



*16PF® Questionnaire
European English version of the 16PF Fifth Edition*

European English Data Supplement 2011

December 2011

© Copyright 2011 Institute for Personality and Ability Testing Inc. (IPAT), PO Box 1188, Champaign, IL 61824-1188, USA. All rights reserved.

No portion of this publication may be translated or reproduced in whole or in part, stored in a retrieval system, or transmitted in any form or by any means (electronic, mechanical, photocopying, recording or otherwise) without the prior written permission of the copyright owner.

This publication may not be resold, rented, lent, leased, exchanged, given or otherwise disposed of to third parties. Neither the purchaser nor any individual test user employed by or otherwise contracted to the purchaser may act as agent, distribution channel or stockist for this publication.

® 16PF is a registered trade mark of the Institute for Personality and Ability Testing, Inc. (IPAT) in the United States, the European Community and other countries.

OPP Ltd, Elsfeld Hall, 15–17 Elsfeld Way, Oxford OX2 8EP

t: +44 (0)845 603 9958

f: +44 (0)1865 557483

www.opp.eu.com

® OPP is a registered trade mark of OPP Ltd.

Table of contents

Introduction	2
Sample.....	2
Descriptive statistics.....	2
Norms	3
Reliability.....	6
Primary scale factor analysis.....	9
Primary Factor intercorrelations	11
Summary.....	12
References.....	12
Appendix 1: Sample description	13

Introduction

This 16PF data supplement for the European English version of the 16PF questionnaire complements the current UK *16PF Manual*. The results reported in the data supplement are based on data gathered from a nationally representative sample of people of working age in the UK and Ireland. Where appropriate, the results obtained from this sample are compared with the findings on the UK standardisation sample (N=1,322) published in the UK *16PF Manual* by M.T. Russell and D.L. Karol (1994). This data supplement provides further evidence of the psychometric soundness and good validity of the instrument.

Sample

The sample consisted of 1,212 respondents (606 female and 606 male) who completed the European English version of the 16PF 5th Edition questionnaire between February and March 2011 via an online data collection platform. The sample was representative of the UK & Ireland working age population. Full sample details are shown in Appendix 1.

Descriptive statistics

Statistical information (mean and standard deviation) for the raw scales for the nationally representative working age sample (N=1,212) can be found in Table 1. The table also provides the same information for the UK standardisation sample (N=1,322). The "Raw score mean difference" column shows the scale mean difference between both samples.

Independent t-tests were conducted to determine if the differences in means between both samples were significant. A significant difference ($p < 0.05$) was found for nine out of the sixteen Primary Factors (Warmth (A), Emotional Stability (C), Social Boldness (H), Vigilance (L), Abstractedness (M), Apprehension (O), Openness to Change (Q1), Self Reliance (Q2), Tension (Q4)). However, when sample sizes are large, as is the case here, even a small difference in means is statistically significant. Statistical significance merely indicates that one can be confident that there is a difference between two samples. This does not necessarily imply that the difference is practically meaningful.

In order to determine if an observed difference is not only statistically significant but also important or meaningful, effect sizes are used. The effect size (d) was calculated by subtracting the mean for the standardisation sample (N=1,322) from the mean for the nationally representative working age sample (N=1,212) and dividing this by the pooled standard deviation. The effect size is a standardised value, ie all effect sizes are calculated on a common scale.

The effect size results in Table 1 show that there is only a small effect size (below the traditional 0.50 cut-off) for all of the Primary Factors. In order to gauge how substantial the differences between both samples are, let us consider how large an effect size is required in order to affect profile interpretation. An effect size of approximately ± 0.50 corresponds to a sten difference of approximately 1 sten. In other words, for all of the Primary Factors, the difference between both samples is within the standard error of measurement, which is approximately 1 sten.

The Primary Factor descriptive statistics reported for the nationally representative working age sample have been used to create new and updated norms for the European English version of the 16PF questionnaire.

Norms

Table 2 presents norms for the 16PF, based on the data gathered for the nationally representative working age sample (N=1,212). They are generally similar, but not identical, to the norms previously generated from the UK standardisation sample data.

