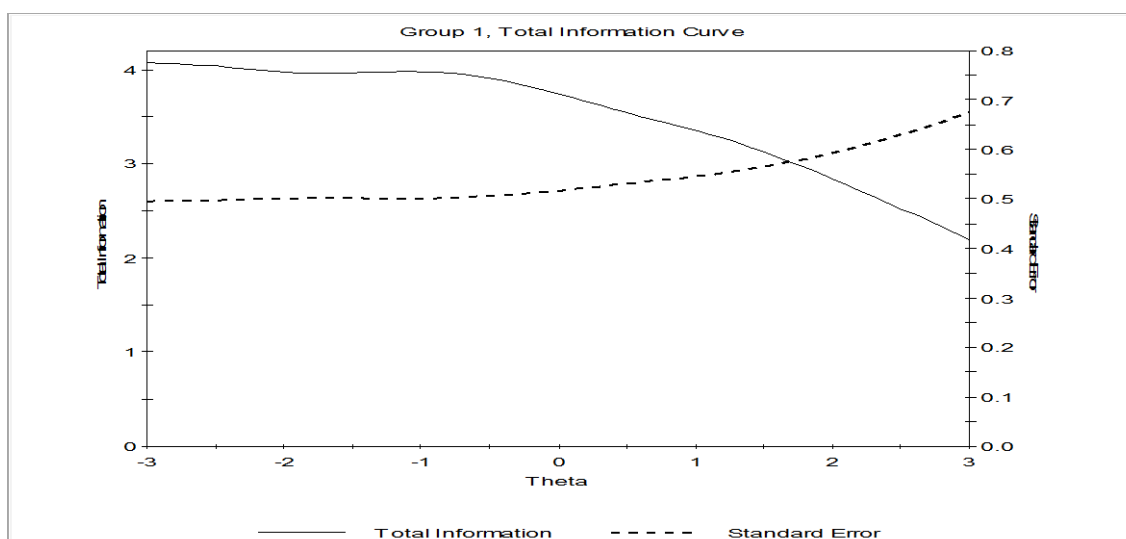


Figura 3.8.

Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para el Factor Apertura a la experiencia

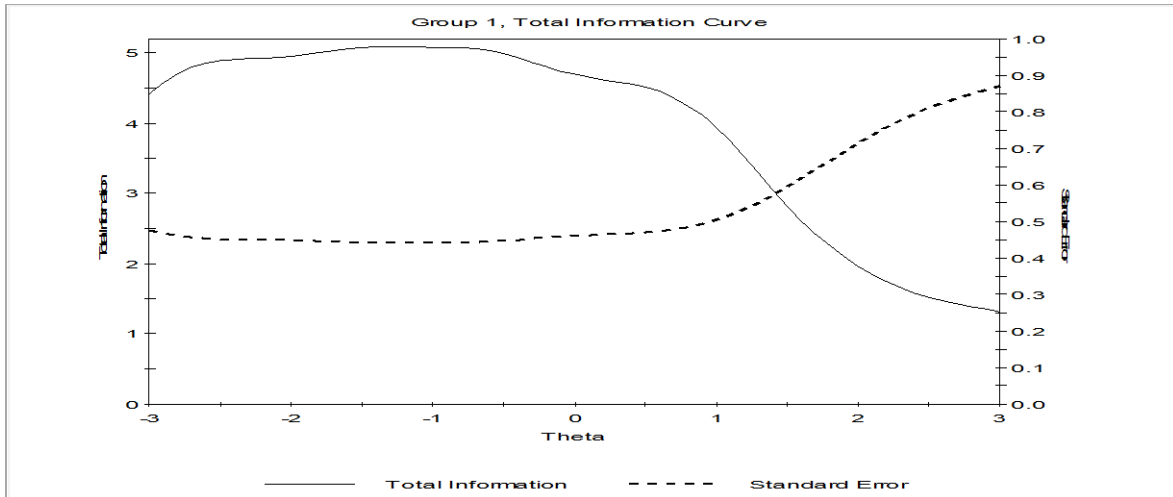


Figura 3.9.

Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para el Factor Estabilidad Emocional

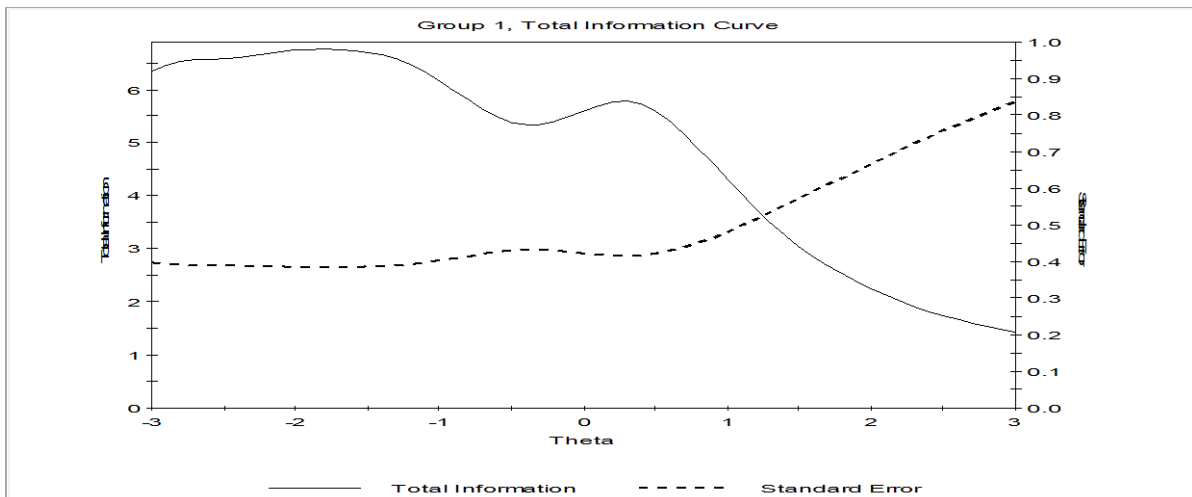


Figura 3.10.

Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para el Factor Extroversión

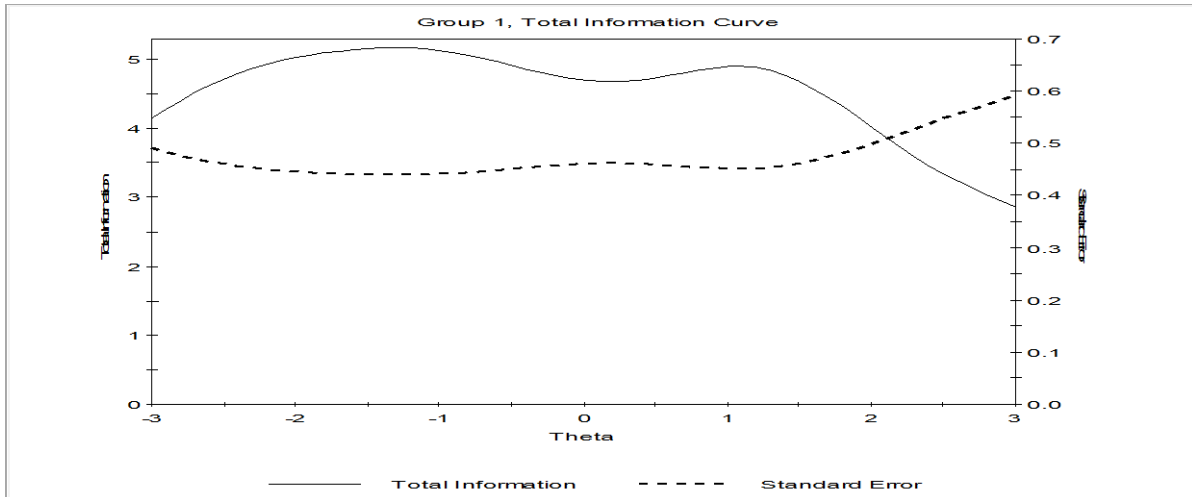


Figura 3.11.

Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para el Factor Responsabilidad

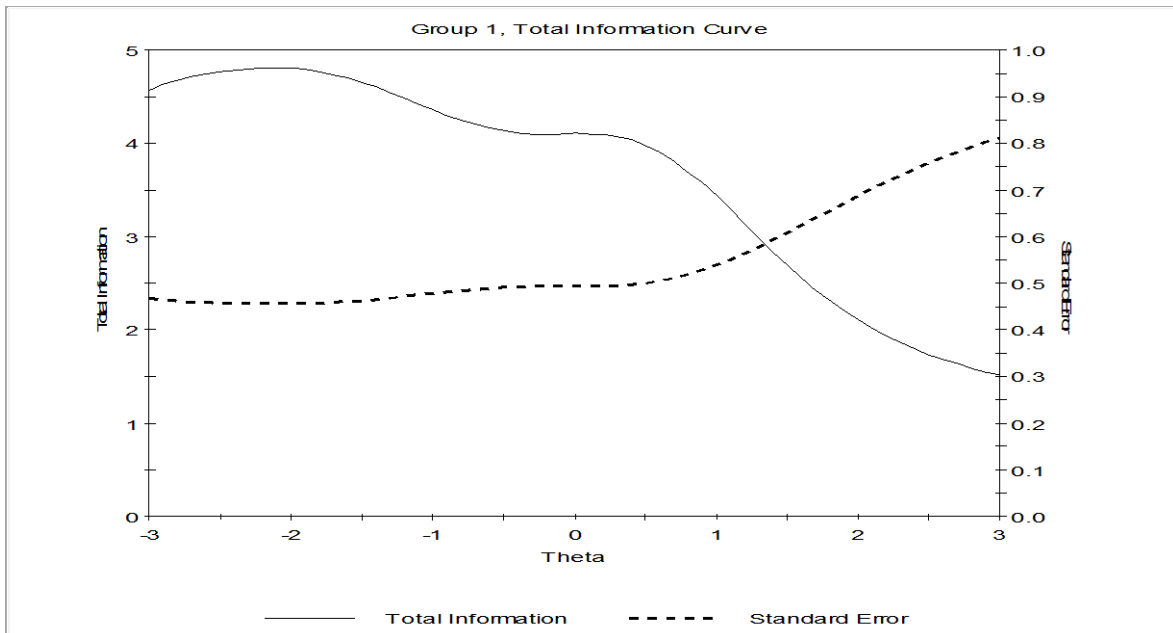
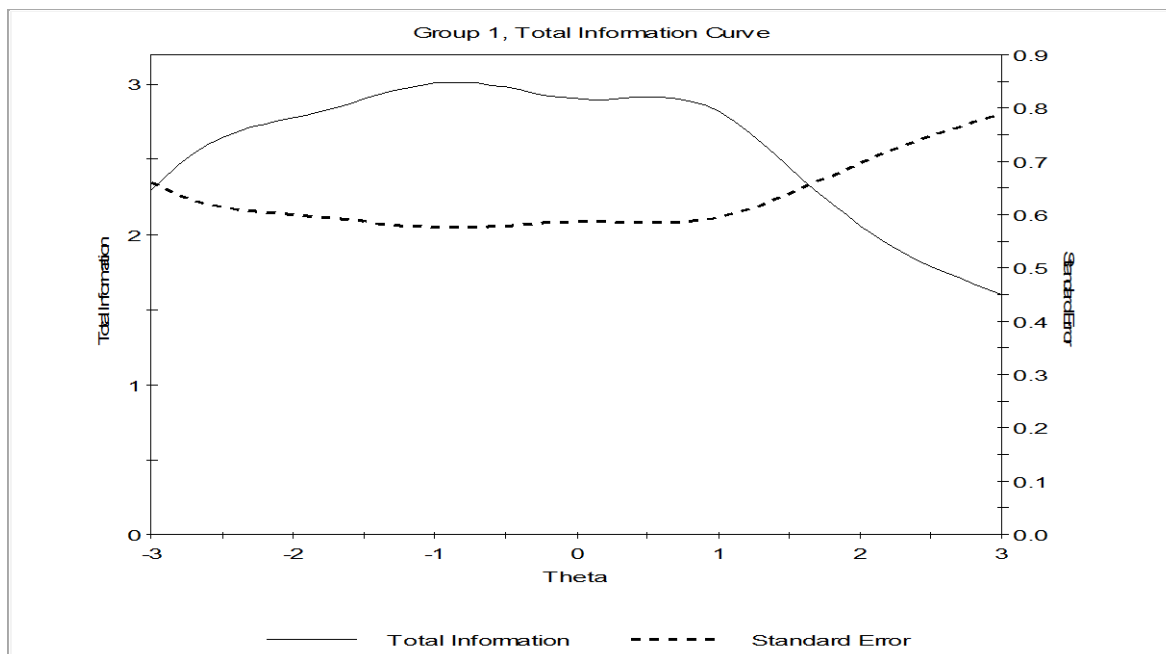


Figura 3.12.

Función de Información (línea continua) y Error Típico de Medida (línea discontinua) para el Factor Deseabilidad social



En la tabla 3.12 se presentan los índices de correlación observados entre las puntuaciones de las dos aplicaciones de OPERAS en euskera, recogidas en el intervalo de un mes. La mediana de los índices de fiabilidad test-retest de los Cinco Grandes fue de .69, y resultó adecuada (Muñiz et al., 2011).

Tabla 3.12.

Fiabilidad test-retest de los factores de OPERAS

	<i>r</i> (N= 151)
Amabilidad	.62
Apertura a la experiencia	.69
Estabilidad	.70
Extroversión	.85
Responsabilidad	.63
Deseabilidad social	.65

3.3.3.5. Funcionamiento diferencial de los ítems según la lengua materna de los y las participantes

En la Tabla 3.13 se pueden observar los resultados de los análisis realizados, y se comprueba que en base a la lengua materna de los y las participantes solamente se obtuvieron valores de ΔR^2 de Nagelkerke superiores a .035 en dos casos: Ítems 9 y 21. Ambos medían Estabilidad y presentaron FDI uniforme, con puntuaciones más altas para las personas con lengua materna vasca en el primer caso y con lengua materna castellana en el segundo. No obstante, según Jodoin y Gierl (2001), el valor del tamaño del efecto de R^2 que presentaron, inferior a .07, es propio de un FDI de tamaño moderado.

Tabla 3.13.

Funcionamiento Diferencial de los Ítems de OPERAS en base a la Lengua Materna de las Personas Participantes

Ítem	FDI Total ΔR^2 Nagelkerke	FDI Uniforme ΔR^2 Nagelkerke	FDI No uniforme ΔR^2 Nagelkerke
2.EX+	.00	.00	.00
3.ES+	.01	.01	.01
4.RE+	.00	.00	.00
5.DS-	.01	.00	.01
6.AM+	.01	.01	.00
7.AP-	.00	-.01	.00
8.EX+	.01	.01	.00
9.ES-	.05	.04	.00
10.RE-	.00	.00	.00
11.DS+	.00	.00	.00
12.AM+	.01	.01	.00
13.AP+	-.01	-.01	.00
14.EX	.00	.00	.00
15.ES-	.02	.01	.01
16.RE-	.01	.00	.00
17.AM+	.02	.00	.01
18.AP-	.00	.00	.00
19.DS+	.03	.03	.00
20.EX+	.00	.00	.00
21.ES+	.06	.06	.00
22.RE-	.00	-.01	.00
23.AM-	.01	.00	.01
24.AP+	.01	.01	.00
25.EX-	.02	.01	.01
26.DS+	.00	.00	.00
27.ES-	.00	.00	.00
28.RE+	.01	.00	.00
29.AM-	.01	.00	.01
30.AP+	.01	.01	.01
31.EX-	.02	.02	.00
32.ES-	.00	.00	.00
33.RE-	.01	.00	.01
34.AM+	.01	.01	.00
35.AP+	.03	.02	.00
36.EX+	.01	.00	.00
37.ES-	.00	.00	.00
38.RE+	.01	.00	.00
39.AM-	.00	.00	.00
40.AP-	.02	.02	.00

Nota. AM= Amabilidad; AP = Apertura a la experiencia; ES=Estabilidad emocional; EX=Extroversión; RE=Responsabilidad; DS=Deseabilidad Social

3.3.4. Conclusiones

El objetivo del presente capítulo era adaptar al euskera un instrumento para la medida de la personalidad y analizar sus propiedades psicométricas. A fin de realizar el estudio de validación del instrumento se utilizaron métodos cualitativos y cuantitativos. Un equipo multidisciplinar se encargó de la traducción de los ítems al euskera empleando un diseño de traducción inversa. Seguidamente se llevaron a cabo entrevistas cognitivas y una aplicación piloto. A continuación, en la fase experimental se emplearon análisis cuantitativos para examinar los ítems y la dimensionalidad del instrumento, así como su consistencia interna y su estabilidad temporal y para obtener evidencias de su validez de constructo y validez diferencial.

Considerando los datos acerca de los ítems de OPERAS en euskera, sus Curvas Características, y la revisión substantiva de los mismos, cabe concluir que la mayoría de ellos presentan características psicométricas aceptables. Los ítems más problemáticos son los ítems 4, 13, 35 y 38. Los ítems 4, 35 y 38 presentan índices de homogeneidad inferiores a .30, saturaciones inferiores a .30 en su factor teórico correspondiente y saturaciones más altas en Aquiescencia, probablemente por la relación entre este sesgo de respuesta y las dimensiones de personalidad Apertura a la experiencia, medida por el ítem 35, y Responsabilidad, medida por los ítems 4 y 38 (Hibbing et al., 2019). El ítem 13, por su parte, presenta un índice de homogeneidad inferior a .30, problemas de asimetría y curtosis, carga factorial inferior a .30 y problemas de ajuste al modelo MRG de la TRI. No obstante, los mencionados ítems muestran aspectos deseables desde la óptica de la TRI, como un nivel de discriminación aceptable (ítems 13 y 38) o un uso razonable de las opciones de respuesta 4 y 5, a excepción del ítem 4.

Asimismo, estos cuatro ítems miden aspectos centrales de los factores Responsabilidad (Ítem 4: *Beti nago prest erantzukizunak nire gain hartzeko / Siempre estoy dispuesto a asumir responsabilidades*; Ítem 38: *Planak egiten ditudanean, bete egiten ditut / Cuando hago planes los mantengo*) y Apertura a la experiencia, (Ítem 13: *Uste dut garrantzitsua dela gure burua kulturaz janztea/ Creo en la importancia de formarse culturalmente*; Ítem 35: *Nire inguruak jakin-mina pizten dit/ Siento curiosidad por el mundo que me rodea*). Por ello, cabe considerarlos de utilidad como elementos del cuestionario, ya que en cuestionarios breves es especialmente relevante asegurar una

cobertura conceptual adecuada del factor a evaluar, de manera que no se comprometa la validez (Torreblanca Murillo, 2017).

Los ítems de OPERAS en euskera presentan saturaciones bajas en las Escalas de Control Deseabilidad Social y Aquiescencia, con un patrón de magnitudes similar al de la versión original en castellano en lo relativo a la Deseabilidad: los ítems de los factores Amabilidad, Estabilidad y Responsabilidad presentaron unas medianas de saturaciones superiores a las de los ítems de Apertura a la experiencia y Extroversión. Este patrón de resultados es similar al encontrado por Calanna et al. (2020) empleando un algoritmo de detección del falseamiento de las respuestas.

En cuanto a la Aquiescencia, ejerce un efecto menor en la versión en euskera, probablemente porque la mayoría de los participantes presentan un nivel educativo alto, lo que correlaciona negativamente con este sesgo de respuesta (Van Vaerenbergh y Thomas, 2013)

El análisis de la dimensionalidad (que se ha basado en un AFE en el que se han corregido las puntuaciones por Deseabilidad y Aquiescencia, además de incluir análisis AFC y ESEM) apoya la estructura interna propuesta de 5 dimensiones: Amabilidad, Apertura a la experiencia, Estabilidad Emocional, Extroversión y Responsabilidad. La varianza explicada por los 5 factores de OPERAS en euskera (49%) es similar a la explicada por la versión original en castellano, 52.88%. Las saturaciones de los ítems en los factores teóricos presentan magnitudes adecuadas ($Mdn = .48$). La magnitud de las correlaciones entre los 5 factores ($Mdn = .23$) es similar a la que cabe esperar según los meta-análisis de Mount et al. (2005) y de Park et al. (2020). Asimismo, la magnitud de las correlaciones entre los factores de personalidad y la Deseabilidad social ($Mdn = .19$) es similar a la que cabe esperar según datos meta-analíticos (Ones et al., 1996).

Por otro lado, la media de las correlaciones entre los ítems de cada factor se corresponde con las recomendaciones de Briggs y Cheek (1986), al haberse observado valores entre .19 y .31. Estos valores son un argumento a favor de la homogeneidad y unidimensionalidad de las dimensiones de OPERAS en euskera.

En cuanto a las evidencias de validez convergente, la mediana de las correlaciones entre las puntuaciones de las dimensiones de OPERAS y las del BFI, que miden las mismas dimensiones, es de .68. Esta magnitud es muy similar a la observada en la validación original en castellano, .69, y puede ser calificada como excelente.

En consecuencia, la evaluación de la validez de constructo y la validez convergente de OPERAS en euskera permite concluir que este instrumento es válido para la toma de decisiones con respecto a la personalidad de las personas evaluadas. Por lo que respecta a la validez diferencial (análisis FDI), únicamente dos ítems (9 y 21) presentan funcionamiento diferencial en base a la lengua materna de los y las participantes, aunque en ambos casos el tamaño del efecto debe considerarse moderado. Asimismo, esos dos ítems suponen un 5% del total de ítems del test, porcentaje muy inferior al que se suele observar en tests adaptados, que tiende a ser ligeramente superior al 20% (Gierl et al., 2004; Hidalgo-Montesinos et al., 2015)

En lo que se refiere a la fiabilidad del instrumento, éste muestra índices aceptables de consistencia interna (tanto desde la óptica de la Teoría Clásica del Test como de la TRI) y de estabilidad temporal. La mediana de los estadísticos de fiabilidad ORION en los 5 factores de personalidad en la versión en euskera de OPERAS, .75, es ligeramente inferior a la de la versión original en castellano, .81. Lo mismo sucede al comparar la fiabilidad test-retest, en la que la mediana de la versión adaptada al euskera, .69, presenta un valor menor que la de la versión original, .73. La diferencia puede estar relacionada con el tamaño de la muestra (678 personas en la adaptación al euskera frente a 3838 en la validación original en castellano), la naturaleza de la muestra (en la versión en euskera estaba compuesta exclusivamente por empleados/as públicos/as mientras que en la versión original se extendió a usuarios/as de Facebook, el ámbito académico y compañías de selección de personal, además de emplearse submuestras de adolescentes y personas jubiladas) y sus características socioculturales. Las funciones de información de los factores de personalidad muestran un mayor nivel de precisión en niveles medios y bajos de las variables medidas, de lo que se deduce que OPERAS en euskera puede ser especialmente útil para descartar candidatos/as con niveles bajos en los cinco grandes factores de personalidad.

Como limitaciones de esta investigación se ha de señalar, al igual que en el capítulo 2, la homogeneidad de la muestra, por lo que sería deseable ampliarla a contextos diferentes al sector público y a otras franjas de edad. Por otro lado, la medición de la personalidad se ha realizado exclusivamente mediante auto-informes, algo que según Sackett y Lievens (2008) debería complementarse con otras metodologías (entrevistas, evaluaciones de informantes sobre la personalidad de terceras

personas). No obstante, la información aportada por los auto-informes es fiable y permite realizar inferencias válidas con respecto a un constructo como es la personalidad (Clark y Watson, 2019). Además, los trabajos de Fronczyk (2019) o Mõttus et al. (2020) apoyan la invarianza en la medida de la personalidad cuando se comparan los auto-informes con las evaluaciones de terceras personas.

En consecuencia y en base a los resultados obtenidos, cabe concluir que la versión de OPERAS en euskera presenta características psicométricas aceptables y, por lo tanto, evalúa de forma adecuada la personalidad. En cuanto a las contribuciones de esta investigación, cabe señalar que la adaptación al euskera de una prueba de personalidad es un avance en la normalización y la garantía de los derechos de las personas con lengua materna vasca ya que supone una alternativa para la evaluación en euskera de la personalidad. Asimismo, poder contar con un instrumento breve es de gran utilidad en procesos selectivos en los que se pretenda emplear un gran número de herramientas de evaluación en una misma sesión.

CAPÍTULO 4.

RELACIONES ENTRE LOS CINCO GRANDES FACTORES DE PERSONALIDAD, LOS CONSTRUCTOS VINCULADOS CON EL DESEMPEÑO LABORAL, EL AJUSTE PERSONA-ORGANIZACIÓN Y LAS VARIABLES SOCIODEMOGRÁFICAS Y LABORALES

4.1. Introducción

Si en los capítulos anteriores se ha abordado la adaptación y validación de la escala de Charbonnier-Voirin y Roussel (2012) al castellano y del cuestionario de personalidad OPERAS (Vigil-Colet et al., 2013) al euskera, en este capítulo se analiza la red nomológica de ambos instrumentos en nuestro contexto social. A tal fin se examina la relación que las dimensiones de ambos instrumentos presentan entre sí y con otros dos constructos, las conductas cívico organizacionales y el ajuste persona-organización. Asimismo, se profundiza en cuestiones aún no resueltas en la literatura que contribuyen a mejorar nuestra comprensión acerca de la relación entre personalidad y desempeño y, por tanto, a incrementar la eficacia de la selección de personal (Gorriti, 2018; Leutner y Chamorro-Premuzic, 2018).

Por otro lado, se analizan las relaciones de variables sociodemográficas y laborales, como la edad, el género, la complejidad de los puestos, el nivel educativo o el desempeño de puestos de mando con los constructos mencionados. Esta cuestión es de capital importancia en el ámbito de selección de personal, ya que está relacionada con las percepciones que tienen los candidatos y las candidatas en procesos selectivos acerca de los instrumentos que miden dichos constructos. En particular, con la percepción acerca de la equidad de los mismos y con los imperativos de la legislación que regula los procesos selectivos, que insta a garantizar la igualdad en la selección.

La información obtenida acerca de la relación de las variables sociodemográficas y laborales con constructos utilizados con frecuencia en la gestión de recursos humanos en las organizaciones, como personalidad, desempeño laboral o ajuste persona-organización es vital para examinar los costes legales y sociales y la aplicabilidad del uso de dichos constructos en el ámbito de la gestión de personas. Asimismo, esta información es de gran utilidad para la auto-evaluación que el personal pueda hacer acerca de su carrera laboral y sus posibilidades de mejorar su desempeño, así como para la toma de decisiones sobre el perfil de las personas que se deben

contratar, promocionar o desarrollar, particularmente en puestos de dirección (Hough et al., 2001; Leutner y Charmorro-Premuzic, 2018; Mount y Barrick, 1995).

4.2. Red de relaciones nomológicas de la escala de Desempeño adaptativo de Charbonnier-Voirin y Roussel en castellano

En relación al desempeño laboral, y como ya se ha recogido en el capítulo 2, hay un consenso generalizado en torno a la existencia de tres subdimensiones de desempeño para cualquier organización y para cualquier trabajo: *desempeño de tarea*, comportamientos exigidos para la ejecución de las tareas de un puesto, como productividad, calidad o eficacia; *desempeño contextual*, conductas que facilitan el desempeño de tarea, como la iniciativa, la ayuda a los demás o la vinculación con la empresa; y las *conductas contraproductivas*, robos, abusos de drogas, etc. (Beck et al., 2014; Gorriti y López, 2009, 2014).

Siguiendo la perspectiva de Katz (1964), si originalmente la investigación en materia de evaluación del desempeño se centró en el desempeño técnico-productivo o *intra-rol* (desempeño de tarea, es decir, las conductas incluidas en las descripciones formales de puestos de trabajo), en los últimos años se han identificado varias dimensiones emergentes *extra-rol* (acciones que van más allá del deber formal del puesto) del desempeño laboral, como el desempeño adaptativo (Johnson, 2003; Ramírez, 2013). Este tipo de desempeño podría definirse (Ramírez, 2013, p. 97) como la “dimensión de desempeño que da cuenta de aquellas conductas de adaptación o ajuste eficaz ante situaciones cambiantes y/o (de la) adquisición de habilidades para ser competente ante nuevas demandas en el trabajo” (Hesketh y Neal, 1999; Ployhart y Bliese, 2006; Pulakos et al., 2000).

Un entorno laboral cambiante como el actual hace que cada vez sea más importante que empleados y empleadas sean flexibles y capaces de adaptarse a los cambios en los sistemas de trabajo o en los roles de los puestos derivados de las nuevas tecnologías, la crisis económica y la consiguiente reestructuración de las empresas o la globalización. Esta tendencia ha provocado que el desempeño adaptativo haya cobrado gran relevancia en el ámbito del desempeño laboral (Koopmans et al., 2011; Landy y Conte, 2016).

Dada la escasez de investigaciones centradas en el constructo desempeño adaptativo (en adelante, DA), el objetivo de este capítulo es situar el DA en su red de relaciones nomológicas, siguiendo el modelo teórico I-ADAPT de Ployhart y Bliese (2006). Para ello se explora su relación con constructos relevantes como las *conductas cívico organizacionales* (Organizational Citizenship Behavior, OCB en inglés) y variables predictoras del desempeño, como el ajuste persona-organización o los cinco grandes factores de la personalidad (este último aspecto dentro del epígrafe que describe la red de relaciones nomológicas de OPERAS).

4.2.1. El modelo I-ADAPT de Ployhart y Bliese (2006) y la conceptualización del desempeño adaptativo

El modelo teórico I-ADAPT de Ployhart y Bliese (2006) se enmarca en una perspectiva que concibe la adaptabilidad como un conjunto de diferencias individuales que predisponen a los/as empleados/as para adaptarse con éxito a los cambios en las tareas. El trabajo de Ployhart y Bliese (2006) es el más destacado en esta área, y teoriza que los conocimientos, las destrezas, los valores, las aptitudes y la personalidad predicen el desempeño adaptativo, y que éste a su vez predice el desempeño de tarea o el contextual, lo que ha sido parcialmente confirmado por diversos estudios que han analizado la relación entre APO y adaptabilidad, o adaptabilidad y desempeño contextual (Almahamid et al., 2010; Chan y Schmitt, 2002; Wang et al., 2011).

Ployhart y Bliese (2006, p. 13) definen la adaptabilidad como “la capacidad, destreza, disposición y/o motivación de un individuo para cambiar o adaptarse a diferentes elementos de las tareas, sociales o ambientales”. Se trata de un constructo estable, una diferencia individual que influye en la manera en que una persona interpreta y responde a una situación cambiante. Los autores la diferencian de las dimensiones tradicionales del desempeño laboral, y la conciben como un conjunto de antecedentes que capturan la varianza compartida de conocimientos, personalidad, valores, etc. que predicen el desempeño. No obstante, señalan que la adaptabilidad es un constructo más relacionado con el desempeño que constructos distales como la personalidad o los valores, que son los determinantes directos de las diferencias individuales en adaptabilidad. En consecuencia, la adaptabilidad no es una medida del desempeño sino

una representación de los antecedentes (personalidad, valores, etc.) necesarios para el desempeño en diferentes contextos. Sin embargo, las investigaciones sobre la conceptualización del predictor (la adaptabilidad) y el criterio (el desempeño) no han sido concluyentes, lo que ha dificultado sintetizar los hallazgos del enfoque que concibe el desempeño adaptativo como diferencia individual (Baard et al., 2014).

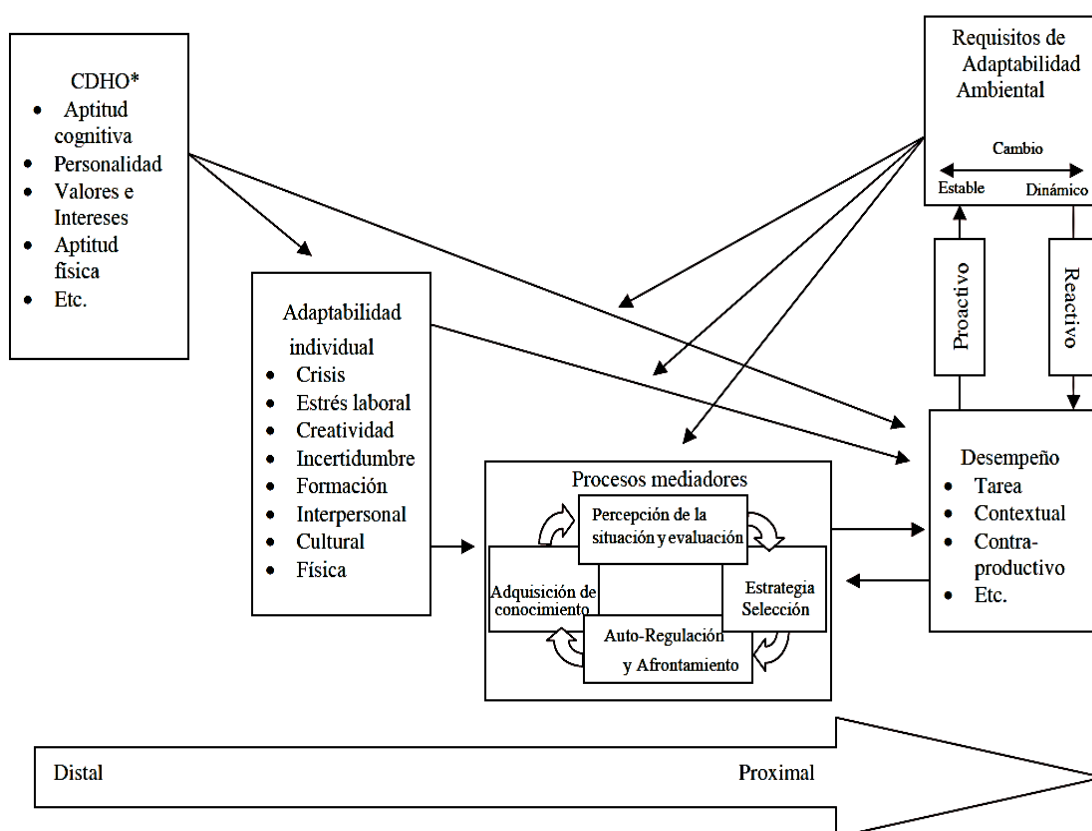
La revisión teórica sobre el desempeño adaptativo de Baard et al. (2014) diferencia el enfoque de la adaptabilidad entendida como diferencias individuales, postulado por Ployhart y Bliese (2006), del enfoque que concibe el desempeño adaptativo como un constructo del desempeño equiparable al desempeño de tarea o al desempeño contextual. Este último enfoque asume que el DA es una dimensión diferenciada del desempeño laboral que puede ser predicha por una serie de diferencias individuales y está limitada por el contexto (Pulakos et al., 2012). No obstante, como señalan varios autores, para extraer conclusiones sobre el DA es recomendable no segregar los hallazgos en diferentes dominios o enfoques sobre el desempeño adaptativo, teniendo en cuenta las conclusiones de estudios con métodos, contextos y objetivos diversos en aras a integrar y sintetizar la investigación en esta materia (Calarco, 2016; Jundt et al., 2015; Ployhart y Bliese, 2006).

Este dialogo entre enfoques diferentes se comprueba en la construcción del instrumento I-ADAPT-M de Ployhart y Bliese (2006), que está basado en la taxonomía de Pulakos et al. (2000) y que es la base del enfoque que concibe el desempeño adaptativo como un constructo del desempeño. El modelo conceptual de Ployhart y Bliese se diseñó para situar al desempeño adaptativo en el marco de una red nomológica de antecedentes (la personalidad, los valores, etc.) y consecuencias (el desempeño laboral) de manera que se pudiesen realizar propuestas de investigación sobre las relaciones y los efectos (Ver Figura 4.1). En el caso de dimensiones relativamente recientes del desempeño laboral, como el desempeño adaptativo, comprender sus relaciones con las dimensiones tradicionales del desempeño laboral es una cuestión vital. Las estimaciones de estas relaciones son necesarias para distinguir empíricamente las dimensiones de desempeño y para mejorar nuestra comprensión de la estructura del desempeño laboral individual (Viswesvaran et al., 2005). El modelo I-ADAPT responde a dicha necesidad y está vinculado conceptualmente con la selección de personal, razón

por la cual se ajusta al estudio de los objetivos de esta tesis y se ha empleado para situar al DA en su red nomológica (Murphy, 2015).

Figura 4.1.

Teoría de la Adaptabilidad Individual (I-ADAPT). Adaptado de Ployhart y Bliese (2006).



*Nota. CDHO = Traducción del acrónimo en inglés KSAO (Knowledge, Skills, Abilities and Other personality characteristics), que se corresponde con las iniciales en castellano de Conocimientos, Destrezas, Habilidades o Aptitudes y Otras características de personalidad.

4.2.2. La relación entre desempeño adaptativo y conductas cívico-organizacionales

Una cuestión en la que apenas se ha profundizado es la relación entre DA y un concepto vinculado al desempeño contextual como son las OCBs (Johnson, 2003). Las OCBs son una faceta del desempeño laboral de gran relevancia, ya que contribuyen a mejorar la efectividad y la gestión de las organizaciones (Organ, 2018; Podsakoff et al., 2009; Vigoda, 2000; Vigoda y Golembiewski, 2001). Como se ha mencionado anteriormente, en la literatura sobre desempeño, este tipo de conductas se

conceptualizan como desempeño *extra-rol*, es decir, diferente al desempeño de tarea o a las conductas *intra-rol*, esto es, aquellas conductas que forman parte de la descripción formal del puesto de trabajo. Las OCBs son conductas que se pueden definir como “contribuciones individuales en el lugar de trabajo que van más allá de los requerimientos de rol y de los logros laborales contractualmente recompensados” (Organ y Ryan, 1995, p. 775).

El meta-análisis de Chiaburu et al. (2017) puso de manifiesto que la importancia relativa de las OCBs era mayor que la del desempeño de tarea a la hora de explicar la varianza de las evaluaciones del desempeño laboral general, siendo su peso específico del 73% frente al 27% del desempeño de tarea. Este hallazgo coincide con lo apuntado por Lievens, Conway et al. (2008), que señalaron que las OCBs son la dimensión del desempeño más relevante en culturas organizacionales basadas en equipos y cuando la evaluación es realizada por compañeros/as de trabajo. Este mayor peso relativo concedido a las OCBs es un reflejo de la creciente importancia que están adquiriendo los constructos de desempeño extra-rol en las últimas décadas (Carpini y Parker, 2017) y supone un cambio conceptual importante que contrasta con la relevancia que investigaciones precedentes concedían al desempeño de tarea (Conway, 1999; Motowidlo y Van Scotter, 1994; Rotundo y Sackett, 2002).

Existen diversas conceptualizaciones de las OCBs, lo que ha provocado que históricamente se hayan identificado diferentes dimensiones de este constructo. Organ (1988) distinguió cinco clusters de conductas: *Deportividad*, *Conciencia*, *Virtud Cívica*, *Altruismo* y *Cortesía*. Podsakoff et al. (2000), por su parte, distinguieron 7 tipos de OCBs: *Conductas de ayuda*, *Deportividad*, *Lealtad organizacional*, *Cumplir con la organización*, *Iniciativa individual*, *Virtud cívica* y *Autodesarrollo*. Sin embargo, el enfoque más parsimonioso y tradicionalmente empleado en la investigación sobre la materia (Chiaburu et al., 2011; Chiaburu et al., 2017; Hoffman y Dilchert, 2012; Ilies et al., 2009; Organ, 2018; Spitzmuller et al., 2008; Tambe y Shanker, 2014) es el de Williams y Anderson (1991), que categorizaron las 5 dimensiones de Organ en dos grandes grupos en función del receptor de la conducta: *OCBs dirigidas a la organización (OCB-O)* y *OCBs dirigidas a otros individuos (OCB-I)*. OCB-I incluye conductas categorizadas por otros autores como cortesía, conductas de ayuda o altruismo, mientras que OCB-O incluye conciencia, virtud cívica o deportividad. Este

enfoque es similar, en buena medida, al de la taxonomía del desempeño contextual de Borman y Motowidlo (1993), que distingue entre las dimensiones *ayuda a los demás*, concepto similar a OCB-I, *vinculación organizacional*, concepto asimilable a OCB-O, y *desempeño más allá del puesto* (Dorsey et al., 2010). Como señalan Tambe y Shanker (2014), una de las escalas más empleadas en la operacionalización de las dos grandes dimensiones de las OCBs es la escala de Lee y Allen (2002).

Además de estas dos dimensiones denominadas OCBs prosociales o afiliativas, existe una tercera dimensión que ha formado parte de las investigaciones sobre las OCBs, las *OCBs orientadas al cambio*, también denominadas proactivas, y vinculadas a la creatividad, innovación y a la introducción de cambios y mejoras de utilidad para la organización (Bettencourt, 2004). Dado que las conductas proactivas están vinculadas al DA y al ajuste persona-organización (Carpini y Parker, 2017; Griffin et al., 2007; Parker y Collins, 2010) y teniendo en cuenta que dentro de las OCBs las conductas orientadas al cambio son las que menor peso específico tienen a la hora de explicar la varianza del desempeño laboral general (Chiaburu et al., 2017), se ha considerado oportuno no incluirlas en esta investigación.

Ployhart y Bliese (2006), en su modelo conceptual del desempeño adaptativo, entienden que cada dimensión del desempeño laboral requiere de adaptabilidad por parte de los/as empleados/as, mencionando como ejemplo, que el desempeño adaptativo contextual puede conllevar ayudar a nuevos/as compañeros/as de diferentes culturas o que prestarse voluntariamente a ayudar a colegas (un aspecto de las OCBs) puede requerir adaptarse a las conductas cambiantes de estos/as compañeros/as de trabajo.

