MBTI® Step II™ instrument

European Data Supplement

January 2018

2nd edition

unlocking potential

OPP Ltd
Elsfield Hall
15–17 Elsfield Way
Oxford OX2 8EP
UK
t: +44 (0)1865 404 500

Contents

Introduction and overview

This European data supplement has been written to provide European MBTI® Step II™ users with a single source of information containing a summary of the research data gathered for European language versions of the MBTI Step II questionnaire. The aim has been to produce a single, easily accessible resource that will better serve multilingual use of the instrument, written in a format that will allow it to be easily updated as additional data become available.

The supplement has been split into discrete chapters, with each language version of the questionnaire having its own dedicated chapter. The aim is that each chapter can be read as a stand-alone document, and hence there is some duplication of text across chapters. The structure of the supplement will allow existing chapters to be updated as more data become available.

What is included in this supplement

The supplement contains updated psychometric information gathered for the European MBTI Step II instrument. As such, it is intended to supplement the data presented in the *MBTI Step II (European Edition) Manual* (Quenk, Hammer and Majors, 2004), and should be read in conjunction with the manual.

Information is provided on Step II facet score distributions, facet intercorrelations, correlations with Step I™ dimensions, out-of-preference score frequencies, reliability and group differences data for the following language versions:

- Danish
- Dutch
- English (European)
- French
- German
- Italian
- Norwegian
- Polish
- Russian
- Spanish
- Swedish.

Overview of findings

What follows is a short summary of several of the key findings. These and other findings are presented in more detail in the relevant language chapters of this supplement. The data described within this supplement show there to be a high degree of consistency across the various European language versions of the Step II questionnaire, and that the psychometric properties of the instrument are highly credible.

Facet scale score distributions

Facet scale descriptive statistics are presented within the supplement for each language version of the questionnaire. In most cases, the means centre quite close to the midpoint of each scale. However, a clear pattern does emerge. The mean scores tend towards the E, S, T and J directions, which is consistent with the most common four-letter type preference amongst the samples of people who have completed the European Step II questionnaire.

Facet scale intercorrelations

Correlations among the Step II facet scales are shown for each language version. Facet scales within each dichotomy consistently correlate more highly (usually substantially more so) with the other scales of the same dichotomy than they do with scales in the other three dichotomies. For example, every E–I facet scale correlates more highly with the other four E–I scales than with any of the S–N, T–F or J–P facet scales. There are, however, a few exceptions to this, for example the T–F scale Questioning–Accommodating, which correlates with many scales across dichotomies. These patterns are similar to those found when the European Step II instrument was originally developed.

Correlations of Step II facet scales with Step I scales

Correlations between Step II facet scales and the continuous scores from the MBTI Step I™ instrument show that in almost all cases the facet scales correlate highly with the corresponding Step I dimension, and considerably lower with the other three dimensions. Again, these correlations are very similar to those found during the original development of the European Step II questionnaire. These correlations, and the consistency with which they have been observed, provide compelling evidence for the theoretical hierarchical structure of the Step II facet scales in relation to the Step I scales.

Out-of-preference scores

It is recognised that the five facet scales relating to each type dimension do not describe the dimension in its entirety; there will not be a precise and exact overlap between, for example, an individual's score on E–I and their total score across the five facet scales that relate to this dimension. For example, it is not uncommon to see an Enthusiastic Introvert or an Intimate Extravert. Such apparent inconsistencies are usually known as 'out-of-preference scores' (OOPS).

Although it is usual to have a number of OOPS in any one profile, it should be unusual to find that OOPS make up more than half of the facet scores relating to any one type dichotomy. Data are presented for all the language versions showing that although this can be found, it is indeed a very infrequent occurrence.

Reliability

The reliability of a test or questionnaire relates to how consistent and precise it is. Internal consistency reliability addresses the question of whether all the questions in a scale measure the same construct. For example, are the Step II facet scales consistent within themselves, and do they hold together well as scales? A common measure of internal consistency reliability, and which is used in this supplement, is coefficient alpha (Cronbach, 1951).

Data presented within this supplement show that, on the whole, the internal consistency reliability is good, especially considering the relatively small number of items in each scale. The three facet scales which do seem to be characterised by lower internal consistency are Practical–Conceptual, Questioning–Accommodating and Critical–Accepting. This pattern is consistent with what was found during the development of the Step II instrument, and is also consistent across language versions.