Table 1. Means, standard deviations, raw score mean differences and effect sizes for 16PF Primary Factors

Primary Factor	Nationally representative working age sample (N = 1,212)		UK standardisation sample (N=1,322)*		Raw score mean difference	Effect size (d)
	Mean	Standard deviation	Mean	Standard deviation		
A Warmth	12.37	4.82	13.5	4.6	-1.13	-0.24
B Reasoning	9.11	3.08	9.1	3.4	0.01	0.00
C Emotional Stability	11.02	5.13	13.0	4.7	-1.98	-0.40
E Dominance	12.43	4.36	12.5	4.4	-0.07	-0.02
F Liveliness	11.76	4.71	11.4	4.9	0.36	0.07
G Rule-Consciousness	11.87	4.15	12.2	4.8	-0.33	-0.07
H Social Boldness	9.01	6.15	10.0	6.4	-0.99	-0.16
I Sensitivity	12.16	5.02	12.2	5.4	-0.04	-0.01
L Vigilance	13.87	3.65	12.5	3.8	1.37	0.37
M Abstractedness	9.06	4.92	8.0	4.9	1.06	0.22
N Privateness	12.60	4.86	12.7	4.7	-0.1	-0.02
O Apprehension	12.17	5.33	11.5	5.3	0.67	0.13
Q1 Openness to Change	15.96	5.15	14.8	5.4	1.16	0.22
Q2 Self-Reliance	11.02	5.30	9.3	5.1	1.72	0.33
Q3 Perfectionism	11.60	4.76	11.7	5.0	-0.1	-0.02
Q4 Tension	11.61	4.68	11.1	4.9	0.51	0.11

* From the UK 16PF Manual by M.T. Russell and D.L. Karol (1994)

European English 16PF Data Supplement 2011

Table 2. Nationally representative working age sample norms (n=1,212)

Primary Factor		Sten										Mean	SD
		1	2	3	4	5	6	7	8	9	10		
A	Warmth	0-2	3-4	5-7	8-9	10-12	13-15	16-17	18-19	20	21-22	12.37	4.82
B	Reasoning	0-3	4	5	6-7	8-9	10-11	12	13	14	15	9.11	3.08
C	Emotional Stability	0	1-2	3-4	5-7	8-11	12-14	15-16	17-18	19-20	-	11.02	5.13
E	Dominance	0-2	3-4	5-7	8-9	10-12	13-15	16-17	18	19-20	-	12.43	4.36
F	Liveliness	0-1	2-3	4-6	7-8	9-11	12-14	15-16	17-18	19-20	-	11.76	4.71
G	Rule-Consciousness	0-3	4-5	6-7	8-9	10-12	13-14	15-16	17-18	19-20	21-22	11.87	4.15
H	Social Boldness	-	-	0-1	2-4	5-9	10-13	14-16	17-18	19-20	-	9.01	6.15
I	Sensitivity	0-2	3-4	5-6	7-9	10-12	13-15	16-17	18-19	20-21	22	12.16	5.02
L	Vigilance	0-5	6-7	8-10	11-12	13-14	15-16	17	18	19-20	-	13.87	3.65
M	Abstractedness	-	0-1	2-3	4-5	6-8	9-11	12-14	15-17	18-19	20-22	9.06	4.92
N	Privateness	0-2	3-4	5-7	8-9	10-13	14-16	17-18	19	20	-	12.60	4.86
O	Apprehension	0-1	2-3	4-5	6-8	9-12	13-16	17-18	19	20	-	12.17	5.33
Q1	Openness to Change	0-5	6-7	8-10	11-12	13-15	16-18	19-21	22-23	24-25	26-28	15.96	5.15
Q2	Self-Reliance	0	1-2	3-5	6-7	8-11	12-14	15-16	17-18	19-20	-	11.02	5.30
Q3	Perfectionism	0-1	2-3	4-5	6-8	9-11	12-14	15-16	17-18	19-20	-	11.60	4.76
Q4	Tension	0-2	3-4	5-6	7-8	9-11	12-14	15-16	17-18	19-20	-	11.61	4.68

Reliability

Reliability gauges the consistency of test results. As a generic term, it relates to a number of different aspects of consistency. Essentially, a reliable test yields the same approximate results when administered repeatedly under similar conditions. Reliability is relevant as it describes how accurately an instrument measures the construct. It is closely related to measurement error. The higher the reliability, the smaller the band width around the observed score. It is within this band width that a person's true score is most likely to be.