No obstante, apenas hay investigaciones acerca de la relación entre DA y OCBs. Entre ellas, cabe mencionar la de Charbonnier-Voirin y Roussel (2012), que hallaron una correlación de .60 entre desempeño contextual y una medida global de Desempeño adaptativo. Mucho menor fue el tamaño de la correlación (.22) encontrada entre una medida de desempeño adaptativo y una medida global de OCBs por Segura-Camacho et al. (2018). Chan y Schmitt (2002), por su parte, empleando un test de juicio situacional concebido como una medida de la adaptabilidad, estudiaron su relación con el desempeño contextual de tipo motivacional e interpersonal, encontrando correlaciones de .38 y .27, respectivamente.

4.2.3. El ajuste persona-organización (APO) y el desempeño laboral

Hasta el momento no se ha profundizado lo suficiente en el estudio del rol del ajuste persona-organización (en adelante, APO) en la selección de personal. Sin embargo, seleccionar a las y los aspirantes a un puesto en base a su ajuste a los valores o la cultura de una empresa es una cuestión a la que se debe atender en selección de personal (Peiró et al., 2020). Esto cobra especial relevancia en la selección interna o promoción de las personas que ya trabajan en una organización y en organizaciones con culturas “fuertes”, en las que hay congruencia entre los valores individuales y organizacionales, como las entidades públicas (Ostroff y Zhan, 2012; Ryan y Ployhart, 2014).

Como indica van Vianen (2018), el APO tiene una relación considerable con las actitudes laborales, menos intensa con el desempeño contextual y baja con el desempeño general (Arthur et al., 2006; Kristof-Brown et al., 2005). Concretamente, la correlación de este constructo con Desempeño General es de .13, y de .27 con Desempeño Contextual (Schmidt et al., 2016). Pese a que no tiene gran capacidad predictiva, consideramos que el APO puede tener relevancia en la selección, en la adaptabilidad de las personas a su entorno de trabajo y en su desempeño laboral. Por ello, nuestra intención ha sido profundizar en el conocimiento acerca de su relación con un constructo relevante del que puede ser antecedente según el modelo de Ployhart y Bliese (2006), el desempeño adaptativo.

El APO fue definido por Kristof (1996) como la similitud entre las características de la persona (p.e. su personalidad, valores, objetivos o actitudes) y las de la organización (p.e. su cultura, clima, valores, objetivos o normas). Este tipo de ajuste es uno de los constructos que forman parte del dominio *ajuste persona-ambiente*, modelo teórico que engloba, además, otros cuatro tipos de ajuste: el *ajuste persona-puesto de trabajo*, el *ajuste persona-grupo*, el *ajuste persona-supervisor/a* y el *ajuste persona-vocación*. El ajuste persona-ambiente fue definido por Kristof-Brown et al. (2005, p.281) como “la compatibilidad que ocurre cuando hay una correspondencia entre las características del individuo y las del ambiente de trabajo”. Para Chatman (1989), el APO es el aspecto más importante a la hora de determinar el ajuste persona-ambiente, y este ajuste puede ser logrado mediante el proceso de selección, en el que las organizaciones contratan a personas que comparten los valores organizativos.

Para los individuos, percibir un mayor grado de APO puede suponer importantes consecuencias a nivel de promoción, como una mayor permanencia en la organización, más satisfacción laboral, un mayor compromiso con la organización, más conductas extra-rol como las OCBs, menor estrés y mayores sentimientos de confort y competencia. Para la organización, los potenciales beneficios son un aumento de la productividad y la eficacia y desempeño grupales (Kristof-Brown y Guay, 2011; Su et al., 2015). Sin embargo, algunos autores han indicado que niveles demasiado altos de APO pueden provocar conductas ineficaces tanto a nivel individual como organizacional, como una mayor conformidad y una menor innovación y adaptabilidad (Chatman, 1989; Schneider, 1987).

Existen dos modalidades de APO: el *ajuste o congruencia suplementaria* y el *ajuste o congruencia complementaria*. El ajuste suplementario se refiere a la correspondencia entre la persona y el grupo de personas que componen un contexto ambiental en base a su similitud. Según Ximénez Gómez (1998, p. 9), ello “implica que una persona decide ingresar y permanecer en una organización porque percibe que sus valores, intereses y preferencias son similares a las de los miembros de esa organización. Es decir, porque percibe que se ajusta a esa organización”. La unidad de análisis es la persona, quien más se beneficia es el individuo, que si se ajusta al ambiente reacciona con satisfacción y permaneciendo en el mismo. Por lo que respecta al ajuste complementario, este se refiere al grado en que las características de una persona sirven para complementar o mejorar las características de un ambiente. Este tipo de ajuste tiene a su vez dos componentes: el ajuste habilidades-demandas, que se produce cuando las habilidades y destrezas de una persona responden a las demandas de la organización, y el ajuste necesidades-recursos, que ocurre cuando los recursos del ambiente responden a las necesidades de las personas. La principal beneficiaria del ajuste complementario es la organización, ya que un buen ajuste redundaría en una mejora de la productividad y efectividad organizacionales (Muchinsky y Monahan, 1987; Su et al., 2015; Ximénez Gómez, 1998). Aunque son diferentes, estos tipos de ajuste tienen ciertas similitudes y suelen co-ocurrir. La mayor parte de las investigaciones se han centrado en el ajuste suplementario (Kristof-Brown y Guay, 2011; Piasentin, 2003; Su et al., 2015).

En la medición del APO se distingue entre medidas “directas” o subjetivas del APO, es decir, instrumentos en los que las personas participantes evalúan su grado de ajuste a la organización y medidas “indirectas” u objetivas, que evalúan por separado las características de la persona y las de la organización y después las comparan, por ejemplo, mediante un algoritmo. Según diversas investigaciones, las medidas directas presentan correlaciones más altas que las medidas indirectas con el desempeño laboral (Kristof-Brown et al., 2005; Verquer et al., 2003).

Por lo que se refiere a los contenidos utilizados en la evaluación del APO, cabe citar la congruencia con la organización en cuanto a valores, objetivos, personalidad, o el conjunto de conocimientos, destrezas y aptitudes de una persona (Knowledge, Skills, Abilities en inglés, KSA), siendo los valores el aspecto empleado con más frecuencia en las investigaciones (Piasentin, 2007; Piasentin y Chapman, 2006; Verquer et al., 2003).

En su análisis sobre los instrumentos empleados para medir el APO, Piasentin (2007) concluyó que la mayoría de investigaciones se orienta a la medición del ajuste suplementario, desde una perspectiva subjetiva y centrada en los valores, parámetros a partir de los cuales diseñó un instrumento que presenta buenas propiedades psicométricas (Cáceres, 2014).

En cuanto a la relación entre el APO y el DA, ha sido menos estudiada que su relación con la personalidad o las OCBs. Una de las escasas investigaciones que han estudiado la relación entre APO y DA es la de Wessel et al. (2008), que encontraron correlaciones en torno a .20 entre ajuste percibido y la adaptabilidad hacia el aprendizaje y a la incertidumbre. Wang et al. (2011), por su parte, encontraron una relación positiva entre APO y adaptabilidad cultural en personas de reciente incorporación al entorno laboral, y propusieron como posibles líneas de investigación analizar si esta relación es potencialmente beneficiosa para el desempeño contextual y si la adaptabilidad hacia el aprendizaje puede influir en el desempeño.

4.3. Red nomológica de OPERAS en euskera: la relación entre personalidad y desempeño laboral

Como se ha indicado en el capítulo 3, la personalidad puede definirse como la predisposición de las personas a pensar, sentir y actuar de una determinada manera en respuesta al entorno (Smith y Smith, 2005). En Psicología del Trabajo y las Organizaciones se refiere a las variables que moderan toda ejecución, que predicen que un/a trabajador/a que posee determinadas aptitudes, destrezas y conocimientos los utilizará correctamente o estará influenciado/a por aspectos emocionales (estabilidad emocional) o predisposiciones personales (extroversión, necesidad de logro, etc.).

Tal y como se ha indicado previamente, en la actualidad, el modelo de personalidad sobre el que existe un mayor consenso es el Modelo de los Cinco Factores (Five Factor Model – FFM), también conocido como los Cinco Grandes -Big Five- (Barrick y Mount, 2003; Soto y John, 2009). Dichos factores son: *Amabilidad* –también conocido como Afabilidad, Agradabilidad o Amigabilidad-, *Apertura a la experiencia* – en algunos casos se han empleado los términos Apertura o Apertura Mental-, *Estabilidad Emocional* –factor que en algunas pruebas se mide mediante su reverso, Neuroticismo-, *Extroversión*, y *Responsabilidad* –en algunos casos se ha traducido como Conciencia o Tesón-.

Shaffer y Postlethwaite (2012) afirman que la evidencia empírica a favor de determinados factores de personalidad como predictores válidos del desempeño laboral es amplia, si bien el debate al respecto continúa, ya que los índices de capacidad predictiva alcanzados por las medidas de personalidad son, en el mejor de los casos, moderados (Morgeson et al., 2007).

Sin embargo, Ones et al. (2007) entienden que estos índices pueden ser relevantes, al aportar las pruebas de personalidad capacidad predictiva añadida, es decir, varianza adicional del desempeño laboral, a la explicada por la mayoría de las pruebas de selección más comunes tanto en el sector público como en el privado (inteligencia general, pruebas de conocimientos relativos al trabajo). Esta capacidad predictiva añadida se debe a que no presentan correlaciones elevadas con dichas pruebas.

Además, los factores de personalidad son predictores relevantes de aspectos importantes para las organizaciones. Prueba de ello es el meta-análisis de Ones et al. (2007) que estudió la relación entre el conjunto de los cinco grandes factores de personalidad y variables como el trabajo en equipo, conductas contraproductivas, la orientación al cliente, la satisfacción laboral y el compromiso con la organización, el liderazgo, o el aprovechamiento de la formación, las conductas cívico organizacionales o el desempeño laboral. Dicho meta-análisis encontró magnitudes de correlación múltiple de tamaño medio y alto, entre .27 y .45.

En general, diversos meta-análisis han demostrado una mayor relación entre la personalidad y la conducta de desempeño cuando esta es más discrecional y volitiva, como en el caso del desempeño contextual, que cuando el desempeño está más monitorizado y estructurado, como en el desempeño de tarea (Barrick y Mount, 2014; Chiaburu et al., 2011). La personalidad es especialmente útil para predecir conductas de tipo “querer hacer”, vinculadas a la motivación y al nivel de atención, dirección, intensidad y persistencia del esfuerzo de una persona ante una situación dada, así como a la adaptación a las normas organizacionales, frente a conductas de tipo “poder hacer”, más vinculadas a las aptitudes y destrezas (Campbell, 1990; Barrick y Mount, 2014).

En consecuencia, ante medidas de desempeño extra-rol, influidas fundamentalmente por la motivación de las personas, como las OCBs o el desempeño adaptativo, la personalidad puede ser un instrumento selectivo de interés para predecir dichos constructos. Esto se deriva de su vinculación con el compromiso con la organización y los/as compañeros/as o el deseo de aprender nuevas maneras de trabajar y mejorar las destrezas individuales (Barrick y Mount, 2014).

En ese sentido, tal y como subrayan Sackett et al. (2017), al evaluar la capacidad predictiva de la personalidad es especialmente relevante tener en cuenta la correspondencia conceptual a nivel de constructo entre los predictores, como los factores de personalidad en el ámbito de selección, y los criterios, es decir, el desempeño laboral y los constructos que lo componen. Por esta razón, para analizar la red de relaciones nomológicas de OPERAS en euskera, a continuación, se examina la relación entre los cinco grandes factores de la personalidad y constructos del dominio del desempeño laboral muy vinculados a los mismos: las subdimensiones del desempeño adaptativo y de las OCBs. Ello contribuirá a incrementar nuestro bagaje de

conocimientos sobre las dimensiones concretas de desempeño que son predichas por factores concretos de personalidad.

4.3.1. La personalidad y el desempeño adaptativo

Como señalan Jundt et al. (2015), los cinco grandes factores de personalidad son uno de los constructos de diferencias individuales más estudiados como antecedentes del desempeño adaptativo (DA). Desde una perspectiva conceptual, Pulakos et al. (2006) examinaron la literatura al respecto y sometieron al juicio de expertos/as el grado de relevancia de potenciales predictores del DA, concluyendo que Estabilidad Emocional y Apertura a la experiencia eran los factores de personalidad más relevantes. No obstante, diversas investigaciones se han centrado en la relación empírica entre los factores de personalidad y el desempeño y han llegado a conclusiones diferentes, encontrando, según el caso, que Estabilidad Emocional, Apertura a la experiencia o Conciencia presentaban relaciones positivas de carácter moderado con el DA (Griffin y Hesketh, 2003; Neal et al., 2012; Pulakos et al., 2002). Pese a ello y con ánimo de aportar nuevos datos al respecto, el meta-análisis de Huang et al. (2014) basado en 71 muestras independientes ($N = 7535$), ha señalado que Estabilidad Emocional y Extroversión son los factores de personalidad que están relacionados positivamente con el DA, con correlaciones corregidas por falta de fiabilidad del criterio y restricción del rango de .16 y .14, respectivamente. Estos valores ascienden a .20 y .26 en el caso de personas que ejercen puestos de mando. No obstante, este meta-análisis se realizó con un único cuestionario de personalidad que no mide los Cinco Grandes *per se*, aunque emplea un modelo próximo al FFM, el Hogan Personality Inventory (Hogan y Hogan, 2007). Este instrumento mide 7 factores de personalidad vinculados conceptualmente al modelo FFM, que presentan una correlación media de .54 entre los factores correspondientes de ambos modelos (Hogan y Holland, 2003).

Las causas de estos resultados poco consistentes pueden estar vinculadas a la insuficiente claridad en la definición y operacionalización del constructo DA que ha sido objeto de crítica por parte de diferentes autores (Baard et al., 2014; Jundt et al., 2015; Stokes et al., 2010), lo que dificulta una mejor comprensión de su red de relaciones nomológicas. No obstante, siguiendo a Baard et al. (2014), el camino a seguir debe partir del modelo teórico más riguroso, el de Pulakos et al. (2000) –descrito en el Capítulo 2-, examinando de manera precisa la estructura factorial subyacente al mismo.

En este sentido, el trabajo de Charbonnier-Vorin y Roussel (2012), basado en el modelo de Pulakos et al. (2000), ha permitido obtener una escala que no solo presenta adecuadas propiedades psicométricas para evaluar el DA, sino que también es conceptualmente sólida (Pini y Mariani, 2014; Walker, 2015). Estas autoras identificaron cinco dimensiones del DA: *Creatividad; Gestión del estrés; Adaptabilidad interpersonal; Reactividad ante emergencias; y Esfuerzo formativo y de aprendizaje*. Su escala ha sido empleada en diferentes investigaciones vinculadas al DA (Kanten et al., 2015; Monteiro, 2015). Dados los resultados poco concluyentes de investigaciones previas, resulta de interés estudiar la correlación entre los cinco grandes factores de personalidad y las cinco dimensiones identificadas por Charbonnier-Voirin y Roussel que se han adaptado al castellano en el marco de esta tesis. El único precedente al respecto es la investigación de Charbonnier-Voirin (2013), que encontró correlaciones estadísticamente significativas y superiores a .20 en una muestra de 238 personas entre los siguientes factores de personalidad y dimensiones de DA: Apertura a la experiencia y Creatividad; Responsabilidad y Esfuerzo de Aprendizaje; Amabilidad y Adaptabilidad Interpersonal; y Estabilidad emocional y Gestión del estrés, que presentaron la magnitud más elevada, .37.

4.3.2. La personalidad y las conductas cívico organizacionales (OCBs)

En cuanto a la relación entre la personalidad y las OCBs, el meta-análisis de Chiaburu et al. (2011) encontró que los factores de personalidad que presentaban correlaciones más altas con las OCBs eran Conciencia, Amabilidad y Apertura a la experiencia, con magnitudes entre .17 y .21 tomando OCB-O y OCB-I como criterios. El meta-análisis de Judge et al. (2013), por su parte, encontró que Conciencia, Extroversión y Amabilidad presentaban unas correlaciones corregidas por atenuación debida a falta de fiabilidad en la predicción del conjunto de las OCBs de .40, .28 y .23, respectivamente (Sackett y Walmsley, 2014). Por su parte, Ilies, et al., en su meta-análisis de 2009, encontraron que Amabilidad estaba más relacionada con las OCB-I y que Responsabilidad presentaba una correlación más elevada con las OCB-O.

4.3.3. La personalidad y el ajuste persona-organización (APO)

Otra cuestión de interés para el análisis de la red nomológica de OPERAS en euskera es la relación entre personalidad y APO. Las escasas investigaciones que se han centrado en este aspecto han encontrado relaciones positivas de los factores de personalidad Extroversión, Amabilidad y Responsabilidad con el ajuste persona-organización (Gardner et al., 2012; Gillespie et al., 2003).

La correlación entre personalidad y valores (Parks, 2007) implica que, para aumentar la eficacia de las prácticas selectivas, podría ser útil complementar la medida de la personalidad con variables vinculadas a los valores, como el APO, ya que sus efectos conjuntos sobre el desempeño podrían ser superiores a los que presentan por separado. De hecho, explorar el efecto de terceras variables en la relación entre personalidad y desempeño es una línea de investigación prometedora (Barrick y Mount, 2014; Hugues y Batey, 2017). Como los valores reflejan lo que es deseable y sostenible por las personas y los grupos (como una unidad de trabajo o una organización), pueden aportar información acerca del ajuste más relevante en selección: el ajuste-persona organización. Por tanto, el APO puede aumentar la capacidad predictiva de la personalidad (Barrick y Mount, 2014; Judge y Cable, 1997; Parks y Guay, 2009). En consecuencia, en el siguiente epígrafe se aborda el papel del APO en la relación entre personalidad y OCBs, por un lado, y entre personalidad y desempeño adaptativo, por otro.

4.3.4. El modelo I-ADAPT y la relación entre personalidad y desempeño laboral. El papel del ajuste persona-organización y del desempeño adaptativo

El APO presenta interés no solo por su relación con la personalidad sino por su relación con el desempeño laboral, y concretamente con las OCBs. En ese sentido, el meta-análisis de Hoffman y Woehr (2006) puso de manifiesto que la relación del APO con las OCBs, sin distinguir entre OCB-O y OCB-I, presenta una magnitud de .21. En su estudio, los autores distinguieron entre medidas de APO objetivas, percibidas y subjetivas, presentando las siguientes correlaciones con las OCBs: .26, .21 y .17, respectivamente. Los cuestionarios basados en el ajuste en valores presentaron

correlaciones más altas (.25) que las de los basados en otros tipos de ajuste (.18). Las correlaciones entre el APO y el desempeño de tarea oscilaron entre .20 y .28.

Estas asociaciones entre APO, personalidad y OCBs han despertado el interés de los/as investigadores/as sobre la capacidad predictiva conjunta de APO y personalidad con las OCBs como criterio, lo cual está relacionado con lo apuntado por el modelo I-ADAPT. Así, las investigaciones de Tsai et al. (2012), con una muestra de empleados/as de una empresa de alta tecnología, y Michaud (2014), con una muestra de empleados/as públicos/as, han mostrado que ambas variables conjuntamente pueden contribuir a explicar las OCBs, indicando que el ajuste persona-organización puede incrementar la capacidad predictiva de la personalidad. No obstante, emplearon en ambos casos muestras pequeñas y obtuvieron tamaños del efecto diferentes ($\Delta R^2 = .02$ y $\Delta R^2 = .33$, respectivamente).

Por otro lado, la conclusión lógica de la aplicación del modelo I-ADAPT al contexto de la selección de personal implica considerar el DA como predictor del desempeño laboral. Ello, junto con las escasas investigaciones que han confirmado la relación entre DA y OCBs, citadas al inicio de este capítulo, plantea una interesante posibilidad: el estudio de la capacidad predictiva añadida por el DA a la personalidad y al APO en la predicción de las OCBs.

Asimismo, la vinculación entre APO y personalidad y su influencia sobre el desempeño laboral es potencialmente extensible al DA. Así, Wang et al. (2011) estudiaron el papel de un concepto vinculado al APO, como es el ajuste persona-ambiente, en la adaptabilidad de las personas de reciente incorporación al entorno laboral chino. Sus resultados sugieren que este ajuste podría influir en la relación entre personalidad y DA, de manera similar a lo apuntado por Avdic (2012), en la relación entre personalidad y OCBs.

Teniendo en cuenta que los valores son un determinante directo del desempeño adaptativo (Ployhart y Bliese, 2006; Tucker et al., 2010) y partiendo de la necesidad de identificar variables que puedan añadir capacidad para predecir el desempeño adaptativo a la de predictores comunes en selección de personal (Lang y Bliese, 2009), resulta interesante explorar la posibilidad de que el ajuste persona-organización pueda incrementar la capacidad de la personalidad para predecir el desempeño adaptativo.

Además, como señalaron Cooper et al. (2013) o Bright (2008), es necesario examinar la relación entre personalidad, APO y desempeño en el sector público.

4.4. Las variables sociodemográficas y laborales

Por otro lado, el estudio de la relación de variables sociodemográficas y laborales tales como el género, la edad, los años de experiencia laboral, el ejercicio de puestos de mando, el nivel educativo, el grado de complejidad del puesto desempeñado o la estabilidad en el puesto, con constructos empleados frecuentemente para la selección o la promoción en el ámbito de recursos humanos, tales como la personalidad, el ajuste persona-organización, las OCBs o el desempeño adaptativo, es una cuestión de especial relevancia por sus implicaciones en la selección y gestión de personas, así como en la evaluación del desempeño laboral.

Esta relevancia se debe a la posibilidad de que un constructo empleado en materia de recursos humanos presente *impacto adverso*, es decir, que, ante ausencia de diferencias en los constructos evaluados, las pruebas que los miden produzcan puntuaciones medias distintas en diferentes subgrupos de variables sociodemográficas (edad, género, condición social, etc.), lo que puede provocar que dichos colectivos tengan ratios de selección desfavorables. Si el subgrupo constituye una clase protegida, ello puede acarrear consecuencias legales negativas para la organización (Shen et al., 2017; Yusko et al., 2017). Como señalan Alonso et al. (2017), la legislación europea insta a que los instrumentos selectivos garanticen la igualdad de oportunidades y eviten la discriminación, tanto directa o intencional, como indirecta. La discriminación directa es intencional, mientras que la indirecta es la que se produce cuando una organización emplea instrumentos aparentemente neutrales en los que un subgrupo protegido (por ejemplo, el de las mujeres) obtiene calificaciones inferiores sin que haya diferencias a nivel de desempeño laboral. Por ello, es preciso investigar sobre la posible relación entre variables sociodemográficas y laborales y los instrumentos que se utilizan en la selección.

Asimismo, como indican Jundt et al. (2015), son muy pocas las investigaciones que se han centrado en el contexto organizacional, social y laboral en el que acontece el DA. En consecuencia, resulta necesario profundizar en el estudio de la relación de variables sociodemográficas y laborales, tales como el ejercicio de puestos de mando o

la complejidad del puesto con este constructo, precisando además las diferencias existentes entre las dimensiones que lo componen.

Una cuestión vinculada a las diferencias entre subgrupos es la *justicia percibida* acerca de las pruebas utilizadas por parte de los candidatos y candidatas a empleos. En los últimos años, esta cuestión ha cobrado especial interés dada la importancia que tiene para el diseño de procedimientos selectivos por parte de las organizaciones y a una demanda social creciente de justicia, imparcialidad y trato justo en la selección de personal. Por consiguiente, los procedimientos selectivos deben ser justos y defendibles legalmente (Carr y Patterson, 2009; Patterson et al., 2011). Asimismo, el estudio de las reacciones de candidatos y candidatas ante la justicia del proceso selectivo es crucial porque pueden influir no solo en la imagen pública de la organización convocante, sino también en el resultado de las pruebas de selección e incluso en el resultado de la evaluación del desempeño una vez estas personas son contratadas (Anderson et al., 2010; McCarthy et al., 2013; Ryan y Ployhart, 2000).

4.4.1. La relación del género con los constructos analizados

De entre las variables mencionadas en este estudio, los factores de personalidad, cuando se emplea el modelo FFM, presentan un impacto adverso de género reducido, ya que las medias son similares entre mujeres y hombres. De ahí que, en términos generales, los cuestionarios de personalidad cumplan con los requisitos de la legislación en materia de igualdad de oportunidades y equidad en la selección y, además, sean percibidos favorablemente por los candidatos y las candidatas (Anderson et al. 2010; Barrick et al., 2011; Barrick y Mount, 2014). No obstante, sí que existen diferencias en función del género, y cabe mencionar que las mujeres presentan puntuaciones superiores a las de los hombres en Amabilidad, pero puntuaciones más bajas que las de los hombres en Estabilidad Emocional. En ambos casos los tamaños del efecto son de magnitud pequeña o media, d de Cohen entre .40 y .50. En el resto de factores las diferencias suelen ser menores, d de Cohen entre 0 y .1 (Barrick y Mount, 2014; Hyde, 2014; Kajonius y Johnson, 2018; Kajonius y McGiolla, 2017; Kaiser et al., 2019).

En lo que se refiere al desempeño adaptativo, tal y como han señalado Jundt et al. (2015) o Dorsey et al. (2010), se han realizado pocos estudios y se han llevado a cabo con medidas globales del DA, sin prestar atención a sus dimensiones. Un ejemplo

es el de O'Connell et al. (2008), en el que se encontraron niveles diferentes de adaptabilidad según el género (mayores puntuaciones para las mujeres, si bien el tamaño del efecto fue pequeño).

Por otra parte, el consenso existente en cuanto al moderado impacto adverso en la personalidad no se puede trasladar a las pruebas que miden las OCBs. Existen resultados discrepantes acerca de la relación entre el género y las puntuaciones en OCB. Así, Akinbode (2011) encontró que el género no es un predictor válido de las puntuaciones en OCB. Sin embargo, Allen y Rush (2001) señalaron que era más probable que los varones realizaran OCBs con el fin de aumentar sus posibilidades de promoción, mientras que Dávila et al. (2011) observaron que las mujeres ponen en práctica con más frecuencia OCBs de tipo interpersonal. No obstante, el meta-análisis de Ng et al. (2016) no ha hallado diferencias de género en las puntuaciones en OCB. Esto coincide con lo apuntado por Dorsey et al. (2010), que señalaron que, en general, las diferencias en subgrupos establecidos en base a variables sociodemográficas (género, raza, etc.) en las OCBs son inferiores a las que existen en el desempeño técnico o de tarea.

En cuanto a la relación entre APO y género, diferentes investigaciones han encontrado que no hay relación entre dichas variables o que el tamaño del efecto es pequeño (Seong et al., 2012; Vanderstikken et al., 2019).

4.4.2. La relación de la edad con los constructos analizados

Una variable sociodemográfica que está adquiriendo gran relevancia en los últimos años es la edad, debido a la importancia que tiene el envejecimiento progresivo de la población en la gestión de recursos humanos (Segura-Camacho et al., 2018). El planteamiento que goza de mayor consenso con respecto al impacto de la edad en la personalidad es que ejerce un efecto modesto y que la personalidad, a partir de los 30 años, se “consolida como la escayola”, mostrándose como un constructo estable que presenta cambios, a nivel de medias, modestos e inconsistentes en diferentes estudios (Barrick y Mount, 2014; Costa y McCrae, 1994; Costa et al., 2019; Wagner et al., 2019). Así, Caspi et al. (2005) encontraron que entre los 18 y los 30 años se reducían las puntuaciones en Neuroticismo, Extroversión y Apertura, mientras que aumentaban los niveles de Amabilidad y Conciencia, y que a partir de los 30 años estas tendencias

proseguían, aunque en una proporción menor. Estos autores acuñaron el *principio de madurez*, según el cual con los años las personas se vuelven más dominantes, amables, responsables y emocionalmente estables. Damian et al. (2018) han encontrado resultados que apoyan esta tendencia hacia la madurez en su estudio longitudinal, que analizó la evolución de la personalidad en una muestra amplia en un intervalo de 50 años de vida, desde los 16 a los 67 años.

Frente a la postura de Caspi et al. (2005), otros autores apuntaron que a partir de los 50 años se daban cambios notables en la personalidad (Ardelt, 2000; Donnellan y Lucas, 2008; Roberts et al., 2006; Specht et al., 2011; y Terracciano et al., 2005). El meta-análisis de Roberts et al. (2006), basado en estudios longitudinales realizados con datos de personas entre 10 y 70 años, encontró que las puntuaciones subían levemente, de manera no significativa hasta los 50 años, y que subían drásticamente entre los 50 y los 60 años para posteriormente estabilizarse. El Neuroticismo bajaba hasta los 40 años para después estabilizarse. La Responsabilidad mostraba un aumento gradual a lo largo de todo el rango de edad y la Apertura a la experiencia mostraba una tendencia curvilínea, ascendente hasta los 20 y descendente a partir de los 50. La Extroversión vinculada a la vitalidad social también descendía a partir de los 50. Ardel (2000) o Specht et al. (2011) coinciden al señalar que los factores de personalidad Estabilidad Emocional, Extroversión, Apertura a la experiencia y Amabilidad registran una relación de U invertida a lo largo de la vida, siendo la edad de 50 años el punto de inflexión a partir del cual se dan puntuaciones más bajas. Soto et al. (2011), en su investigación transversal, encontraron puntuaciones significativamente más altas en Responsabilidad y Amabilidad en las personas de más edad.

Partiendo de estos trabajos que indicaban variaciones en la personalidad al final de la mediana edad, Marsh et al. (2013), empleando técnicas basadas en ESEM longitudinal, hallaron lo que denominaron *efecto la dulce vida*. Según este efecto, las personas de avanzada edad se muestran más felices (más agradables y menos neuróticas), más centradas en sí mismas (menos extrovertidas y abiertas), más relajadas y satisfechas con ellas mismas y con su vida (menos responsables, extrovertidas y abiertas), así como menos preocupadas por la productividad.

Por otro lado, los estudios de Kandler et al. (2015), Möttus et al. (2012), o Wortman et al. (2012), realizados con muestras de personas de más de 70 años, han

encontrado cambios en diferente sentido, registrando puntuaciones más altas en Neuroticismo y más bajas en Amabilidad, Extroversión o Responsabilidad.

La cuestión del cambio en la personalidad a lo largo de la vida ha cobrado especial relevancia en los últimos años y ha dado lugar al modelo DATA del desarrollo de la personalidad en el trabajo de Woods et al. (2018). Este modelo plantea que las personas adaptan sus rasgos de personalidad para alcanzar un mayor nivel de ajuste a su entorno laboral, lo que implica posibles cambios en factores como Responsabilidad o Amabilidad para conseguir promociones en la organización u objetivos prosociales. No obstante, el cambio intencionado de la personalidad ha presentado tamaños del efecto pequeños (Hudson et al., 2020)

En cuanto al resto de constructos, O'Connell et al. (2008) no encontraron niveles diferentes de adaptabilidad en función de la edad. Sobre este aspecto DeArmond et al. (2006) señalaron que existe la percepción de que los trabajadores mayores son menos adaptables, si bien cabría pensar que su experiencia les puede ayudar en algunas dimensiones del DA. El meta-análisis de Ng y Feldman (2012) mostró que las personas de más edad realizan menos acciones formativas, aspecto vinculado a la dimensión de DA Esfuerzo Formativo. Por lo que se refiere a las OCBs, Akinbode (2011) observó que la edad no era un predictor válido de las puntuaciones en OCB, pero Chou y Pearson (2011) encontraron resultados opuestos.

4.4.3. La relación de la experiencia laboral con los constructos analizados

Una variable vinculada con la edad es la experiencia laboral, entendida como el número de años en un mismo puesto de trabajo. Su relación con el desempeño laboral es de tamaño pequeño: según el meta-análisis de Schmidt et al. (2016) es de .16 y según el de Van Iddekinge et al. (2019) es de .06. Ng y Feldman (2013), por su parte, en su meta-análisis encontraron correlaciones prácticamente nulas entre experiencia en el puesto y Creatividad o las OCBs. Woods et al. (2018), en el marco de su modelo DATA, señalan que es necesario investigar si las personas con poca experiencia en la organización adaptan sus perfiles de personalidad para aumentar su nivel de ajuste a la organización. Por consiguiente, resulta interesante comparar las puntuaciones medias en los constructos analizados entre personas con poca experiencia y personas con un elevado nivel de experiencia.

Con respecto al resto de constructos, existen escasas investigaciones. El estudio de O'Connell et al. (2008) halló una correlación negativa entre años de experiencia laboral y DA. Su et al., en su revisión de 2015, plantearon como hipótesis que la experiencia laboral debería estar relacionada con el APO, cuestión que ellos no contrastaron empíricamente pero que confirmaron Vleugels et al. (2019).

Diversas investigaciones han prestado atención a la relación de esta variable con las OCBs. Así, Akinbode (2011) encontró que la experiencia laboral no es un predictor válido de las puntuaciones en OCB. Sin embargo, Chou y Pearson (2011) observaron que la experiencia, entendida como años de antigüedad en el puesto, sí es un predictor significativo de las OCBs.

4.4.4. La relación del desempeño de puestos de mando con los constructos analizados

Otra materia de interés es la relación entre personalidad y desempeño de puestos de mando, aspecto vinculado al estudio de las características del liderazgo. El meta-análisis más importante sobre esta cuestión indica que la personalidad explica más del 50% de la variabilidad tanto de la emergencia del liderazgo, es decir, el grado en que alguien es percibido como posible líder, como de la efectividad del liderazgo, la capacidad de un/a líder para influir y orientar a su unidad hacia el logro de sus objetivos (Judge et al., 2002).

La literatura señala que la Extroversión es el factor más relacionado con el ejercicio del liderazgo. Las personas extrovertidas son enérgicas, dominantes, sociables y tienden a aspirar a puestos de líder, siendo mayor su probabilidad de éxito como líderes. En este sentido, Seibert y Kraimer (2001), hallaron que la Extroversión estaba relacionada con el nivel salarial y con las promociones. De los cinco factores del FFM, este es el factor que presenta la correlación más alta con emergencia del liderazgo (junto a Responsabilidad), así como con efectividad en el ejercicio del liderazgo (junto a Apertura a la experiencia), si bien su estatus como principal predictor del liderazgo no ha sido confirmado en el sector público (Bono et al., 2014; Judge et al., 2002; t'Hart, 2014).

Por lo que se refiere a la Amabilidad, características de personas con puntuaciones altas en este factor, como el altruismo, la cooperación o la tendencia a evitar conflictos, en principio adaptativas, pueden suponer dificultades para tomar decisiones complicadas o para decidir entre lo que es positivo para las personas y lo que es beneficioso para la organización. De hecho, se ha encontrado una correlación negativa entre este factor y emergencia del liderazgo, así como con efectividad en el liderazgo, esta última de tamaño muy reducido (Bono et al, 2014; Judge et al., 2002; Judge et al., 2009; Smith y Smith, 2005).

Un factor de personalidad que ha recibido menos atención en cuanto a su vinculación con el liderazgo es Apertura a la experiencia, pese a su relación con la creatividad, la solución de problemas, la capacidad de análisis o el pensamiento divergente o el hecho de que autores como Niess y Zacher (2015) hayan subrayado la importancia de la Apertura en el ascenso hacia puestos de jefatura. No obstante, el meta-análisis de Judge et al. (2002) no encontró una relación positiva significativa de este factor ni con emergencia ni con efectividad del liderazgo en el sector gubernamental (Bono et al., 2014).

Judge et al. (1999), por su parte, encontraron que el éxito en el estatus ocupacional estaba vinculado a Estabilidad Emocional y a Responsabilidad. Sin embargo, Judge et al. (2002) no observaron una relación positiva entre Estabilidad y emergencia o efectividad del liderazgo. Por lo que se refiere a Responsabilidad, estos autores encontraron una relación positiva en el sector público entre este factor y el liderazgo, algo que no se ha trasladado al sector privado (Bono et al., 2014).

El meta-análisis de Huang et al. (2014) mostró que el nivel del puesto (con o sin funciones de mando) ejercía influencia sobre la relación entre personalidad y DA, siendo las personas con puestos de responsabilidad quienes tienen más oportunidades para desarrollar conductas adaptativas, lo que coincide con lo apuntado por Berg et al. (2010). Crowley (2012) también indicó que el nivel de exigencia de un puesto en materia de desempeño adaptativo podría ser un moderador de la relación entre el factor Conciencia de personalidad y el DA, y propuso, al igual que otros autores (Ployhart & Bliese, 2006; Pulakos et al., 2000; Pulakos et al., 2002), investigar la efectividad de los predictores del DA en diferentes tipos de puestos.

El vínculo entre liderazgo y OCB ha sido más estudiado, y como señalan Piccolo et al. (2018), el liderazgo se puede medir mediante las conductas extra-rol (OCBs) de los subordinados, es decir, mediante su deseo de ir más allá de los requerimientos del puesto. Incluso los propios líderes van más allá de las tareas instrumentales de supervisión, es decir, del mero ejercicio de un puesto de mando, comprometiéndose en acciones, emociones y actividades que mejoran las relaciones y que sobrepasan los requerimientos de su puesto (Babcock-Roberson & Strickland, 2010; Podsakoff et al., 2000). Un estudio sobre esta cuestión es el de Ye et al. (2012), que hallaron que jefas y jefes exhibían más OCBs que el resto de trabajadores/as. En consecuencia, resulta de interés comparar las puntuaciones medias en OCBs de las personas que desempeñan puestos de mando y las de sus subordinados/as en el sector público.

Asimismo, las posibles diferencias existentes en el APO entre responsables y personas que no ejercen puestos de mando son de interés, dada la escasez de investigaciones al respecto y la importancia concedida a este constructo en el modelo DATA de Woods et al. (2018).

4.4.5. La relación entre nivel educativo, nivel de complejidad del puesto y los constructos analizados

En comparación con el desempeño de puestos de mando, se ha prestado menos atención a la relación entre personalidad y nivel educativo. No obstante, cabe mencionar el trabajo de Van Eijck y De Graaf (2004), que encontraron que Estabilidad Emocional y Apertura a la experiencia tienen relación con el éxito educativo, o el de Goldberg et al. (1998), que observaron que las personas con mayor nivel de estudios presentaban mayores niveles de Apertura a la experiencia.

Por otra parte, en los últimos años el estudio de la complejidad del puesto ha cobrado interés, tanto en lo referido a su relación con la personalidad, como a su papel en la relación entre personalidad y desempeño, con resultados que remarcan la importancia de los factores Conciencia, Estabilidad Emocional y Extroversión en puestos de baja complejidad (Lado y Alonso, 2017). Por otro lado, se ha hallado que la Amabilidad es un predictor de utilidad en puestos de complejidad media y que Apertura a la experiencia es un predictor válido en puestos de complejidad alta si se emplean tests de elección forzada (Salgado, 2017). No obstante, el estudio de Niess y Zacher (2015)

halló un tamaño del efecto pequeño ($d = .21$) al comparar las medias en Apertura a la experiencia de personas en puestos de alta complejidad con las de personas en puestos de menor complejidad.

En cuanto a la relación de estas variables con DA, el estudio de O'Connell et al. (2008), encontró diferentes niveles de adaptabilidad en función del nivel educativo (mayor adaptabilidad para aquellas personas con niveles de estudios más elevados), si bien con diferencias de tamaño reducido. Sung et al. (2017), por su parte, encontraron que la creatividad, concepto vinculado tangencialmente al DA, presentaba una correlación de tamaño medio con complejidad laboral y con nivel educativo, aunque con una muestra menor a la de O'Connell et al. En todo caso, Dorsey et al. (2010) subrayaron la importancia de estudiar la relación de la complejidad de las tareas de un puesto con el DA, especialmente en el caso de puestos que requieran liderazgo.

Por otra parte, el estudio de Pavalalache-Ilie (2014) mostró que las personas con un nivel de educación secundaria presentaban más OCBs que las que tenían un nivel de educación superior. Ng y Feldman (2009), en su meta-análisis, observaron una relación de tamaño pequeño entre el nivel educativo y las OCBs, hallando además que la complejidad del puesto no era un factor moderador en la relación entre ambas variables.

Tal y como sucede con el resto de variables, se ha prestado escasa atención a la relación entre el nivel educativo y el APO. Son una excepción los trabajos de Tsai et al. (2012), o Xu (2014), que no han encontrado relación entre el nivel educativo y el APO. Por todo ello, en la presente investigación se examinará la relación entre esta variable y el APO.

4.4.6. La relación entre la estabilidad laboral y los constructos analizados

Asimismo, son escasas las investigaciones acerca de la relación entre la estabilidad en el puesto y constructos como la personalidad, el APO, el DA o las OCBs. Autores como Guest (1997) o Shaw et al. (2005) han sugerido que la estabilidad puede ser un factor que motive mejoras en el desempeño. Sin embargo, Broschak et al. (2008) o Greene et al. (2018) han observado que su relación con el desempeño puede ser positiva o negativa dependiendo de las circunstancias. Wright (2001), por su parte, examinó investigaciones precedentes realizadas en el sector público concluyendo que,

en comparación con el sector privado, en la Administración Pública existe una relación más débil entre estabilidad laboral y desempeño. No obstante, el meta-análisis de Sverke et al. (2002) puso de manifiesto que la inestabilidad laboral no presenta una correlación significativa con el desempeño laboral. Ye et al. (2012), por su parte, hallaron que la estabilidad laboral no tenía relación con las OCBs, resultado que coincide con lo apuntado por la revisión de De Bree (2010). En la presente investigación se analizará la relación entre esta variable sociodemográfica y las variables de personalidad, el APO, el DA y las OCBs.