The aspect of reliability addressed here is that of internal consistency, or homogeneity, of the test items, as measured by Cronbach's coefficient alpha (Cronbach, 1951). Internal consistency of the 16 factors measured by the 16PF questionnaire reflects the degree to which that set of scale items is sampling the same personality domain. In statistical terms, internal consistency reliability displays how large the intercorrelation is between the items that make up each of the 16 personality scales. Cronbach's coefficient alpha essentially calculates the average value of all possible split-half reliabilities. Internal consistency can be viewed as reliability estimated from a single test administration. As the intercorrelations among items within a scale increase, reliability of the scale itself increases. Internal consistency is lowered to the degree that items on the same scale measure different traits, or to the extent that scale items are not intercorrelated. However, it needs to be noted that – even though a high reliability coefficient is desirable – it can also lead to a scale that is too narrow in measuring a construct.

Cronbach alpha coefficients for the 16PF questionnaire were calculated based on the nationally representative working age sample of respondents. Table 3 presents a comparison with the UK standardisation data of the coefficients for each primary scale.

Two aspects need to be taken into consideration when judging the obtained reliability coefficients.

Firstly, as mentioned earlier, reliability is dependent on the breadth of the measured construct. The 16PF Primary Factors measure distinct constructs as confirmed by factor analysis (for details see the Factor Analysis section of this data supplement). Nevertheless, when measuring the factors of personality, several behavioural preferences related to the construct are covered in the items constituting one factor. For example, the items of the factor Warmth (A) do not only cover the extent to which a person cares about others, but also how much a person is interested in spending time with others and, additionally, talking about other people's personal concerns. Making the scale more reliable by removing aspects of the constructs (ie including items about only a restricted number of behavioural preferences per scale) would mean that the construct is not accounted for in its whole range. In other words, one would increase reliability at the expense of measuring the construct adequately.

Secondly, the 16PF instrument is a personality questionnaire. Personality questionnaires are usually based on self-reports and thus measure typical behaviour. Reporting on typical behaviour rather than displayed behaviour, as is the case with ability tests, introduces more variance in the responses provided by an individual (Chernyshenko, Stark, Chan, Drasgow, & Williams, 2001), thus lowering the reliability. In addition, when judging the reliability coefficients, the

application of the 16PF questionnaire needs to be considered. The results obtained when administering the 16PF instrument are validated in a feedback process where a practitioner discusses the profile with the respondent. Due to this process, it is possible to explore a person's true score by reflecting on the person's preferences, events that may have impacted on the responses given and by finding practical evidence. Taking this into consideration, the obtained coefficients are highly satisfactory.

The reliability coefficients can also be assessed based on the criteria for rating the technical qualities of an instrument defined by the European Federation of Psychologists' Associations (EFPA). Using EFPA's rating system, four Primary Factors achieve adequate reliability (ie coefficients between 0.60 and 0.70), ten factors show good reliability (ie coefficients between 0.70 and 0.80) and two display excellent reliability (ie coefficient larger than 0.80). A comparison with the reliability coefficients based on the UK standardisation sample shows that the differences between both samples with regard to reliability are small, thus confirming that the psychometric properties of the European English 16PF questionnaire are stable over time.

Table 3. Internal consistency for 16PF Primary Factors

		Nationally representative working age sample (N = 1,212)		UK standardisation sample (N=1,322)*	
	Primary Factor	Number of items	Cronbach's alpha	Cronbach's alpha	
A	Warmth	11	0.70	0.69	
B	Reasoning	15	0.71	0.80	
C	Emotional Stability	10	0.77	0.73	
E	Dominance	10	0.68	0.68	
F	Liveliness	10	0.74	0.74	
G	Rule-Consciousness	11	0.62	0.70	
H	Social Boldness	10	0.87	0.87	
I	Sensitivity	11	0.71	0.76	
L	Vigilance	10	0.64	0.60	
M	Abstractedness	11	0.71	0.71	
N	Privateness	10	0.76	0.72	
O	Apprehension	10	0.79	0.77	
Q1	Openness to Change	14	0.63	0.65	
Q2	Self-Reliance	10	0.80	0.75	
Q3	Perfectionism	10	0.73	0.74	
Q4	Tension	10	0.73	0.73	

* From the UK 16PF Manual by M.T. Russell and D.L. Karol (1994)

Primary scale factor analysis

Exploratory factor analysis is a statistical technique for discovering, within a large set of variables, a smaller set of variables that can explain much of the larger domain. Raymond Cattell's original development of the 16PF questionnaire used factor analysis to identify 16 Primary Factors. Based on the sample described above (N=1,212), an analysis was conducted to examine if the same 16 Primary Factors could be replicated from the European English 16PF questionnaire.

The factor structure of the final set of items was examined for the sample using the procedure discussed by Conn and Rieke (1994). Items within each factor were grouped into 'parcels' based upon the strength of their correlations with items within the same scale. Hence the term 'parcels' refers to small groupings of items within a scale. For each Primary Factor, three or four items were summed within each parcel in order to achieve a parcel score. Each scale was partitioned into three to four parcels, resulting in a total of 49 parcels.

These parcels, rather than separate items, were factor analysed, because it has been shown that parcels are more reliable (Berstein & Teng, 1989; Cattell & Burdsal, 1975; Gorsuch, 1983). In addition, item responses on the 16PF instrument tend to follow a bimodal rather than normal distribution, which violates one of the statistical assumptions of normal theory factor analysis. By grouping three or four items together, the distribution better approximates a normal distribution, thus providing a better estimate of the factor structure.

In accordance with Cattell's theoretical basis for the original development of the 16PF questionnaire, an oblique rather than an orthogonal factor analysis was conducted of the parcels. Principal Axis Factoring was conducted using the statistical package SPSS. This was followed by an oblique rotational method (Promax) with the Kappa value set at 3. The Primary Factor correlation structure is reported in Table 4 with absolute loadings <0.20 excluded.

Overall, the pattern shows a very good, simple structure for the 16PF Primary Factors. All but four of the 49 parcels exhibit the highest loading onto the factor to which they were assigned. The factor loadings of the parcels onto their respective factors range from 0.20 to 0.84 (median of 0.64 and mean of 0.63); and 44 of the 49 parcels (90%) showed a loading of 0.50 or higher, thus suggesting strong links between the parcels and their assigned factor. In addition, as can be seen in Table 4, there are only six cross-loadings equal to or larger than +/-0.20. All other parcels display close-to-zero loadings onto other factors, demonstrating that these parcels represent distinct constructs that are only represented in their assigned factor, and not in the remaining factors that measure other traits.

The parcels displaying cross-loadings equal to or larger than +/-0.20 showed higher or equal loadings onto their respective factors, with four exceptions; Parcel A1 loaded at -0.38 on Privatness (N), and Parcels C1, C2 and C3 loaded at -0.52, -0.67 and -0.63 on Apprehension (O). This corresponds with the fact that the Emotional Stability and Apprehension Primary Factors correlate highly with each other. The results as a whole confirm empirically the strong conceptual links between the item parcels and their assigned factors.

In summary, 16 factors are clearly defined, corresponding to Cattell's 16 Primary Factors in the US 16PF questionnaire and many other language versions of the instrument. Such a clear Primary Factor structure of the European English 16PF questionnaire provides excellent evidence for its construct validity.

Table 4. Rotated factor pattern loadings of 16PF Primary Factors (N=1,212; 606 females, 606 males)

Parcel	Factor															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
	O	Q2	N	B	I	Q3	H	E	Q1	Q4	F	L	M	G	A	C
A1			-38												23	
A2					26										42	
A3															55	
B1				79												
B2				51												
B3				54												
B4				74												
C1	-52															35
C2	-67															20
C3	-63															31
E1								63								
E2								59								
E3								64								
F1											61					
F2											59					
F3											66					
G1														50		
G2														67		
G3														55		
H1								77								
H2								70								
H3								82								
I1					60											
I2					76											
I3					64											
L1												66				
L2												57				
L3												65				
M1													55			
M2													64			
M3													79			
N1			70													
N2			84													
N3			57													
O1	66															
O2	78															
O3	73															
Q1_1									64							
Q1_2									54							
Q1_3									66							
Q2_1		77														
Q2_2		63														
Q2_3		84														
Q3_1					72											
Q3_2					66											
Q3_3					66											
Q4_1										60						32
Q4_2										75						
Q4_3										69						

Note. Decimals omitted. Factor loading less than absolute value 0.2 deleted.