4.5. Objetivos e hipótesis

4.5.1. Objetivos

Objetivo 3: Analizar la relación de los cinco grandes factores de personalidad con constructos vinculados con el desempeño laboral (desempeño adaptativo y OCBs) y el APO. Para ello se examinarán las redes de relaciones nomológicas de las adaptaciones tanto del cuestionario de Charbonnier-Voirin y Roussel como de OPERAS.

Objetivo 4: Examinar la posible relación existente entre constructos relevantes en el ámbito de recursos humanos (personalidad, desempeño adaptativo, OCBs y APO) y un conjunto de variables sociodemográficas y laborales.

4.5.2. Hipótesis

4.5.2.1. Relación entre las dimensiones de desempeño adaptativo y las OCBs

Diferentes estudios han hallado relaciones positivas entre el desempeño adaptativo (empleando una medida genérica de dicho constructo que no distingue dimensiones) y las OCBs (Chan y Schmitt, 2002; Charbonnier-Voirin y Roussel, 2012; Segura-Camacho et al., 2018). En base a la proximidad conceptual entre las dimensiones de ambos constructos, cabe formular las siguientes hipótesis:

- H1. Las dimensiones de DA Creatividad y Esfuerzo Formativo tendrán una relación positiva con las OCB-O.

- H2. Cabe pronosticar una relación positiva de la dimensión de DA Adaptabilidad Interpersonal y las OCB-I.

4.5.2.2. Relación entre las dimensiones de desempeño adaptativo y el APO

La investigación de Wessel et al. (2008) mostró una relación positiva del APO con la adaptabilidad al aprendizaje y a la incertidumbre. En base a dicho estudio, es posible plantear la siguiente hipótesis:

- H3. Se espera encontrar una relación positiva entre las dimensiones de DA Esfuerzo Formativo y Reactividad ante emergencias y el APO.

4.5.2.3. Relación entre factores de personalidad y dimensiones de desempeño adaptativo

El estudio de Charbonnier-Voirin (2013) halló una relación positiva entre factores de personalidad y dimensiones de DA. Las magnitudes de relación más elevadas correspondieron a Apertura a la experiencia y Creatividad, Responsabilidad y Esfuerzo Formativo, Amabilidad y Adaptabilidad interpersonal y Estabilidad emocional y Gestión del estrés. Todo ello nos lleva a plantear las siguientes hipótesis:

- H4. El factor de personalidad Apertura a la experiencia tendrá una relación positiva con la dimensión de DA Creatividad.

- H5. Cabe pronosticar una relación positiva del factor de personalidad Responsabilidad con la dimensión de DA Esfuerzo Formativo.

- H6. Se espera encontrar una relación positiva entre Amabilidad y Adaptabilidad interpersonal.

- H7. Cabe esperar que Estabilidad Emocional presente una relación positiva con Gestión del estrés.

4.5.2.4. Relación entre factores de personalidad y dimensiones de las conductas cívico-organizacionales (OCBs)

Dado que en investigaciones previas se ha encontrado que Responsabilidad y Amabilidad se relacionan de forma positiva con dos dimensiones de las OCBs, conductas dirigidas a la organización (OCB-O) y conductas de tipo interpersonal (OCB-

I), respectivamente (Chiaburu et al., 2011; Cooper et al., 2013; Ilies et al., 2009), cabe plantear las siguientes hipótesis:

-H8. La Responsabilidad presentará una relación positiva con la dimensión OCB-O.

-H9. La Amabilidad presentará una relación positiva con la dimensión OCB-I.

4.5.2.5. Relación entre factores de personalidad y el ajuste persona-organización (APO)

En base a los resultados de trabajos anteriores en los que se encontraron relaciones positivas entre las dimensiones del modelo de los cinco grandes factores de personalidad Responsabilidad y Extroversión y la puntuación en APO (Gardner et al., 2012; Gillespie et al., 2003):

-H10. Se espera encontrar una correlación positiva de los factores de personalidad Responsabilidad y Extroversión con el APO.

4.5.2.6. Relación entre la personalidad y el desempeño: el rol del APO y del desempeño adaptativo

El meta-análisis de Hoffman y Woehr (2006) halló una correlación significativa entre el APO y las OCBs. Por otro lado, en las escasas investigaciones que se han centrado en la capacidad predictiva conjunta del APO y la personalidad con las OCBs como criterio, Tsai et al. (2012) y Michaud (2014) han señalado que la personalidad y el APO conjuntamente pueden contribuir a explicar las OCBs. El DA, a su vez, puede añadir capacidad predictiva a la personalidad y el APO en la predicción de las OCBs, según el modelo I-ADAPT de Ployart y Bliese (2006). No se plantearán hipótesis en lo relativo a la relación de la personalidad y el APO con DA por carecer de base suficiente en la literatura, aunque se prestará atención a dicha cuestión en los resultados. En consecuencia, cabe plantear que:

-H11. Se espera una correlación positiva entre el APO y los dos tipos de OCBs.

-H12. El APO y la personalidad conjuntamente tendrán capacidad predictiva para explicar los dos tipos de OCBs.

-H13. El DA añadirá varianza a la explicada por la personalidad y el APO para explicar ambos tipos de OCBs.

4.5.2.7. Relación entre el género y los constructos analizados

En base a investigaciones precedentes (Kajonius y Johnson, 2018; Kajonius y McGiolla, 2017; Ng et al., 2016; O'Connell et al., 2008; Seong et al., 2012; Vanderstukken et al. 2019), cabe plantear las siguientes hipótesis:

-H14. Cabe esperar que, en comparación con los varones, las mujeres presenten puntuaciones más altas en Amabilidad y más bajas en Estabilidad Emocional.

-H15. Cabe esperar que el género no presente relación con el desempeño adaptativo.

-H16. Se espera que el género no muestre relación con las OCBs.

-H17. Se espera que el género no muestre relación con el APO.

4.5.2.8. Relación entre la edad y los constructos analizados

Dado que solo confluyen algunos resultados (Costa et al., 2019; Wagner et al., 2019) en lo que respecta a la relación de esta variable con la personalidad, solo se establecerá una hipótesis para la edad. Asimismo, se examinarán de forma exploratoria las relaciones de la edad con el resto de constructos.

-H18. No se esperan diferencias en los factores de personalidad en función de la edad.

4.5.2.9. Relación entre la experiencia laboral y los constructos analizados

Dada la escasez de investigaciones al respecto (Woods et al., 2018), solo se planteará una hipótesis acerca de la relación entre experiencia laboral y OCBs y DA (Ng y Feldman, 2013), si bien se compararán las puntuaciones medias de personas con poca experiencia en la organización para la que trabajan y en el puesto que desempeñan con las de personas con amplia experiencia, para el conjunto de los constructos objeto de interés de esta investigación.

-H19. No se esperan diferencias en OCBs o DA en función de la experiencia laboral.

4.5.2.10. Relación entre el ejercicio de puestos de mando y los constructos analizados

En base a investigaciones anteriores (Judge et al., 2002; Ye et al., 2012), cabe plantear las siguientes hipótesis:

-H20. Cabe esperar que las personas que desempeñan puestos de mando presenten puntuaciones más altas en Extroversión que las personas en puestos sin ejercicio de mando.

-H21. Cabe esperar que las personas que desempeñan puestos de mando presenten puntuaciones más altas en OCBs que las personas en puestos sin ejercicio de mando.

4.5.2.11. Relación entre el nivel educativo y la complejidad del puesto y los constructos analizados

Ante la carencia de estudios previos acerca de la relación entre las dimensiones de DA y el nivel de complejidad del puesto, se explorarán las relaciones entre estas variables. Por otro lado, estudios precedentes (Goldberg et al., 1998; Niess y Zacher, 2015; O'Connell et al., 2008) nos llevan a plantear las siguientes hipótesis:

-H22. No se esperan diferencias en Apertura a la experiencia en función de la complejidad del puesto.

-H23. Cabe esperar puntuaciones más altas en Apertura a la experiencia para personas con un nivel educativo superior.

-H24. No se esperan diferencias en las puntuaciones en DA en función del nivel educativo.

4.5.2.12. Relación entre la estabilidad laboral y los constructos analizados

Teniendo en cuenta el meta-análisis de Sverke et al. (2002), únicamente se planteará una hipótesis referida a la relación entre la estabilidad en el puesto de trabajo y el desempeño laboral:

-H25. No cabe esperar diferencias en las puntuaciones en desempeño laboral en función de la estabilidad en el puesto.

4.6. Metodología

4.6.1. Participantes

La muestra estuvo compuesta por los 678 empleados/as públicos/as (444 mujeres y 234 varones) cuyas características han sido descritas en el Capítulo 2.

La edad media fue de 44.63 años ($SD = 7.66$), con un rango que osciló entre 21 y 63 años. Para estudiar la relación de la edad con las variables analizadas, se distinguieron 3 grupos de edad cuyas puntuaciones medias se compararon: 207 personas estaban entre 21 y 40 años, 304 entre 41 y 50 años, y 166 entre 51 y 63 años.

La media de tiempo en la organización fue de 11.46 años ($SD = 9.71$) y la media de tiempo en el puesto fue de 7.71 años ($SD = 7.82$). Para estudiar la relación de la experiencia en la organización y en el puesto con las variables analizadas se distinguieron 2 grupos en ambos casos: personas con menos de dos años de experiencia en la organización y en el puesto, 129 y 179, respectivamente, y personas con dos o más años de experiencia en la organización y en el puesto, 549 y 479, respectivamente. El punto de corte se estableció en dos años ya que la legislación en materia de función pública determina que se necesitan dos años para poder participar en concursos que permiten progresar en la carrera profesional (Congreso de los Diputados, 2015).

Un total de 213 personas ocupaban puestos que incluían funciones de mando sobre otros/as empleados/as, frente a 456 que no desempeñaban ese tipo de funciones de mando.

Los empleados y las empleadas participantes desempeñaban diferentes tipos de puestos que requerían diversas titulaciones de acceso: 238 desempeñaban puestos que requerían Licenciatura o Grado (Grupo A); 146 tenían puestos que requerían Diplomatura (Grupo B); 203 desempeñaban puestos que requerían Educación Secundaria (Grupo C); 62 prestaban servicio en puestos que exigían tener acreditados

estudios de Educación Primaria (Grupo D); y 17 desempeñaban puestos que requerían Certificado de Escolaridad (Grupo E). En base al tipo de titulación requerida para el acceso, cabe distinguir entre puestos de alta complejidad, aquellos que requerían estudios universitarios ($N = 384$), los pertenecientes a los Grupos A y B, frente a puestos de baja complejidad ($N = 294$), los correspondientes a los Grupos C, D y E.

En cuanto a la titulación académica que poseían, la mayoría estaba en posesión de una Licenciatura ($N = 420$), seguidos de 120 diplomados y diplomadas, 113 personas con Educación Secundaria, 20 con Educación Primaria y 5 con Certificado de Escolaridad. En consecuencia, había 540 personas con estudios universitarios (nivel de educación alto) frente a 138 que no los poseían (nivel de educación bajo).

En lo que se refiere a la estabilidad laboral de los y las participantes, 245 personas tenían plaza fija (bien como funcionarios/as de carrera o personal laboral fijo), mientras que 426 desempeñaban sus puestos como personal interino o temporal.

4.6.2. Instrumentos

Para medir el desempeño adaptativo se empleó la versión en castellano del cuestionario originalmente desarrollado por Charbonnier y Roussel (2012), cuyo proceso de adaptación al castellano se ha descrito en el Capítulo 2. Este instrumento está compuesto por 22 ítems que miden 5 dimensiones (Creatividad, Gestión del estrés, Adaptabilidad interpersonal, Reactividad ante emergencias, y Esfuerzo formativo) con formato de respuesta de escala tipo Likert de 7 puntos, que va de 1 (*muy en desacuerdo*) a 7 (*muy de acuerdo*). Como se ha señalado en el Capítulo 2, la mediana de los coeficientes de fiabilidad de las dimensiones de la adaptación, medida mediante la Rho de Jöreskog, Alfa de Cronbach y Omega fue, respectivamente de .77, .73 y .75.

Los cinco grandes factores de personalidad (Amabilidad, Apertura a la experiencia, Estabilidad emocional, Extroversión, y Responsabilidad) se midieron mediante la versión en euskara del cuestionario OPERAS (Vigil-Colet et al., 2013) ya descrita en el capítulo 3. Dicho cuestionario consta de 40 ítems con cinco alternativas de respuesta (de 1, “*muy en desacuerdo*”, a 5, “*muy de acuerdo*”) y se puede aplicar en múltiples ámbitos: en clínica, recursos humanos, educación, investigación, etc. Una de sus principales características es que permite eliminar los efectos de la deseabilidad

social y la aquiescencia, aumentando el nivel de precisión de la medida de los factores de interés. Tal y como se ha señalado en el capítulo 3, la consistencia interna de los factores fue evaluada mediante el estadístico ORION (Ferrando y Lorenzo-Seva, 2016), Alfa de Cronbach y Omega. Las medianas respectivas de dichos estadísticos en la adaptación al euskera fueron de .75, .70 y .68, respectivamente.

Las OCBs fueron evaluadas mediante la escala de Lee y Allen (2002), en su versión en castellano adaptada por Dávila y Finkelstein (2010). La escala está compuesta por 16 ítems, con formato de respuesta de escala tipo Likert de 5 puntos, que va de 1 (*nunca*) a 5 (*siempre*). Este instrumento permite evaluar dos dimensiones: comportamientos dirigidos a la organización (OCB-O), y comportamientos de tipo interpersonal (OCB-I). Su fiabilidad en esta investigación presentó, respectivamente, valores de alfa de Cronbach de .80 y .77, mientras que en la adaptación original al castellano fueron de .86 y .89. Los valores obtenidos en el estadístico Omega fueron de .80 y .78, respectivamente.

El APO se midió mediante la adaptación al castellano realizada por Cáceres (2014) del cuestionario de 5 ítems desarrollado por Piasentin (2007) con formato de respuesta de escala tipo Likert de 5 puntos, que va de 1 (*muy en desacuerdo*) a 5 (*muy de acuerdo*). La consistencia interna de la escala original en castellano fue de .66 (alfa de Cronbach), y en esta investigación presentó un valor de .69, con un valor de Omega de .68.

Para evaluar variables sociodemográficas y laborales se utilizó un cuestionario elaborado ad hoc para la investigación. Dicho cuestionario incluyó preguntas acerca de variables que, a tenor de la revisión de la literatura sobre el modelo de los Cinco Grandes y los constructos de desempeño, están relacionadas con las distintas dimensiones que miden las pruebas empleadas (Bernerth y Aguinis, 2016). El cuestionario incluía 11 preguntas con diferentes formatos de respuesta: género, edad, tiempo de permanencia en el lugar de trabajo, tiempo de permanencia en el puesto de trabajo actual, tipo de relación laboral (fijo o temporal), nivel educativo, titulación exigida por el puesto, nivel de conocimiento de euskera, lengua materna, idioma de estudios, y una pregunta sobre la exigencia del puesto de trabajo desempeñado de supervisión de otros puestos.

4.6.3. Procedimiento

Se contó con el informe favorable del Comité de Ética para las Investigaciones relacionadas con Seres Humanos (CEISH/GIEB) de la UPV/EHU, por lo que previamente se recogieron los consentimientos informados de las personas participantes. La aplicación de los cuestionarios fue grupal y se realizó en los respectivos centros de trabajo de las personas que participaron en el estudio. El procedimiento para garantizar el anonimato fue el descrito en el Capítulo 2.

Los instrumentos se cumplimentaron según el siguiente orden: Cuestionario de variables sociodemográficas y laborales, cuestionario de personalidad OPERAS (2013), Escala de Lee y Allen para evaluar OCBs (2002), Escala de Piasentin (2007) para medir APO, y Escala de Charbonnier-Voirin y Roussel (2012) para evaluar DA.

Las escalas utilizadas fueron cumplimentadas por el mismo informante (los empleados/as públicos/as que participaron en la investigación) en la misma sesión, lo que puede generar una inflación artificial de los resultados debido a la varianza común (Common method variance, CMV, en inglés). Para reducir este sesgo potencial, se siguieron las recomendaciones de tipo estadístico y procedimental de Podsakoff et al. (2003 y 2012), que consisten en separar psicológicamente las variables en el cuestionario camuflando la relevancia de las variables criterio y su vinculación a las variables predictoras, el uso de escalas de respuesta con características diferentes, la utilización de tests breves y la garantía de anonimato de las personas participantes.

4.6.4. Análisis de datos

Se realizaron los siguientes análisis estadísticos:

a) Modelos de ecuaciones estructurales para testar la varianza debida al método común (CMV). A fin de analizar el posible efecto de la varianza común se añadió una variable latente que representaba el método común al modelo inicial con 13 variables (5 variables de personalidad, 2 de OCBs, 1 de APO, y 5 de DA) para medir la mejora potencial en el ajuste del modelo relacionada con la adición de este factor de método.

De manera complementaria se realizó el test *Harman's single factor*, el procedimiento más utilizado para estudiar la presencia de CMV y que indica presencia de sesgos en caso de que emerja un factor general que explique más del 50% de la varianza de las variables empleadas (Podsakoff y Organ, 1986) o incluso del 70% siempre que se utilicen variables con resultados típicos en fiabilidad (Fuller et al., 2016).

b) Estadísticos descriptivos (media, desviación típica, asimetría y curtosis), así como fiabilidad de los constructos analizados (alfa de Cronbach).

c) Correlaciones entre las distintas variables. Se utilizaron correlaciones de Pearson para medir la asociación entre los constructos de interés.

d) Modelos de regresión múltiple jerárquica para estudiar la varianza añadida por APO a la personalidad al predecir las puntuaciones en desempeño adaptativo. Asimismo, se examinó si el desempeño adaptativo contribuía una cantidad significativa de la varianza de las OCBs más allá de la explicada conjuntamente por la personalidad y el APO.

De manera previa, se analizaron los supuestos que requiere la regresión: linealidad, normalidad en la distribución de los errores, homocedasticidad, independencia de errores y ausencia de multicolinealidad.

e) Análisis de las diferencias en las puntuaciones medias de los distintos constructos en base a variables sociodemográficas y laborales. Se calculó la *d* de Cohen para las variables con dos niveles y la *g* de Hedges para la variable edad, que presentaba tres niveles (personas entre 21 y 40 años, personas entre 41 y 50 años y personas entre 51 y 65 años).

Dado que debido al tamaño muestral la probabilidad de obtener resultados estadísticamente significativos era muy alta aun cuando los tamaños del efecto fuesen pequeños, en el contraste de hipótesis se tuvo en cuenta como criterio preferente el tamaño del efecto de los resultados hallados. A tal fin se tomaron como punto de corte las sugerencias de Ferguson (2009) con respecto a los tamaños del efecto mínimos recomendados que representan un efecto a tener en cuenta a nivel práctico en las ciencias sociales. En lo que se refiere a la intensidad de las correlaciones, el valor

mínimo recomendado es de .20. En cuanto a los índices ΔR^2 , el valor mínimo recomendado es de .04, mientras que en lo referido a las diferencias entre grupos (d de Cohen, g de Hedges), el valor mínimo es de .40.

4.7. Resultados

4.7.1. Varianza del método común

El modelo en el que se añadió una variable latente representativa del método común presentó un ajuste (χ^2 [49] = 209.02, CFI = .95, TLI = .91, RMSEA = .07) aceptable y ligeramente superior al del modelo que no incluyó el factor de método (χ^2 [61] = 394.23, CFI = .90, TLI = .84, RMSEA = .09), lo que indicaría presencia de CMV (Richardson et al., 2009). No obstante, añadir el factor de método supone un aumento del 10% de la varianza total, que es inferior al porcentaje generalmente relacionado con cuestionarios auto-administrados relativos a las percepciones sobre el contexto laboral (25%). Por tanto, cabría inferir que el CMV presente en los datos no es suficiente para sesgar los resultados (Williams et al., 1989). Asimismo, las correlaciones entre las variables examinadas son básicamente las mismas ($\Delta \leq .08$), tanto cuando el factor de método se incluye como cuando no se incluye en el modelo de prueba.

Por otro lado, para medir el *Harman's single factor* se llevó a cabo un análisis de componentes principales cuyo resultado puso de manifiesto que no había un factor dominante que explicara la mayoría de la varianza, ya que el primer factor explicaba únicamente el 35.73% de la varianza, inferior al 50% que sería indicativo de CMV.

4.7.2. Estadísticos descriptivos y fiabilidad de los constructos analizados

En la tabla 4.1 se presentan las medias y desviaciones típicas de los constructos analizados. En cuanto a la fiabilidad, se ha optado por el estadístico más empleado en las investigaciones originales de los instrumentos utilizados, Alfa de Cronbach, lo que permite comparar los resultados con los obtenidos en dichos estudios. La fiabilidad de las medidas empleadas fue adecuada (Nunnally, 1978; Smith y Smith, 2005), salvo en lo que respecta a Adaptabilidad interpersonal, que sin embargo presentó un nivel de fiabilidad aceptable en el resto de estadísticos de fiabilidad utilizados en el Capítulo 2 (Omega, Fiabilidad marginal, Rho de Jöreskog).

Algunas de las dimensiones que miden DA presentaron problemas de asimetría y curtosis, si bien todos los constructos, con la excepción de Adaptabilidad Interpersonal, mostraron valores inferiores a los límites que indican problemas de normalidad, 2 para asimetría y 7 para curtosis (Curran et al., 1996).

Tabla 4.1.

Estadísticos Descriptivos

	N	Fiabilidad	Media	Desviación típica	Asimetría		Curtosis	
					Estadístico	Error estándar	Estadístico	Error estándar
Amabilidad	674	.69	27.55	3.36	-.60	.09	1.08	.19
Apertura	674	.62	28.42	3.64	-.43	.09	.07	.19
Estabilidad	674	.72	27.20	3.78	-.62	.09	.53	.19
Extroversión	674	.76	21.83	4.25	-.09	.09	-.03	.19
Responsabilidad	674	.62	28.04	3.39	-.49	.09	.63	.19
CRE	678	.72	21.06	3.84	-1.29	.09	4.22	.19
GE	678	.74	29.89	5.60	-.79	.09	2.71	.19
AI	678	.37	23.52	2.95	-2.62	.09	17.23	.19
REA	678	.78	21.91	3.45	-1.42	.09	6.37	.19
EF	678	.81	22.45	3.91	-1.54	.09	4.97	.19
OCB-O	655	.80	29.90	4.05	-.07	.10	.04	.19
OCB-I	655	.77	28.60	4.22	-.05	.10	.68	.19
APO	655	.69	19.03	3.04	-.79	.10	2.60	.19

Nota. CRE = Creatividad; GE = Gestión del estrés; AI = Adaptabilidad Interpersonal; REA = Reactividad ante emergencias; EF = Esfuerzo Formativo; OCB-O = OCBs de tipo organizativo; OCB-I = OCBs de tipo interpersonal; APO = Ajuste persona-organización.

4.7.3. Correlaciones entre los constructos analizados

En lo que se refiere a las correlaciones (Tabla 4.2) entre las dimensiones que miden DA y las que miden OCBs, fueron de tamaño medio-alto, con medianas de .42 y .40 con OCB-O y OCB-I respectivamente. Creatividad presentó la correlación más elevada con OCB-O, .54, mientras que Creatividad, Adaptabilidad interpersonal y Reactividad ante emergencias presentaron las magnitudes más elevadas con OCB-I, próximas a .40.

Por otro lado, la mediana de las correlaciones de APO con las dimensiones de DA fue de .21, tamaño entre reducido y medio. La correlación más alta fue la que presentó con la dimensión de DA Adaptabilidad interpersonal, .28.

En cuanto a la relación entre los cinco factores de personalidad y el resto de constructos, las medianas de las correlaciones con DA, OCBs y APO fueron, respectivamente, de .25, .25 y .21. Específicamente en lo referido a la relación entre personalidad y DA, Apertura a la experiencia y Responsabilidad presentaron los valores más elevados con Creatividad, .25, Estabilidad Emocional con Gestión del Estrés, .47, Amabilidad con Adaptabilidad interpersonal, .34, Estabilidad Emocional y Responsabilidad con Reactividad ante emergencias, .32 y Responsabilidad con Esfuerzo Formativo, .28. Los factores de personalidad que presentaron correlaciones más elevadas con las OCBs fueron Apertura a la experiencia, Extroversión y Responsabilidad, con valores entre .25 y .27. Por otro lado, Amabilidad y Estabilidad Emocional presentaron las magnitudes más elevadas con APO, .25. Desde el punto de vista del tamaño del efecto de las correlaciones, estas fueron de tamaño entre reducido y medio según Cohen (1988), y de tamaño medio según los estándares de Bosco et al. (2015), Funder y Ozer (2019), Gignac y Szodorai (2016)². La mayoría de las correlaciones registradas superaron el umbral de .20 establecido por Ferguson (2009).

Deseabilidad social, por su parte, presentó correlaciones estadísticamente significativas con todos los constructos a excepción de Extroversión. No obstante, el tamaño de las correlaciones fue pequeño en todos los casos, siendo las más altas, de tamaño reducido-medio, las que presentó con Amabilidad, Responsabilidad y Esfuerzo Formativo, entre .21 y .28.

Por otro lado, la correlación de APO con los dos tipos de OCBs fue similar, .25 con OCB-O y .26 con OCB-I. En ambos casos el tamaño de las correlaciones fue medio.

² Para Cohen (1988) una correlación de tamaño reducido es de .1, una de tamaño medio es de .3 y una grande es de .5. Según Bosco et al. (2015), una correlación de tamaño medio es la que se encuentra entre .09 y .26. Para Funder y Ozer (2019) y Gignac y Szodorai (2016) una correlación es pequeña si es de .10, media si es de .20 y grande si es de .30.

Tabla 4.2.

Correlaciones entre los Constructos Estudiados

	AMA	APE	EST	EXT	RES	DS	APO	CRE	GE	AI	REA	EF	OCB-O	OCB-I
AMA	1													
APE	.24**	1												
EST	.40**	.18**	1											
EXT	.14**	.20**	.22**	1										
RES	.35**	.21**	.37**	.11**	1									
DS	.28**	.17**	.19**	.07	.26**	1								
APO	.25**	.12**	.25**	.05	.21**	.16**	1							
CRE	.17**	.25**	.18**	.24**	.25**	.12**	.20**	1						
GE	.32**	.18**	.47**	.25**	.26**	.14**	.26**	.52**	1					
AI	.34**	.27**	.25**	.22**	.24**	.16**	.28**	.59**	.56**	1				
REA	.23**	.22**	.32**	.27**	.32**	.14**	.21**	.59**	.61**	.63**	1			
EF	.21**	.24**	.17**	.20**	.28**	.21**	.20**	.62**	.46**	.66**	.50**	1		
OCB-O	.24**	.27**	.20**	.25**	.25**	.16**	.25**	.54**	.38**	.42**	.45**	.40**	1	
OCB-I	.24**	.26**	.13**	.25**	.25**	.16**	.22**	.41**	.31**	.40**	.40**	.34**	.73**	1

Nota. **. La correlación es significativa en el nivel 0,01 (bilateral). AMA = Amabilidad; APE = Apertura a la experiencia; EST = Estabilidad emocional; EXT = Extroversión; RES = Responsabilidad; DS = Deseabilidad Social; APO = Ajuste persona-organización; CRE = Creatividad; GE = Gestión del estrés; AI = Adaptabilidad Interpersonal; REA = Reactividad ante emergencias; EF = Esfuerzo Formativo; OCB-O = OCBs de tipo organizativo; OCB-I = OCBs de tipo interpersonal.

4.7.4. Modelos de regresiones múltiples jerárquicas

En lo que se refiere a la Regresión Múltiple Jerárquica (Tablas 4.3 y 4.4), se cumplieron los supuestos que requiere la regresión: media del estadístico Durbin-Watson para analizar la independencia de los casos de 1.97; los resultados de los factores de inflación de varianza (*Variance-inflation factor*, VIF, en inglés) oscilaron entre 1.13 y 2.15, lo que indicaba que la multicolinealidad no era problemática; la inspección visual de los gráficos de dispersión determinó que se cumplían la linealidad, la normalidad en la distribución de los errores y la homocedasticidad.

El ajuste persona-organización y los cinco factores de personalidad explicaron el 19% de la varianza de los comportamientos cívicos organizacionales de tipo organizativo y el 20 % de los de tipo interpersonal. Añadir el DA a la personalidad y el APO supuso que el modelo explicase el 35% de la varianza de las OCB-Os y el 27% de la varianza de las OCB-Is.

APO añadió capacidad predictiva a la personalidad tanto con OCB-O como con OCB-I como criterios. Con OCB-O como variable dependiente, la R^2 ajustada pasó de .15 a .18, lo que supone un 20 % de varianza adicional a la explicada por la personalidad. Con OCB-I, la varianza adicional fue del 26,67%, pasando la R^2 ajustada de .15 a .19.

El DA, por su parte, también añadió varianza a la explicada por la personalidad y el APO con ambos tipos de OCBs como criterios. El incremento en la varianza explicada de las OCB-O fue del 94,44% (la R^2 ajustada pasó de .18 a .35), mientras que, con las OCB-I como criterio, la varianza adicional que supuso añadir DA al modelo fue del 42%. La R^2 ajustada pasó de .19 a .27, superando el umbral de ΔR^2 (.04) establecido por Ferguson (2009).

Tabla 4.3.

Regresión Múltiple Jerárquica con OCB-O como Variable Criterio

Modelo	Variables	R	R^2 ajustado	ΔR^2	Cambio en F
1	Personalidad	.40	.15		24.35*
2	Personalidad APO	.43	.18	.03	20.45*
3	Personalidad APO Desempeño Adaptativo	.60	.35	.17	35.83*

* $p < .05$

Tabla 4.4.

Regresión Múltiple Jerárquica con OCB-I como Variable Criterio

Modelo	Variables	R	R^2 ajustado	ΔR^2	Cambio en F
1	Personalidad	.40	.15		24.20*
2	Personalidad APO	.44	.19	.04	28.60*
3	Personalidad APO Desempeño Adaptativo	.53	.27	.08	16.24*

* $p < .05$

En lo que se refiere a la Regresión múltiple jerárquica con las dimensiones que miden DA como criterio, en términos generales, la R^2 ajustada explicada conjuntamente por la personalidad y el APO fue menor que la registrada con las OCBs como criterio (Tablas 4.5 a 4.9), oscilando entre el 27% de la varianza de Gestión del estrés y el 11% de Adaptabilidad interpersonal.

Lo mismo sucedió con los incrementos asociados a la inclusión de APO en los modelos predictivos. La varianza adicional más elevada explicada por APO fue del 22.22% en el caso de Adaptabilidad interpersonal como criterio. Con el resto de criterios la varianza adicional no superó el 7.69%.

Tabla 4.5.

Regresión Múltiple Jerárquica con Creatividad como Variable Criterio

Modelo	Variables	R	R ² ajustado	ΔR^2	Cambio en F
1	Personalidad	.37	.13		20,15*
2	Personalidad APO	.39	.14	.01	11,48*

* $p < .05$

Tabla 4.6.

Regresión Múltiple Jerárquica con Gestión del Estrés como Variable Criterio

Modelo	Variables	R	R ² ajustado	ΔR^2	Cambio en F
1	Personalidad	.52	.26		47,42*
2	Personalidad APO	.53	.27	.01	10,73*

* $p < .05$

Tabla 4.7.

Regresión Múltiple Jerárquica con Adaptabilidad Interpersonal como Variable Criterio

Modelo	Variables	R	R ² ajustado	ΔR^2	Cambio en F
1	Personalidad	.31	.09		14,13*
2	Personalidad APO	.34	.11	.02	11,73*

* $p < .05$

Tabla 4.8.

Regresión Múltiple Jerárquica con Reactividad ante Emergencias como Variable Criterio

Modelo	Variables	R	R ² ajustado	ΔR^2	Cambio en F
1	Personalidad	.45	.20		32,41*
2	Personalidad APO	.46	.20	.01	7,52*

* $p < .05$

Tabla 4.9.

Regresión Múltiple Jerárquica con Esfuerzo Formativo como Variable Criterio

Modelo	Variables	R	R ² ajustado	ΔR^2	Cambio en F
1	Personalidad	.37	.13		20,02*
2	Personalidad APO	.39	.14	.01	10,28*

* $p < .05$

4.7.5. Relación entre los constructos evaluados y las variables sociodemográficas y laborales

En la Tabla 4.10 se muestran las puntuaciones medias y desviaciones típicas de los y las participantes en esta investigación en las variables sociodemográficas de tipo dicotómico para cada uno de los constructos analizados. Las medias y desviaciones típicas de los 3 grupos de edad en los constructos analizados se muestran en la Tabla 4.11.

En cuanto a las diferencias en personalidad en función de las variables sociodemográficas y laborales, las mujeres presentaron medias más elevadas en Apertura a la experiencia en comparación con los varones, $t(418,48) = -3,24, p = .01, d = .54$, así como en Extroversión, $t(412,62) = -4.44, p = .01, d = .37$, mientras que los varones tuvieron medias más elevadas que las mujeres en Estabilidad emocional, $t(664) = 3,32, p = .01, d = .27$. No se observaron diferencias estadísticamente significativas en el resto de factores.

En lo que se refiere a la relación entre las variables sociodemográficas y el DA, el género se relacionó con Esfuerzo Formativo, con puntuaciones más altas para las mujeres en comparación con los varones, $t(676) = -1.29, p = .01, d = .20$, aunque el tamaño del efecto fue pequeño. En la misma línea, los participantes presentaron diferencias en los OCBs en función del género con puntuaciones más altas para los varones en OCB-O en comparación con las mujeres, $t(645) = 2.11, p = .03, d = .17$, si bien el tamaño del efecto fue inferior al umbral de .40 señalado por Ferguson (2009). Asimismo, no se registraron diferencias estadísticamente significativas en APO en base al género, y en coherencia con ello, los tamaños del efecto fueron pequeños.

Tabla 4.10.

Puntuaciones Medias en función de las Variables Sociodemográficas y Laborales

	Género		Experiencia Organización		Experiencia puesto		Nivel puesto		Nivel educativo		Complejidad puesto		Estabilidad laboral	
	Mujeres	Hombres	Alta	Baja	Alta	Baja	Superior	Subordinado	Alto	Bajo	Alta	Baja	Fijo	Temporal
	(N=437)	(N=229)	(N=549)	(N=125)	(N=480)	(N=198)	(N=213)	(N=456)	(N=540)	(N=138)	(N=384)	(N=294)	(N=245)	(N=426)
AMA	27.69 (3.10)	27.28 (3.77)	27.44 (3.28)	28.01 (3.65)	27.56 (3.29)	27.51 (3.54)	28.07 (2.88)	27.29 (3.54)	27.63 (3.19)	27.21 (3.90)	27.56 (3.12)	27.53 (3.65)	27.32 (3.22)	27.65 (3.43)
APE	28.77 (3.44)	27.79 (3.87)	28.42 (3.67)	28.44 (3.51)	28.42 (3.73)	28.43 (3.41)	28.80 (3.82)	28.27 (3.52)	28.57 (3.64)	27.86 (3.47)	28.69 (3.65)	28.08 (3.60)	28.14 (3.67)	28.60 (3.59)
EST	26.87 (3.72)	27.89 (3.78)	27.04 (3.84)	27.91 (3.41)	27.21 (3.85)	27.18 (3.60)	27.44 (4.00)	27.11 (3.66)	27.32 (3.74)	26.80 (3.86)	27.16 (3.77)	27.26 (3.81)	27.15 (3.69)	27.22 (3.82)
EXT	22.39 (3.95)	20.82 (4.52)	21.77 (4.32)	22.06 (3.92)	21.69 (4.32)	22.16 (4.05)	22.47 (4.32)	21.56 (4.14)	21.978 (4.23)	21.36 (4.11)	22.02 (4.36)	21.58 (4.10)	21.76 (4.40)	21.89 (4.12)
RES	28.21 (3.32)	27.75 (3.47)	27.94 (3.38)	28.49 (3.42)	27.96 (3.40)	28.22 (3.38)	28.43 (3.46)	27.89 (3.31)	28.08 (3.33)	27.93 (3.58)	28.09 (3.47)	27.97 (3.30)	28.17 (3.48)	28.00 (3.31)
DS	14.02 (2.83)	13.58 (3.03)	13.79 (2.91)	14.38 (2.93)	13.80 (2.95)	14.14 (2.82)	13.72 (2.95)	13.93 (2.89)	13.85 (2.94)	13.94 (2.81)	13.70 (2.93)	14.09 (2.87)	13.70 (2.819)	13.96 (2.96)
APO	19.10 (3.08)	18.89 (2.99)	18.95 (3.00)	19.35 (3.17)	19.04 (2.90)	18.98 (3.35)	19.13 (3.04)	18.99 (3.07)	19.02 (3.16)	19.04 (2.57)	19.03 (3.11)	19.03 (2.97)	19.23 (2.72)	18.89 (3.22)
CRE	20.94 (3.67)	21.29 (4.14)	21.14 (3.72)	20.73 (4.32)	21.14 (3.80)	20.87 (3.95)	21.91 (3.72)	20.67 (3.82)	21.31 (3.66)	20.10 (4.35)	21.58 (3.62)	20.38 (4.02)	21.40 (3.45)	20.85 (4.03)
GE	29.59 (5.46)	30.46 (5.82)	29.69 (5.60)	30.74 (5.56)	29.78 (5.70)	30.15 (5.36)	30.29 (5.42)	29.71 (5.69)	29.98 (5.46)	29.54 (6.11)	29.87 (5.47)	29.92 (5.77)	30.35 (5.05)	29.61 (5.88)
AI	23.72 (2.65)	23.15 (3.43)	23.41 (2.93)	23.99 (3.02)	23.43 (2.98)	23.74 (2.88)	23.99 (5.12)	23.43 (2.97)	23.70 (3.90)	23.28 (3.28)	23.59 (4.36)	23.64 (2.86)	23.61 (2.43)	23.59 (4.38)
REA	22.03 (3.29)	21.67 (3.72)	21.85 (3.38)	22.17 (3.74)	21.87 (3.42)	22.00 (3.54)	22.06 (3.47)	21.85 (3.45)	21.89 (3.44)	21.97 (3.49)	21.73 (3.50)	22.14 (3.37)	22.19 (3.07)	21.73 (3.65)
EF	22.72 (3.63)	21.92 (4.35)	22.31 (3.94)	23.02 (3.75)	22.30 (4.05)	22.80 (3.54)	22.68 (4.01)	22.34 (3.90)	22.59 (3.76)	21.86 (4.42)	22.51 (3.98)	22.35 (3.82)	22.08 (3.80)	22.65 (3.93)
OCB-O	29.65 (3.99)	30.35 (4.08)	29.93 (4.10)	29.77 (3.81)	30.10 (4.05)	29.43 (4.01)	30.74 (3.93)	29.50 (4.02)	29.98 (3.92)	29.54 (4.42)	30.09 (4.03)	29.63 (4.02)	30.31 (4.07)	29.64 (4.01)
OCB-I	28.53 (4.11)	28.65 (4.41)	28.64 (4.29)	28.39 (3.91)	28.75 (4.21)	28.21 (4.23)	29.17 (3.96)	28.29 (4.32)	28.70 (4.21)	28.08 (4.21)	28.74 (4.36)	28.36 (4.02)	28.68 (4.31)	28.51 (4.19)

Nota. Las medias señaladas en negrita presentan diferencias estadísticamente significativas en función de los respectivos grupos. Las desviaciones típicas están entre paréntesis. AMA = Amabilidad; APE = Apertura a la experiencia; EST = Estabilidad emocional; EXT = Extroversión; RES = Responsabilidad; DS = Deseabilidad Social; APO = Ajuste persona-organización; CRE = Creatividad; GE = Gestión del estrés; AI = Adaptabilidad Interpersonal; REA = Reactividad ante emergencias; EF = Esfuerzo Formativo; OCB-O = OCBs de tipo organizativo; OCB-I = OCBs de tipo interpersonal.

Por otro lado, se llevó a cabo un ANOVA intersujetos para estudiar la relación de la edad con los 5 factores de personalidad comparando las puntuaciones de 3 grupos de edad: Grupo 1 (21 a 40 años), Grupo 2 (41 a 50 años) y Grupo 3 (51 a 65 años). Hubo diferencias estadísticamente significativas en función de la edad en los factores de personalidad Amabilidad, $F(2,662) = 6.21, p = .01$, y Estabilidad Emocional, $F(2,662) = 3.20, p = .04$ en sentido contrario al hipotetizado. Las comparaciones post-hoc empleando el test de Tukey HSD indicaron que el Grupo 3 presentó una media significativamente más baja ($M=26.78, SD=3.15$) que los grupos 2 ($M=27.73, SD=3.45$) y 1 ($M=27.91, SD=3.27$) en Amabilidad. Del mismo modo, las personas del Grupo 3 presentaron una media inferior en Estabilidad ($M=26.57, SD=3.63$) a las de los grupos 2 ($M=27.36, SD=3.74$) y 1 ($M=27.50, SD=3.87$). Sin embargo, el tamaño del efecto medido mediante el test g de Hedges para Amabilidad y Estabilidad fue respectivamente de .36 y .20, tamaños pequeños en ambos casos.

Tabla 4.11.