Primary Factor intercorrelations

Although the factor pattern shows that the 16PF items tend to associate with their own scale and not with others, the Primary Factor scales do clearly show a predictable pattern of intercorrelations, because the factors are oblique. Table 5 presents intercorrelations of the Primary Factor scales for the sample of respondents (N=1,212) described above.

Due to an absence of primary factor intercorrelations based on the UK Standardisation sample, a comparison with the results obtained for the US standardisation sample, as reported in the US 16PF Manual, was used. This comparison shows that the differences are relatively minor (absolute difference: mean 0.00 median 0.01.) This suggests that the relationships between the factors of the 16PF questionnaire are stable across different samples and over time.

Table 4. 16PF Primary Factor intercorrelations (N=1,212; 606 females, 606 males)

	A	B	C	E	F	G	H	I	L	M	N	O	Q1	Q2	Q3
A															
B	-21														
C	12	03													
E	05	03	23												
F	41	-10	25	18											
G	09	-02	07	-04	-17										
H	35	-10	41	36	51	-03									
I	34	-04	-19	-22	10	01	-06								
L	-19	-07	-29	17	-15	-11	-18	-07							
M	-08	-02	-36	-04	04	-33	-10	16	22						
N	-48	07	-16	-10	-37	02	-45	-11	28	04					
O	01	03	-62	-25	-13	06	-42	29	25	28	14				
Q1	15	09	11	31	24	-21	25	09	02	23	-13	-04			
Q2	-46	09	-26	-04	-53	-04	-37	-04	30	18	40	13	-14		
Q3	04	-05	07	16	-10	36	01	-01	07	-33	05	05	-09	10	
Q4	-25	07	-42	09	-17	-13	-28	03	35	20	19	32	-14	28	-00

Note. Decimals omitted.

Summary

The results reported in this 16PF data supplement provide excellent evidence of the instrument's psychometric qualities. They demonstrate that the European English version of the 16PF questionnaire is a reliable and valid tool for the assessment of personality. A comparison of our current findings with the results obtained from the UK standardisation in 1994 shows compelling similarity in the psychometric properties between the data from the two different points in time. This means that what we knew then about the European English version of the 16PF questionnaire is concurrent with what we found in our more recent data. We can therefore conclude that the 16PF instrument is just as current now as it was when the data for the standardisation was gathered. Users of the European English version of the 16PF questionnaire can be confident that the instrument remains a reliable and valid tool for the objective assessment of personality.

References

- Bernstein, I.H., & Teng, G. (1989). Factoring items and factoring scales are different: Spurious evidence of multidimensionality due to item categorization. *Psychological Bulletin*, *105*, 467–477.
- Cattell, R.B., & Burdsal, C.A. (1975). The radial parcel double factoring design: A solution to the item-vs.-parcel controversy. *Multivariate Behavioral Research*, *10*, 165–179.
- Chernyshenko, O. S., Stark, S., Chan, K.-Y., Drasgow, F., & Williams, B. (2001). Fitting Item Response Theory Models to Two Personality Inventories: Issues and Insights. *Multivariate Behavioral Research*, *36*(4), 523–562.
- Conn, S.R., & Rieke, M.L. (1994). *The 16PF Fifth Edition Technical Manual*. Champaign, IL: Institute for Personality and Ability Testing, Inc.
- Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297–334.
- EFPA (2008). EFPA review model for the description and evaluation of psychological tests. Test review form and notes for reviewers – Version 3.42. Brussels: European Federation of Psychologists' Associations.
- Gorsuch, R.L. (1983). *Factor Analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Russell, M.T., & Karol, D.L. (1994). *16PF5 User's Manual*. Champaign, IL: Institute for Personality and Ability Testing, Inc.

Appendix 1: Sample description

The sample consists of 1,212 individuals, specifically sampled to be representative of the UK & Irish working age population. Respondents completed the European English version of the 16PF 5th Edition questionnaire between February and March 2011 via an online data collection platform.

606 of the respondents (50%) were female and 606 (50%) were male.

The mean age was 39 years, and the range was as follows:

Age (years)	Number	Percentage
16-19	51	4.2%
20-24	99	8.2%
25-34	331	27.3%
35-49	449	37.0%
50-65	282	23.3%

A broad range of educational levels was represented, with the largest single group being Bachelors degree for the UK respondents (24%) and School – Leaving Certificate for the Republic of Ireland sample (33%), as shown below:

UK Educational level	Number	Percentage
School – Pre GCSE (or equivalent)	47	5.1%
School – GCSE (or equivalent)	219	23.6%
School – A'Level (or equivalent)	194	20.9%
Uni/College (1 st year)	55	5.9%
Uni/College (2 nd year)	100	10.8%
Bachelors Degree	225	24.3%
Masters Degree	73	7.9%
Doctorate	9	1.0%
Post Doctorate	5	0.5%

Republic of Ireland Educational level	Number	Percentage
School – Junior Certificate	35	12.3%
School – Leaving Certificate	94	33.0%
Uni/College (1 st year)	29	10.2%
Uni/College (2 nd year)	40	14.0%
Bachelors Degree	58	20.4%
Masters Degree	22	7.7%
Doctorate	5	1.8%
Post Doctorate	2	0.7%

European English 16PF Data Supplement 2011

927 respondents lived in the UK & 285 in the Republic of Ireland, with 88% having been born in either the UK or Ireland. Respondents lived in different regions of the UK & Ireland, as follows:

UK Region	Number	Percentage
North East	37	4.0%
North West	85	9.2%
Yorkshire & the Humber	63	6.8%
West Midlands	63	6.8%
East Midlands	40	4.3%
South West	60	6.5%
South East	93	10.0%
East	54	5.8%
London	85	9.2%
Northern Ireland	125	13.5%
Scotland	140	15.1%
Wales	82	8.8%

Republic of Ireland Region	Number	Percentage
Border Region	22	7.7%
West Region	32	11.2%
Midlands Region	26	9.1%
Mid-East Region	18	6.3%
Dublin Region	120	42.1%
South East Region	31	10.9%
South West Region	26	9.1%
Mid West Region	10	3.5%

The majority of the group were in full-time employment:

Employment status	Number	Percentage
Working full-time (for an employer)	660	54.5%
Working full-time (self-employed)	87	7.2%
Working part-time (for an employer)	188	15.5%
Working part-time (self-employed)	50	4.1%
Unemployed	68	5.6%
Working not for income	5	0.4%
Student	77	6.4%
Homemaker	43	3.5%
Retired	34	2.8%

Amongst the 985 people in paid employment, the occupational levels were as follows:

Occupational level	Number	Percentage
Executive level management/owner	37	3.8%
Senior management	53	5.4%
Middle management	200	20.3%
First level management/supervisor	179	18.2%
Employee	461	46.8%
Other	55	5.6%

And a range of work areas were represented:

Work area	Number	Percentage
Agriculture, forestry and fishing	16	1.3%
Mining and quarrying	6	0.5%
Manufacturing	81	6.7%
Electricity, gas, steam and air conditioning supply	11	0.9%
Water supply, sewerage, waste management and remediation activities	2	0.2%
Construction	59	4.9%
Wholesale and retail trade; repair of motor vehicles and motor cycles	66	5.4%
Accommodation and food service activities	54	4.5%
Transport and storage	53	4.4%
Information and communication	75	6.2%
Financial and insurance activities	73	6.0%
Real estate activities	8	0.7%
Professional, scientific and technical activities	90	7.4%
Administrative and support service activities	113	9.3%
Public administration and defence; compulsory social security	53	4.4%
Education	102	8.4%
Human health and social work activities	99	8.2%
Arts, entertainment and recreation	42	3.5%
Other service activities	194	16%
Activities of households as employers; undifferentiated goods and services producing activities of households for own use	11	0.9%
Activities of extraterritorial organisations and bodies	4	0.3%