Puntuaciones Medias en los Constructos Analizados según los Grupos de Edad

	21-40 (<i>N</i> = 207)	41-50 (<i>N</i> = 304)	51-65 (<i>N</i> = 166)
	27.91	27.72	26.76
AMA	(3.27)	(3.45)	(3.15)
	28.21	28.76	28.14
APE	(3.56)	(3.64)	(3.59)
	27.50	27.36	26.57
EST	(3.87)	(3.74)	(3.63)
	21.75	22.03	21.6
EXT	(3.76)	(4.33)	(4.51)
	28.46	28.01	27.67
RES	(3.21)	(3.49)	(3.32)
	13.83	13.64	14.36
DS	(2.97)	(2.90)	(2.79)
	19.08	18.94	19.09
APO	(2.90)	(3.27)	(2.80)
	20.97	21.08	21.11
CRE	(4.02)	(3.90)	(3.50)
	29.83	30.07	29.61
GE	(5.91)	(5.57)	(5.27)
	23.59	23.81	23.23
AI	(2.96)	(4.74)	(2.47)
	21.97	21.95	21.73
REA	(3.26)	(3.67)	(3.27)
	22.80	22.39	22.08
EF	(3.72)	(4.02)	(3.93)
	30.01	29.68	30.11
OCB-O	(3.91)	(4.00)	(4.24)
	28.78	28.39	28.65
OCB-I	(4.01)	(4.04)	(4.78)

Nota. Las diferencias significativas en las puntuaciones medias aparecen señaladas en negrita. Las desviaciones típicas están entre paréntesis. AMA = Amabilidad; APE = Apertura a la experiencia; EST = Estabilidad emocional; EXT = Extroversión; RES = Responsabilidad; DS = Deseabilidad Social; APO = Ajuste persona-organización; CRE = Creatividad; GE = Gestión del estrés; AI = Adaptabilidad Interpersonal; REA = Reactividad ante emergencias; EF = Esfuerzo Formativo; OCB-O = OCBs de tipo organizativo; OCB-I = OCBs de tipo interpersonal.

En lo que respecta a la experiencia laboral, los participantes con menos experiencia en la organización presentaron puntuaciones medias superiores a las de las personas con más experiencia en el factor de personalidad Estabilidad Emocional, $t(672) = 2.24, p = .02, d = .24$, en Deseabilidad Social, $t(672) = 2.24, p = .02, d = .20$, y en la dimensión de DA Adaptabilidad interpersonal, $t(672) = 2.24, p = .02, d = .19$. Como puede observarse, los tamaños del efecto fueron inferiores a .40 en los tres casos.

En cuanto al desempeño de puestos de mando y su relación con la personalidad, quienes ejercían puestos de jefatura presentaron puntuaciones significativamente más elevadas en Amabilidad, en comparación con quienes no ejercían puestos de responsabilidad, $t(497.18) = 3.04, p = .01, d = .24$, con un tamaño del efecto pequeño. Lo mismo sucedió en el factor Extroversión, con medias más elevadas para responsables que para subordinados, $t(656) = 2.61, p = .01, d = .21$, si bien el tamaño del efecto fue pequeño.

Por otro lado, las personas que ejercen puestos de mando presentaron puntuaciones más altas en Creatividad que las personas sin puestos de responsabilidad, $t(667) = 3.92, p = .01, d = .32$. Se comprueba que el tamaño del efecto obtenido es pequeño.

El desempeño de funciones de mando también mostró diferencias de tamaño reducido en las puntuaciones en OCBs, con puntuaciones más altas en OCB-O para las personas que ocupaban puestos de mando frente a las personas sin puestos de responsabilidad, $t(638) = 3.64, p = .01, d = .31$. En OCB-I se dieron resultados similares, con medias superiores para las personas con puestos de mando en comparación con las de las personas sin puestos de mando, $t(638) = 2.45, p = .01, d = .21$, aunque con un tamaño del efecto pequeño.

Por otro lado, en lo relativo a las diferencias en personalidad en base al Nivel educativo, se registraron diferencias estadísticamente significativas en el factor de personalidad Apertura a la experiencia, con medias más elevadas para las personas con estudios universitarios que para quienes tenían un nivel educativo inferior, $t(672) = 2.13, p = .03, d = .20$, aunque con un tamaño del efecto pequeño. De manera similar, quienes desempeñaban puestos de mayor complejidad presentaron medias más elevadas en Apertura a la experiencia, que quienes desempeñaban puestos menos complejos, $t(672) = 2.16, p = .03, d = .16$, pero con un tamaño del efecto pequeño.

Por otra parte, se observaron puntuaciones más elevadas en Creatividad en las personas con Niveles de estudios más elevados (universitarios) en comparación con personas sin titulación universitaria, $t(676) = 3.31, p = .01, d = .30$, aunque con un tamaño del efecto pequeño. Otro tanto sucedió con quienes desempeñan puestos con mayor Nivel de complejidad, cuyas medias en Creatividad fueron superiores a las de personas con puestos de menor complejidad, $t(676) = 4.08, p = .01, d = .31$, si bien el tamaño del efecto fue pequeño.

No se hallaron puntuaciones medias diferentes en los constructos examinados en función del desempeño de puestos de carácter fijo o temporal, a excepción de las puntuaciones en OCB-O, en las que las personas con puesto fijo presentaron medias significativamente más altas que las personas en puestos temporales, $t(639) = 2.03, p = .04, d = .16$, aunque con un tamaño del efecto pequeño.

Tabla 4.12.

Resumen de las Diferencias de Medias Estadísticamente Significativas en las Puntuaciones según Variables Sociodemográficas y Laborales

Variable sociodemográfica	Diferencia	Tamaño del efecto
Mujeres	> Apertura a la experiencia	$d = .54$
Mujeres	> Extroversión	$d = .37$
Hombres	> Estabilidad	$d = .27$
Mujeres	> Esfuerzo Formativo	$d = .20$
Hombres	> OCB-O	$d = .17$
> 51 años	< Amabilidad	$g = .36$
> 51 años	< Estabilidad	$g = .20$
> 2 años experiencia en la organización	< Estabilidad	$d = .24$
> 2 años experiencia en la organización	< Deseabilidad	$d = .20$
> 2 años experiencia en la organización	< Adaptabilidad interpersonal	$d = .19$
Puestos de mando	> Amabilidad	$d = .24$
Puestos de mando	> Extroversión	$d = .21$
Puestos de mando	> Creatividad	$d = .32$
Puestos de mando	> OCB-O	$d = .31$
Puestos de mando	> OCB-I	$d = .21$
Estudios superiores	> Apertura a la experiencia	$d = .20$
Estudios superiores	> Creatividad	$d = .30$
Puestos complejos	> Apertura a la experiencia	$d = .16$
Puestos complejos	> Creatividad	$d = .31$
Puesto fijo	> OCB-O	$d = .16$

Por último, ninguna de las variables sociodemográficas o laborales consideradas en este estudio se relacionó con el APO ni presentó puntuaciones medias diferentes en base a los subgrupos definidos por las mismas.

Tabla 4.13.

Resumen de las Hipótesis Formuladas y Contraste desde el Punto de Vista del Tamaño del Efecto (según Ferguson, 2009)

Hipótesis	Literatura	Resultado	Contraste
<i>Relación entre Desempeño adaptativo y OCBs</i>			
1. Relación positiva de Creatividad y Esfuerzo Formativo con OCB-O	Mediana Correlaciones > .20 (Charbonnier-Voirin y Roussel, 2012)	Mediana correlaciones > .20	Confirmada
2. Relación positiva entre Adaptabilidad interpersonal y OCB-I	Correlación >.20 (Charbonnier-Voirin y Roussel, 2012)	Correlación > .20	Confirmada
<i>Relación entre Desempeño adaptativo y APO</i>			
3. Relación positiva de Esfuerzo Formativo y Reactividad ante emergencias con APO	Correlación >.20 (Wessel et al., 2008)	Correlación > .20	Confirmada
<i>Relación entre Personalidad y Desempeño adaptativo</i>			
4. Relación positiva entre Apertura a la experiencia y Creatividad	$r = .34$ (Charbonnier-Voirin, 2013)	$r = .25$	Confirmada
5. Relación positiva entre Responsabilidad y Esfuerzo Formativo	$r = .29$ (Charbonnier-Voirin, 2013)	$r = .28$	Confirmada
6. Relación positiva entre Amabilidad y Adaptabilidad interpersonal	$r = .34$ (Charbonnier-Voirin, 2013)	$r = .34$	Confirmada
7. Relación positiva entre Estabilidad Emocional y Gestión del Estrés	$r = .37$ (Charbonnier-Voirin, 2013)	$r = .47$	Confirmada

Tabla 4.13 (continuación).

Resumen de las Hipótesis Formuladas y Contraste desde el Punto de Vista del Tamaño del Efecto (según Ferguson, 2009)

Hipótesis	Literatura	Resultado	Contraste
<i>Relación entre Personalidad y OCBs</i>			
8.Relación positiva entre Responsabilidad y OCB-O	Correlaciones entre .17 y .30 (Chiaburu et al., 2011; Ilies et al., 2009)	$r = .25$	Confirmada
9.Relación positiva entre Amabilidad y OCB-I	Correlaciones entre .18 y .30 (Chiaburu et al., 2011; Ilies et al., 2009)	$r = .24$	Confirmada
<i>Relación entre Personalidad y Ajuste Persona-Organización</i>			
10.Relación positiva de Responsabilidad, Amabilidad y Extroversión con Ajuste persona-organización	Correlaciones $> .20$ (Gardner et al., 2012; Gillespie et al., 2003)	$r > .20$ con Amabilidad y Responsabilidad, pero $r = .05$ con Extroversión	Confirmada parcialmente
<i>El papel del Ajuste Persona-Organización y del Desempeño Adaptativo en la relación entre Personalidad y Desempeño laboral</i>			
11. Relación positiva entre APO y los dos tipos de OCBs	$r = .21$ (Hoffmann y Woehr, 2006)	Mediana correlaciones $> .20$	Confirmada
12. APO y personalidad tienen capacidad predictiva para explicar los dos tipos de OCBs	$\Delta R^2 = .02$ (Tsai et al., 2012); $\Delta R^2 = .33$ (Michaud, 2014)	$\Delta R^2 = .04$ con OCB-I como criterio. En el resto de casos, $\Delta R^2 < .04$	Confirmada parcialmente
13. DA añade capacidad predictiva a la personalidad y el APO para explicar los dos tipos de OCBs	El modelo I-ADAPT sugiere esta posibilidad (Ployhart y Bliese, 2006)	$\Delta R^2 > .08$ con ambas OCB como criterio	Confirmada
<i>Relación entre el género y los constructos analizados</i>			
14. Puntuaciones más altas en Amabilidad y más bajas en Estabilidad Emocional en mujeres	$d > .40$ (Kajonius y Johnson, 2018; Kajonius y McGiolla, 2017)	$d < .40$ en ambos casos	No confirmada
15. El género no presenta relación con el Desempeño Adaptativo	$d = .22$ (O'Connell et al., 2008)*	La única diferencia significativa es para Esfuerzo formativo, pero $d = .20$	Confirmada

Tabla 4.13 (continuación).

Resumen de las Hipótesis Formuladas y Contraste desde el Punto de Vista del Tamaño del Efecto (según Ferguson, 2009)

Hipótesis	Literatura	Resultado	Contraste
16.El género no presenta relación con las OCBs	$d = 0$ (Ng et al., 2016)*	Los hombres presentan medias superiores en OCB-O, pero con un tamaño del efecto de .17	Confirmada
17.El género no presenta relación con el APO	$d = .22$ (Seong et al., 2012)*	No hay diferencias estadísticamente significativas y $d < .40$	Confirmada
<i>Relación entre la edad y los constructos analizados</i>			
18.No se esperan diferencias en la personalidad en función de la edad	Aunque algunos estudios muestran diferencias sustanciales (Soto 2011; Roberts y Mroczek, 2008) la personalidad se ha mostrado como un constructo estable a lo largo de la vida (Costa et al., 2019; Wagner et al., 2019)	$g < .40$ en todos los casos	Confirmada
<i>Relación entre experiencia laboral y los constructos analizados</i>			
19.No habrá diferencias en OCBs o DA en función de la experiencia laboral	$d < .40$ (Ng y Feldman, 2013)*	$d < .40$ en todos los casos	Confirmada
<i>Relación entre el ejercicio de puestos de mando y los constructos analizados</i>			
20. Puntuaciones más altas en Extroversión en personas con puestos de mando	$d = .65$ (Judge et al., 2002)*	$d = .21$	No confirmada
21. Puntuaciones más altas en OCBs en personas con puestos de mando	$d = .89$ (Ye et al., 2012)*	$d < .40$ con ambos tipos de OCBs	No confirmada
<i>Relación entre nivel educativo y complejidad del puesto y los constructos analizados</i>			
22. No habrá diferencias en Apertura a la experiencia para personas con puestos más complejos	$d = .21$ (Niess y Zacher, 2015)	$d = .16$	Confirmada

Tabla 4.13 (continuación).

Resumen de las Hipótesis Formuladas y Contraste desde el Punto de Vista del Tamaño del Efecto (según Ferguson, 2009)

Hipótesis	Literatura	Resultado	Contraste
23. Puntuaciones más altas en Apertura a la experiencia para personas con mayor nivel educativo	$d = .72$ (Goldberg et al., 1998)*	$d = .20$	No confirmada
24. No habrá diferencias en las puntuaciones en DA en función del nivel educativo	$d = .24$ (O'Connell et al., 2012)*	$d < .40$ en todas las dimensiones de DA	Confirmada
<i>Relación entre estabilidad en el puesto y desempeño laboral</i>			
25. No habrá diferencias en desempeño laboral en función de la estabilidad en el puesto	Sverke et al. (2002) no encontraron diferencias significativas	Media superior en OCB-O para personas con puestos fijos, pero $d = .16$	Confirmada

*Los valores de r de las investigaciones revisadas en la literatura se han transformado en d a los efectos de la presente investigación

4.8. Conclusiones

El estudio presentado en este capítulo aporta evidencias con respecto a la red de relaciones nomológicas del DA y la personalidad, y particularmente acerca de la relación entre la personalidad y el desempeño laboral. Asimismo, subraya el papel que pueden jugar el APO y el desempeño adaptativo en la predicción de las OCBs en selección de personal, además de examinar la relación de estos constructos con variables sociodemográficas y laborales.

En lo que se refiere a la relación entre el DA y las OCBs, las correlaciones fueron de tamaño medio-alto, lo que coincide con lo apuntado por Charbonnier-Voirin y Roussel (2012), que hallaron una correlación de .60 entre una medida global de DA y las OCBs. Según Murphy (2015), esto se debe a que ambos constructos incluyen competencias de evaluación de la situación y utilización de la información para ajustar las conductas de manera efectiva. Por subdimensiones, la correlación entre Creatividad y OCB-O fue de .54, y puede explicarse por la importancia que conceden las personas con puntuaciones elevadas en OCB-O al compromiso con la organización, lo que puede llevarles a hacer un esfuerzo especial a la hora de resolver problemas atípicos, mal definidos o complejos que impliquen creatividad. La vinculación conceptual entre Adaptabilidad interpersonal y OCBs dirigidas a otras personas es elevada y explica la correlación obtenida entre ambos constructos, .32. Sin embargo, esta magnitud fue inferior a la que las OCB-I presentaron con dimensiones de desempeño adaptativo menos vinculadas a priori con la interacción con los demás, como Creatividad o Reactividad ante emergencias. No obstante, diferentes estudios han mostrado relación entre interacción social y creatividad o innovación (Schmidt y Rosenberg, 2014; Tripathi y Burleson, 2012).

Por otro lado, en lo referido a la relación entre APO y DA, la correlación más elevada se observó con Adaptabilidad interpersonal, lo que conceptualmente tiene sentido ya que en APO se mide la congruencia en valores entre las personas que conforman una organización.

En cuanto a la relación entre personalidad y DA, se confirmaron todas las hipótesis planteadas, siendo el tamaño de las correlaciones medio, ligeramente superior a lo apuntado por el meta-análisis de Huang et al. (2014), muy similar al observado por

Ramos-Villagrasa et al. (2020), e inferior al de las correlaciones encontradas por Charbonnier-Voirin (2013). De acuerdo con los hallazgos de esta autora, Apertura a la experiencia presentó las correlaciones más elevadas con Creatividad, lo cual es lógico dada la tendencia a la innovación de las personas con puntuaciones altas en Apertura a la experiencia. También presentó correlaciones altas con Creatividad el factor de personalidad Responsabilidad, algo que cabía esperar dado que las personas con puntuaciones altas en este factor de personalidad son perseverantes y se adaptan fácilmente a situaciones nuevas (Le Pine et al., 2000).

Del mismo modo, Estabilidad emocional fue el factor de personalidad que presentó una correlación más elevada con Gestión del estrés, con una magnitud similar a la encontrada por Charbonnier-Voirin (2013), lo que era esperable dada la similitud conceptual entre ambos constructos. Lo mismo cabe decir acerca de la correlación observada entre Adaptabilidad interpersonal y Amabilidad, lo que corrobora que este factor de personalidad es relevante en contextos laborales que requieren interacciones frecuentes con otras personas.

Responsabilidad, factor relacionado con la constancia y la realización de esfuerzos adicionales para cumplir con los objetivos, y Apertura a la experiencia, constructo vinculado al placer de estudiar y formarse, presentaron las correlaciones más altas con Esfuerzo formativo (Motowidlo y Van Scotter, 1994). Por otro lado, mientras Charbonnier-Voirin (2013) encontró que Reactividad ante emergencias presentaba correlaciones más elevadas con los factores de personalidad Apertura a la experiencia, Responsabilidad y Extroversión, en nuestro caso, y de manera similar a lo apuntado por Oprins et al. (2018), los factores más relacionados con esta dimensión de DA fueron Responsabilidad y Estabilidad emocional, seguidas de Extroversión. La capacidad de superar obstáculos de las personas responsables, así como la capacidad de permanecer en calma de las personas con puntuaciones más altas en Estabilidad emocional pueden explicar el vínculo entre estas variables. La Extroversión, por su parte, implica mayor rapidez en la toma de decisiones y en pasar a la acción para manejar situaciones imprevistas, aspecto clave en la Reactividad ante emergencias (Pulakos et al., 2006).

Las correlaciones entre los factores de personalidad y las dimensiones de desempeño que miden OCBs, como cabía esperar en base a meta-análisis o investigaciones previas, fueron de tamaño medio (Chiaburu et al., 2011; Judge et al., 2013; Sackett y Walmsley, 2014). El factor de personalidad que presentó una

correlación más alta con las OCBs fue Apertura a la experiencia, lo que coincide con el meta-análisis de Chiaburu et al. (2011) en cuanto a las OCB organizacionales. Tras Apertura a la experiencia, los factores de personalidad con correlaciones de magnitudes más elevadas fueron Conciencia y Amabilidad, los predictores más establecidos de las OCBs (Chiaburu et al., 2011). Los resultados guardan similitudes con la investigación que llevaron a cabo Cooper et al. (2013) en una muestra de empleados públicos en la que encontraron que las OCB-O estaban vinculadas principalmente a Responsabilidad y Extroversión y que las OCB-I presentaban correlaciones más elevadas con Responsabilidad y Apertura a la experiencia.

Las magnitudes de las correlaciones del APO con Conciencia y Amabilidad coincidieron con lo hallado en investigaciones precedentes (Gardner et al., 2012; Gillespie et al. 2003). No así la correlación obtenida con Estabilidad emocional, más elevada de lo que cabía esperar. A diferencia de lo apuntado por Gillespie et al. (2003), no se observó una correlación estadísticamente significativa entre Extroversión y APO, presentando además un tamaño del efecto pequeño. Una explicación a este hecho la ofrecen Gardner et al. (2012), quienes señalaron que la relación entre Extroversión y APO en culturas con fuerte componente jerárquico, como la que caracteriza a la Administración Pública, es menos intensa que en culturas de tipo “clan”, en las que se enfatiza el trabajo en equipo y la cohesión. Además, el APO presentó una correlación positiva tanto con las OCB de tipo organizativo como con las de tipo interpersonal, también de tamaño entre reducido y moderado, con valores similares a los del meta-análisis de Hoffman y Woehr (2005).

En definitiva, se encontraron relaciones estadísticamente significativas entre todos los constructos estudiados salvo entre APO y extroversión. Desde el punto de vista del tamaño del efecto, se confirmaron las Hipótesis 1 a 11 con la única excepción de la Hipótesis 10, que se confirmó parcialmente.

Cabe destacar la capacidad predictiva que el APO añadió a la personalidad en la predicción de las OCB-I, lo cual indica que este constructo puede ser de utilidad en procesos selectivos para gestionar la promoción profesional en las organizaciones. El tamaño del efecto encontrado es ligeramente superior al observado por Tsai et al. (2012), e inferior al hallado por Michaud (2014), lo que puede deberse a que este último empleó instrumentos de características diferentes a los empleados en la presente investigación: una medida global de las OCBs, un test de personalidad de tipo ipsativo y

una herramienta para medir el APO de tipo indirecto. En consecuencia, la Hipótesis 12 se cumple parcialmente, dado que solo se confirma en lo referido a OCB-I.

Mayor es la capacidad predictiva que el DA añadió a la personalidad y el APO en la predicción de ambos tipos de OCBs, siendo especialmente destacable su aportación con las OCB-O como criterio. En este caso la capacidad predictiva múltiple fue elevada ($R = .60$) y comparable a la que se puede alcanzar con las mejores combinaciones de tres predictores del desempeño laboral (Salgado y Moscoso, 2008). Asimismo, esta capacidad predictiva múltiple es superior a la obtenida por Murphy (2015) empleando la Adaptabilidad, el factor Conciencia de personalidad y la Inteligencia como predictores de las OCBs. Estos resultados conllevan que se pueda confirmar la Hipótesis 13 y suponen un apoyo al modelo I-ADAPT de Ployhart y Bliese (2006).

Otro aspecto a señalar es que la capacidad predictiva añadida a la personalidad por APO fue ligeramente superior cuando el criterio empleado eran las OCBs que cuando el criterio eran las dimensiones de DA, aunque los tamaños del efecto fueron muy pequeños en todos los casos, por lo que estos resultados deberían ser confirmados en investigaciones futuras. Una explicación plausible la ofrecen Chatman (1989) o Schneider (1987), quienes señalan que niveles elevados de APO son incompatibles con la Adaptabilidad, lo que puede haber afectado negativamente a la validez incremental del APO.

Asimismo, cabe destacar que la Deseabilidad social no parece haber afectado a los constructos analizados ya que las correlaciones encontradas fueron de tamaño pequeño y similares a lo que cabe esperar en base al meta-análisis de Ones et al. (1996) sobre personalidad y desempeño laboral. Las correlaciones entre las puntuaciones de las OCBs (desempeño contextual) y Deseabilidad social fueron muy inferiores a las apuntadas por Delgado-Rodríguez et al. (2018).

Por lo que se refiere a las diferencias debidas a variables sociodemográficas y laborales, han sido pequeñas en términos de Cohen (1988), al presentar valores de d entre .16 y .37, a excepción de la diferencia observada en Apertura a la experiencia en función del género, $d = .54$, de tamaño medio. Siguiendo a Murphy y Jacobs (2012), una d de Cohen de tamaño pequeño indica que las diferencias entre grupos observadas

son una evidencia de ausencia de impacto adverso al utilizar los instrumentos empleados en esta investigación.

El tamaño de las diferencias de género encontradas fue ligeramente inferior a lo apuntado por resultados de investigaciones anteriores. En personalidad las mujeres presentaron puntuaciones más altas en Apertura a la experiencia y Extroversión, frente a puntuaciones más elevadas para los varones en Estabilidad Emocional. Ello se corresponde con lo hallado por Kajonius y Johnson (2018), si bien en la presente investigación las diferencias han sido menores que en el meta-análisis de estos autores. La excepción fue Apertura a la experiencia, con un tamaño del efecto medio y próximo a lo apuntado por Kajonius y Johnson con respecto a las facetas de personalidad del mencionado factor de personalidad. Atendiendo al tamaño del efecto, se rechaza la Hipótesis 14 ya que los efectos fueron de magnitud entre pequeña y media y no superaron el umbral de .40 en los factores hipotetizados (Estabilidad y Amabilidad).

En cuanto a las puntuaciones más elevadas de las mujeres en Esfuerzo Formativo, se podrían explicar en base a su vinculación con el factor de personalidad Apertura a la experiencia, donde las mujeres obtuvieron puntuaciones medias superiores a los varones. No obstante, el tamaño del efecto fue pequeño, tal y como hallaron O'Connell et al., (2008) para el conjunto del desempeño adaptativo, lo que confirma la Hipótesis 15.

La ausencia de relación entre el género y las OCBs confirma la Hipótesis 16 y concuerda con lo observado en el meta-análisis de Ng et al. (2016), quienes, no obstante, señalan que los roles de género, con los varones más orientados a progresar en su carrera, pueden afectar a la relación entre OCBs y género. Por otro lado, la ausencia de puntuaciones sustancialmente diferentes en función del género en APO, corrobora lo hallado por Seong et al. (2012). A la vista de los resultados, la ausencia de diferencias en función del género en las puntuaciones medias en APO sustenta la Hipótesis 17.

De acuerdo con lo esperado, se confirmó la Hipótesis 18: las personas con más edad presentaron puntuaciones significativamente más bajas en Amabilidad y Estabilidad Emocional, si bien los tamaños del efecto encontrados fueron pequeños. Estos resultados confirman que las diferencias en las puntuaciones en personalidad son reducidas cuando se comparan personas de diferentes edades, especialmente cuando se miden los cinco grandes factores, y no sus facetas o subdimensiones (Costa et al., 2019;

Mõttus y Rozgonjuk, 2019; Wagner et al., 2019). La diferencia más cercana al umbral de .40 fue la registrada en Amabilidad, $g = .36$, siendo las puntuaciones de personas con más de 51 años inferiores a las de los grupos de edad entre 21 y 40 años y entre 41 y 50 años.

Las puntuaciones más bajas en Amabilidad y Estabilidad Emocional obtenidas por las personas mayores con respecto a los menores de 50 años coinciden con lo hallado por estudios precedentes con personas de avanzada edad (Kandler et al., 2015; McAdams y Olson, 2010; Mroczek y Almeida, 2004; Schwaba y Bleidorn, 2018; Teachman, 2006; Wortman et al., 2012). Una posible explicación a los resultados del presente estudio la proporcionaron Kandler et al. (2015), al postular que las personas mayores, de manera adaptativa, tienden a ser más conscientes de los peligros de la vida diaria y los problemas de salud, lo que incrementa el Neuroticismo, y además son más selectivos/as en sus actividades y relaciones sociales, lo que reduce la Amabilidad y la Extroversión. Por otro lado, no se obtuvieron puntuaciones diferentes en función de la edad en el resto de los constructos, lo cual es un apoyo a la estabilidad de la personalidad en la vida adulta, como indicaron Wagner et al. (2019), quienes señalaron que Amabilidad es el factor menos estable, probablemente por ser el más afectado por características ambientales fluctuantes.

Asimismo, los resultados obtenidos en la presente investigación se pueden interpretar a la luz de los observados en lo que respecta a la Experiencia en la organización y Deseabilidad Social. En este caso son las personas con menos experiencia, que son también las de menos edad, las que presentan mayores puntuaciones en Estabilidad Emocional, Adaptabilidad interpersonal, y Deseabilidad Social, dimensiones que se perciben vinculadas a puestos con salarios elevados y a un buen ajuste a la organización para la que se trabaja (Jokela et al., 2017; Ng y Feldman, 2015; Salgado, 2016; Woods et al., 2018). En todo caso, dada la naturaleza transversal del presente diseño de investigación y el pequeño tamaño de las diferencias halladas, es preciso confirmar los presentes resultados con futuros estudios de tipo longitudinal.

No se observaron puntuaciones diferentes en ninguno de los constructos en función de la Experiencia en el puesto, algo que cabía esperar dado que la correlación entre Experiencia y desempeño laboral es muy pequeña según los meta-análisis de Schmidt et al. (2016) y de Van Iddekinge et al. (2019). En consecuencia, se confirma la Hipótesis 19, en línea con lo apuntado por Ng y Feldman (2013).

Por otro lado, en lo que se refiere al perfil mostrado por las personas que desempeñan puestos de mando, este colectivo presentó puntuaciones más elevadas en Amabilidad y Extroversión, en ambos casos con tamaños del efecto inferiores a .40 y similares a los hallados por Lounsbury et al. (2016). El tamaño de la relación entre Extroversión y el ejercicio de puestos de mando, .21, fue menor a lo apuntado por el meta-análisis de Judge et al. (2002) y no permite confirmar la Hipótesis 20. Una posible explicación para este tamaño del efecto menor de lo previsto la ofrecen Ng et al. (2008), quienes señalaron que la mayor capacidad de decisión en puestos con alta autonomía ofrece más oportunidades para que los/as líderes extrovertidos/as pongan en práctica sus tendencias a comunicar, influir en otros o asumir responsabilidades. Dado que en el sector público los puestos de jefatura presentan diferentes niveles en la jerarquía de la organización vinculados a una mayor o menor autonomía, una posibilidad es que la autonomía, variable no contemplada en esta investigación, haya moderado la relevancia de la Extroversión en el caso que nos ocupa.

Asimismo, las personas con puestos de mando obtuvieron puntuaciones más altas en OCB-O y OCB-I, aunque con un tamaño del efecto reducido, por lo que los resultados muestran una asociación menor de lo señalado por la literatura entre liderazgo y OCB e implican el rechazo de la Hipótesis 21 (Ye et al., 2012). Una posible explicación es que los empleados y empleadas del sector público destacan por su elevado nivel de OCBs en comparación con el personal del sector privado, algo que se puede vincular a la motivación de servicio público presente en gran medida en el colectivo funcional, lo que podría haber reducido el tamaño de la diferencia de medias (Kim, 2006; Sharma et al., 2011).

La diferencia de medias vinculada al ejercicio de puestos de mando más cercana a .40 fue la registrada en Creatividad, con puntuaciones más elevadas para los y las responsables. Este resultado está en línea con lo señalado por Huang et al. (2014) o Berg et al. (2010), quienes subrayaron la importancia que para el ejercicio de este tipo de puestos tiene la capacidad de resolver problemas atípicos o complejos que requieran ser creativo/a (Klemm, 2017).

El factor de personalidad Apertura a la experiencia ha presentado una vinculación con Complejidad del puesto similar a la indicada por Niess y Zacher (2015), lo que confirma la Hipótesis 22 y era esperable dado el nivel de sobretitulación

de las personas que trabajan en el sector público desempeñando trabajos de poca complejidad.

Por lo que respecta a la asociación entre Apertura a la experiencia y Nivel educativo, la relación fue estadísticamente significativa, lo que coincide con lo apuntado por investigaciones precedentes, si bien el tamaño del efecto fue menor de lo apuntado por dichos estudios (Goldberg, et al., 1998; Möttus et al., 2017). Estos resultados suponen que no se pueda confirmar la Hipótesis 23 desde el punto de vista del tamaño del efecto. Una posible explicación es que la muestra estaba compuesta por empleados y empleadas del sector público, y la motivación para trabajar en el sector público está vinculada a la Apertura a la experiencia (Van Witteloostuijn et al., 2016), lo que habría podido provocar que las diferencias hayan sido menores de lo que cabría esperar.

En cuanto a la relación entre Nivel educativo y DA, todas las dimensiones de DA presentaron diferencias inferiores a .40, en línea con lo hallado por O'Connell et al. (2008) con una medida global de DA. En consecuencia, la Hipótesis 24 se confirma en base a los tamaños del efecto hallados.

La estabilidad laboral no presentó relaciones significativas con ninguno de los constructos analizados ya que no se registraron diferencias superiores al umbral establecido entre las medias de las personas participantes en personalidad, APO, DA u OCBs en función de si tenían o no tenían un puesto fijo. Los resultados coinciden con lo hallado por Sverke et al. (2002) y permiten confirmar la Hipótesis 25. Este hallazgo, especialmente en lo relativo al desempeño laboral, tiene interés de cara a la gestión de organizaciones públicas, aún más en un contexto de crisis como el actual en el que hay una preocupante tasa de temporalidad en el sector público (Randstad, 2018; Gimeno, 2020).

Por último, cabe destacar que ninguna variable sociodemográfica o laboral presentó una relación a tener en cuenta con el APO, lo que coincide con lo hallado por Tsai et al (2012) o Xu (2014).

La presente investigación tiene varias limitaciones. La más importante es que los datos se han obtenido mediante un diseño de investigación transversal, es decir, un diseño en el que todas las medidas se han recogido a la vez, en un contexto no selectivo. Esto suele elevar el tamaño de las correlaciones entre .05 y .10 en comparación con los diseños de validez predictiva, en los que existe una distancia temporal entre la recogida

de las variables predictoras susceptibles de ser empleadas en un proceso selectivo, como la personalidad, y la recogida de las variables criterio, como puede ser la evaluación del desempeño laboral (Van Iddekinge y Ployhart, 2008).

Adaptabilidad Interpersonal ha presentado problemas de asimetría y curtosis, algo habitual en la medición del desempeño laboral (Beck et al., 2014). No obstante, la presencia de distribuciones sesgadas no afecta a la robustez de los análisis realizados en esta investigación (Martínez Arias et al., 2006).

Además, las medidas de desempeño se han basado en auto-evaluaciones de las personas participantes. No obstante, Demerouti et al. (2014) han hallado que las auto-evaluaciones explican más varianza del desempeño contextual que las evaluaciones de compañeros/as de trabajo. Además, el meta-análisis de Carpenter et al. (2014) en el que compararon auto-evaluaciones de OCBs con evaluaciones de jefes/as y compañeros/as, puso de manifiesto que las diferencias de medias eran muy pequeñas y pueden reflejar que las auto-evaluaciones recogen mejor las OCBs realizadas y que compañeros/as o superiores no pueden observar todos los aspectos de las conductas ajenas. Estos autores encontraron también una correlación de tamaño medio entre auto y hetero-evaluaciones, lo que indica convergencia entre las medidas, pero a la vez una singularidad considerable. Por último, estos autores encontraron que ambos tipos de medidas presentan magnitudes y patrones de relaciones de carácter similar con los correlatos comunes. De todo ello dedujeron que las auto-evaluaciones no solo son una manera viable de medir las conductas laborales, sino que también pueden ser la manera más adecuada de medirlas. Asimismo, Conway y Lance (2010) señalaron que las relaciones estimadas usando diferentes métodos (o evaluaciones de terceras personas) tienden a ser más atenuadas y menos precisas que las obtenidas con auto-evaluaciones, que son una representación adecuada y precisa de las correlaciones reales. Además, debemos considerar que la muestra empleada no formaba parte de un proceso selectivo, por lo que cabe esperar que el efecto de la deseabilidad social fuese pequeño.

En la misma línea, el meta-análisis de Kim et al. (2019) pone de manifiesto que las diferencias entre la auto-evaluación de la personalidad y las evaluaciones realizadas por terceras personas son mínimas. Cabe mencionar, además, que en diversas investigaciones transculturales las puntuaciones en deseabilidad social muestran que en la mayor parte de las culturas las personas se evalúan a sí mismas de manera menos favorable que como son percibidas por otras personas (Allik y Realo, 2017)

Por otro lado, esta investigación se ha realizado con una muestra de empleados públicos, lo que no permite una comparación empírica con el sector privado y, por tanto, es preciso contextualizar los resultados a las características de la muestra empleada ya que la literatura señala que el personal del sector público está motivado por diferentes factores y los distintos tipos de personalidad pueden interactuar con las estructuras organizacionales de diferentes maneras (Cooper et al., 2013).

En este sentido, sería interesante comparar los sectores público y privado en lo relativo a la deseabilidad en las respuestas en base al estudio de Roulin y Krings (2019). Según estos autores, en organizaciones con características similares a las de las Administraciones Públicas (menor competitividad y menor nivel de innovación), los candidatos y las candidatas tienden a exagerar sus puntuaciones en Amabilidad y a fingir que tienen menos Apertura a la experiencia que la que tienen realmente para adaptarse a la cultura organizacional. Peiró y Ramos (2012), por su parte, hallaron que quienes trabajan en el sector público presentan valores más elevados en el desempeño percibido.

Asimismo, futuras investigaciones deberían emplear medidas del desempeño vinculadas al desempeño de tarea, la dimensión de desempeño laboral que más atención ha recibido en el marco de la psicología del trabajo y las organizaciones. Ello contribuiría a tener una visión más holística del impacto de los resultados obtenidos en esta investigación a nivel de desempeño laboral.

Como implicaciones prácticas de este trabajo, cabe mencionar que se ha contribuido a precisar las relaciones entre los cinco grandes factores de personalidad y las cinco dimensiones de DA del modelo de Charbonnier-Voirin y Roussel (2012), así como la relación entre estas dimensiones de DA y los dos grandes tipos de OCBs, aspecto poco investigado hasta la fecha, lo que supone una importante contribución para establecer la red nomológica de las relaciones de DA con otros constructos.

Asimismo, es preciso destacar que el APO puede ser una herramienta complementaria de interés para ser utilizada junto con la personalidad en la predicción de las OCB-I. Además, este constructo no se ve afectado por las variables sociodemográficas y laborales estudiadas, lo que facilita su posible uso en contextos selectivos en el sector público.

Más relevante, si cabe, es concebir al DA como un constructo predictor del desempeño laboral, en aplicación del modelo I-ADAPT al ámbito de la selección de personal. En vista de los resultados obtenidos, incluir el DA en los diseños selectivos podría contribuir a aumentar de manera considerable la eficacia de la selección de personal.

Por otro lado, el reducido tamaño de las diferencias observadas en las puntuaciones medias en los factores del FFM en función de variables sociodemográficas y laborales es un argumento a favor de la inclusión de la personalidad en contextos selectivos en base a su reducido impacto adverso. Asimismo, la ausencia de impacto adverso en las puntuaciones medias del DA y de las OCBs es un argumento a favor de emplear pruebas que miden estos constructos en los procesos selectivos. Ello contribuiría a satisfacer uno de los grandes valores en materia de selección de personal, la igualdad.

En cualquier caso, a fin de completar el bagaje de conocimientos en este ámbito de estudio, se considera necesario realizar investigaciones similares en un contexto de selección con un diseño de validez predictiva, con medidas de desempeño tomadas tiempo después de la selección, y empleando personal no solo del sector público sino también del sector privado.

CAPÍTULO 5.

CONCLUSIONES GENERALES

En este capítulo se presentan las principales conclusiones extraídas del estudio realizado. En primer lugar, cabe afirmar que se han alcanzado los objetivos establecidos utilizando, para ello, una metodología rigurosa. Las ciencias se basan en la medición de los fenómenos que analizan, y, por ello, la psicología debe redoblar sus esfuerzos en el proceso de elaboración y adaptación de escalas objetivas, válidas y fiables que sean precisas, eficaces y estén imbricadas en el marco de una teoría bien articulada (Clark y Watson, 2019).

En este contexto, el objetivo general de la tesis era proporcionar instrumentos que permitan una evaluación rigurosa de los factores de personalidad y el desempeño laboral en el ámbito organizacional. A tal fin, en el Capítulo 1 se han presentado las cuestiones metodológicas más relevantes que deben tomarse en consideración en la adaptación de instrumentos de evaluación psicológica. Así, se han descrito los aspectos a tener en cuenta en la traducción de los ítems, los métodos para establecer equivalencias de traducción, la evaluación de la equivalencia métrica a nivel de ítem y a nivel de test o escala, así como los sesgos más habituales y la manera de afrontarlos. La última parte de este capítulo se ha centrado en las directrices que deben seguirse en la adaptación, en particular en el marco de la investigación en las ciencias del comportamiento y de las organizaciones (Balluerka et al., 2007; Brislin, 1986; Carretero-Dios y Pérez, 2005; Hambleton y Patsula, 1999; Heggstad et al., 2019; Matthews-López, 2003).

La aplicación de la metodología descrita en el primer capítulo ha contribuido a la consecución del primer objetivo específico de la presente tesis: adaptar a la lengua y cultura castellana el cuestionario para evaluar el desempeño adaptativo de Charbonnier-Voirin y Roussel (2012). El instrumento adaptado en el marco de esta investigación ha presentado índices psicométricos adecuados y similares a los del instrumento original, lo que se ha descrito en el Capítulo 2. Todo ello hace que estemos ante un instrumento que puede ser empleado para evaluar de forma válida y fiable el desempeño adaptativo en el contexto laboral.

La importancia de la medición del desempeño adaptativo es cada día mayor, dado que debe ser tenido en cuenta tanto en procesos de selección como de promoción, desde distintos ámbitos de la Psicología del Trabajo y las Organizaciones. Cullen et al. (2014) señalaron que la selección se debe realizar teniendo en cuenta la adaptabilidad de los/as candidatos/as. En ese sentido, Barrick y Parks-Leduc (2019) observaron que la adaptabilidad puede ser un factor esencial para conocer el rango de puestos a los que puede optar un/a candidato/a y, por tanto, puede resultar decisivo para incrementar su ajuste a la organización. Oprins et al. (2018), por su parte, han señalado que resulta necesario medir los requisitos de adaptabilidad de los puestos de trabajo, ya que cada tipo de puesto, en función de sus características, demanda concretar la relevancia que deben tener las diferentes subdimensiones del desempeño adaptativo. Asimismo, las personas pueden ser formadas para mejorar en adaptabilidad, por lo que medir este constructo tiene implicaciones en lo relativo a la gestión y desarrollo de la carrera laboral en una organización.

La relevancia de este constructo hace que la adaptación al castellano de la escala de Charbonnier-Voirin y Roussel (2012) suponga una contribución notable, dada la carencia de instrumentos que permitan evaluar el desempeño adaptativo en castellano. Asimismo, supone un avance en la investigación sobre la teoría del desempeño adaptativo, cuestión trascendental en el ámbito de las organizaciones (Baard et al., 2014).

La metodología que se ha descrito en el primer capítulo también se ha utilizado para cumplir con el segundo objetivo específico de la presente tesis: adaptar a la lengua y cultura vascas el cuestionario de personalidad OPERAS (Vigil-Colet et al., 2013). La aplicación de dicha metodología ha dado lugar a una herramienta a partir de la cual se pueden tomar decisiones fundamentadas en evidencias de fiabilidad y validez, tal y como se ha comprobado en el Capítulo 3. La principal aplicación de este nuevo instrumento disponible en euskera puede ser la selección de personal, particularmente en el sector público, muy necesitado de instrumentos de evaluación en euskera que permitan cumplir con el requisito legal de garantizar los derechos lingüísticos de la ciudadanía. La medición de la personalidad cobra especial importancia en un contexto VUCA (acrónimo en inglés de Volatilidad, Incertidumbre, Complejidad y Ambigüedad) como el actual, dado que las diferencias individuales en personalidad tienden a

acentuarse en contextos caracterizados por la novedad, la ambigüedad y la incertidumbre (Caspi y Moffitt, 1993).

Asimismo, es preciso destacar que OPERAS en euskera es un instrumento breve y de rápida aplicación, lo que es relevante en la evaluación de la personalidad. Poder disponer de instrumentos aplicables en un espacio de tiempo corto, siempre que presenten propiedades psicométricas adecuadas, es muy útil para obtener una información similar a la aportada por instrumentos con más ítems reduciendo los efectos adversos de los instrumentos con un tiempo de administración más extenso, como fatiga o falta de motivación (Balluerka y Gorostiaga, 2012).

El proceso de adaptación se ha focalizado en tres de las cinco fuentes de evidencia a emplear para evaluar si la interpretación de las puntuaciones de un test se ha realizado correctamente según los *Standards for Educational and Psychological Testing*: evidencias basadas en los procesos de respuesta (mediante entrevistas cognitivas), evidencias basadas en la estructura interna y evidencias basadas en las relaciones con otras variables, estas últimas abordadas en el Capítulo 4 (Sireci y Padilla-García, 2014). No se ha profundizado en las evidencias basadas en el contenido, más propias del desarrollo inicial de un test que de su adaptación a otro idioma o cultura, ni en evidencias basadas en las consecuencias de la evaluación puesto que todavía no se aplicado más allá del contexto de la presente investigación.

En la adaptación de ambos instrumentos se han aplicado análisis estadísticos recomendados por expertos (Carretero-Dios y Pérez, 2005; Muñiz et al., 2011): correlaciones medias entre los ítems de las dimensiones, correlaciones entre las dimensiones, análisis factorial para estudiar la dimensionalidad de los instrumentos, fiabilidad entendida como consistencia interna (índices ORION, omega o de fiabilidad marginal mediante la Teoría de Respuesta al Ítem), fiabilidad test-retest, evidencias de validez convergente y/o estudio de la red nomológica de los instrumentos, así como el estudio del funcionamiento diferencial de los ítems según variables sociodemográficas y laborales. Este último aspecto es crucial para garantizar la equivalencia métrica en la adaptación de instrumentos (Zumbo, 2003). Es importante destacar que en tales análisis se han obtenido resultados adecuados, lo que garantiza la validez y fiabilidad en la medición de los constructos evaluados por ambos instrumentos.

Otro aspecto en común de los Capítulos 2 y 3 que queremos destacar es la metodología ESEM empleada para obtener evidencias de validez de constructo en la adaptación de OPERAS y de la escala de Charbonnier-Voirin y Roussel (2012). Dicha metodología se está utilizando cada vez con mayor frecuencia y debe ser tenida en cuenta para evaluar el ajuste de los datos empíricos a la teoría dada la flexibilidad que presenta frente a los análisis factoriales confirmatorios (AFCs). Los AFCs no permiten que los ítems saturen en más de un factor y, en consecuencia, no se adecúan a la naturaleza de algunos constructos. Un ejemplo de ello son los problemas que presenta la validación del modelo de los Cinco Grandes de personalidad mediante los mencionados AFCs, a diferencia de lo que ocurre con la metodología ESEM (Marsh et al., 2014).

Pese a sus problemas con los AFCs, el rigor metodológico en la construcción de los cuestionarios de personalidad conlleva que, los actuales problemas de replicabilidad que se observan en las ciencias de la conducta, entre ellas en la psicología (Button et al., 2013; Camerer et al., 2016; Cova et al., 2018; Open Science Collaboration, 2015; Wingen et al., 2020), apenas hayan afectado al campo de la psicología de la personalidad. El reciente trabajo de Soto (2019) muestra que en este ámbito el 87% de investigaciones son replicadas con éxito, frente, por ejemplo, al 25% de la psicología social.

Esta replicabilidad de las pruebas de personalidad ha sido un factor decisivo para que los resultados del Capítulo 4, los relativos a las relaciones entre los constructos estudiados, sean similares a los de estudios anteriores. En este capítulo se abordaba el tercer objetivo específico de la tesis: analizar la relación de los cinco grandes factores de personalidad con constructos vinculados con el desempeño laboral (OCBs y desempeño adaptativo) y el Ajuste Persona-Organización (APO). Este tercer objetivo estaba vinculado a la eficacia en la predicción del desempeño laboral, centrándose en la personalidad y el desempeño adaptativo a fin de identificar con más precisión a las personas más adecuadas para desempeñar las tareas del puesto para el que se seleccionan. Mediante los análisis realizados se ha obtenido información sobre la red de relaciones nomológicas de los dos instrumentos adaptados en los capítulos 2 y 3: la escala de desempeño adaptativo de Charbonnier-Voirin y Roussel (2012) en castellano y OPERAS en euskera. A tal fin se ha empleado el marco teórico del modelo I-ADAPT de Ployhart y Bliese (2006), que plantea que la personalidad y los valores (constructo

vinculado al APO) son antecedentes del desempeño adaptativo, que a su vez predice las OCBs.

En el análisis de la red nomológica de la adaptación de la escala de Charbonnier-Voirin y Roussel (2012), cabe destacar la intensidad de la relación entre las dimensiones de Desempeño adaptativo y las de las OCBs, aspecto poco investigado previamente. Esta cuestión es relevante para avanzar en modelos teóricos del desempeño laboral, ya que a nivel conceptual no hay consenso entre los diferentes enfoques del Desempeño Adaptativo sobre su relación con el resto de componentes del desempeño (Baard et al., 2014). Así, por un lado, está el enfoque que concibe el Desempeño adaptativo como un componente del desempeño laboral al mismo nivel que el desempeño de tarea o el contextual. Los resultados de Charbonnier-Voirin y Roussel (2012), que se basan en este enfoque, son similares a los obtenidos en la presente investigación. No obstante, existe un enfoque alternativo, el del modelo teórico I-ADAPT de Ployhart y Bliese (2006), que se enmarca en una perspectiva que concibe la adaptabilidad como un conjunto de diferencias individuales que predisponen a las personas a adaptarse con éxito a los cambios en las tareas. Según estos autores, el desempeño adaptativo predice el desempeño de tarea o el contextual (como las OCBs), por lo que las conductas adaptativas juegan un papel relevante a la hora de explicar el desempeño laboral. Dado que Baard et al. (2014) proponen que puede resultar recomendable estudiar el desempeño adaptativo desde diferentes enfoques, futuras investigaciones deberían profundizar en la relación entre desempeño adaptativo y OCBs de la mano del modelo I-ADAPT. Dicho modelo identificó procesos mediadores entre ambas variables que no se han incluido en el presente estudio, como la evaluación y percepción de la situación, la estrategia de selección adaptativa para la situación, la auto-regulación y el afrontamiento, y la adquisición de conocimientos.

Otra cuestión de interés abordada en el análisis de la red de relaciones nomológicas del desempeño adaptativo ha sido la relación entre la adaptabilidad y el APO, cuestión poco estudiada hasta la fecha. Los resultados muestran una correlación de tamaño medio-bajo, igual que la observada en estudios precedentes.

Un aspecto que comparten las redes de relaciones nomológicas de las adaptaciones de la escala de Charbonnier-Voirin y Roussel (2012) y de OPERAS es la relación entre los factores de personalidad y las dimensiones de desempeño laboral, cuestión escasamente abordada con anterioridad (Murphy, 2015). Las correlaciones

obtenidas tienen sentido dada la proximidad conceptual entre los constructos: Apertura a la experiencia está vinculada a Creatividad y Esfuerzo Formativo; Estabilidad Emocional se relaciona con Gestión del estrés; Amabilidad con Adaptabilidad interpersonal. Todo ello contribuye a mejorar el bagaje de conocimientos en torno a un constructo relativamente nuevo como es el desempeño adaptativo.

En el estudio de las relaciones de los factores de personalidad de OPERAS con el resto de constructos, también se ha analizado la relación de los cinco grandes factores de personalidad con la dimensión del desempeño laboral que más atención está recibiendo en los últimos años, las OCBs. Los resultados obtenidos se corresponden con los hallados por investigaciones meta-analíticas previas.

Otra cuestión de interés en lo relativo a la red de relaciones de OPERAS en euskera es el rol de una variable a menudo olvidada, el APO. Las correlaciones de dicha variable con los Cinco Grandes factores de la personalidad se han correspondido con lo hallado por estudios precedentes. Una cuestión más innovadora es la que se refiere al potencial del APO como predictor que aporta validez incremental a la personalidad en la predicción del desempeño. De hecho, la validez añadida o incremental es una cuestión vital a considerar en el desarrollo de un instrumento de medida (Clark y Watson, 2019; Hunsley y Meyer, 2003). Los resultados coinciden con lo hallado por Tsai et al. (2012) e indican que el APO puede aumentar, de manera moderada, la capacidad predictiva de la personalidad con las OCB-I como criterio en mayor medida que con el desempeño adaptativo como criterio. Por otro lado, cabe señalar que la validez añadida del APO a la personalidad con el DA como criterio, cuestión poco estudiada previamente, ha sido reducida.

Más relevante, si cabe, ha sido el papel del DA como predictor que añade capacidad predictiva a la personalidad y al APO con las OCBs como criterio. El desempeño adaptativo incrementa notablemente dicha capacidad predictiva conjunta, de modo que los tres predictores conjuntamente explican hasta el 35% de la varianza de las OCBs de tipo organizativo. Dicha capacidad predictiva conjunta está a la altura de las mejores combinaciones de tres predictores selectivos (Salgado y Moscoso, 2008). Asimismo, cabe subrayar que se ha obtenido empleando tres instrumentos breves, lo que supone un diseño selectivo de gran eficiencia que podría aplicarse junto con otras pruebas incluso en un mismo día (Ramió, 2011). En consecuencia, estos resultados

subrayan la conveniencia de incluir la evaluación del desempeño adaptativo en los procesos de selección de personal.

El enfoque I-ADAPT remarca la importancia de considerar la doble naturaleza de la adaptabilidad (como variable resultado y como variable predictora) y subraya su utilidad en el contexto de la selección de personal. El DA es un predictor del desempeño laboral más proximal (más cercano conceptualmente al desempeño) que la personalidad porque está más relacionado con las diferentes situaciones que se pueden dar en el trabajo. La adaptabilidad refleja los cambios en las personas para adecuarse a su entorno, lo que es vital en el ámbito laboral actual. Diversas investigaciones han mostrado que el desempeño fluctúa a lo largo del tiempo debido en buena medida a cambios intrapersonales. Los cambios de las personas para adaptarse podrían estar vinculados a las variaciones del desempeño, lo que convierte al DA en un predictor prometedor del desempeño fluctuante (Murphy, 2015).

En conjunto, a la vista de los resultados obtenidos en el análisis de las redes de relaciones nomológicas de los instrumentos adaptados en el marco de esta tesis, cabe colegir que dichas redes nomológicas son una evidencia de su validez de constructo y de su validez externa. Asimismo, los resultados hallados, especialmente en lo relativo a las relaciones del DA con las OCBs y a la capacidad predictiva del desempeño laboral que el DA aporta a la personalidad y al APO, han contribuido a ampliar el bagaje de conocimientos en materia de psicología de las organizaciones.

Por otro lado, en el capítulo 4 se ha profundizado en un aspecto de gran relevancia social, legal y organizacional en la actualidad, las diferencias entre subgrupos, cuestión vinculada a la igualdad en materia de selección de personal (Gorriti, 2018; Guion, 2011). Ello ha permitido dar cumplimiento al cuarto objetivo específico de la tesis: analizar la posible relación existente entre los constructos de interés en selección de personal (personalidad, DA, APO y OCBs) y un conjunto de variables sociodemográficas y laborales.

La mayoría de las diferencias entre subgrupos han sido de tamaño reducido, lo que minimiza el riesgo potencial de impacto adverso de los instrumentos empleados y contribuye a señalar posibles líneas de investigación futuras. La excepción en cuanto al tamaño de las diferencias observadas ha sido la relativa al factor Apertura a la experiencia en función del género, cuyo tamaño ha sido medio. En lo que se refiere a la

relación del género con el resto de variables, las mujeres han presentado puntuaciones más altas en Apertura a la experiencia, Extroversión y Esfuerzo Formativo, mientras que los varones han presentado puntuaciones superiores en Estabilidad Emocional y en ambos tipos de OCBs, lo que se corresponde con investigaciones precedentes, si bien las diferencias han sido menores de lo esperado en cuanto a los factores de personalidad.

La presente investigación ha aportado evidencias acerca de la estabilidad de la personalidad, dadas las reducidas diferencias halladas entre personas de distintos grupos de edad, aunque no debemos olvidar que en este estudio se ha empleado un diseño transversal. En este sentido, sería interesante realizar investigaciones longitudinales para estudiar una posible evolución en los rasgos de personalidad en función de la edad. En particular, para comprobar si las generaciones más jóvenes presentan puntuaciones más altas en rasgos como Amabilidad y Estabilidad, en línea con lo apuntado por Jokela et al. (2017), o para estudiar posibles cambios en la personalidad en distintas etapas de la vida laboral (Bleidorn et al., 2019). Asimismo, un planteamiento longitudinal permitiría aplicar la escala en castellano de Charbonnier-Voirin y Roussel (2012) en diferentes momentos de la vida laboral del personal. De este modo se podría comprobar el enfoque de Ramos-Villagrasa et al. (2019), según el cual la adaptación es un proceso mediante el que las conductas adaptativas pueden cambiar y mejorar.

Por otro lado, se ha estudiado el perfil de las personas que ejercen puestos de mando en el sector público vasco, con resultados que resaltan la importancia del factor de personalidad Extroversión, aunque con efectos más pequeños que los obtenidos en investigaciones precedentes. En menor medida, los resultados también remarcan el papel de un factor no asociado tradicionalmente ni a la emergencia del liderazgo ni a su eficacia, la Amabilidad (Judge et al., 2002; Hart, 2014). Dada la importancia de este factor en el sector público y en organizaciones no competitivas (Buurman et al., 2012; Roulin y Krings, 2019), se debería profundizar tanto en su evaluación para el acceso a puestos de jefatura como en su impacto en el desempeño de los mismos.

En esta investigación se ha prestado especial atención a variables sociodemográficas y laborales como el nivel educativo y la complejidad del puesto desempeñado y a su vinculación con constructos de interés en materia de selección de personal: la personalidad, el desempeño laboral y el ajuste persona-organización. Los resultados subrayan que dichas variables sociodemográficas y laborales presentan una

relación de tamaño reducido con los constructos mencionados. Ello indica que la aplicación de instrumentos que evalúen la personalidad, el APO, las OCBs o el DA presenta un reducido potencial de impacto adverso y, por consiguiente, escasa probabilidad de ser origen de litigios judiciales. Todo ello es relevante de cara al diseño de procesos selectivos y de gestión de recursos humanos.

En cualquier caso, y pese al interés de las diferencias entre grupos encontradas (de tamaño pequeño, en su mayor parte), como señalan Hanel et al. (2019), las investigaciones tienden a centrarse en mayor medida en las diferencias que en las similitudes entre grupos. Estos autores proponen el uso del Coeficiente de superposición (Overlap Coefficient, OVL, en inglés) para medir el grado de superposición entre las distribuciones de los datos de grupos diferentes, frente a la d de Cohen, que está basada en la media. Futuras investigaciones deberían emplear este estadístico junto a las medidas de tamaño del efecto de las diferencias intergrupos.

Una de las limitaciones de esta investigación ya mencionada en los capítulos precedentes, es que la muestra ha estado formada exclusivamente por personal empleado público. No obstante, esta característica de la muestra también aporta información de interés para la investigación sobre la realidad del sector público. Así, cabe remarcar que dos factores de personalidad a los que se ha prestado menos atención que al factor Responsabilidad en la literatura científica, por su menor capacidad predictiva del desempeño, Apertura a la experiencia y Amabilidad, tienen un rol relevante en el ámbito público. Su vinculación en la presente investigación con las OCBs y el APO, o las dimensiones de desempeño adaptativo Creatividad y Adaptación Interpersonal, así como su relación, aunque de tamaño reducido, con el desempeño de puestos de mando (Amabilidad) o con variables sociodemográficas y laborales, tales como el nivel educativo o la complejidad del puesto (Apertura a la experiencia), subrayan su papel en la Administración. Ambas variables están relacionadas con el constructo Motivación para el Servicio Público (Van Witteloostuijn et al., 2016), por lo que una conclusión de esta tesis es que deberían tener un peso más importante en el diseño de los procesos selectivos en el ámbito público. Asimismo, la importancia de la Apertura a la experiencia cuestiona la imagen habitual del funcionariado como colectivo poco dado a la innovación, flexibilidad o a la asunción de riesgos (Buurman et al., 2012; De la Nuez, 2018; Demmke, 2007).

Amabilidad y Apertura a la experiencia, a su vez, son predictores de algunas de las destrezas más relevantes de los puestos de las Administraciones Públicas que serán más demandados las próximas décadas según la base de datos O*NET del Departamento de Trabajo del Gobierno de Estados Unidos: Pensamiento crítico; Escucha Activa o Aprendizaje Activo. Asimismo, están vinculadas conceptualmente a varias de las destrezas críticas de los puestos del futuro identificadas por la Singularity University: Pensamiento crítico; Colaboración en red; Curiosidad e imaginación (Bidshahri, 2017; Gorriti, 2018).

Otra de las destrezas clave identificadas por Bidshahri (2017) es la Agilidad y Adaptabilidad, lo que subraya la importancia de medir el desempeño adaptativo en la gestión de recursos humanos en el ámbito público. El instrumento de Charbonnier-Voirin y Roussel (2012) adaptado al castellano en el marco de esta tesis puede responder a tal objetivo. Asimismo, su modelo de 5 dimensiones puede servir de herramienta de evaluación de las 5 competencias clave de los trabajadores del siglo XXI: Logro de objetivos; Adaptación al cambio; Analizar y resolver problemas; Aprendizaje y auto-desarrollo; y Trabajar bien con otras personas (Herde et al., 2019). En ese sentido, sería interesante llevar a cabo la adaptación al euskera de dicho instrumento, de manera que se pudiera utilizar en nuestro contexto bilingüe.

Una potencial utilidad adicional del modelo de Charbonnier-Voirin y Roussel (2012) se encuentra en el ámbito de la formación y el desarrollo de recursos humanos y se deriva de su paralelismo con modelos centrados en destrezas socio-emocionales, es decir, las destrezas esenciales para abrirse camino en los ámbitos de la educación y el trabajo (Binkley et al., 2012). Así, según Abrahams et al. (2019) existen cinco dominios de destrezas socioemocionales de amplio espectro que guardan similitudes con las dimensiones de desempeño adaptativo: Auto-Gestión; Interacción con los demás; Amistad; Gestión de emociones negativas (como el estrés o la frustración); y Apertura a la experiencia (vinculada a la creatividad). Guerra et al. (2014), por su parte, proponen el modelo PRACTICE, de 8 destrezas para el éxito laboral, de las que cuatro presentan similitudes con el modelo de Charbonnier-Voirin y Roussel: Solución de Problemas (relacionada con la creatividad); Resiliencia (por ejemplo, ante el estrés); Motivación para el logro (que incluye la motivación para el aprendizaje); y Trabajo en equipo (vinculada a la interacción y relaciones con los demás).

Los mencionados dominios de destrezas socio-emocionales presentan, a su vez, una relación conceptual con los Cinco Grandes factores de personalidad. La identificación de destrezas modificables alineadas con rasgos duraderos (como la personalidad) puede dar lugar a programas de aprendizaje de destrezas socioemocionales. Ello puede contribuir al bienestar y a las conductas adaptativas de las personas en el trabajo (Abrahams et al., 2019; Guerra et al. 2014; Primi et al., 2017, Soto et al., in press).

En consecuencia, los instrumentos adaptados en el marco de esta tesis pueden ser de utilidad no solo para el ámbito de la selección de personal, sino también para la gestión del desempeño, el desarrollo de la carrera o la identificación de necesidades formativas (Murphy, 2015).

Los resultados no parecen haber sido afectados negativamente por la Deseabilidad Social, variable que ha presentado correlaciones de tamaño reducido con el resto de constructos. Lo mismo cabe decir con respecto a la Varianza del método común, cuestión abordada en el Capítulo 4 tanto desde un punto de vista procedimental como desde un punto de vista estadístico, con resultados que indican que este potencial problema no ha tenido incidencia significativa en los resultados obtenidos.

En conjunto, cabe concluir que se han cumplido los objetivos establecidos al inicio de esta investigación y que los resultados obtenidos suponen importantes contribuciones para la Psicología del Trabajo y las Organizaciones y la gestión de recursos humanos, especialmente desde la óptica del cumplimiento de los dos valores principales en materia de selección de personal: la eficacia y la igualdad.

Referencias

- Abrahams, L., Pancorbo, G., Primi, R., Santos, D., Kyllonen, P., John, O. P., & De Fruyt, F. (2019). Social-emotional skill assessment in children and adolescents: Advances and challenges in personality, clinical, and educational contexts. *Psychological Assessment, 31*(4), 460-473. <http://dx.doi.org/10.1037/pas0000591>
- Aguado, D., Lucia, B., Ponte, G., y Arranz, V. (2008). Análisis inicial de las propiedades psicométricas del cuestionario BFCP Internet para la evaluación de Big-Five. *Revista Electrónica de Metodología Aplicada, 13*(2), 1-15. <https://www.unioviado.es/reunido/index.php/Rema/article/download/9781/9524>
- Aguilar, A., Tous, J. M., y Andrés-Pueyo, A. (1990). Adaptación y estudio psicométrico del EPQ-R. *Anuario de psicología/The UB Journal of Psychology, 46*, 101-118. <https://revistes.ub.edu/index.php/Anuario-psicologia/article/download/9351/12026>
- Akinbode, G. A. (2011). Demographic and dispositional characteristics as predictors of organizational citizenship behavior. *Ife Psychologia, 19* (1), 376-403. <https://doi.org/10.4314/ife.v19i1.64609>
- Allen, T.D., & Rush, M.C. (2001). The influence of ratee gender on ratings of organizational citizenship behavior. *Journal of Applied Social Psychology, 31* (12), 2561-2587. <https://doi.org/10.1111/j.1559-1816.2001.tb00191.x>
- Allik, J., & McCrae, R. R. (2004). Escapable conclusions: Toomela (2003) and the universality of trait structure. *Journal of Personality and Social Psychology, 87*(2), 261-265. <http://dx.doi.org/10.1037/0022-3514.87.2.261>
- Allik, J., & Realo, A. (2017). Universal and specific in the Five Factor Model of personality. In T. A. Widiger (Ed.), *The Oxford handbook of the five factor model* (pp.173-190). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199352487.013.23>
- Allik, J., & Realo, A. (2019). Culture and personality. In D. Matsumoto (Ed.), *The Handbook of Culture and Psychology* (pp. 401-430). Oxford University Press. <https://doi.org/10.1093/oso/9780190679743.003.0013>
- Allport, G. W., & Odbert, H. S. (1936). Trait names. A psycho-lexical study. *Psychological Monographs, 47*(211), 1-171. <http://dx.doi.org/10.1037/h0093360>
- Allworth, E., & Hesketh, B. (1999). Construct-oriented biodata: Capturing change-related and contextually relevant future performance. *International Journal of Selection and Assessment, 7*, 97-111. <http://dx.doi.org/10.1111/1468-2389.00110>
- Almahamid, S., McAdams, A. C., & Kalaldehy, T. (2010). The relationships among organizational knowledge sharing practices, employees' leaning commitments, employees' adaptability, and employees' job satisfaction: An empirical investigation of the listed manufacturing companies in Jordan. *Interdisciplinary Journal of Information, Knowledge, and Management, 5*, 327-356. <https://doi.org/10.28945/1225>

- Alonso, P., Moscoso, S., & Salgado, J. F. (2017). Structured behavioral interview as a legal guarantee for ensuring equal employment opportunities for women: A meta-analysis. *The European Journal of Psychology Applied to Legal Context*, 9(1), 15-23. <http://dx.doi.org/10.1016/j.ejpal.2016.03.002>
- Aluja, A., y Blanch, A. (2002). Relaciones entre la estructura del 16PF-5 y el modelo de cinco grandes factores de personalidad. *Boletín de Psicología*, 75, 7-18. <https://bit.ly/2YomXkI>
- Aluja, A., Kuhlman, M., & Zuckerman, M. (2010). Development of the Zuckerman–Kuhlman–Aluja personality questionnaire (ZKA–PQ): A factor/facet version of the Zuckerman–Kuhlman personality questionnaire (ZKPQ). *Journal of Personality Assessment*, 92(5), 416-431. <http://dx.doi.org/10.1080/00223891.2010.497406>
- American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (1985). *Standards for Educational and Psychological Testing*. American Psychological Association.
- Anderson, N., Salgado, J. F., & Hülsheger, U. R. (2010). Applicant Reactions in Selection: Comprehensive meta-analysis into reaction generalization versus situational specificity. *International Journal of Selection and Assessment*, 18, 291-304. <https://doi.org/10.1111/j.1468-2389.2010.00512.x>
- Angleitner, A., Ostendorf, F., & John, O. P. (1990). Towards a taxonomy of personality descriptors in German: A psycho-lexical study. *European Journal of Personality. Special Issue: Personality language*, 4(2), 89-118. <https://doi.org/10.1002/per.2410040204>
- Angoff, W., & Cook, L. (1988). *Equating the scores of the Prueba de Aptitud Académica and the Scholastic Aptitude Test* (College Board Report No. 88-2). College Entrance Examination Board. <https://doi.org/10.1002/j.2330-8516.1988.tb00259.x>
- Aoki, T. (1971). A psycho-lexical study of personality trait words: Selection, classification and desirability ratings of 455 words. *Japanese Journal of Psychology*, 42(1), 1-13. <https://doi.org/10.4992/jjpsy.42.1>
- Aparicio, M. E., y Sánchez López, M. P. (1999). Los estilos de personalidad: su medida a través del inventario Millon de estilos de personalidad. *Anales de Psicología*, 15(2), 191-211. <https://revistas.um.es/analesps/article/download/30071/29261/>
- Aramburu-Zabala, L. (2005). Seleccionar sin discriminar. ¿Posibilidad o utopía? *Lan Harremanak: Revista de Relaciones Laborales*, 13, 153-166. https://www.ehu.es/ojs/index.php/Lan_Harremanak/article/download/3472/3106
- Ardelt, M. (2000). Intellectual versus wisdom-related knowledge: The case for a different kind of learning in the later years of life. *Educational Gerontology*, 26(8), 771-789. <https://doi.org/10.1080/036012700300001421>
- Arribas, D. (2014). Algunos apuntes sobre los tests psicológicos y su uso para la selección de personal en la administración pública. *Revista Vasca de Gestión de Personas y Organizaciones Públicas*, 7, 24-35. <https://bit.ly/3hkUN2v>

- Arribas, D., y Pereña, J. (2009). *CompeTEA. Manual*. TEA Ediciones.
- Arribas-Águila, D. (2011). Psychometric properties of the TEA Personality Test: Evidence of reliability and construct validity. *European Journal of Psychological Assessment*, 27(2), 121-126. <https://doi.org/10.1027/1015-5759/a000057>
- Arthur, W., Bell, S.T., Villado, A.J., & Doverspike, D. (2006). The use of person-organization fit in employment decision making: an assessment of its criterion-related validity. *Journal of Applied Psychology*, 91, 786–801. <https://doi.org/10.1037/0021-9010.91.4.786>
- Ashton, M. C., Lee, K., & Son, C. (2000). Honesty as the sixth factor of personality: Correlations with Machiavellianism, primary psychopathy, and social adroitness. *European Journal of Personality*, 14, 359–368. [https://doi.org/10.1002/1099-0984\(200007/08\)14:4<359::aid-per382>3.0.co;2-y](https://doi.org/10.1002/1099-0984(200007/08)14:4<359::aid-per382>3.0.co;2-y)
- Auchter, J., & Stansfield, C. (1997). *Developing parallel tests across languages: Focus on the translation and adaptation process*. GED Testing Service. https://www.researchgate.net/publication/234568653_Developing_Parallel_Tests_across_Languages_Focus_on_the_Translation_and_Adaptation_Process
- Avdic, A. (2012). *Criterion-related validity of narrow-trait personality for predicting job performance, and the test of mediating mechanisms* (Order No. 3514364) [Doctoral dissertation, Southern Illinois University at Carbondale]. ProQuest Dissertations & Theses Global. (1027904984).
- Baard, S. K., Rench, T. A., & Kozlowski, S. W. (2014). Performance adaptation: A theoretical integration and review. *Journal of Management*, 40(1), 48-99. <https://doi.org/10.1177/0149206313488210>
- Babcock-Roberson, M. E., & Strickland, O. J. (2010). The relationship between charismatic leadership, work engagement, and organizational citizenship behaviors. *The Journal of Psychology, Interdisciplinary and Applied*, 144, 313-326. <https://doi.org/10.1080/00223981003648336>
- Baker, F. B. (2001). *The basics of item response theory*. ERIC. <https://eric.ed.gov/?id=ED458219>
- Balluerka, N., y Gorostiaga, A. (2012). Elaboración de versiones reducidas de instrumentos de medida: una perspectiva práctica. *Psychosocial Intervention*, 21(1), 103-110. <https://doi.org/10.5093/in2012v21n1a7>
- Balluerka, N., Gorostiaga, A., Alonso-Arbiol, I., y Haranburu, M. (2007). La adaptación de instrumentos de medida de unas culturas a otras: Una perspectiva práctica. *Psicothema*, 19 (1), 124–133. <http://www.psicothema.com/pdf/3338.pdf>
- Balluerka, N., Gorostiaga, A., Gómez-Benito, J., & Hidalgo-Montesinos, M. D. (2010). Use of multilevel logistic regression to identify the causes of differential item functioning. *Psicothema*, 22(4), 1018-1025. <http://www.psicothema.com/pdf/3835.pdf>

- Balluerka, N., Gorostiaga, A., Isasi, X., & Huizi, P. (2009, June 29- July 3). *Key issues for a proper adaptation of Psychological Tests from Spanish into Basque* [Symposium presentation]. Symposium “Quantitative and qualitative approaches to validate psychological and educational questionnaires in cross-cultural research”, 3rd Conference of the European Survey Research Association, Warsaw, Poland. <https://www.europeansurveyresearch.org/sites/default/files/abstracts.pdf>
- Balluerka, N., Plewis, I., Gorostiaga, A., & Padilla, J. L. (2014). Examining sources of DIF in psychological and educational assessment using multilevel logistic regression. *Methodology*, 10(2), 71-79. <https://doi.org/10.1027/1614-2241/a000076>
- Bangerter, A., Roiulin, N., & Köning, C. J. (2012). Personnel selection as a signaling game. *Journal of Applied Psychology*, 97, 719–739. <https://doi.org/10.1037/a0026078>
- Barrick, M. R., Feild, H. S. & Gatewood, R. (2011). The process of personality assessment. In M. R. Barrick, H. S. Feild, & R. Gatewood, (Eds.), *Selection in Human Resource Management* (7th ed., pp. 451-475). South-Western Cengage Learning.
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1–26. <https://doi.org/10.1111/j.1744-6570.1991.tb00688.x>
- Barrick, M. R., & Mount, M. K. (2003). Impact of meta-analysis methods on understanding personality-performance relations. In K. R. Murphy (Ed.), *Validity generalization: A critical review* (pp. 197-221). Lawrence Erlbaum Associates Publishers.
- Barrick, M. R., & Mount, M. K. (2014). Nature and use of personality in selection. In Neal Schmitt (Ed.), *The Oxford handbook of personnel assessment and selection* (pp. 225-251). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199732579.013.0011>
- Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, 9, 9-30. <https://doi.org/10.1111/1468-2389.00160>
- Barrick, M. R., & Parks-Leduc, L. (2019). Selection for fit. *Annual Review of Organizational Psychology and Organizational Behavior*, 6, 171-193. <http://dx.doi.org/10.1146/annurev-orgpsych-012218-015028>
- Bartram, D. (2005). The Great Eight competencies: a criterion-centric approach to validation. *Journal of Applied Psychology*, 90(6), 1185-1203. <https://doi.org/10.1037/0021-9010.90.6.1185>
- Baumgarten, F. (1933). Die Charaktereigenschaften. In *Beitrage zur Charakter- und Personlichkeitsforschung*. A. Francke.
- Beck, J.W., Beatty, A.S., & Sackett P.R. (2014). On the distribution of job performance: the role of measurement characteristics in observed departures from normality. *Personnel Psychology*, 67, 531–566. <https://doi.org/10.1111/peps.12060>

- Bell, B. S., & Kozlowski, S. W. (2002). Adaptive guidance: Enhancing self-regulation, knowledge, and performance in technology-based training. *Personnel Psychology*, 55(2), 267-306. <https://doi.org/10.1111/j.1744-6570.2002.tb00111.x>
- Benet-Martinez, V., & John, O. P. (1998). Los Cinco Grandes across cultures and ethnic groups: Multitrait-multimethod analyses of the Big Five in Spanish and English. *Journal of Personality and Social Psychology*, 75 (3), 729-750. <https://doi.org/10.1037/0022-3514.75.3.729>
- Berg, J. M., Wrzesniewski, A., & Dutton, J. E. (2010). Perceiving and responding to challenges in job crafting at different ranks: When proactivity requires adaptivity. *Journal of Organizational Behavior*, 31, 158–186. <https://doi.org/10.1002/job.645>
- Bernardin, H.J., & Beatty, R. (1984). *Performance appraisal: Assessing human behaviour at work*. Kent-PWS.
- Bernerth, J. B., & Aguinis, H. (2016). A critical review and best-practice recommendations for control variable usage. *Personnel Psychology*, 69(1), 229-283. <https://doi.org/10.1111/peps.12103>
- Bettencourt, L. A. (2004). Change-oriented organizational citizenship behaviors: The direct and moderating influence of goal orientation. *Journal of Retailing*, 80, 165–180. <https://doi.org/10.1016/j.jretai.2003.12.001>
- Bidshahri, R. (2017). *7 Critical Skills for the Jobs of the Future*. <https://singularityhub.com/2017/07/04/7-critical-skills-for-the-jobs-of-the-future/#sm.0000ryujjgncdixim2da4e5tinz>
- Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining twenty-first century skills. In P. Griffin, B. McGaw, & E. Care (Eds.), *Assessment and teaching of 21st century skills* (pp. 17–66). Springer Netherlands. https://doi.org/10.1007/978-94-007-2324-5_2
- Bleidorn, W., Hill, P., Back, M., Denissen, J., Hennecke, M., Hopwood, C., Jokela, M., Kandler, C, Lucas R.E., Luhmann, M., Orth, U., Wagner, J., Wrzus, C., Zimmermann, J, & Roberts, B. (2019). The Policy Relevance of Personality Traits. *American Psychologist*, 74(9), 1056–1067. <https://doi.org/10.7892/boris.136811>
- Boies, K., Lee, K., Ashton, M. C., Pascal, S., & Nicol, A. A. M. (2001). The structure of the French personality lexicon. *European Journal of Personality*, 15(4), 277-295. <https://doi.org/10.1002/per.411>
- Bolden, R., & Gosling, J. (2006). Leadership competencies: time to change the tune? *Leadership*, 2(2), 147-163. <https://doi.org/10.1177/1742715006062932>
- Bono, J. E., Shen, W., Yoon, D. J., & Day, D. (2014). Personality and leadership: Looking back, looking ahead. In David V. Day (Ed.), *The Oxford handbook of leadership and organizations* (pp. 99-218). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199755615.013.042>

- Booth, T., & Hughes, D. J. (2014). Exploratory structural equation modeling of personality data. *Assessment*, 21(3), 260–271. <https://doi.org/10.1177/1073191114528029>
- Borgatta, E. F. (1964). The structure of personality characteristics. *Behavioral Science*, 9, 8-17. <http://dx.doi.org/10.1002/bs.3830090103>
- Borkenau, P., & Ostendorf, F. (1990). Comparing exploratory and confirmatory factor analysis: A study on the 5-factor model of personality. *Personality and Individual Differences*, 11(5), 515-524. [https://doi.org/10.1016/0191-8869\(90\)90065-y](https://doi.org/10.1016/0191-8869(90)90065-y)
- Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt, & W.C. Borman (Ed.), *Personnel selection in organizations* (pp. 71-98). Jossey-Bass. https://scholarcommons.usf.edu/psy_facpub/1111
- Borman, W. C., Penner, L. A., Allen, T. D., & Motowidlo, S. J. (2001). Personality predictors of citizenship performance. *International Journal of Selection and Assessment*, 9, 52-69. <https://doi.org/10.1111/1468-2389.00163>
- Bosco, F. A., Aguinis, H., Singh, K., Field, J. G., & Pierce, C. A. (2015). Correlational effect size benchmarks. *Journal of Applied Psychology*, 100(2), 431-449. <https://doi.org/10.1037/a0038047>
- Bourdeau, N. R., & Lock, J. D. (2005, April 16). *Evaluating Applicant Faking via “Bright” and “Dark-Side” Measures of Personality* [Conference presentation]. 19th Annual Conference of the Society for Industrial and Organizational Psychology, Los Angeles, CA. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.123.1164&rep=rep1&type=pdf>
- Bowling, N. A., & Burns, G. N. (2010). A comparison of work-specific and general personality measures as predictors of work and non-work criteria. *Personality and Individual Differences*, 49 (2), 95-101. <https://doi.org/10.1016/j.paid.2010.03.009>
- Briggs, S.R., & Cheek, J.M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality*, 54, 106-148. <https://doi.org/10.1111/j.1467-6494.1986.tb00391.x>
- Bright, L. (2008). Does public service motivation really make a difference on the job satisfaction and turnover intentions of public employees? *American Review of Public Administration*, 38, 149-166. <https://doi.org/10.1177/0275074008317248>
- Brislin, R. W. (1970). Back-translation for cross-cultural research. *Cross-Cultural Psychology*, 1, 185-216. <https://doi.org/10.1177/135910457000100301>
- Brislin, R.W. (1986). The wording and translation of research instruments. In W.J. Lonner & J.W. Berry (Eds.), *Field methods in cross-cultural research* (pp. 137-164). Sage.
- Brokken, F. B. (1978). *The language of personality* [Unpublished doctoral dissertation, University of Groningen].

- Broschak, J. P., Davis-Blake, A., & Block, E. S. (2008). Nonstandard, not substandard: The relationship among work arrangements, work attitudes, and job performance. *Work and Occupations*, 35(1), 3-43. <https://doi.org/10.1177/0730888407309604>
- Brown, A., & Maydeu-Olivares, A. (2011). Item response modeling of forced-choice questionnaires. *Educational and Psychological Measurement*, 71(3), 460-502. <https://doi.org/10.1177/0013164410375112>
- Burisch, M. (1984). Approaches to personality inventory construction: A comparison of merits. *American Psychologist*, 39, 214–227. <https://doi.org/10.1037/0003-066x.39.3.214>
- Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14, 365-376. <http://dx.doi.org/10.1038/nrn3475>
- Buurman, M., Delfgaauw, J., Dur, R., & Van den Bossche, S. (2012). Public sector employees: Risk averse and altruistic? *Journal of Economic Behavior & Organization*, 83(3), 279-291. <http://dx.doi.org/10.1016/j.jebo.2012.06.003>
- Cáceres, G. A. (2014). *Interacción de las variables nivel jerárquico en la organización y percepción de distancia al poder sobre el ajuste persona-organización. Análisis del personal administrativo y docente de la PUCE* [Bachelor Thesis, Pontificia Universidad Católica de Ecuador]. Repositorio Digital PUCE <http://repositorio.puce.edu.ec/handle/22000/9749>
- Caicedo, E., & Zalazar-Jaime, M. F. (2018). Entrevistas cognitivas: revisión, directrices de uso y aplicación en investigaciones psicológicas. *Avaliação Psicológica*, 17(3), 362-370.
- Calanna, P., Lauriola, M., Saggino, A., Tommasi, M., & Furlan, S. (2020). Using a supervised machine learning algorithm for detecting faking good in a personality self-report. *International Journal of Selection and Assessment*, 28(2), 176-185. <https://doi.org/10.1111/ijsa.12279>
- Calarco, H. (2016). *Measuring the Relationship between Adaptive Performance and Job Satisfaction* (Order No. 10146852) [Doctoral dissertation, Middle Tennessee State University]. ProQuest Dissertations & Theses Global. (1829548928).
- Camerer, C. F., Dreber, A., Forsell, E., Ho, T. H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T, Razen, M, & Wu H. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, 351(6280), 1433-1436. <https://doi.org/10.1126/science.aaf0918>
- Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (pp. 687–732). Consulting Psychologists Press.

- Campbell, J. P. (1999). The definition and measurement of performance in the new age. In D. R. Ilgen & E. D. Pulakos (Eds.), *The changing nature of performance: Implications for staffing, motivation, and development* (pp. 399–430). Jossey-Bass.
- Campbell, J. P., Gasser, M. B., & Oswald, F. L. (1996). The substantive nature of performance variability. In K. R. Murphy (Ed.), *Individual differences and behavior in organizations* (pp. 258-299). Jossey-Bass.
- Campbell, J. P., & Wiernik, B. M. (2015). The modeling and assessment of work performance. *Annual Review of Organizational Psychology and Organizational Behavior*, 2(1), 47–74. <http://dx.doi.org/10.1146/annurev-orgpsych-032414-111427>
- Caprara, G.V., Barbaranelli, C., & Borgogni, L. (2007). *BFQ. Cuestionario “Big Five”*. TEA Ediciones.
- Caprara, G. V., & Perugini, M. (1994). Personality described by adjectives: The generalizability of the Big Five to the Italian lexical context. *European Journal of Personality*, 8(5), 357-369. <https://doi.org/10.1002/per.2410080502>
- Caprara, G. V., Vecchione, M., Barbaranelli, C., & Borgogni, L. (2008). *BFQ-2: Big five Questionnaire-2*. Giunti Organizzazioni Speciali.
- Cardoso-Ribeiro, C., Gómez-Conesa, A., y Hidalgo-Montesinos, M.D. (2010). Metodología para la adaptación de instrumentos de evaluación. *Fisioterapia*, 32, 6, 264-270. <http://dx.doi.org/10.1016/j.ft.2010.05.001>
- Carlo, G., Knight, G. P., Roesch, S. C., Opal, D., & Davis, A. (2014). Personality across cultures: A critical analysis of Big Five research and current directions. In F. T. L. Leong, L. Comas-Diaz, G. C. Nagayama Hall, V. C. McLoyd, & J. E. Trimble(Eds.), *APA handbook of multicultural psychology, vol. 1: Theory and research* (pp. 285–298). American Psychological Association. <https://doi.org/10.1037/14189-015>
- Carpenter, N. C., Berry, C. M., & Houston, L. (2014). A meta-analytic comparison of self-reported and other-reported organizational citizenship behavior. *Journal of Organizational Behavior*, 35(4), 547-574. <https://doi.org/10.1002/job.1909>
- Carpini, J. A., & Parker, S. K. (2017). The bigger picture: How OCBs fit within a broader conceptualization of work performance. *Oxford Handbook of Organizational Citizenship Behavior*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190219000.013.3>
- Carr, V., & Patterson, F. (2009, January 14-16). *Developing assessment centre methodologies for high-stakes selection: Implications for policy* [Conference presentation]. British Psychological Society, Division of Occupational Psychology Conference, Blackpool, UK. <https://www.workpsychologygroup.com/wp-content/uploads/2016/12/Library-developing-AC-for-high-stakes-selection-DOP.pdf>
- Carretero-Dios, H. y Pérez, C. (2005). Normas para el desarrollo y revisión de estudios instrumentales. *International Journal of Clinical and Health Psychology*, 5, 521-551.

- Cascio, W. F. (1995). Whither industrial and organizational psychology in a changing world of work. *American Psychologist*, 50, 928-939. <https://doi.org/10.1037/0003-066x.50.11.928>
- Caspi, A., & Moffitt, T. E. (1993). When do individual differences matter? A paradoxical theory of personality coherence. *Psychological Inquiry*, 4(4), 247-271. https://doi.org/10.1207/s15327965pli0404_1
- Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology*, 56, 453-484. <http://dx.doi.org/10.1146/annurev.psych.55.090902.141913>
- Cattell, R. B. (1943). The description of personality: basic traits resolved into clusters. *The Journal of Abnormal and Social Psychology*, 38(4), 476-506. <http://dx.doi.org/10.1037/h0054116>
- Cattell, R. B. (1945). The description of personality: principles and findings in a factor analysis. *American Journal of Psychology*, 58, 69-90. <http://dx.doi.org/10.2307/1417576>
- Cattell, R. B. (1947). Confirmation and clarification of primary personality factors. *Psychometrika*, 12, 197-220. <http://dx.doi.org/10.1007/BF02289253>
- Cattell, R. B., Cattell, A. K., & Cattell, H. E. P. (1993). *16PF Fifth Edition Questionnaire*. Institute for Personality and Ability Testing.
- Chan, D., & Schmitt, N. (2002). Situational judgment and job performance. *Human Performance*, 15, 233-254. http://dx.doi.org/10.1207/S15327043HUP1503_01
- Charbonnier-Voirin, A. (2013). L'influence des traits de personnalité sur les composantes de la performance adaptative des salariés. *Revue de Gestion des Ressources Humaines*, 88(2), 17-32. <http://dx.doi.org/10.3917/grhu.088.0017>
- Charbonnier-Voirin, A., El Akremi, A., & Vandenberghe, C. (2010). A multilevel model of transformational leadership and adaptive performance and the moderating role of climate for innovation. *Group & Organization Management*, 35(6), 699-726. <http://dx.doi.org/10.1177/1059601110390833>
- Charbonnier-Voirin, A., & Roussel, P. (2012). Adaptive performance: a new scale to measure individual performance in organizations. *Canadian Journal of Administrative Sciences*, 29(3), 280-293. <http://dx.doi.org/10.1002/cjas.232>
- Chatman, J. A. (1989). Improving interactional organizational research: A model of person-organization fit. *Academy of Management Review*, 14, 333-349. <https://doi.org/10.5465/amr.1989.4279063>
- Chen, G., Thomas, B., & Graig, W. (2005). A multilevel examination of the relationships among training outcomes, mediating regulatory processes, and adaptive performance. *Journal of Applied Psychology*, 90(5), 827-841. <https://doi.org/10.1037/0021-9010.90.5.827>

- Chiaburu, D. S., Oh, I. S., Berry, C. M., Li, N., & Gardner, R. G. (2011). The five-factor model of personality traits and organizational citizenship behaviors: a meta-analysis. *Journal of Applied Psychology*, 96 (6), 1140-1166. <https://doi.org/10.1037/a0024004>
- Chiaburu, D. S., Oh, I. S., Wang, J., & Stoverink, A. C. (2017). A bigger piece of the pie: The relative importance of affiliative and change-oriented citizenship and task performance in predicting overall job performance. *Human Resource Management Review*, 27(1), 97-107. <https://doi.org/10.1016/j.hrmr.2016.09.006>
- Choragwicka, B. (2010). *Las facetas de los Big Five y la predicción del desempeño* [Tesis doctoral, Universidad de Santiago de Compostela]. Minerva. <http://hdl.handle.net/10347/2884>
- Chou, S., & Pearson, J. (2011). A demographic study of information technology professionals' organizational citizenship behavior. *Journal of Management Research*, 3(2), 1-15. <https://doi.org/10.5296/jmr.v3i2.625>
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309-319. <http://dx.doi.org/10.1037/1040-3590.7.3.309>
- Clark, L.A., & Watson, D. (2019) Constructing validity: New developments in creating objective measuring instruments. *Psychological Assessment*, 31(12), 1412–1427. <http://dx.doi.org/10.1037/pas0000626>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. LEA. <https://doi.org/10.4324/9780203771587>
- Congreso De Los Diputados. (2007, 13 de abril). *Ley 7/2007, de 12 de abril, del Estatuto Básico del Empleado Público*. Boletín Oficial de las Cortes Generales, 94-14, 331-367. <https://www.boe.es/buscar/pdf/2007/BOE-A-2007-7788-consolidado.pdf>
- Congreso de los Diputados. (2015, 31 de octubre). *Real Decreto Legislativo 5/2015, de 30 de octubre, por el que se aprueba el texto refundido de la Ley del Estatuto Básico del Empleado Público*. Boletín Oficial del Estado 261, 103105-103159. <https://www.boe.es/buscar/doc.php?id=BOE-A-2015-11719>
- Conway, J. M. (1999). Distinguishing contextual performance from task performance for managerial jobs. *Journal of Applied Psychology*, 84(1), 3-13. <https://doi.org/10.1037/0021-9010.84.1.3>
- Conway, J. M., & Lance, C. E. (2010). What reviewers should expect from authors regarding common method bias in organizational research. *Journal of Business and Psychology*, 25(3), 325-334. <https://doi.org/10.1007/s10869-010-9181-6>
- Cooper, C. A., Knotts, H. G., McCord, D. M., & Johnson, A. (2013). Taking Personality Seriously. The Five-Factor Model and Public Management. *The American Review of Public Administration*, 43(4), 397-415. <https://doi.org/10.1177/0275074012446509>
- Corral, S., Pereña, J., Pamos, A., y Seisdedos, N. (2002). *TPT. Test de personalidad de TEA*. TEA Ediciones.

- Cortina, J. M., & Luchman, J. N. (2012). Personnel selection and employee performance. In I. B. Weiner, N. W. Schmitt & S. Highhouse (Eds.), *Handbook of psychology: Industrial and organizational psychology* (2nd ed., Vol. 12, pp. 143-183). Wiley. <https://doi.org/10.1002/9781118133880.hop212007>
- Costa, P. T., Herbst, J. H., McCrae, R. R., & Siegler, I. C. (2000). Personality at midlife: Stability, intrinsic maturation, and response to life events. *Assessment*, 7(4), 365–378. <http://dx.doi.org/10.1177/107319110000700405>
- Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO personality inventory manual*. Psychological Assessment Resources.
- Costa, P. T., Jr., & McCrae, R. R. (1989). *The NEO-PI/NEO-FFI manual supplement*. Psychological Assessment Resources.
- Costa, P. T., & McCrae, R. R. (1992a). Four ways five factors are basic. *Personality and Individual Differences*, 135, 653–665. [http://dx.doi.org/10.1016/0191-8869\(92\)90236-I](http://dx.doi.org/10.1016/0191-8869(92)90236-I)
- Costa, P. T., & McCrae, R. R. (1992b). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Psychological Assessment Resources, Inc.
- Costa, P. T., Jr., & McCrae, R. R. (1994). Set like plaster? Evidence for the stability of adult personality. In T. F. Heatherton & J. L. Weinberger (Eds.), *Can personality change?* (pp. 21–40). American Psychological Association. <https://doi.org/10.1037/10143-002>
- Costa Jr., P. T., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of Personality Assessment*, 64(1), 21-50. http://dx.doi.org/10.1207/s15327752jpa6401_2
- Costa Jr., P. T., & McCrae, R. R. (2006). Trait and factor theories. In *Comprehensive handbook of personality and psychopathology* (Vol. 1, pp. 96-114). Wiley.
- Costa, P.T., & McCrae, R. R. (2008). *Inventario de personalidad NEO revisado (NEO-PI-R) y Inventario NEO reducido de cinco factores (NEOFFI). Manual profesional* (3ª ed.). TEA Ediciones.
- Costa Jr., P. T., McCrae, R. R., & Löckenhoff, C. E. (2019). Personality across the life span. *Annual Review of Psychology*, 70, 423-448. <http://dx.doi.org/10.1146/annurev-psych-010418-103244>
- Cova, F., Strickland, B., Abatista, A., Allard, A., Andow, J., Attie, M., Beebe, J., Berniunas, R., Boudesseul, J., Colombo, M., Cushman, F., Diaz, R., van Dongen, N., Dranseika, V., Earp, B.D., Gaitán, A., Hannikainen, I., Hernández-Conde, J.V., Hu, W., ... & Zhou, Y. (2018). Estimating the reproducibility of experimental philosophy. *Review of Philosophy and Psychology*, 1-36. <http://dx.doi.org/10.1007/s13164-018-0400-9>
- Credé, M., Harms, P., Niehorster, S., & Gaye-Valentine, A. (2012). An evaluation of the consequences of using short measures of the big five personality traits. *Journal of Personality and Social Psychology*, 102, 874–888. <https://doi.org/10.1037/a0027403>

- Cronbach, L. J., & Gleser, C. G. (1957). *Psychological tests and personnel decisions*. University of Illinois Press.
- Crowley, M. L. (2012). *Predicting job adaptability: A facet-level examination of the relationship between conscientiousness and adaptive performance with autonomy as a moderator* (Order No. 10159473) [Doctoral dissertation, Purdue University]. ProQuest Dissertations & Theses Global. (1840894210).
- Cubiks. (2014). *PAPI3 User Manual*. Cubiks.
- Cullen, K. L., Edwards, B. D., Casper, W. C., & Gue, K. R. (2014). Employees' adaptability and perceptions of change-related uncertainty: Implications for perceived organizational support, job satisfaction, and performance. *Journal of Business and Psychology*, 29(2), 269-280. <https://doi.org/10.1007/s10869-013-9312-y>
- Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1(1), 16-29. <https://doi.org/10.1037/1082-989x.1.1.16>
- Damian, R. I., Spengler, M., Sutu, A., & Roberts, B. W. (2018). Sixteen going on sixty-six: A longitudinal study of personality stability and change across 50 years. *Journal of Personality and Social Psychology*, 117(3), 674-695. <https://doi.org/10.1037/pspp0000210>
- Dávila, M. C., & Finkelstein, M. A. (2010). Predicting organizational citizenship behavior from the functional analysis and role identity perspectives: Further evidence in Spanish employees. *The Spanish Journal of Psychology*, 13(1), 277-283. <https://doi.org/10.1017/s1138741600003851>
- Dávila, M. C., Finkelstein, M. A., & Castien, J. I. (2011). Gender differences in prosocial behavior: Organizational citizenship behavior. *Anales de Psicología*, 27(2), 498-506. <https://revistas.um.es/analesps/article/view/123111>
- De Bree, I. (2010). *Job insecurity and OCBs* [Bachelor Thesis, Tilburg University]. <http://arno.uvt.nl/show.cgi?fid=111156>
- De Jong, A., & De Ruyter, K. (2004). Adaptive versus proactive behavior in service recovery: The role of self-managing teams. *Decisions Sciences*, 35(3), 457-491. <https://doi.org/10.1111/j.0011-7315.2004.02513.x>
- De la Nuez, E. (2018). Acceso a la función pública: atraer talento y cambiar el modelo. *Revista Vasca de Gestión de Personas y Organizaciones Públicas, nº especial 2*, 86-97. <https://bit.ly/37mx8KB>
- De Raad, B. (1992). The replicability of the big five dimensions in three word-classes of the Dutch language. *European Journal of Personality*, 6(1)15-29. <https://doi.org/10.1002/per.2410060103>

- De Raad, B. (2000). *The big five personality factors: the psycholexical approach to personality*. Hogrefe and Huber. <https://www.researchgate.net/publication/232576768> The Big Five Personality Factors The psycholexical approach to personality
- De Raad, B., Hendriks, A. A. J., & Hofstee, W. K. (1992). Towards a refined structure of personality traits. *European Journal of Personality*, 6(4), 301-319. <https://doi.org/10.1002/per.2410060405>
- DeArmond, S., Tye, M., Chen, P. Y., Krauss, A., Apryl Rogers, D., & Sintek, E. (2006). Age and gender stereotypes: New challenges in a changing workplace and workforce. *Journal of Applied Social Psychology*, 36(9), 2184-2214. <https://doi.org/10.1111/j.0021-9029.2006.00100.x>
- Delamare le Deist, F., & Winterton, J. (2005). What Is Competence? *Human Resource Development International*, 8(1), 27-46. <http://dx.doi.org/10.1080/1367886042000338227>
- Delgado-Rodríguez, N., Hernández-Fernaud, E., Rosales, C., Díaz-Vilela, L., Isla-Díaz, R., & Díaz-Cabrera, D. (2018). Contextual Performance in Academic Settings: The Role of Personality, Self-efficacy, and Impression Management. *Revista de Psicología del Trabajo y de las Organizaciones*, 34(2), 63-68. <http://dx.doi.org/10.5093/jwop2018a8>
- Demerouti, E., Xanthopoulou, D., Tsaousis, I., & Bakker, A. B. (2014). Disentangling task and contextual performance: A multitrait-multimethod approach. *Journal of Personnel Psychology*, 13(2), 59-69. <https://doi.org/10.1027/1866-5888/a000104>
- Demmke, C. (2007). *Els funcionaris són diferents? Qui són els funcionaris? I com són?* Escola d'Administració Pública de Catalunya. http://eapc.gencat.cat/web/.content/home/publicacions/varia/25._els_funcionaris_s_n_diferents_qui_s_n_els_funcionaris_i_com_s_n/els_funcionaris.pdf
- DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, 93, 880-896. <https://doi.org/10.1037/e633942013-310>
- Digman, J. M. (1990). Personality Structure: Emergence of the Five-Factor Model. *Annual Review of Psychology*, 41(1), 417-440. <http://dx.doi.org/10.1146/annurev.ps.41.020190.002221>
- Digman, J. M., & Takemoto-Chock, N. K. (1981). Factors in the natural language of personality: Re-analysis, comparison, and interpretation of six major studies. *Multivariate Behavioral Research*, 16(2), 149-170. https://doi.org/10.1207/s15327906mbr1602_2
- DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research & Evaluation*, 14(20), 1-11. <https://doi.org/10.7275/da8t-4g52>

- Donnellan, M. B., & Lucas, R. E. (2008). Age differences in the Big Five across the life span: Evidence from two national samples. *Psychology and Aging, 23*, 558–566. <https://doi.org/10.1037/a0012897>
- Dorsey, D. W., Cortina, J. M., & Luchman, J. (2010). Adaptive and citizenship-related behaviors at work. In J. L. Farr y N. T. Tippins (Eds.), *Handbook of personnel selection* (pp. 463-487). Routledge. <https://doi.org/10.4324/9781315690193-21>
- Drasgow, F. (1984). Scrutinizing psychological test: Measurement equivalence and equivalent relations with external variables are the central issues. *Psychological Bulletin, 95*, 134–135. <http://dx.doi.org/10.1037/0033-2909.95.1.134>
- Drasgow, F., & Kanfer, R. (1985). Equivalence of psychological measurement in heterogeneous populations. *Journal of Applied Psychology, 70*, 662–680. <http://dx.doi.org/10.1037/0021-9010.70.4.662>
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist, 41*, 1040–1048. <https://doi.org/10.1037/0003-066X.41.10.1040>
- Dweck, C. S. (2002). Beliefs that make smart people dumb. In R. J. Sternberg (Ed.), *Why smart people can be so stupid* (pp.24–41). Yale University Press.
- ECPA - Les Editions du Centre de Psychologie Appliquée. (2016). *IPV. Inventario de Personalidad para Vendedores. Manual*. TEA Ediciones.
- Edwards, B. D., Bell, S. T., Arthur Jr., W., & Decuir, A. D. (2008). Relationships between facets of job satisfaction and task and contextual performance. *Applied Psychology: An International Review, 57*, 441-465. <https://doi.org/10.1111/j.1464-0597.2008.00328.x>
- Ellingson, J.E., Sackett, P.R., & Connelly, B.S. (2007). Personality assessment across selection and development contexts: insights into response distortion. *Journal of Applied Psychology, 2*, 386-395. <https://doi.org/10.1037/0021-9010.92.2.386>
- Ellingson, J. E., Smith, D. B., & Sackett, P. R. (2001). Investigating the influence of social desirability on personality factor structure. *Journal of Applied Psychology, 86*(1), 122-133. <https://doi.org/10.1037/0021-9010.86.1.122>
- Elosua, P. (2012). Tests publicados en España: usos, costumbres y asignaturas pendientes. *Papeles del Psicólogo, 33*(1), 12-21. <http://www.papelesdelpsicologo.es/pdf/2031.pdf>
- Emons, W. H., Sijtsma, K., & Meijer, R. R. (2007). On the consistency of individual classification using short scales. *Psychological Methods, 12*, 105-120. <https://doi.org/10.1037/1082-989x.12.1.105>
- Epstein, J., Santo, R. M., & Guillemín, F. (2015). A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *Journal of Clinical Epidemiology, 68*(4), 435-441. <https://doi.org/10.1016/j.jclinepi.2014.11.021>
- Etxebarria, M. (1992). *La variación lingüística en comunidades bilingües y su incidencia en el léxico, en Bilingüismo y admisión del español*. Inst. Horizonte.
- Etxebarria, M. (2003). El español y el euskera en contacto. *LINRED Revista de Lingüística, 1*, 36-54.

- Eysenck, H. J. y Eysenck, S. B. G. (2001). *EPQ-R. Cuestionario revisado de Personalidad de Eysenck*. TEA Ediciones.
- Feist, G. J., & Barron, F. X. (2003). Predicting creativity from early to late adulthood: Intellect, potential, and personality. *Journal of Research in Personality*, 37(2), 62–88. [https://doi.org/10.1016/s0092-6566\(02\)00536-6](https://doi.org/10.1016/s0092-6566(02)00536-6)
- Ferguson, C. J. (2009). An effect size primer: a guide for clinicians and researchers. *Professional Psychology: Research and Practice*, 40(5), 532-538. <https://doi.org/10.1037/14805-020>
- Fernández-Seara, J. L., Seisdedos, N., y Mielgo, M. (2008). *CPS. Cuestionario de personalidad situacional. Manual*. TEA Ediciones.
- Ferrando, P. J., y Demestre, J. (2008). Características de forma y contenido que predicen la capacidad discriminativa en ítems de personalidad: un análisis basado en la Teoría de Respuesta a los Ítems. *Psicothema*, 20(4), 851-856. <http://www.psicothema.com/pdf/3566.pdf>
- Ferrando, P. J., & Lorenzo-Seva U. (2016). A note on improving EAP trait estimation in oblique factor-analytic and item response theory models. *Psicologica*, 37, 235-247. <https://www.uv.es/revispsi/articulos2.16/7Ferrando.pdf>
- Ferrando, P.J., Lorenzo-Seva, U., & Chico, E. (2009). A general factor analytic procedure for assessing response bias in questionnaire measures. *Structural Equation Modeling*, 16, 364-381. <http://dx.doi.org/10.1080/10705510902751374>
- Fiske, D. W. (1949). Consistency of the factorial structures of personality ratings from different sources. *The Journal of Abnormal and Social Psychology*, 44(3), 329-344. <https://doi.org/10.1037/h0057198>
- Floyd, F.J., & Widaman, K.F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7, 286-299. <https://doi.org/10.1037/1040-3590.7.3.286>
- Frame, M., Roberto, K., & Rigdon, W. (2006, September). *The UT-Arlington adaptive performance scale: A new measure for a not-so-new concept* [Conference presentation]. Dallas Area Industrial and Organizational Psychologists, Plano, TX.
- Fronczyk, K. (2019). Congruence and measurement invariance of self-report and informant-ratings of the Big Five dimensions. *Personality and Individual Differences*, 139, 7-12. <https://doi.org/10.1016/j.paid.2018.10.036>
- Fuller, C. M., Simmering, M. J., Atinc, G., Atinc, Y., & Babin, B. J. (2016). Common methods variance detection in business research. *Journal of Business Research*, 69(8), 3192-3198. <https://doi.org/10.1016/j.jbusres.2015.12.008>
- Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, 2(2), 156-168. <https://doi.org/10.1177/2515245919847202>

- Furnham, A., Guenole, N., Levine, S. Z., & Chamorro-Premuzic, T. (2013). The NEO Personality Inventory–Revised: Factor structure and gender invariance from exploratory structural equation modeling analyses in a high-stakes setting. *Assessment*, 20(1), 14-23. <http://dx.doi.org/10.1177/1073191112448213>
- García, O., Aluja, A., & García, L. F. (2004). Psychometric Properties of Goldberg's 50 Personality Markers for the Big Five Model: A Study in the Spanish Language. *European Journal of Psychological Assessment*, 20(4), 310-319. <https://doi.org/10.1027/1015-5759.20.4.310>
- Gardner, W. L., Reithel, B. J., Coglisier, C. C., Walumbwa, F. O., & Foley, R. T. (2012). Matching personality and organizational culture: Effects of recruitment strategy and the Five-Factor Model on subjective person–organization fit. *Management Communication Quarterly*, 26(4), 585-622. <https://doi.org/10.1177/0893318912450663>
- Gatewood, R., Feild, H. S., & Barrick, M. (2015). *Human resource selection*. Nelson Education.
- Gierl, M. J., Gotzmann, A., & Boughton, K. A. (2004). Performance of SIBTEST when the percentage of DIF items is large. *Applied Measurement in Education*, 17(3), 241-264. https://doi.org/10.1207/s15324818ame1703_2
- Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, 102, 74-78. <https://doi.org/10.1016/j.paid.2016.06.069>
- Gillespie, M. A., Oswald, F. L., Schmitt, N., Kim, B. H. & Ramsay, L. J. (2003, April 11-13). *The influence of personality traits on direct measures of PO fit* [Symposium presentation]. 18th Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL, USA.
- Gimeno, R. (2020, Febrero 5). El sector público ya supera en temporalidad al sector privado: 28% frente al 26%. *NIUS*. https://www.niusdiario.es/economia/empleo/sector-publico-supera-temporalidad-privado-precariedad-laboral-contratos-publicos-funcionarios-interinos_18_2894070364.html
- Gismero, E. (2000). *EHS: Escala de Habilidades Sociales*. TEA ediciones.
- Gobierno Vasco (2017). *VI Encuesta Sociolingüística*. Gobierno Vasco. https://www.euskadi.eus/contenidos/noticia/eas_mas_noticias/es_def/adjuntos/INKEST_A_ESt.pdf
- Goffin, R. D., & Christiansen, N. D. (2003). Correcting personality tests for faking: A review of popular personality tests and an initial survey of researchers. *International Journal of Selection and Assessment*, 11(4), 340-344. <https://doi.org/10.1111/j.0965-075x.2003.00256.x>
- Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology*, 59(6), 1216-1229. <https://doi.org/10.1037/0022-3514.59.6.1216>

- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4(1), 26-42. <https://doi.org/10.1037/1040-3590.4.1.26>
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48(1), 26-34. <https://doi.org/10.1037/0003-066X.48.1.26>
- Goldberg, L. R., Sweeney, D., Merenda, P. F., & Hughes Jr, J. E. (1998). Demographic variables and personality: The effects of gender, age, education, and ethnic/racial status on self-descriptions of personality attributes. *Personality and Individual Differences*, 24(3), 393-403. [http://dx.doi.org/10.1016/S0191-8869\(97\)00110-4](http://dx.doi.org/10.1016/S0191-8869(97)00110-4)
- Gómez-Benito, J., Hidalgo, M. D., & Guilera, G. (2010). El sesgo de los instrumentos de medición. Test justos. *Papeles del Psicólogo*, 31(1), 75-84. <http://papelesdelpsicologo.es/pdf/1798.pdf>
- Gómez-Benito, J., Sireci, S., Padilla, J. L., Hidalgo, M. D., & Benitez, I. (2018). Differential item functioning: Beyond validity evidence based on internal structure. *Psicothema*, 30(1), 104-109. <http://www.psicothema.com/pdf/4457.pdf>
- Gordon, L.V. (2001). *PPG-IPG. Perfil e inventario de personalidad*. TEA Ediciones.
- Gordon, L.V., ECPA, & TEA Ediciones. (1997). *SOSIA. Gestión por competencias. Manual*. TEA Ediciones.
- Gorostiaga, A., Balluerka, N., Alonso-Arbiol, I., & Haranburu, M. (2011). Validation of the Basque Revised NEO Personality Inventory (NEO PI-R). *European Journal of Psychological Assessment*, 27(3), 193–205. <https://doi.org/10.1027/1015-5759/a000067>
- Gorriti, M. (2007). La evaluación del desempeño en las administraciones públicas españolas. *Revista de Psicología del Trabajo y de las Organizaciones*, 23, 3, 367-387. <https://journals.copmadrid.org/jwop/files/105905.pdf>
- Gorriti, M. (2010). *Modelo de proceso selectivo por grupos de clasificación en la Administración pública*. CEMICAL. http://cemical.diba.cat/publicacions/fitxers/GORRITI_M_SeleccionAAPP.pdf
- Gorriti, M. (2018). Innovar en selección desde la evidencia empírica y las nuevas competencias. *Revista Vasca de Gestión de Personas y Organizaciones Públicas, Número especial 2*, 66-85. <https://bit.ly/2XSLpM0>
- Gorriti, M., y López, J. (2009). *Identificación de criterios de desempeño a partir del análisis de puestos*. IVAP
- Gorriti, M., y López, J. (2010). *Análisis de puestos en la Administración Pública*. IVAP.
- Gorriti, M., y López, J. (2014). *Guía para la validación de sistemas de acceso a la Función Pública*. IVAP.
- Gough, H. G., & Bradley, P. (1996). *California personality inventory manual*. Consulting Psychologists.

- Green, B. F., Bock, R. D., Humphreys, L. G., Linn, R. L., & Reckase, M. D. (1984). Technical guidelines for assessing computerized adaptive tests. *Journal of Educational Measurement*, 21(4), 347-360. <https://doi.org/10.1111/j.1745-3984.1984.tb01039.x>
- Greene, J., Mero, N., & Werner, S. (2018). The negative effects of job embeddedness on performance. *Journal of Managerial Psychology*, 33(1), 58-73. <https://doi.org/10.1108/jmp-02-2017-0074>
- Griffin, B., & Hesketh, B. (2003). Adaptable behaviours for successful work and career adjustment. *Australian Journal of Psychology*, 55, 65-73. <https://doi.org/10.1080/00049530412331312914>
- Griffin, B., & Hesketh, B. (2004). Why openness to experience is not a good predictor of job Performance. *International Journal of Selection and Assessment*, 12 (3), 243-251. https://doi.org/10.1111/j.0965-075x.2004.278_1.x
- Griffin, B., & Hesketh, B. (2005). Are conscientious workers adaptable? *Australian Journal of Management*, 30(2), 245-259. <https://doi.org/10.1177/031289620503000204>
- Griffin, M. A., Neal, A., & Parker, S. K. (2007). A new model of work role performance: Positive behavior in uncertain and interdependent contexts. *Academy of Management Journal*, 50(2), 327-347. <https://doi.org/10.5465/amj.2007.24634438>
- Grisay, A., de Jong, J.H.A.L., Gebhardt, E., Berezner, A., & Halleux-Monseur, B. (2007). Translation equivalence across PISA countries. *Journal of Applied Measurement*, 8(3), 249-266.
- Grisay, A., & Monseur, C. (2007). Measuring the equivalence of item difficulty in the various versions of an international test. *Studies in Educational Evaluation*, 33(1), 69-86. <https://doi.org/10.1016/j.stueduc.2007.01.006>
- Guerra, N., Modecki, K., & Cunningham, W. (2014). *Developing social-emotional skills for the labor market: The PRACTICE model*. The World Bank. <https://doi.org/10.1596/1813-9450-7123>
- Guest, D. E. (1997). Human resource management and performance: a review and research agenda. *International Journal of Human Resource Management*, 8(3), 263-276. <https://doi.org/10.1080/095851997341630>
- Guion, R. M. (1997). Criterion measures and the criterion dilemma. In N. Anderson & y P. Herriot (Eds.), *International handbook of selection and assessment* (pp. 267-286). Wiley.
- Guion, R. M. (2011). *Assessment, measurement, and prediction for personnel decisions*. Routledge. <https://doi.org/10.4324/9780203836767>
- Hall, J. A., Schlegel, K., Castro, V. L., & Back, M. (2019). What laypeople think the Big Five trait labels mean. *Journal of Research in Personality*, 78, 268-285. <https://doi.org/10.1016/j.jrp.2018.12.007>
- Hambleton, R. K. (1993). Translating achievement tests for use in cross-national studies. *European Journal of Psychological Assessment*, 9, 57-68

- Hambleton, R. K. (1994). Guidelines for adapting educational and psychological test: a progress report. *European Journal of Psychological Assessment, 10*, 229-244.
- Hambleton, R. K. (1996). Adaptación de test para su uso en diferentes idiomas y culturas: fuentes de error, posibles soluciones y directrices prácticas. En Muñiz J., (Eds.) *Psicometría*. Universitas.
- Hambleton, R. K., & De Jong, J. H. A. L. (2003). Advances in translating and adapting educational and psychological tests. *Language Testing, 20*(2), 127-134. <http://dx.doi.org/10.1191/0265532203lt247xx>
- Hambleton, R.K., & Kanjee, A. (1993, April 12-16). *Enhancing the Validity of Cross-Cultural Studies: Improvements in Instrument Translation Methods* [Conference presentation]. Annual Meetings of the American Educational Research Association, Atlanta, GA. ERIC.
- Hambleton, R. K., & Kanjee, A. (1995). Increasing the validity of cross-cultural assessments: Use of improved methods for test adaptations. *European Journal of Psychological Assessment, 11*(3), 147-157. <https://doi.org/10.1027/1015-5759.11.3.147>
- Hambleton, R.K., & 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*. <http://www.jattjournal.com/index.php/atp/article/view/48345>
- Hamtiaux, A., Houssemand, C., & Vrignaud, P. (2013). Individual and career adaptability: Comparing models and measures. *Journal of Vocational Behavior, 83*(2), 130-141. <https://doi.org/10.1016/j.jvb.2013.03.006>
- Han, T. Y., & Williams, K. J. (2008). Multilevel investigation of adaptive performance: Individual- and team-level relationships. *Group & Organization Management, 33*(6), 657-684. <https://doi.org/10.1177/1059601108326799>
- Hanel, P. H., Maio, G. R., & Manstead, A. S. (2019). A new way to look at the data: Similarities between groups of people are large and important. *Journal of Personality and Social Psychology, 116*(4), 541–562. <https://doi.org/10.1037/pspi0000154>
- Hanson, M. A., & Borman, W. C. (2006). Citizenship performance: An integrative review and motivational analysis. In W. Bennett, Jr., D. J. Woehr, & C. E. Lance (Eds.), *Performance measurement: Current perspectives and future challenges* (pp. 141- 173). Lawrence Erlbaum Associates.
- Haranburu, M., Gorostiaga, A., Alonso-Arbiol, I., & Balluerka, N. (2007, may 3-6). *The NEO FFI short version of the Basque NEO PI-R inventory of personality* [Poster presentation]. 9th European Conference on Psychological Assessment, Tesalónica, Grecia. https://www.researchgate.net/publication/292335342_The_NEO_FFI_short_version_of_the_Basque_NEO_PI-R_inventory_of_Personality
- Heene, M., Bollmann, S., & Bühner, M. (2014). Much ado about nothing, or much to do about something? Effects of scale shortening on criterion validity and mean differences. *Journal of Individual Differences, 35*, 245–249. <https://doi.org/10.1027/1614-0001/a000146>

- Heggestad, E. D., Morrison, M., Reeve, C. L., & McCloy, R. A. (2006). Forced-choice assessments of personality for selection: Evaluating issues of normative assessment and faking resistance. *Journal of Applied Psychology*, *91*, 9-24. <https://doi.org/10.1037/0021-9010.91.1.9>
- Heggestad, E. D., Scheaf, D. J., Banks, G. C., Monroe Hausfeld, M., Tonidandel, S., & Williams, E. B. (2019). Scale Adaptation in Organizational Science Research: A Review and Best-Practice Recommendations. *Journal of Management*, *45*(6), 2596-2627. <https://doi.org/10.1177/0149206319850280>
- Hendriks, A. A. J. (2002). *The construction of the Five-Factor Personality Inventory* [Unpublished Doctoral Thesis, University of Groningen]. UMCG research database. <http://hdl.handle.net/11370/092402c2-8a4d-4976-8b62-9e7a5d0d00ec>
- Hendriks, A. A.J., Perugini, M., Angleitner, A., Ostendorf, F., Johnson, J. A., De Fruyt, F., Hrebickova, M, Kreitler, S, Murakami T., Bratko, D, Conner, M., Nagy J., Rodríguez-Fornells, A., & Ruisel, I. (2003). The five-factor personality inventory: Cross-cultural generalizability across 13 countries. *European Journal of Personality*, *17*(5), 347- 373. <https://doi.org/10.1002/per.491>
- Henry, R. A., & Hulin, C. L. (1987). Stability of skilled performance across time: Some generalization and limitations on utilities. *Journal of Applied Psychology*, *72*, 457-462. <https://doi.org/10.1037/0021-9010.72.3.457>
- Herde, C. N., Lievens, F., Solberg, E. G., Harbaugh, J. L., Strong, M. H., & Burkholder, G. J. (2019). Situational Judgment Tests as Measures of 21st Century Skills: Evidence across Europe and Latin America. *Revista de Psicología del Trabajo y de las Organizaciones*, *35*(2), 65-74. <https://doi.org/10.5093/jwop2019a8>
- Herrmann, A., & Pfister, H. R. (2013). Simple measures and complex structures: Is it worth employing a more complex model of personality in Big Five inventories? *Journal of Research in Personality*, *47*(5), 599-608. <https://doi.org/10.1016/j.jrp.2013.05.004>
- Hesketh, B., & Feiler, H. (1994, April 6-9). *Using fuzzy rating scales for Graphic Distributional Performance Ratings* [Conference presentation]. 9th Annual Conference of the Society for Industrial and Organisational Psychology, Nashville, Tennessee.
- Hesketh, B., & Neal, A. (1999). Technology and performance. In D. R. Ilgen y E. D. Pulakos (Eds.), *The changing nature of performance: Implications for staffing, motivation, and development* (pp. 21-55). Jossey-Bass.
- Hibbing, M. V., Cawvey, M., Deol, R., Bloeser, A. J., & Mondak, J. J. (2019). The relationship between personality and response patterns on public opinion surveys: The Big Five, extreme response style, and acquiescence response style. *International Journal of Public Opinion Research*, *31*(1), 161-177. <https://doi.org/10.1093/ijpor/edx005>
- Hidalgo-Montesinos, M. D. y French, B. F. (2016). Una introducción didáctica a la Teoría de Respuesta al Ítem para comprender la construcción de escalas. *Revista de Psicología Clínica con Niños y Adolescentes*, *3*(2), 13-21. <https://www.revistapcna.com/sites/default/files/16-11.pdf>

- Hidalgo-Montesinos, M. D., Galindo-Garre, F., & Gómez-Benito, J. (2015). Differential item functioning and cut-off scores: Implications for test score interpretation. *Anuario de Psicología*, 45(1), 55-69. <http://diposit.ub.edu/dspace/bitstream/2445/99100/1/653303.pdf>
- Hoffman, B. J., & Dilchert, S. (2012). A review of citizenship and counterproductive behaviors in organizational decision-making. In Neal Schmitt (Ed.), *The Oxford Handbook of Personnel Selection and Selection* (pp. 543-569). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199732579.013.0024>
- Hoffman, B. J., & Woehr, D. J. (2006). A quantitative review of the relationship between person–organization fit and behavioral outcomes. *Journal of Vocational Behavior*, 68(3), 389-399. <https://doi.org/10.1016/j.jvb.2005.08.003>
- Hofstee, W. K. B., ten Berge, J. M. F., & Hendriks, A. A. J. (1998). How to score questionnaires. *Personality and Individual Differences*, 25, 897–909. [http://dx.doi.org/10.1016/S0191-8869\(98\)00086-5](http://dx.doi.org/10.1016/S0191-8869(98)00086-5)
- Hogan, J., Barrett, P., & Hogan R. (2007). Personality measurement, faking, and employment selection. *Journal of Applied Psychology*, 92, 1270-1285. <https://doi.org/10.1037/0021-9010.92.5.1270>
- Hogan, J., & Holland, B. (2003). Using theory to evaluate personality and job–performance relations: A socioanalytic perspective. *Journal of Applied Psychology*, 88, 100 –112. <https://doi.org/10.1037/0021-9010.88.1.100>
- Hogan, R. T. (1991). Personality and personality measurement. *Handbook of industrial and organizational psychology* (pp. 873-919). Consulting Psychologists Press.
- Hogan, R.T., & Hogan, J. (2007). *Hogan Personality Inventory manual* (3rd ed.). Hogan Assessment Systems.
- Holtrop, D., Born, M. P., de Vries, A., & de Vries, R. E. (2014). A matter of context: A comparison of two types of contextualized personality measures. *Personality and Individual Differences*, 68, 234-240. <https://doi.org/10.1016/j.paid.2014.04.029>
- Hossiep, R., & Paschen, M. (2008). *BIP. Inventario Bochum de Personalidad y Competencias. Manual*. TEA Ediciones.
- Hough, L. M., Oswald, F. L., & Ployhart, R. E. (2001). Determinants, detection and amelioration of adverse impact in personnel selection procedures: Issues, evidence and lessons learned. *International Journal of Selection and Assessment*, 9(1-2), 152-194. <https://doi.org/10.1111/1468-2389.00171>
- Howe, M. (2014). *Fool me once, shame on you; fool me twice, shame on me: An investigation of individual differences, goals, and adaptive performance in a multiple change context* (Order No. 3623712) [Doctoral Dissertation, Michigan State University]. ProQuest Dissertations & Theses Global. (1550352982).

- Hrebickova, M. (1995). *The structural model of personality based on the lexical analysis: A Czech replication study of the five-factor model based on a comprehensive taxonomy of personality-descriptive adjectives*. Unpublished manuscript, Institute of Psychology, Academy of Sciences of the Czech Republic, Brno.
- Hu, L., & Bentler, P. M. (1999). Cutoff criterion for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55. <https://doi.org/10.1080/10705519909540118>
- Huang, J. L., Ryan, A. M., Zabel, K. L., & Palmer, A. (2014). Personality and adaptive performance at work: A meta-analytic investigation. *Journal of Applied Psychology*, 99(1), 162-179. <https://doi.org/10.1037/a0034285>
- Hudson, N. W., Fraley, R. C., Chopik, W. J., & Briley, D. A. (2020). Change Goals Robustly Predict Trait Growth: A Mega-Analysis of a Dozen Intensive Longitudinal Studies Examining Volitional Change. *Social Psychological and Personality Science*, Article 1948550619878423. <https://doi.org/10.1177/1948550619878423>
- Hughes, D. J., & Batey, M. (2017). Using personality questionnaires for selection. In H. Goldstein, E. Pulakos, J. Passmore, & C. Semedo (Eds.), *The Wiley Blackwell Handbook of the Psychology of Recruitment, Selection and Employee Retention*, (pp. 151-181). Wiley-Blackwell. <https://doi.org/10.1002/9781118972472.ch8>
- Hui, C.H. & Triandis, H.C. (1985). Measurement in cross-cultural psychology. *Journal of Cross-Cultural Psychology*, 16, 131-152. <https://doi.org/10.1177/0022002185016002001>
- Hulin, C.L. (1987). A psychometric theory of evaluations of item and scale translation: Fidelity across languages. *Journal of Cross-Cultural Psychology*, 18, 115-142. <https://doi.org/10.1177/0022002187018002001>
- Hunsley, J., & Meyer, G. J. (2003). The incremental validity of psychological testing and assessment: conceptual, methodological, and statistical issues. *Psychological Assessment*, 15(4), 446-455. <https://doi.org/10.1037/1040-3590.15.4.446>
- Hutz, C. S., Nunes, C. H., Silveira, A. D., Jovana, S., Anton, M., & Wieczorek, L. S. (1998). O desenvolvimento de marcadores para a avaliação da personalidade no modelo dos cinco grandes fatores. *Psicologia: Reflexão e Crítica*, 11(2), 395-409. <https://doi.org/10.1590/S0102-79721998000200015>
- Hyde, J. S. (2014). Gender similarities and differences. *Annual Review of Psychology*, 65, 373-398. <http://dx.doi.org/10.1146/annurev-psych-010213-115057>
- Ibañez, M. (1997). *Primeros datos de la versión en castellano del cuestionario revisado de personalidad de Eysenck (EPQ-R)*. <https://core.ac.uk/download/pdf/61427966.pdf>
- Ilgen, D. R., & Hollenbeck, J. R. (1991). The structure of work: Job design and roles. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 2, pp. 165-207). Consulting Psychologists Press.

- Iliev, R., Fulmer, I. S., Spitzmuller, M. & Johnson, M. D. (2009). Personality and citizenship behavior: the mediating role of job satisfaction. *Journal of Applied Psychology*, 94 (4), 945-959. <https://doi.org/10.1037/a0013329>
- International Test Commission. (2017). The ITC Guidelines for Translating and Adapting Tests (Second edition). *International Journal of Testing*, 18(2), 101-134 <https://doi.org/10.1080/15305058.2017.1398166>
- Isaka, H. (1990). Factor analysis of trait terms in everyday Japanese language. *Personality and Individual Differences*, 11(2), 115-124. [https://doi.org/10.1016/0191-8869\(90\)90003-a](https://doi.org/10.1016/0191-8869(90)90003-a)
- Isasi, X., Balluerka, N., y Gorostiaga, A. (2000). La utilización de instrumentos de medida en situaciones de contacto de lenguas y de culturas: una reflexión metodológica. *Psicothema*, 12, 305-310. <http://www.psicothema.com/pdf/571.pdf>
- Jenkins, G. D., Mitra, A., Gupta, N., & Shaw, J. D. (1998). Are financial incentives related to performance? A meta-analytic review of empirical research. *Journal of Applied Psychology*, 83, 777-787. <https://doi.org/10.1037/0021-9010.83.5.777>
- Jin, L., & Zhong, Y. (2014). Contextual factors affecting the influence of perceived organizational support on team innovative performance. *Social Behavior and Personality: an International Journal*, 42(3), 517-528. <https://doi.org/10.2224/sbp.2014.42.3.517>
- Jodoin, M. G., & Gierl, M. J. (2001). Evaluating type I error and power rates using an effect size measure with the logistic regression procedure for DIF detection. *Applied Measurement in Education*, 14(4), 329-349. https://doi.org/10.1207/s15324818ame1404_2
- John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66-100). Guilford Press.
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five inventory—versions 4a and 54*. University of California, Berkeley, Institute of Personality and Social Research.
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm Shift to the Integrative Big- Five Trait Taxonomy: History, Measurement, and Conceptual Issues. In O. P. John, R. W. Robins & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 114-158). Guilford Press.
- John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102-138). Guilford Press.
- Johnson, J. W. (2001). The relative importance of task and contextual performance dimensions to supervisor judgments of overall performance. *Journal of Applied Psychology*, 86, 984-996. <https://doi.org/10.1037/0021-9010.86.5.984>

- Johnson, J. W. (2003). Toward a better understanding of the relationship between personality and individual job performance. In M. R. Barrick, A. M. Ryan (Eds.), *Personality and work: Reconsidering the role of personality in organizations* (pp. 83-120). Jossey-Bass.
- Jokela, M., Pekkarinen, T., Sarvimäki, M., Terviö, M., & Uusitalo, R. (2017). Secular rise in economically valuable personality traits. *Proceedings of the National Academy of Sciences*, *114*(25), 6527-6532. <https://doi.org/10.1073/pnas.1609994114>
- Jöreskog, K. G. (1999, June 22). *How large can a standardized coefficient be*. <http://www.statmodel.com/download/Joreskog.pdf>
- Joung, W., Hesketh, B., & Neal, A. (2006). Using “war stories” to train for adaptive performance. It is better to learn from error or success? *Applied Psychology*, *55*(2), 282–302. <https://doi.org/10.1111/j.1464-0597.2006.00244.x>
- Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: a qualitative and quantitative review. *Journal of Applied Psychology*, *87*(4), 765-780. <https://doi.org/10.1037/0021-9010.87.4.765>
- Judge, T. A., & Cable, D. M. (1997). Applicant personality, organizational culture, and organization attraction. *Personnel Psychology*, *50*, 354–359. <https://doi.org/10.1111/j.1744-6570.1997.tb00912.x>
- Judge, T. A., Higgins, C. A., Thoresen, C. J., & Barrick, M. R. (1999). The big five personality traits, general mental ability, and career success across the life span. *Personnel Psychology*, *52*(3), 621-652. <https://doi.org/10.1111/j.1744-6570.1999.tb00174.x>
- Judge, T. A., Piccolo, R. F., & Kosalka, T. (2009). The bright and dark sides of leader traits: A review and theoretical extension of the leader trait paradigm. *The Leadership Quarterly*, *20*(6), 855-875. <https://doi.org/10.1016/j.leaqua.2009.09.004>
- Judge, T. A., Rodell, J. B., Klinger, R. L., Simon, L. S., & Crawford, E. R. (2013). Hierarchical representations of the five-factor model of personality in predicting job performance: integrating three organizing frameworks with two theoretical perspectives. *Journal of Applied Psychology*, *98*(6), 875-925. <https://doi.org/10.1037/a0033901>
- Jundt, D. K., Shoss, M. K. & Huang, J. L. (2015). Individual adaptive performance in organizations: A review. *Journal of Organizational Behavior*, *36* (S1), 53-71. <https://doi.org/10.1002/job.1955>
- Kaiser, T. (2019). Nature and evoked culture: Sex differences in personality are uniquely correlated with ecological stress. *Personality and Individual Differences*, *148*, 67-72. <https://doi.org/10.1016/j.paid.2019.05.011>
- Kaiser, T., Del Giudice, M, & Booth, T. (2019). Global Sex Differences in Personality: Replication with an Open Online Dataset. *Journal of Personality*, *88*(3), 415-429. <https://doi.org/10.1111/jopy.12500>
- Kajonius, P. J., & Johnson, J. (2018). Sex differences in 30 facets of the five factor model of personality in the large public (N= 320,128). *Personality and Individual Differences*, *129*, 126-130. <https://doi.org/10.1016/j.paid.2018.03.026>

- Kajonius, P., & Mac Giolla, E. (2017). Personality traits across countries: Support for similarities rather than differences. *PLoS ONE*, 12 (6), Article e0179646. <https://doi.org/10.1371/journal.pone.0179646>
- Kandler, C., Kornadt, A. E., Hagemeyer, B., & Neyer, F. J. (2015). Patterns and sources of personality development in old age. *Journal of Personality and Social Psychology*, 109(1), 175-191. <https://doi.org/10.1037/pspp0000028>
- Kanten, P., Kanten, S., & Gurlek, M. (2015). The effects of organizational structures and learning organization on job embeddedness and individual adaptive performance. *Procedia Economics and Finance*, 23, 1358-1366. [https://doi.org/10.1016/s2212-5671\(15\)00523-7](https://doi.org/10.1016/s2212-5671(15)00523-7)
- Katz, D. (1964). The motivational basis of organizational behavior. *Behavioral science*, 9(2), 131-146. <https://doi.org/10.1002/bs.3830090206>
- Kim, H., Di Domenico, S. I., & Connelly, B. S. (2019). Self–Other Agreement in Personality Reports: A Meta-Analytic Comparison of Self-and Informant-Report Means. *Psychological Science*, 30(1), 129-138. <https://doi.org/10.1177/0956797618810000>
- Kim, S. (2006). Public service motivation and organizational citizenship behavior in Korea. *International Journal of Manpower*, 27(8), 722-740. <https://doi.org/10.1108/01437720610713521>
- Klemm, W.R. (2017) Leadership and Creativity. In Marques J., Dhiman S. (Eds.), *Leadership Today*. Springer. <https://doi.org/10.1007/978-3-319-31036-7>
- Komar, S., Komar, J. A., Robie, C., & Taggar, S. (2010). Speeding personality measures to reduce faking. *Journal of Personnel Psychology*, 9(3), 126-137. <https://doi.org/10.1027/1866-5888/a000016>
- Koopmans, L., Bernaards, C. M., Hildebrandt, V. H., Schaufeli, W. B., de Vet Henrica, C. W., & van der Beek, A. J. (2011). Conceptual frameworks of individual work performance: a systematic review. *Journal of Occupational and Environmental Medicine*, 53(8), 856-866. <https://doi.org/10.1097/jom.0b013e318226a763>
- Kostick, M.M. (1976). *PAPI. Perception and Preference Inventory. Technical manual*. Cubiks.
- Kozlowski, S. W.J., Gully, S. M., Brown, K. G., Salas, E., Smith, E. M., & Nason, E. R. (2001). Effects of training goals and goal orientation traits on multidimensional training outcomes and performance adaptability. *Organizational Behavior and Human Decision Processes*, 85(1), 1-31. <https://doi.org/10.1006/obhd.2000.2930>
- Kozlowski, S. W. J., Gully, S. M., Nason, E. R., & Smith, E. M. (1999). Developing adaptive teams: A theory of compilation and performance across levels and time. In D. R. Ilgen & E. D. Pulakos (Eds.), *The changing nature of work and performance: Implications for staffing, personnel actions, and development* (pp. 240-292). Jossey-Bass.

- Kristof, A. L. (1996). Person-organization fit: An integrative review of its conceptualizations, measurement, and implications. *Personnel Psychology*, 49(1), 1-49. <https://doi.org/10.1111/j.1744-6570.1996.tb01790.x>
- Kristof-Brown, A. L., & Guay, R. P. (2011). Person-environment fit. In S. Zedeck, (Ed.), *APA Handbook of industrial and organizational psychology* (Vol. 3, pp. 3–50). APA. <http://dx.doi.org/10.1037/12171-001>
- Kristof-Brown, A. L., Zimmerman, R. D., & Johnson, E. C. (2005). Consequences of individuals' fit at work: a meta-analysis of person–job, person–organization, person–group, and person–supervisor fit. *Personnel Psychology*, 58(2), 281-342. <https://doi.org/10.1111/j.1744-6570.2005.00672.x>
- Kröger, U., & Staufenbiel, T. (2012). Entwicklung und Validierung eines Fragebogens zur Erfassung von “Adaptive Performance”. *Zeitschrift für Arbeits- und Organisationspsychologie A&O*, 56(2), 55-69. <https://doi.org/10.1026/0932-4089/a000069>
- Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology*, 50, 537–567. <http://dx.doi.org/10.1146/annurev.psych.50.1.537>
- Kruyen, P. M. (2012). *Making decisions about individuals using tests and questionnaires: When is short too short?* [Doctoral Dissertation, Tilburg University]. Ridderprint. https://pure.uvt.nl/ws/portalfiles/portal/1467991/Kruyen_using_14-12-2012.pdf
- Kruyen, P. M., Emons, W. H., & Sijtsma, K. (2012). Test length and decision quality in personnel selection: When is short too short? *International Journal of Testing*, 12, 321–344. <https://doi.org/10.1080/15305058.2011.643517>
- Kruyen, P. M., Emons, W. H., & Sijtsma, K. (2013). On the shortcomings of shortened tests: A literature review. *International Journal of Testing*, 13, 223–248. <https://doi.org/10.1080/15305058.2012.703734>
- Kurz, R., & Bartram, D. (2002). Competency and individual performance: Modeling the world of work. In I. T. Robertson, M. Callinan, & D. Bartram (Eds.), *Organizational effectiveness: The role of psychology* (pp. 227–255). Wiley. <https://doi.org/10.1002/9780470696736.ch10>
- Laajaj, R., Macours, K., Hernandez, D. A. P., Arias, O., Gosling, S. D., Potter, J., Rubio-Codina, M. & Vakis, R. (2019). Challenges to capture the big five personality traits in non-WEIRD populations. *Science Advances*, 5(7), eaaw5226. <https://doi.org/10.1126/sciadv.aaw5226>
- Lado, M., & Alonso, P. (2017). The Five-Factor model and job performance in low complexity jobs: A quantitative synthesis. *Revista de Psicología del Trabajo y de las Organizaciones*, 33(3), 175-182. <https://doi.org/10.1016/j.rpto.2017.07.004>
- Landy, F. J., & Conte, J. M. (2016). *Work in the 21st century, Binder ready version: An introduction to Industrial and Organizational Psychology*. John Wiley & Sons. https://doi.org/10.1111/j.1744-6570.2008.00119_2.x

- Lang, J. W. B., & Bliese, P. D. (2009). General mental ability and two types of adaptation to unforeseen change: Applying discontinuous growth models to the task-change paradigm. *Journal of Applied Psychology*, 94, 411-428. <https://doi.org/10.1037/a0013803>
- Lee, K., & Allen, N. J. (2002). Organizational citizenship behavior and workplace deviance: the role of affect and cognitions. *Journal of Applied Psychology*, 87(1), 131-142. <https://doi.org/10.1037/0021-9010.87.1.131>
- LePine, J. A. (2003). Team adaptation and postchange performance: Effects of team composition in terms of members' cognitive ability and personality. *Journal of Applied Psychology*, 88, 27-39. <https://doi.org/10.1037/0021-9010.88.1.27>
- LePine J.A., Colquitt J.A., & Erez, A. (2000). Adaptability to changing task contexts: effects of general cognitive ability, conscientiousness, and openness to experience. *Personnel Psychology*, 53, 563-593. <https://doi.org/10.1111/j.1744-6570.2000.tb00214.x>
- Leppink, J., & Pérez-Fuster, P. (2017). We need more replication research—A case for test-retest reliability. *Perspectives on medical education*, 6(3), 158-164. <https://doi.org/10.1007/s40037-017-0347-z>
- Leutner, F., & Chamorro-Premuzic, T. (2018). Stronger Together: Personality, Intelligence and the Assessment of Career Potential. *Journal of Intelligence*, 6(4), Article 49. <https://doi.org/10.3390/jintelligence6040049>
- Lievens, F., Conway, J. M., & De Corte, W. (2008). The relative importance of task, citizenship and counterproductive performance to job performance ratings: Do rater source and team-based culture matter? *Journal of Occupational and Organizational Psychology*, 81(1), 11-27. <https://doi.org/10.1348/096317907x182971>
- Lievens, F., De Corte, W., & Schollaert, E. (2008). A closer look at the frame-of-reference effect in personality scale scores and validity. *Journal of Applied Psychology*, 93, 268–279. <https://doi.org/10.1037/0021-9010.93.2.268>
- Lillard, R. M., Watts, L. L., Frame, M. C., Hein, M. B., Rigdon, W. D., & Orsak Robinson, K. (2012, April 26-28). *Initial development and validation of a measure of adaptive performance* [Poster presentation]. 27th annual Conference of the Society for Industrial/Organizational Psychology, San Diego, CA. <https://www.siop.org/Portals/84/SIOP%20Docs/SIOP%20Conference%20Presentations/2012ACProgram.pdf>
- Lodhi, P. H., Deo, S., & Belhekar, V. M. (2002). The Five-Factor model of personality in Indian context: measurement and correlates. In R. R. McCrae & J. Allik (Eds.), *The Five-Factor model of personality across cultures* (pp. 227–248). Kluwer Academic Publisher.
- Lounsbury, J. W., Sundstrom, E. D., Gibson, L. W., Loveland, J. M., & Drost, A. W. (2016). Core personality traits of managers. *Journal of Managerial Psychology*, 31(2), 434-450. <http://dx.doi.org/10.1108/JMP-03-2014-0092>

- Mariani, M. G. (2011). *Valutare le prestazioni. Come gestire e migliorare la performance lavorativa*. Il Mulino.
- Marlow, K. K., Calarco, H. N., Frame, M. C., & Hein, M. B., (2015, October 23-24). *Building a better adaptive performance measure: Factor analysis and scale validation* [Poster presentation]. 11th annual River Cities Industrial/Organizational Psychology conference, Chattanooga, TN. <https://www.utc.edu/psychology/rcio/rcioarchive/rcio2015.php>
- Marques-Quinteiro, P., Ramos-Villagrasa, P. J., Passos, A. M., & Cural, L. (2015). Measuring adaptive performance in individuals and teams. *Team Performance Management*, 21(7/8), 339-360. <https://doi.org/10.1108/tpm-03-2015-0014>
- Marsh, H. W., Hau, K. T., & Grayson, D. (2005). Goodness of fit evaluation in structural equation modeling. In A. Maydeu-Olivares & J. McArdle (Eds.), *Psychometrics. A festschrift to Roderick P. McDonald*. Lawrence Erlbaum Associates, Inc.
- Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J. S., Trautwein, U., & Nagengast, B. (2010). A new look at the Big Five factor structure through exploratory structural equation modeling. *Psychological Assessment*, 22, 471–491. <https://doi.org/10.1037/a0019227>
- Marsh, H. W., Morin, A. J., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology*, 10, 85-110. <https://doi.org/10.1146/annurev-clinpsy-032813-153700>
- Marsh, H. W., Nagengast, B., & Morin, A. J. (2013). Measurement invariance of big-five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and la dolce vita effects. *Developmental Psychology*, 49(6), 1194-1218. <https://doi.org/10.1037/a0026913>
- Martínez Arias, R., Hernández, M.J., y Hernández, M.V. (2006). *Psicometría*. Alianza Editorial.
- Matthews-López, J. L. (2003). *Best practices and technical issues in cross-lingual, crosscultural assessments* (Order No. 3120524) [Doctoral dissertation, Ohio University]. ProQuest Dissertations & Theses Global. (305315896).
- McAdams, D. P., & Olson, B. D. (2010). Personality development: Continuity and change over the life course. *Annual Review of Psychology*, 61, 517-542. <https://doi.org/10.1146/annurev.psych.093008.100507>
- McCarthy, J. M., Van Iddekinge, C. H., Lievens, F., Kung, M.-C., Sinar, E. F., & Campion, M. A. (2013). Do candidate reactions relate to job performance or affect criterion-related validity? A multistudy investigation of relations among reactions, selection test scores, and job performance. *Journal of Applied Psychology*, 98, 701-709. <https://doi.org/10.1037/a0034089>
- McCrae, R.R., & Allik, J. (2002). *The Five-Factor Model of personality across cultures*. Kluwer Academic/Plenum Publishers. <https://doi.org/10.1007/978-1-4615-0763-5>

- McCrae, R. R., & Costa, P. T. (1985). Updating Norman's "adequacy taxonomy": Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology*, 49(3), 710-721. <https://doi.org/10.1037/0022-3514.49.3.710>
- McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52(1), 81-90. <https://doi.org/10.1037/0022-3514.52.1.81>
- McCrae, R. R., & Costa Jr., P. T. (1996). Toward a new generation of personality theories. In J. S. Wiggins (Ed.), *Theoretical contexts for the five-factor model* (pp. 51-87). Guilford Press.
- McCrae, R.R., & Costa Jr., P.T. (1997). Personality trait structure as a human universal. *American Psychologist*, 52, 509–516. <https://doi.org/10.1037/0003-066x.52.5.509>
- McCrae, R. R., & Costa Jr., P. T. (1999). A five-factor theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 139-153). Guilford.
- McCrae, R. R., & Costa, P. T., Jr. (2003). *Personality in adulthood: A five-factor theory perspective* (2nd ed.). Guilford. <https://doi.org/10.4324/9780203428412>
- McCrae, R.R., Costa P.T., Jr., & Martin, T.A. (2005). The NEOPI- 3: A more readable Revised NEO Personality Inventory. *Journal of Personality Assessment*, 84, 261–270. https://doi.org/10.1207/s15327752jpa8403_05
- McCrae, R. R., & Sutin, A. R. (2018). A five-factor theory perspective on causal analysis. *European Journal of Personality*, 32(3), 151-166. <https://doi.org/10.1002/per.2134>
- McCrae, R. R., Terracciano, A., & 78 members of the Personality Profiles of Cultures Project. (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology*, 88(3), 547-561. <https://doi.org/10.1037/0022-3514.88.3.547>
- McCrae, R. R., Zonderman, A. B., Costa, P. T., Bond, M. H., & 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(3), 552-566. <https://doi.org/10.1037/0022-3514.70.3.552>
- Mehl, M. R., Gosling, S. D., & Pennebaker, J. W. (2006). Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology*, 90(5), 862-877. <https://doi.org/10.1037/0022-3514.90.5.862>
- Michaud, R. (2014). *La validité prédictive et incrémentielle de l'appariement personne-organisation pour les activités de sélection* [Doctoral dissertation, Université Laval]. <http://hdl.handle.net/20.500.11794/24991>

- Millon, T. (2001). *MIPS: Inventario de Estilos de Personalidad de Millon*. Manual. TEA Ediciones.
- Mischel, W. (2003). *Introduction to personality: toward an integration* (7th Ed). Wiley.
- Mlinarić, V., & Podlesek, A. (2013). Item context effects on Big Five personality measures. *Review of Psychology*, *20*(1-2), 23-28.
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. & De Vet, H. C. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research*, *19*(4), 539-549. <https://doi.org/10.1007/s11136-010-9606-8>
- Monteiro, A. R. D. O. V. (2015). *The impact of job insecurity on adaptive performance via burnout* [Doctoral dissertation, NOVA –School of Business and Economics]. NSBE: Nova SBE - MA Dissertations. <http://hdl.handle.net/10362/15567>
- Morgeson, F.P., Campion, M.A., Dipboye, R.L., Hollenbeck, J.R., Murphy, K., & Schmitt, N. (2007). Are we getting fooled again? Coming to terms with limitations in the use of personality tests for personnel selection. *Personnel Psychology*, *60*, 1029– 1049. <https://doi.org/10.1111/j.1744-6570.2007.00100.x>
- Morizot, J. (2014). Construct validity of adolescents’ self-reported Big Five personality traits: Importance of conceptual breadth and initial validation of a short measure. *Assessment*, *21*, 580–606. <https://doi.org/10.1177/1073191114524015>
- Motowidlo, S. J., & Van Scotter, J. R. (1994). Evidence that task performance should be distinguished from contextual performance. *Journal of Applied Psychology*, *79*, 475–480. <https://doi.org/10.1037/0021-9010.79.4.475>
- Mõttus, R., Allik, J., & Realo, A. (2020). Do self-reports and informant-ratings measure the same personality constructs? *European Journal of Psychological Assessment*, *36*(2), 289-295. <https://doi.org/10.1027/1015-5759/a000516>
- Mõttus, R., Johnson, W., & Deary, I. J. (2012). Personality traits in old age: Measurement and rank-order stability and some mean-level change. *Psychology and Aging*, *27*, 243-249. <https://doi.org/10.1037/a0023690>
- Mõttus, R., Realo, A., Vainik, U., Allik, J., & Esko, T. (2017). Educational attainment and personality are genetically intertwined. *Psychological Science*, *28*(11), 1631-1639. <https://doi.org/10.1177/0956797617719083>
- Mõttus, R., & Rozgonjuk, D. (2019). Development is in the details: Age differences in the Big Five domains, facets, and nuances. *Journal of Personality and Social Psychology*. Advance online publication. <https://doi.org/10.1037/pspp0000276>
- Mount, M. K., & Barrick, M. R. (1995). The Big Five personality dimensions: Implications for research and practice in human resources management. *Research in Personnel and Human Resources Management*, *13*(3), 153–200.

- Mount, M. K., Barrick, M. R., Scullen, S. M., & Rounds, J. (2005). Higher-order dimensions of the big five personality traits and the big six vocational interest types. *Personnel Psychology*, *58*(2), 447-478. <https://doi.org/10.1111/j.1744-6570.2005.00468.x>
- Mroczek, D. K., & Almeida, D. M. (2004). The effect of daily stress, personality, and age on daily negative affect. *Journal of Personality*, *72*(2), 355-378. <https://doi.org/10.1111/j.0022-3506.2004.00265.x>
- Muchinsky, P. M., & Monahan, C. J. (1987). What is person–environment congruence? Supplementary versus complementary models of fit. *Journal of Vocational Behavior*, *31*, 268–277. [https://doi.org/10.1016/0001-8791\(87\)90043-1](https://doi.org/10.1016/0001-8791(87)90043-1)
- Muñiz, J. y Fernández-Hermida, J. R. (2010). La opinión de los psicólogos españoles sobre el uso de los tests. *Papeles del Psicólogo*, *31*(1), 108-121. <http://www.papelesdelpsicologo.es/pdf/1801.pdf>
- Muñiz, J., Fernández-Hermida, J. R., Fonseca-Pedrero, E., Campillo-Álvarez, Á., & Peña-Suárez, E. (2011). Evaluación de tests editados en España. *Papeles del Psicólogo*, *32*(2), 113-128. <http://www.papelesdelpsicologo.es/pdf/1947.pdf>
- Muñiz, J., y Fonseca-Pedrero, E. (2019). Diez pasos para la construcción de un test. *Psicothema*, *31*(1), 7-16. <https://doi.org/10.7334/psicothema2018.291>
- Muñiz J., y Hambleton R. (1996) Directrices para la traducción y adaptación de los tests. *Papeles del Psicólogo*, *66*, 63-70. <http://www.papelesdelpsicologo.es/resumen?pii=737>
- Muñiz, J., Hernández, A., y Fernández-Hermida, J. R. (2020). Utilización de los test en España: el punto de vista de los psicólogos. *Papeles del Psicólogo*, *41*(1), 1-15. <https://doi.org/10.23923/pap.psicol2020.2921>
- Murphy, K.R. (1989). Dimensions of job performance. In R. Dillon y J. Pellegrino (Eds.), *Testing: Applied and theoretical perspectives* (pp. 218-247). Praeger.
- Murphy, K. R., & Jacobs, R. R. (2012). Using effect size measures to reform the determination of adverse impact in equal employment litigation. *Psychology, Public Policy, and Law*, *18*(3), 477-499. <https://doi.org/10.1037/a0026350>
- Murphy, P. R., & Jackson, S. E. (1999). Managing work-role performance: Challenges for 21st century organizations and employees. In D. R. Ilgen & E. D. Pulakos (Eds.), *The changing nature of work performance* (pp. 325–365). Jossey-Bass.
- Murphy, S. L. (2015). *Individual adaptability as a predictor of job performance* [Doctoral dissertation, Louisiana Tech University]. Louisiana Tech Digital Commons. <https://digitalcommons.latech.edu/dissertations/209>
- Naami, A., Behzadi, E., Parisa, H., & Charkhabi, M. (2014). A study on the personality aspects of adaptive performance among governmental hospitals nurses: a conceptual model. *Procedia-Social and Behavioral Sciences*, *159*, 359-364. <https://doi.org/10.1016/j.sbspro.2014.12.388>

- Neal, A., Yeo, G., Koy, A., & Xiao, T. (2012). Predicting the form and direction of work role performance from the Big 5 model of personality traits. *Journal of Organizational Behavior*, 33, 175–192. <https://doi.org/10.1002/job.742>
- Neill, J. A., & Jackson, D. N. (1970). An evaluation of item selection strategies in personality scale construction. *Educational and Psychological Measurement*, 30, 647–661. <https://doi.org/10.1177/001316447003000312>
- Ng, K. Y., Ang, S., & Chan, K. Y. (2008). Personality and leader effectiveness: A moderated mediation model of leadership self-efficacy, job demands, and job autonomy. *Journal of Applied Psychology*, 93(4), 733-743. <https://doi.org/10.1037/0021-9010.93.4.733>
- Ng, T. W., & Feldman, D. C. (2009). How broadly does education contribute to job performance? *Personnel Psychology*, 62(1), 89-134. <https://doi.org/10.1111/j.1744-6570.2008.01130.x>
- Ng, T. W., & Feldman, D. C. (2012). Evaluating six common stereotypes about older workers with meta-analytical data. *Personnel Psychology*, 65(4), 821-858. <https://doi.org/10.1111/peps.12003>
- Ng, T. W., & Feldman, D. C. (2013). Does longer job tenure help or hinder job performance? *Journal of Vocational Behavior*, 83(3), 305-314. <https://doi.org/10.1016/j.jvb.2013.06.012>
- Ng, T. W., Lam, S. S., & Feldman, D. C. (2016). Organizational citizenship behavior and counterproductive work behavior: Do males and females differ? *Journal of Vocational Behavior*, 93, 11-32. <https://doi.org/10.1016/j.jvb.2015.12.005>
- Niess, C., & Zacher, H. (2015). Openness to experience as a predictor and outcome of upward job changes into managerial and professional positions. *PloS one*, 10(6), Article e0131115. <https://doi.org/10.1371/journal.pone.0131115>
- Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *The Journal of Abnormal and Social Psychology*, 66(6), 574-583. <https://doi.org/10.1037/h0040291>
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.
- O’Connell, D. J., McNeely, E. & Hall, D. T. (2008). Unpacking personality adaptability at work. *Journal of Leadership & Organizational Studies*, 14, 248–259. <https://doi.org/10.1177/1071791907311005>
- O’Connor, B. P. (2002) A quantitative review of the comprehensiveness of the five-factor model in relation to popular personality inventories. *Assessment*, 9, 188–203. <https://doi.org/10.1177/1073191102092010>
- Olaru, G., Witthöft, M., & Wilhelm, O. (2015). Methods matter: testing competing models for designing short-scale Big-Five assessments. *Journal of Research in Personality*, 59, 56-68. <https://doi.org/10.1016/j.jrp.2015.09.001>

- Ones, D. S., Dilchert, S., Viswesvaran, C., & Judge, T. A. (2007). In support of personality assessment in organizational settings. *Personnel Psychology*, *60*(4), 995-1027. <https://doi.org/10.1111/j.1744-6570.2007.00099.x>
- Ones, D. S., & Viswesvaran, C. (1996). Bandwidth-fidelity dilemma in personality measurement for personnel selection. *Journal of Organizational Behavior*, *17*, 609–626. [https://doi.org/10.1002/\(sici\)1099-1379\(199611\)17:6<609::aid-job1828>3.0.co;2-k](https://doi.org/10.1002/(sici)1099-1379(199611)17:6<609::aid-job1828>3.0.co;2-k)
- Ones, D. S., Viswesvaran, C., & Dilchert, S., (2005). Personality at work: raising awareness and correcting misconceptions. *Human Performance*, *18*(4), 389-404. https://doi.org/10.1207/s15327043hup1804_5
- Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology*, *81*(6), 660-679. <https://doi.org/10.1037/0021-9010.81.6.660>
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349* (6251), Article aac4716. <https://doi.org/10.31219/osf.io/447b3>
- Oprins, E. A., Van den Bosch, K., & Venrooij, W. (2018). Measuring adaptability demands of jobs and the adaptability of military and civilians. *Military Psychology*, *30*(6), 576-589. <https://doi.org/10.1080/08995605.2018.1521689>
- Organ, D. W. (1988). *Organizational citizenship behavior: The good soldier syndrome*. Lexington Books.
- Organ, D. W. (1997). Organizational citizenship behavior: It's construct clean-up time. *Human Performance*, *10*, 85-97. https://doi.org/10.1207/s15327043hup1002_2
- Organ, D. W. (2018). Organizational Citizenship Behavior: Recent Trends and Developments. *Annual Review of Organizational Psychology and Organizational Behavior*, *80*, 295-306. <http://dx.doi.org/10.1146/annurev-orgpsych-032117-104536>
- Organ, D. W., & Ryan K. (1995). A meta-analytic review of attitudinal and dispositional predictors of organizational citizenship behavior. *Personnel Psychology*, *48*, 775- 802. <https://doi.org/10.1111/j.1744-6570.1995.tb01781.x>
- Orlando, M., & Thissen, D. (2000). Likelihood-based item fit indices for dichotomous item response theory models. *Applied Psychological Measurement*, *24*, 50-64. <https://doi.org/10.1177/01466216000241003>
- Orlando, M., & Thissen, D. (2003). Further investigation of the performance of the S-X2: An item fit index for use with dichotomous item response theory models. *Applied Psychological Measurement*, *27*, 289-298. <https://doi.org/10.1177/0146621603027004004>
- Ortet, G., Martínez, T., Mezquita, L., Morizot, J., & Ibáñez, M. I. (2017). Big Five Personality Trait Short Questionnaire: Preliminary Validation with Spanish Adults. *The Spanish Journal of Psychology*, *20*, 1-11. <https://doi.org/10.1017/sjp.2017.8>

- Ostendorf, F. (1990). *Sprache und Persönlichkeitsstruktur: Zur Validität des Fünf-Faktoren-Modells der Persönlichkeit*. S. Roderer Verlag.
- Ostroff, C. & Zhan, Y. (2012). Person–Environment Fit in the Selection Process. In N. Schmitt (Ed.), *Oxford Handbook of Personnel Assessment and Selection* (pp. 252-273). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199732579.013.0012>
- Pamos, A. (2016). *Diseño y validación de un cuestionario de evaluación de competencias con soporte informático* [Tesis doctoral, Universidad Complutense de Madrid]. E-prints Complutense. <https://eprints.ucm.es/37591/>
- Park, H. H., Wiernik, B. M., Oh, I. S., Gonzalez-Mulé, E., Ones, D. S., & Lee, Y. (2020). Meta-analytic five-factor model personality intercorrelations: Eeny, meeny, miney, moe, how, which, why, and where to go. *Journal of Applied Psychology*. Advance online publication. <https://doi.org/10.1037/apl0000476>
- Parker, S. K., & Collins, C. G. (2010). Taking stock: Integrating and differentiating multiple proactive behaviors. *Journal of Management*, 36, 633–662. <https://doi.org/10.1177/0149206308321554>
- Parks, L. (2007, April 26-29). *Personality and values: A meta-analysis* [Conference presentation]. 22nd Annual Conference for the Society of Industrial and Organizational Psychology, New York, New York.
- Parks, L., & Guay, R. P. (2009). Personality, values, and motivation. *Personality & Individual Differences*, 47, 675–684. <http://dx.doi.org/10.1016/j.paid.2009.06.002>
- Parlamento Vasco. (1982, 24 de noviembre). *Ley 10/1982, Básica de normalización del uso del Euskera*. BOPV 160, 16 de Diciembre de 1982, 3138-3146. https://www.euskadi.eus/web01-a2berrik/es/contenidos/ley/bopv198201955/es_def/index.shtml
- Parlamento Vasco. (1989, 6 de julio). *Ley 6/1989, de la Función Pública Vasca*. BOPV 144 de 28 de Julio de 1989, 4954-5001. <https://www.euskadi.eus/y22-bopv/es/bopv2/datos/1989/07/8902092a.pdf>
- Patterson, F., Zibarras, L., Carr, V., Irish, B., & Gregory, S. (2011). Evaluating candidate reactions to selection practices using organisational justice theory. *Medical Education*, 45, 289-297. <https://doi.org/10.1111/j.1365-2923.2010.03808.x>
- Paulhus, D.L. (1991). Measurement and control of response bias. In J.P. Robinson, P.R. Shaver, & L.S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). Academic Press. <https://doi.org/10.1016/b978-0-12-590241-0.50006-x>
- Paunonen, S. V., Haddock, G., Forsterling, F., & Keinonen, M. (2003). Broad versus narrow personality measures and the prediction of behaviour across cultures. *European Journal of Personality*, 17(6), 413-433. <https://doi.org/10.1002/per.496>
- Pavalache-Ilie, M. (2014). Organizational citizenship behaviour, work satisfaction and employees' personality. *Procedia-Social and Behavioral Sciences*, 127, 489-493. <https://doi.org/10.1016/j.sbspro.2014.03.296>

- Peiró, J. M., Bayona, J. A., Caballer, A., & Di Fabio, A. (2020). Importance of work characteristics affects job performance: The mediating role of individual dispositions on the work design-performance relationships. *Personality and Individual Differences*, 157, 109808. <https://doi.org/10.1016/j.paid.2019.109808>
- Peiró, J. M., y Ramos, J. (2012). Contrato psicológico y prácticas de recursos humanos en el sector de la educación. Una comparación entre empleados públicos y privados en cinco países europeos. *Revista Vasca de Gestión de Personas y Organizaciones Públicas*, 3, 8-29. <https://bit.ly/3n9BBI1>
- Penfield, R. D. (2014). An NCME Instructional Module on Polytomous Item Response Theory Models. *Educational Measurement: Issues and Practice*, 33, 36-48. <https://doi.org/10.1111/emip.12023>
- Perera, H. N., McIlveen, P., Burton, L. J., & Corser, D. M. (2015). Beyond congruence measures for the evaluation of personality factor structure replicability: An exploratory structural equation modeling approach. *Personality and Individual Differences*, 84, 23-29. <https://doi.org/10.1016/j.paid.2015.01.004>
- Perry-Smith, J. E., & Shalley, C. E. (2003). The social side of creativity: A static and dynamic social network perspective. *Academy of Management Review*, 28, 89-106. <https://doi.org/10.2307/30040691>
- Piasentin, K. A. (2003). *Fitting in and standing out: an empirical test of supplementary and complementary person-organization fit* (Order No. MQ87395) [Master Dissertation, University of Calgary]. ProQuest Dissertations & Theses Global. (305338899)
- Piasentin, K. A. (2007). *How do employees conceptualize fit? Development of a multidimensional measure of subjective Person-Organization fit* (Order No. NR25707) [Doctoral dissertation, University of Calgary]. ProQuest Dissertations & Theses Global. (304899363)
- Piasentin, K. A., & Chapman, D. S. (2006). Subjective person–organization fit: Bridging the gap between conceptualization and measurement. *Journal of Vocational Behavior*, 69(2), 202-221. <https://doi.org/10.1016/j.jvb.2006.05.001>
- Piccolo, R. F., Buengeler, C., & Judge, T. A. (2018). Leadership [Is] Organizational Citizenship Behavior. In P. M. Podsakoff & S.B. McKenzie (Eds.), *The Oxford Handbook of Organizational Citizenship Behavior*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190219000.013.39>
- Pini, A., & Mariani, M. G. (2014). Performance adattiva nei contesti organizzativi: verso un modello di analisi integrato. *Psicologia sociale*, 1, 29-47. <https://doi.org/10.1482/76367>
- Pinillos, J.L. (1964). Validez interna del cuestionario de personalidad C.E.P. *Saitabi*, 14, 215-237. <https://ojs.uv.es/index.php/saitabi/article/view/5537/5296>

- Ployhart, R. E., & Bliese, P. D. (2006). Individual adaptability (I-ADAPT) theory: Conceptualizing the antecedents, consequences, and measurement of individual differences in adaptability. In *Understanding adaptability: A prerequisite for effective performance within complex environments* (pp. 3-39). Emerald Group Publishing Limited. [https://doi.org/10.1016/s1479-3601\(05\)06001-7](https://doi.org/10.1016/s1479-3601(05)06001-7)
- Podsakoff, N. P., Whiting, S. W., Podsakoff, P. M., & Blume, B. D. (2009). Individual- and organizational-level consequences of organizational citizenship behaviors: A meta-analysis. *Journal of Applied Psychology*, *94*, 122–141. <https://doi.org/10.1037/e518442013-522>
- Podsakoff, P. M., MacKenzie, S.B, Lee, J.Y., & Podsakoff, N.P. (2003). Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology*, *88*, 879-903. <https://doi.org/10.1037/0021-9010.88.5.879>
- Podsakoff, P. M., MacKenzie, S. B., Paine, J. B., & Bachrach, D. G. (2000). Organizational citizenship behaviours: A critical review of the theoretical and empirical literature and suggestions for future research. *Journal of Management*, *26*, 513–563. <https://doi.org/10.1177/014920630002600307>
- Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, *63*, 539-569. <https://doi.org/10.1146/annurev-psych-120710-100452>
- Podsakoff, P.M., & Organ, D.W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management*, *12*, 69–82. <https://doi.org/10.1177/014920638601200408>
- Primi, R., John, O. P., Santos, D., & De Fruyt, F. (2017). *SENNA inventory*. Institute Ayrton Senna
- Psicólogos Empresariales. (2008). *PRISMA: Cuestionario de personalidad laboral. Manual Técnico*. Psicólogos Empresariales.
- Pulakos, E. D., Arad, S., Donovan, M. A., & Plamondon, K. E. (2000). Adaptability in the workplace: Development of a taxonomy of adaptive performance. *Journal of Applied Psychology*, *85*, 612-624. <https://doi.org/10.1037/0021-9010.85.4.612>
- Pulakos, E. D., Dorsey, D. W., & White, S. S. (2006). Adaptability in the workplace: Selecting an adaptive workforce. In C. S. Burke, L. G. Pierce & E. Salas (Eds.), *Understanding adaptability: A prerequisite for effective performance within complex environments* (pp. 41-71). Emerald. [https://doi.org/10.1016/s1479-3601\(05\)06002-9](https://doi.org/10.1016/s1479-3601(05)06002-9)
- Pulakos, E. D., Mueller-Hanson, R., & Nelson, J. (2012). Adaptive performance and trainability as criteria in selection research. In N. Schmitt (Ed.), *The Oxford handbook of personnel assessment and selection* (pp. 595-613). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199732579.013.0026>

- Pulakos, E. D., Schmitt, N., Dorsey, D. W., Arad, S., Hedge, J. W., & Borman, W. C. (2002). Predicting adaptive performance: Further tests of a model of adaptability. *Human Performance, 15*, 299-323. https://doi.org/10.1207/s15327043hup1504_01
- Quevedo-Aguado, M. P., Iraegui, A., Anivarro, E. M., & Ross, P. (1996). Linguistic descriptors of personality in the Spanish language: A first taxonomic study. *European Journal of Personality, 10*(1), 25-34. [https://doi.org/10.1002/\(sici\)1099-0984\(199603\)10:1<25::aid-per243>3.0.co;2-e](https://doi.org/10.1002/(sici)1099-0984(199603)10:1<25::aid-per243>3.0.co;2-e)
- Quijano, S. (1997). *Sistemas Efectivos de Evaluación del Rendimiento: Resultados y desempeño. Técnicas y Sistemas para la Gestión y Desarrollo del Personal*. EUB.
- Raju, N. S., Laffitte, L. J., & Byrne, B. M. (2002). Measurement equivalence: A comparison of methods based on confirmatory factor analysis and item response theory. *Journal of Applied Psychology, 87*(3), 517-529. <https://doi.org/10.1037/0021-9010.87.3.517>
- Ramió, C. (2011). Estrategias para la modernización del acceso a la Función Pública. *Nuevas Políticas Publicas, 7*, 129-144. <https://bit.ly/2XTCYjC>
- Ramírez, R. G. (2013). *Diseño del trabajo y desempeño laboral individual* [Tesis doctoral, Universidad Autónoma de Madrid]. Biblos e-Archivo. <http://hdl.handle.net/10486/13039>
- Rammstedt, B., & Beierlein, C. (2014). Can't we make it any shorter? The limits of personality assessment and ways to overcome them. *Journal of Individual Differences, 35*, 212–220. <https://doi.org/10.1027/1614-0001/a000141>
- Rammstedt, B., Goldberg, L.R., & Borg, I. (2010). The measurement equivalence of Big-Five factor markers for persons with different levels of education. *Journal of Research in Personality, 44*, 53-61. <https://doi.org/10.1016/j.jrp.2009.10.005>
- Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality, 41*, 203–212. <https://doi.org/10.1016/j.jrp.2006.02.001>
- Rammstedt, B., Kemper, C.J., & Borg, I. (2013). Correcting Big Five personality measurements for acquiescence: An 18-country cross-cultural study. *European Journal of Personality, 27*, 71–81. <https://doi.org/10.1002/per.1894>
- Ramos-Villagrasa, P. J., Fernández-del-Río, E., & Barrada, J. R. (2020). Does Evil Prevail? The “Bright” and “Dark” Sides of Personality as Predictors of Adaptive Performance. *Sustainability, 12*(2), 474. <https://doi.org/10.3390/su12020474>
- Ramos-Villagrasa, P. J., Passos, A. M., & García-Izquierdo, A. L. (2019). From Planning to Performance: The Adaptation Process as a Determinant of Outcomes. *The Journal of Applied Behavioral Science, 55*(1), 96–114. <https://doi.org/10.1177/0021886318807484>
- Randstad. (2018). *El 24'9% de los trabajadores públicos son temporales, la tasa más elevada desde 2009*. (Nota de prensa) <https://www-randstad-es.s3.amazonaws.com/wp-content/uploads/2018/10/El-249-de-los-trabajadores-p%C3%BAblicos-son-temporales-la-tasa-m%C3%A1s-elevada-desde-2009.pdf>

- Reuterberg, S. E., & Gustafsson, J. E. (1992). Confirmatory factor analysis and reliability: Testing measurement model assumptions. *Educational and Psychological Measurement*, 52, 795–811. <https://doi.org/10.1177/0013164492052004001>
- Richardson, H. A., Simmering, M. J., & Sturman, M. C. (2009). A tale of three perspectives: Examining post hoc statistical techniques for detection and correction of common method variance. *Organizational Research Methods*, 12(4), 762-800. <https://doi.org/10.1177/1094428109332834>
- Rigdon, E. E. (1996). CFI versus RMSEA: A comparison of two fit indexes for structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 3(4), 369-379. <https://doi.org/10.1080/10705519609540052>
- Roberts, B. W., & Mroczek, D. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science*, 17(1), 31-35. <https://doi.org/10.1111/j.1467-8721.2008.00543.x>
- Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, 132, 1–25. <https://doi.org/10.1037/0033-2909.132.1.1>
- Robinson, M. D., & Sedikides, C. (2018). Personality and The Self. In P. Corr & G. Matthews (Eds.), *Cambridge University Press handbook of personality* (2nd ed.). Cambridge University Press. https://www.researchgate.net/publication/329828901_Personality_and_the_self
- Rodríguez-Fornells, A., Lorenzo-Seva, U., & Andrés-Pueyo, A. (2001). Psychometric properties of the Spanish adaptation of the Five Factor Personality Inventory. *European Journal of Psychological Assessment*, 17(2), 145-153. <https://doi.org/10.1027//1015-5759.17.2.145>
- Rodríguez González, J. (2004). La tecnología de evaluación ipsativa y el sistema APP-APT de Thomas Internacional. *Revista de Psicología del Trabajo y de las Organizaciones*, 20(2), 169-197. <https://journals.copmadrid.org/jwop/files/93978.pdf>
- Roivainen, E. (2020). Generational changes in personality: the evidence from corpus linguistics. *Psychological Reports*, 123(2), 325-340. <https://doi.org/10.1177/0033294118805937>
- Rotundo, M., & Sackett, P. R. (2002). The relative importance of task, citizenship, and counterproductive performance to global ratings of job performance: A policy capturing approach. *Journal of Applied Psychology*, 87, 66-80. <https://doi.org/10.1037/0021-9010.87.1.66>
- Roulin, N., & Krings, F. (2019). Faking to Fit in: Applicants' Response Strategies to Match Organizational Culture. *Journal of Applied Psychology*, 105(2), 130-145. <https://doi.org/10.1037/apl0000431>

- Ryan, A. M., & Ployhart, R. E. (2000). Applicants' perceptions of selection procedures and decisions: A critical review and agenda for the future. *Journal of Management*, 26, 565-606. <http://dx.doi.org/10.1177/014920630002600308>
- Ryan, A. M., & Ployhart, R. E. (2014). A century of selection. *Annual Review of Psychology*, 65, 693-717. <http://dx.doi.org/10.1146/annurev-psych-010213-115134>
- Sackett, P. R., y Lievens, F. (2008). Personnel selection. *Annual Review of Psychology*, 59, 419–450. <http://dx.doi.org/10.1146/annurev.psych.59.103006.093716>
- Sackett, P. R., Lievens, F., Van Iddekinge, C. H., & Kuncel, N. R. (2017). Individual differences and their measurement: A review of 100 years of research. *Journal of Applied Psychology*, 102(3), 254-273. <http://dx.doi.org/10.1037/apl0000151>
- Sackett, P. R., & Walmsley, P. T. (2014). Which personality attributes are most important in the workplace? *Perspectives on Psychological Science*, 9(5), 538-551. <http://dx.doi.org/10.1177/1745691614543972>
- Salgado, J. F. (1998). *Manual Técnico del Inventario de Personalidad 5F (IP/5F)*. Tórculo Edicions.
- Salgado, J.F. (2003). Predicting job performance by FFM and non-FFM personality measures. *Journal of Occupational and Organizational Psychology*, 76, 323-346. <https://doi.org/10.1348/096317903769647201>
- Salgado, J. F. (2005). Personalidad y deseabilidad social en contextos organizacionales: implicaciones para la práctica de la psicología del trabajo y las organizaciones. *Papeles del Psicólogo*, 26, 115-128. <http://www.papelesdelpsicologo.es/pdf/1252.pdf>
- Salgado, J. F. (2014). *Technical manual for the QI5F*. Metis.
- Salgado, J. F. (2016). A Theoretical Model of Psychometric Effects of Faking on Assessment Procedures: Empirical findings and implications for personality at work. *International Journal of Selection and Assessment*, 24(3), 209-228. <https://doi.org/10.1111/ijsa.12142>
- Salgado, J. F. (2017). Moderator effects of job complexity on the validity of forced-choice personality inventories for predicting job performance. *Revista de Psicología del Trabajo y de las Organizaciones*, 33(3), 229-238. <https://doi.org/10.1016/j.rpto.2017.07.001>
- Salgado, J. F., Anderson, N., & Tauriz, G. (2015). The validity of ipsative and quasi-ipsative forced-choice personality inventories for different occupational groups: A comprehensive meta-analysis. *Journal of Occupational and Organizational Psychology*, 88(4), 797-834. <https://doi.org/10.1111/joop.12098>
- Salgado, J. F., Gorriti M., y Moscoso, S. (2007). La entrevista estructurada y el desempeño laboral en la administración pública española: propiedades psicométricas y reacciones de justicia. *Revista de Psicología del Trabajo y de las Organizaciones*, 23(1), 151-178. <https://journals.copmadrid.org/jwop/files/103493.pdf>

- Salgado, J. F., & Lado, M. (2018). Faking Resistance of a Quasi-ipsative Forced-Choice Personality Inventory without Algebraic Dependence. *Revista de Psicología del Trabajo y las Organizaciones*, 34(3), 213-216. <http://dx.doi.org/10.5093/jwop2018a23>
- Salgado, J. F., y Moscoso, S. (2008). Selección de personal en la empresa y las Administraciones Públicas: de la visión tradicional a la visión estratégica. *Papeles del Psicólogo*, 29(1), 16-24. <http://www.papelesdelpsicologo.es/pdf/1534.pdf>
- Salgado, J. F., Moscoso, S., Sanchez, J. I., Alonso, P., Choragwicka, B., & Berges, A. (2015). Validity of the five-factor model and their facets: The impact of performance measure and facet residualization on the bandwidth-fidelity dilemma. *European Journal of Work and Organizational Psychology*, 24(3), 325-349. <http://dx.doi.org/10.1080/1359432X.2014.903241>
- Samale, S. L. (2016). *Investigating the factor structure of adaptability: a measure of adaptive performance* (Order No. 10146816) [Doctoral dissertation, Middle Tennessee State University]. ProQuest Dissertations & Theses Global. (1829637027).
- Sanz, J., Gil, F., Barrasa, Á., & García-Vera, M. P. (2006). Self-assessment of needs and behavior patterns at work: Psychometric properties of the Personality and Preference Inventory-Normative (PAPI-N). *Personality and Individual Differences*, 41(5), 837-847. <https://doi.org/10.1016/j.paid.2006.03.019>
- Schipolowski, S., Schroeders, U., & Wilhelm, O. (2014). Pitfalls and challenges in constructing short forms of cognitive ability measures. *Journal of Individual Differences*, 35, 190–200. <https://doi.org/10.1027/1614-0001/a000134>
- Schlueter, J. E., Carlson, J. F., Geisinger, K. F., & Murphy, L. L. (2018). *Pruebas publicadas en Español II: An index of Spanish tests in print*. Buros Center for Testing. <https://buros.org/pruebas-publicadas-en-espanol#esp>
- Schmidt, E., & Rosenberg, J. (2014). *How Google works*. Grand Central Publishing.
- Schmidt, F. L. Oh, I. -S., & Shaffer, J. A. (2016). *The validity and utility of selection methods in personnel psychology: practical and theoretical implications of 100 years of research findings*. Fox School of Business Research Paper. https://www.researchgate.net/publication/309203898_The_VValidity_and_Utility_of_Selection_Methods_in_Personnel_Psychology_Practical_and_Theoretical_Implications_of_100_Years_of_Research_Findings
- Schmitt, D. P., Allik, J., McCrae, R. R., & Benet-Martínez, V. (2007). The geographic distribution of Big Five personality traits: Patterns and profiles of human self-description across 56 nations. *Journal of Cross-Cultural Psychology*, 38(2), 173-212. <https://doi.org/10.1177/0022022106297299>
- Schmitt, N., & Oswald F.L. (2006) The impact of corrections for faking on the validity of noncognitive measures in selection settings. *Journal of Applied Psychology*, 3, 613-621. <https://doi.org/10.1037/0021-9010.91.3.613>

- Schmitt, N., Cortina, J. M., Ingerick, M. J., & Wiechmann, D. (2003). Personnel selection and employee performance. In W. C. Borman, D. R. Ilgen & R. J. Klimoski (Eds.), *Handbook of psychology* (pp. 77-106). Wiley. <https://doi.org/10.1002/0471264385.wei1205>
- Schneider, B. (1987). The people make the place. *Personnel Psychology*, 40, 437–453. <https://doi.org/10.1111/j.1744-6570.1987.tb00609.x>
- Schwaba, T., & Bleidorn, W. (2018). Personality Trait Development Across the Transition to Retirement. *Journal of Personality and Social Psychology*, 116, (4), 651–665. <https://doi.org/10.31219/osf.io/j6byv>
- Segura-Camacho, A., Rodríguez-Cifuentes, F., Sáenz De la Torre, L. C., & Topa, G. (2018). Successful Aging at Work: Psychometric Properties of the Spanish Version of Selection, Optimization and Compensation Questionnaire. *Frontiers in Psychology*, 9, Article 410. <https://doi.org/10.3389/fpsyg.2018.00410>
- Seibert, S. E., & Kraimer, M. L. (2001). The five-factor model of personality and career success. *Journal of Vocational Behavior*, 58(1), 1-21. <https://dx.doi.org/10.1006/jvbe.2000.1757>
- Seiseddos, N. (1993). Personnel selection, questionnaires, and motivational distortion: An intelligent attitude of adaptation. In H. Schuler, J. L. Farr, & M. Smith (Eds.), *Personnel selection and assessment* (pp. 91–108). Erlbaum.
- Seong, J. Y., Hong, D. S., & Park, W. W. (2012). Work status, gender, and organizational commitment among Korean workers: The mediating role of person-organization fit. *Asia Pacific Journal of Management*, 29(4), 1105-1129. <https://doi.org/10.1007/s10490-011-9248-6>
- Shaffer, J. A., & Postlethwaite B. E. (2012). A matter of context: a meta-analytic investigation of the relative validity of contextualized and noncontextualized personality measures. *Personnel Psychology*, 65, 445–494. <https://doi.org/10.1111/j.1744-6570.2012.01250.x>
- Sharma, J. P., Bajpai, N., & Holani, U. (2011). Organizational citizenship behavior in public and private sector and its impact on job satisfaction: A comparative study in Indian perspective. *International Journal of Business and Management*, 6(1), 67-75. <https://doi.org/10.5539/ijbm.v6n1p67>
- Shaw, J. D., Gupta, N., & Delery, J. E. (2005). Alternative conceptualizations of the relationship between voluntary turnover and organizational performance. *Academy of Management Journal*, 48(1), 50-68. <https://doi.org/10.5465/amj.2005.15993112>
- Shen, W., Sackett, P. R., Lievens, F., Schollaert, E., Van Hove, G., Steiner, D.D., Rolland, F., Georgior, K., Nikolaou, I., Tomprou, M., Tzafirir, S., Bamberger, P., Bertolino, M., Mariani, M. Fraccaroli, F., Sekiguchi, T., Onyura, B., Yang, H., Oostrom, J.K., ... & Cook, M. (2017). Updated perspectives on the international legal environment for selection. In *Handbook of employee selection* (pp. 659-677). Routledge. <https://doi.org/10.4324/9781315690193-29>

- SHL. (2006). *OPQ32 Technical Manual*. SHL Group.
- Shmelyov, A. G., & Pokhil'ko, V. I. (1993). A taxonomy-oriented study of Russian personality-trait names. *European Journal of Personality*, 7(1), 1-17. <https://doi.org/10.1002/per.2410070102>
- Shoss, M. K., Witt, L. A., & Vera, D. (2012). When does adaptive performance lead to higher task performance? *Journal of Organizational Behavior*, 33, 910-924. <https://doi.org/10.1002/job.780>
- Silva, F., Avia, M.D., Sanz, J., Martínez-Arias, R.M., Graña, J.L., & Sánchez-Bernardos, M.L. (1994). The Five Factor Model: I. Contributions to the structure of the NEO-PI. *Personality and Individual Differences*, 17, 741-753. [http://dx.doi.org/10.1016/0191-8869\(94\)90043-4](http://dx.doi.org/10.1016/0191-8869(94)90043-4)
- Sireci, S. G., & Allalouf, A. (2003). Appraising item equivalence across multiple languages and cultures. *Language Testing*, 20(2), 148-166. <https://doi.org/10.1191/0265532203lt249oa>
- Sireci, S., & Padilla-García, J. L. (2014). Validating assessments: introduction to the special section. *Psicothema*, 26(1), 97-99. <https://doi.org/10.7334/psicothema2013.255>
- Sireci, S. G., Thissen, D., & Wainer, H. (1991). On the reliability of testlet-based tests. *Journal of Educational Measurement*, 28(3), 237-247. <https://doi.org/10.1111/j.1745-3984.1991.tb00356.x>
- Slocum, S. L., Gelin, M. N., & Zumbo, B. D. (2004). *Statistical and Graphical Modeling to Investigate Differential Item Functioning for Rating Scale and Likert Item Formats*. Edgeworth Laboratory, University of British Columbia. http://brunozumbo.com/wp-content/uploads/2016/12/Slocum_Gelin_Zumbo.pdf
- Smaldino, P. E., Lukaszewski, A., von Rueden, C., & Gurven, M. (2019). Niche diversity can explain cross-cultural differences in personality structure. *Nature Human Behaviour*, 3(12), 1276-1283. <https://doi.org/10.1038/s41562-019-0730-3>
- Smith, G. M. (1967). Usefulness of peer ratings of personality in educational research. *Educational and Psychological Measurement*, 27, 967-984. <https://doi.org/10.1177/001316446702700445>
- Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment*, 12(1), 102-111. <https://doi.org/10.1037/1040-3590.12.1.102>
- Smith M., & Smith P. (2005). *Testing People at Work: Competencies in Psychometric Testing*. Wiley-Blackwell.
- Solano-Flores, G., Contreras-Niño, L. A., y Backhoff-Escudero, E. (2006). Traducción y adaptación de pruebas: Lecciones aprendidas y recomendaciones para países participantes en TIMSS, PISA y otras comparaciones internacionales. *Revista Electrónica de Investigación Educativa*, 8, 2-21. <https://redie.uabc.mx/redie/article/view/143>

- Solé, M. D. (2006). *Validació i estandarització espanyola del NEO-PI-R, NEO-FFI, NEO-FFI-R i escales de Schinka, en mostres universitàries i població general*. [Tesis Doctoral, Universitat de Lleida]. Repositori Obert UDL. <http://hdl.handle.net/10459.1/63798>
- Somer, O., & Goldberg, L. R. (1999). The structure of Turkish trait-descriptive adjectives. *Journal of Personality and Social Psychology*, 76(3), 431-450. <https://doi.org/10.1037/0022-3514.76.3.431>
- Soto, C. J. (2019). How replicable are links between personality traits and consequential life outcomes? The Life Outcomes of Personality Replication Project. *Psychological Science*, 30(5), 711-727. <https://doi.org/10.1177/0956797619831612>
- Soto, C. J., & John, O. P. (2009). Ten facet scales for the Big Five Inventory: Convergence with NEO PI-R facets, self-peer agreement, and discriminant validity. *Journal of Research in Personality*, 43(1), 84-90. <https://doi.org/10.1016/j.jrp.2008.10.002>
- Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, 113(1), 117-143. <https://doi.org/10.1037/pspp0000096>
- Soto, C. J., & John, O. P. (2019). Optimizing the length, width, and balance of a personality scale: How do internal characteristics affect external validity? *Psychological Assessment*, 31(4), 444-459. <https://doi.org/10.1037/pas0000586>
- Soto, C.J., John, O.P., Gosling, S.D., & Potter, J. (2008). The developmental psychometrics of Big Five self-reports: Acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology*, 94, 718-737. <https://doi.org/10.1037/0022-3514.94.4.718>
- Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology*, 100(2), 330-348. <https://doi.org/10.1037/a0021717>
- Soto, C. J., Napolitano, C. M., & Roberts, B. W. (in press). Taking skills seriously: Toward an integrative model and agenda for social, emotional, and behavioral skills. *Current Directions in Psychological Science*. http://www.colby.edu/psych/wp-content/uploads/sites/50/2020/08/Soto_et_al_in_press.pdf
- Specht, J., Egloff, B., & Schmukle, S. C. (2011). Stability and change of personality across the life course: The impact of age and major life events on mean-level and rank-order stability of the Big Five. *Journal of Personality and Social Psychology*, 101(4), 862-882. <https://doi.org/10.1037/a0024950>
- Spitzmuller, M., Van Dyne, L., & Ilies, R. (2008). Organizational citizenship behavior: A review and extension of its nomological network. In J. Barling & C. L. Cooper (Eds.) *The SAGE handbook of organizational behavior* (pp. 106-123). SAGE. <https://doi.org/10.4135/9781849200448.n7>

- Srivastava, S. (2019). *Measuring the big five personality domains*. Consultado el 11 de octubre de 2020. <https://pages.uoregon.edu/sanjay/bigfive.html>
- Stokes, C. K., Schneider, T. R., & Lyons, J. B. (2010). Adaptive performance: A criterion problem. *Team Performance Management: An International Journal*, 16(3/4), 212-230. <https://doi.org/10.1108/13527591011053278>
- Streiner, D. L. (2003). Starting at the beginning: an introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*, 80(1), 99-103. https://doi.org/10.1207/s15327752jpa8001_18
- Su, R., Murdock, C. D., & Rounds, J. (2015). Person-environment fit. In P. Hartung, M. Savickas, & B. Walsh (Eds.), *APA Handbook of career intervention Vol. 1. Foundations* (pp. 81-98). American Psychological Association. <https://doi.org/10.1037/14438-005>
- Su-ju, L. (2006). On how to achieve functional equivalence in translation between Chinese and English. *Sino-US English Teaching*, 3, 71-75.
- Sung, S. Y., Antefelt, A., & Choi, J. N. (2017). Dual effects of job complexity on proactive and responsive creativity: Moderating role of employee ambiguity tolerance. *Group & Organization Management*, 42(3), 388-418. <https://doi.org/10.1177/1059601115619081>
- Sverke, M., Hellgren, J., & Näswall, K. (2002). No security: a meta-analysis and review of job insecurity and its consequences. *Journal of Occupational Health Psychology*, 7(3), 242-264. <https://doi.org/10.1037/1076-8998.7.3.242>
- Swaminathan, H., & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement*, 27(4), 361-370. <https://doi.org/10.1111/j.1745-3984.1990.tb00754.x>
- Swisher, L. L., Beckstead, J. W., & Bebeau, M. J. (2004). Factor analysis as a tool for survey analysis using a professional role orientation inventory as an example. *Physical Therapy*, 84(9), 784-799. <https://doi.org/10.1093/ptj/84.9.784>
- Szarota, P. (1996). Taxonomy of the Polish personality-descriptive adjectives of the highest frequency of use. *Polish Psychological Bulletin*, 27, 343-352.
- Szirmák, Z., & De Raad, B. (1994). Taxonomy and structure of Hungarian personality traits. *European Journal of Personality*, 8(2), 95-117. <https://doi.org/10.1002/per.2410080203>
- ’t Hart, P. T. (2014). *Understanding public leadership*. Macmillan International Higher Education. <https://doi.org/10.1007/978-1-137-45057-9>
- Tambe, S. & Shanker, M. (2014). A study of organizational citizenship behaviour (OCB) and its dimensions: a literature review. *International Research Journal of Business and Management*, 1, 67-73.
- Teachman, B.A. (2006). Aging and negative affect: the rise and fall and rise of anxiety and depressive symptoms. *Psychology and Aging*, 21, 201-207. <https://doi.org/10.1037/0882-7974.21.1.201>

- Terracciano, A., McCrae, R. R., Brant, L. J., & Costa, P. T., Jr. (2005). Hierarchical linear modeling analyses of the NEO-PI-R scales in the Baltimore Longitudinal Study of Aging. *Psychology and Aging, 20*, 493–506. <https://doi.org/10.1037/0882-7974.20.3.493>
- Thalmayer, A. G., Saucier, G., & Eigenhuis, A. (2011). Comparative validity of brief to medium-length Big Five and Big Six personality questionnaires. *Psychological Assessment, 23*, 995–1009. <https://doi.org/10.1037/a0024165>
- Thalmayer, A. G., Saucier, G., Ole-Kotikash, L., & Payne, D. (2019). Personality structure in east and west Africa: Lexical studies of personality in Maa and Supyire-Senufo. *Journal of Personality and Social Psychology. https://doi.org/10.1037/pspp0000264*
- Thurstone, L. L. (1934). The vectors of mind. *Psychological Review, 41*, 1-32. <https://doi.org/10.1037/h0075959>
- Toops, H. A. (1944). The criterion. *Educational and Psychological Measurement, 4*, 271- 297. <http://dx.doi.org/10.1177/001316444400400402>
- Tornimbeni, S., Pérez, E, Olaz, F., y Fernández, A. (2004). *Introducción a los Tests Psicológicos* (3ª Edición revisada y aumentada). Ed. Brujas.
- Torreblanca Murillo, L. (2017). *Desarrollo de un Nuevo Instrumento de Evaluación: El Cuestionario Breve de Personalidad (CBP)* [Tesis doctoral, Universidad de Málaga]. RIUMA. <https://hdl.handle.net/10630/15686>
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511819322>
- Trask, R. L. (1997). *The history of Basque*. Routledge. <https://doi.org/10.4324/9781315004358>
- Tripathi, P., & Burleson, W. (2012, February 11-15). *Predicting creativity in the wild: Experience sample and sociometric modeling of teams* [Conference presentation]. ACM 2012 Conference on Computer Supported Cooperative Work, Seattle, WA. <https://doi.org/10.1145/2145204.2145386>
- Trull, T. J., & Geary, D. C. (1997). Comparison of the big-five factor structure across samples of Chinese and American adults. *Journal of Personality Assessment, 69*(2), 324–341. https://doi.org/10.1207/s15327752jpa6902_6
- Tsai, W. C., Chen, H. Y., & Chen, C. C. (2012). Incremental validity of person-organization fit over the big five personality measures. *The Journal of Psychology, 146*(5), 485-509. <https://doi.org/10.1080/00223980.2012.656154>
- Tubré, T., Arthur, W., Jr., & Bennett, W., Jr. (2006). General models of job performance: Theory and practice. In W. Bennett, Jr., D. J. Woehr, & C. E. Lance (Eds.), *Performance measurement: Current perspectives and future challenges* (pp. 175-203). Lawrence Erlbaum Associates.
- Tucker, J. S., Pleban, R. J., & Gunther, K. M. (2010). The mediating effects of adaptive skill on values-performance relationships. *Human Performance, 23*, 81-99. <https://doi.org/10.1080/08959280903400275>

- Tupes, E. C., & Christal, R. C. (1961). *Recurrent personality factors based on trait ratings*. (USAF ASD Technical Report N° 61-97). U.S. Air Force. <https://doi.org/10.21236/ad0267778>
- Valdivia, J.A. (2014). *Test adaptation activities across languages and cultures*. (Order No. 3640085) [Doctoral dissertation, Washington State University]. ProQuest Dissertations & Theses Global. (1625043213).
- Valentine, A. (2013). *Is translation enough? A study of the item characteristics that influence equivalence between English and Spanish versions of a selection test* (Order No. 3568766) [Doctoral dissertation, State University of New York]. ProQuest Dissertations & Theses Global. (1418807195).
- Van de Vijver, F. J. R., & Hambleton, R. K. (1996). Translating tests: Some practical guidelines. *European Psychologist*, 1(2), 89-99. <https://doi.org/10.1027/1016-9040.1.2.89>
- Van de Vijver, F. J. R., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology*, 54(2), 119-135. <https://doi.org/10.1016/j.erap.2003.12.004>
- Van Dyne, L., Cummings, L. L., & Parks, J. M. (1995). Extra-role behaviors: In pursuit of construct and definitional clarity (a bridge over muddied waters). *Research in Organizational Behavior*, 17, 215-285.
- van Eijck, C. J. M., & de Graaf, P. M. (2004). The Big Five at school: The impact of personality on educational attainment. *The Netherlands' Journal of Social Sciences*, 40, 24-40. <https://repository.ubn.ru.nl/bitstream/handle/2066/64541/64541.pdf?sequence=1>
- Van Hemert, D. A., van de Vijver, F. J. R., Poortinga, Y. H., & Georgas, J. (2002). Structural and functional equivalence of the Eysenck Personality Questionnaire within and between countries. *Personality and Individual Differences*, 33, 1229-1249. [https://doi.org/10.1016/s0191-8869\(02\)00007-7](https://doi.org/10.1016/s0191-8869(02)00007-7)
- Van Iddekinge, C. H., Arnold, J. D., Frieder, R. E., & Roth, P. L. (2019). A meta-analysis of the criterion-related validity of prehire work experience. *Personnel Psychology*, 72(4), 571-598. <https://doi.org/10.1111/peps.12335>
- Van Iddekinge, C. H., & Ployhart, R. E. (2008). Developments in the criterion-related validation of selection procedures: a critical review and recommendations for practice. *Personnel Psychology*, 61(4), 871-925. <https://doi.org/10.1111/j.1744-6570.2008.00133.x>
- Van Vaerenbergh, Y., & Thomas, T. D. (2013). Response styles in survey research: A literature review of antecedents, consequences, and remedies. *International Journal of Public Opinion Research*, 25(2), 195-217. <https://doi.org/10.1093/ijpor/eds021>
- van Vianen, A. E. (2018). Person–Environment Fit: A Review of Its Basic Tenets. *Annual Review of Organizational Psychology and Organizational Behavior*, 5, 75-101. <https://dx.doi.org/10.1146/annurev-orgpsych-032117-104702>

- Van Witteloostuijn, A., Esteve, M., & Boyne, G. (2016). Public sector motivation ad fonts: Personality traits as antecedents of the motivation to serve the public interest. *Journal of Public Administration Research and Theory*, 27(1), 20-35. <https://doi.org/10.1093/jopart/muw027>
- Vanderstukken, A., Proost, K., & Van Den Broeck, A. (2019). Subjective PO fit in recruitment: is it always really 'O'? Organizational values may be industry values, depending on temporal distance. *European Journal of Work and Organizational Psychology*, 28(5), 602-615. <https://doi.org/10.1080/1359432x.2018.1427143>
- Vassend, O., & Skrandal, A. (2011). The NEO personality inventory revised (NEO-PI-R): Exploring the measurement structure and variants of the five-factor model. *Personality and Individual Differences*, 50, 1300–1304. <https://doi.org/10.1016/j.paid.2011.03.002>
- Verquer, M. L., Beehr, T. A., & Wagner, S. H. (2003). A meta-analysis of relations between person–organization fit and work attitudes. *Journal of Vocational Behavior*, 63(3), 473-489. [https://doi.org/10.1016/s0001-8791\(02\)00036-2](https://doi.org/10.1016/s0001-8791(02)00036-2)
- Vigil-Colet, A., Morales-Vives, F., Camps, E., Tous, J., & Lorenzo-Seva, U. (2013). Development and validation of the overall personality assessment scale (OPERAS). *Psicothema*, 25(1), 100-106. <https://doi.org/10.7334/psicothema2011.411>
- Vigoda, E. (2000). Internal politics in public administration systems. *Public Personnel Management*, 29, 185-210. <https://doi.org/10.1177/009102600002900203>
- Vigoda, E., & Golembiewski, R. T. (2001). Citizenship behavior and the spirit of new managerialism: A theoretical framework and challenge for governance. *American Review of Public Administration*, 31, 273-295. <https://doi.org/10.1177/02750740122064956>
- Viladrich, C., Angulo-Brunet, A., & Doval, E. (2017). A journey around alpha and omega to estimate internal consistency reliability. *Anales de Psicología*, 33(3), 755-782. <https://doi.org/10.6018/analesps.33.3.268401>
- Viswesvaran, C. (1993). *Modeling job performance: Is there a general factor?* (Order No. 9404550) [Doctoral dissertation, University of Iowa]. ProQuest Dissertations & Theses Global. (304038896).
- Viswesvaran, C. (2001). Assessment of individual job performance: a review of the past century and a look ahead. In N. Anderson, D. S. Ones, H. K. Sinangil y C. Viswesvaran (Eds.), *Handbook of Industrial and Organizational Psychology* (Vol. 1, pp. 110-126). SAGE. <http://dx.doi.org/10.4135/9781848608320.n7>
- Viswesvaran, C., Ones, D.S., & Schmidt, F.L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology*, 81, 557–574. <https://doi.org/10.1037/0021-9010.81.5.557>

- Viswesvaran, C., Schmidt, F. L., & Ones, D. S. (2005). Is There a General Factor in Ratings of Job Performance? A Meta-Analytic Framework for Disentangling Substantive and Error Influences. *Journal of Applied Psychology, 90*(1), 108–131. <https://doi.org/10.1037/0021-9010.90.1.108>
- Vleugels, W., Tierens, H., Billsberry, J., Verbruggen, M., & De Cooman, R. (2019). Profiles of fit and misfit: a repeated weekly measures study of perceived value congruence. *European Journal of Work and Organizational Psychology, 28*(5), 616-630. <https://doi.org/10.5465/ambpp.2018.16188abstract>
- Wagner, J., Lüdtke, O., & Robitzsch, A. (2019). Does personality become more stable with age? Disentangling state and trait effects for the big five across the life span using local structural equation modeling. *Journal of Personality and Social Psychology, 116*(4), 666–680. <https://doi.org/10.1037/pspp0000203>
- Walker, J. C. (2015). *Adaptability in the workplace: An exploratory study on adaptive performance in the workplace using a scenario-based tool* (Order No. 3723123) [Doctoral dissertation, University of Pennsylvania]. ProQuest Dissertations & Theses Global. (1728882484).
- Wang, M., Zhan, Y., McCune, E., & Truxillo, D. (2011). Understanding newcomers' adaptability and work-related outcomes: Testing the mediating roles of perceived P-E fit variables. *Personnel Psychology, 64*, 163-189. <https://doi.org/10.1111/j.1744-6570.2010.01205.x>
- Wanous, J. P., Reichers, A. E., & Hudy, M. J. (1997). Overall job satisfaction: How good are single-item measures? *Journal of Applied Psychology, 82*, 247–252. <https://doi.org/10.1037/0021-9010.82.2.247>
- Ware Jr., J. E., & Gandek, B. (1998). Methods for testing data quality, scaling assumptions, and reliability: the IQOLA Project approach. *Journal of Clinical Epidemiology, 51*(11), 945-952. [https://doi.org/10.1016/s0895-4356\(98\)00085-7](https://doi.org/10.1016/s0895-4356(98)00085-7)
- Watts, L. L., Frame, M., Rigdon, W. D., & Orsak-Robinson, K. (2011, October). *Exploring the factor structure of adaptive performance: Finding a more parsimonious fit with four dimensions* [Poster presentation]. 7th annual River Cities I/O Psychology Conference, Chattanooga, TN. <https://www.utc.edu/psychology/rcio/rcioarchive/rcio2011.php>
- Welbourne, T. M., Johnson, D. E., & Erez, A. (1998). The Role-Based Performance Scale: Validity analysis of a theory-based measure. *Academy of Management Journal, 41*, 540-555. <https://doi.org/10.5465/256941>
- Wessel, J. L., Ryan, A. M., & Oswald, F. L. (2008). The relationship between objective and perceived fit with academic major, adaptability, and major-related outcomes. *Journal of Vocational Behavior, 72*(3), 363-376. <https://doi.org/10.1016/j.jvb.2007.11.003>
- Wherry, R. J. (1957). The past and future of criterion evaluation. *Personnel Psychology, 10*, 1-5. <http://dx.doi.org/10.1111/j.1744-6570.1957.tb00760.x>

- Whiteman, H. J. (2011). *Differential item functioning and adverse impact: a comparison of Mantel-Haenszel and logistic regression* [Master of Arts Thesis, California State University]. Sacramento Masters Theses. <http://hdl.handle.net/10211.3/138427>
- Williams, L. J., & Anderson, S. E. (1991). Job satisfaction and organizational commitment as predictors of Organizational citizenship and in-role behaviors. *Journal of Management*, *17*, 601-617. <https://doi.org/10.1177/014920639101700305>
- Williams, L. J., Cote, J.A., & Buckley, R. (1989). Lack of method variance in self-reported affect and perceptions at work: reality or artifact? *Journal of Applied Psychology*, *74*, 462-468. <https://doi.org/10.1037/0021-9010.74.3.462>
- Wingen, T., Berkessel, J., & English, B. (2020). No Replication, no Trust? How Low Replicability Influences Trust in Psychology. *Social Psychological and Personality Science*, *11*(4), 454-463. <https://doi.org/10.1177/1948550619877412>
- Woodcock, R. W., & Muñoz-Sandoval, A. (1999, May 20-22). *Test translation guidelines: A "consensus translation" approach* [Conference presentation]. International Conference on Test Adaptation, Georgetown University, Washington, DC.
- Woods, S. A., Wille, B., Wu, C. H., Lievens, F., & De Fruyt, F. (2018). The influence of work on personality trait development: The demands-affordances Transactional (DATA) model, an integrative review, and research agenda. *Journal of Vocational Behavior*, *110*, 258-271. <https://doi.org/10.1016/j.jvb.2018.11.010>
- Wortman, J., Lucas, R. E., & Donnellan, M. B. (2012). Stability and change in the Big Five personality domains: Evidence from a longitudinal study of Australians. *Psychology and Aging*, *27*, 867-874. <https://doi.org/10.1037/a0029322>
- Wright, B. E. (2001). Public-sector work motivation: A review of the current literature and a revised conceptual model. *Journal of Public Administration Research and Theory*, *11*(4), 559-586. <https://doi.org/10.1093/oxfordjournals.jpart.a003515>
- Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behavior Research Methods*, *51*, 409-428. <https://doi.org/10.3758/s13428-018-1055-2>
- Ximénez Gómez, C. (1998). *El ajuste de la persona a la organización. Revisión, planteamiento de una medida y de un estudio empírico en una muestra española* [Tesis Doctoral, Universidad Autónoma de Madrid]. Biblos e-Archivo. <http://hdl.handle.net/10486/4463>
- Xu, M. (2014). *Cross-Cultural Differences in Person-Environment Fit and Organizational Citizenship Behavior* [Master of Science Thesis, Emporia State University]. ESIRC. <http://hdl.handle.net/123456789/3284>
- Yamagata, S., Suzuki, A., Ando, J., Ono, Y., Kijima, N., Yoshimura, K., Ostendorf, F., Angleitner, A., Riemann, R., Spinath, F. M., Livesley, W. J., & Jang, K. L. (2006). Is the genetic structure of human personality universal? Across-cultural twin study from North America, Europe, and Asia. *Journal of Personality and Social Psychology*, *90*, 987-998. <https://doi.org/10.1037/0022-3514.90.6.987>

- Yang, K. S., & Bond, M. H. (1990). Exploring implicit personality theories with indigenous or imported constructs: The Chinese case. *Journal of Personality and Social Psychology*, 58(6), 1087-1095. <https://doi.org/10.1037/0022-3514.58.6.1087>
- Ye, J., Cardon, M. S., & Rivera, E. (2012). A mutuality perspective of psychological contracts regarding career development and job security. *Journal of Business Research*, 65(3), 294-301. <https://doi.org/10.1016/j.jbusres.2011.03.006>
- Yusko, K. P., Bellenger, B. L., Larson, E. C., Hanges, P. J., & Aiken, J. R. (2017). Legal and Fairness Considerations in Employee Selection. In H. W. Goldstein, E. D. Pulakos, J. Passmore, C. Semedo (Eds.) *The Wiley Blackwell Handbook of the Psychology of Recruitment, Selection and Employee Retention*, (pp. 422-441). Wiley. <https://doi.org/10.1002/9781118972472.ch20>
- Zhang, X., Gao, Y., Chen, H., Sun, Y., & de Pablos, P. O. (2015). *Enhancing Creativity or Wasting Time? The Mediating Role of Adaptability on Social Media-Job Performance Relationship* [Conference presentation]. Pacific Asia Conference on Information Systems (PACIS). <https://pdfs.semanticscholar.org/73c5/d4be8c817e8ff60d40151f1da35e5c246bba.pdf>
- Ziegler, M., Kemper, C.J., & Kruyen, P. (2014). Short Scales – Five misunderstandings and ways to overcome them. *Journal of Individual Differences*, 35, 185–189. <https://doi.org/10.1027/1614-0001/a000148>
- Zuazo, K. (1998). Euskalkiak, gaur. *Fontes Linguae Vasconum*, 78, 191-234. <https://dialnet.unirioja.es/descarga/articulo/26274.pdf>
- Zuckerman, M. (1991). *Psychobiology of personality*. Cambridge University press. <https://doi.org/10.1017/CBO9780511813733>
- 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. [https://doi.org/10.1016/0191-8869\(91\)90182-b](https://doi.org/10.1016/0191-8869(91)90182-b)
- Zumbo, B. D. (2003). Does item-level DIF manifest itself in scale-level analyses? Implications for translating language tests. *Language Testing*, 20(2), 136-147. <https://doi.org/10.1191/0265532203lt248oa>