Group differences

Analyses were conducted to explore links between facet scale scores and various demographic variables. The latter included gender, age, age at which the person left full-time education, occupational level, work area, nationality and employment status. There were many interesting findings but the results were too numerous to summarise in this introduction. The reader is therefore invited to refer to the individual chapters for further details.

MBTI® Step II™ instrument

European Data Supplement

English (European)

January 2018

2nd edition

OPP Ltd
Elsfield Hall
15–17 Elsfield Way
Oxford OX2 8EP
UK
t: +44 (0)1865 404 500

Contents

Introduction

Data collected for the European MBTI® Step II™ instrument were analysed to produce the findings in this supplement. This is the second data supplement produced for this instrument. A brief description of the sample is given below. Further details of the sample are provided in Appendix 1.

- The sample consisted of 136,837 individuals who completed the MBTI Step II instrument in European English via the OPPassessment™ system between 2005 and mid-2016.[1] This sample is considered to be representative of the groups of people with whom the European English MBTI Step II instrument has been and will be used for applications such as management development, coaching, counselling and teambuilding. As such, it is likely to represent a cross-section of the European English-speaking professional and managerial population.

The results of the analyses are outlined below.

[1] OPPassessment allows personality questionnaires such as the MBTI instrument to be administered via email and/or completed online.

Type distribution

Type tables are a way of illustrating the proportion of each type within a particular group. Below is a type table taken from the sample described on the previous page.

For each of the 16 different types, the number of cases, the percentage of the total that this represents and the self-selection ratio (SSR) are shown. The SSR (Myers et al., 1998) is a way of demonstrating whether a given type appears more or less often in a particular group than would be expected given its frequency in a reference group. The UK general population is used as the reference group when calculating the SSRs in this chapter.

An SSR of greater than 1 indicates that a type is over-represented, and an SSR of less than 1 denotes that it is under-represented. Asterisks are used to denote whether the over- or under-representations are statistically significant, based on the results of chi-square analysis.[2]

[2] Chi-square analysis (often abbreviated to χ^2) is a technique used to explore whether observed frequency distributions differ significantly from other, predefined, distributions. In this case, the UK general population group is used as the reference group, and the chi-square analysis indicates whether the proportion of people of each type within a particular sample differs significantly from the proportion of people reporting the same type within the reference group.

OPPassessment data (representative European English-speaking professional and managerial sample)

Table 2.1: Type table for OPPassessment data (reported type, n=136,837)

ISTJ	ISFJ	INFJ	INTJ
n=17,426 12.7% SSR=0.93**	n=3,601 2.6% SSR=0.20**	n=1940 1.4% SSR=1.21**	n=8,670 6.3% SSR=4.50**
ISTP	**ISFP**	**INFP**	**INTP**
n=4,377 3.2% SSR=0.50**	n=978 0.7% SSR=0.11**	n=2,377 1.7% SSR=0.53**	n=7,116 5.2% SSR=2.17**
ESTP	**ESFP**	**ENFP**	**ENTP**
n=6,968 5.1% SSR=0.88**	n=2,226 1.6% SSR=0.18**	n=7,126 5.2% SSR=0.83**	n=14,608 10.7% SSR=3.82**
ESTJ	**ESFJ**	**ENFJ**	**ENTJ**
n=27,689 20.2% SSR=1.94**	n=7,162 5.2% SSR=0.41**	n=5,679 4.2% SSR=1.50**	n=18,894 13.8% SSR=4.76**

Type	n	%
E	90,352	66.0%**
I	46,485	34.0%**
S	70,427	51.5%**
N	66,410	48.5%**
T	105,748	77.3%**
F	31,089	22.7%**
J	91,061	66.5%**
P	45,776	33.5%**

*Difference significant at p<0.05, based on chi-square results.
**Difference significant at p<0.01, based on chi-square results.

The most common single type preference is ESTJ (20% of the total); this is a common finding with managerial groups in other countries. The SSR results suggest that, in comparison with the UK general population, those with preferences for NT are over-represented, and those with preferences for SF are under-represented. Again, this is a common finding with managerial groups.

Properties of the Step II facet scales

Facet scale score distribution

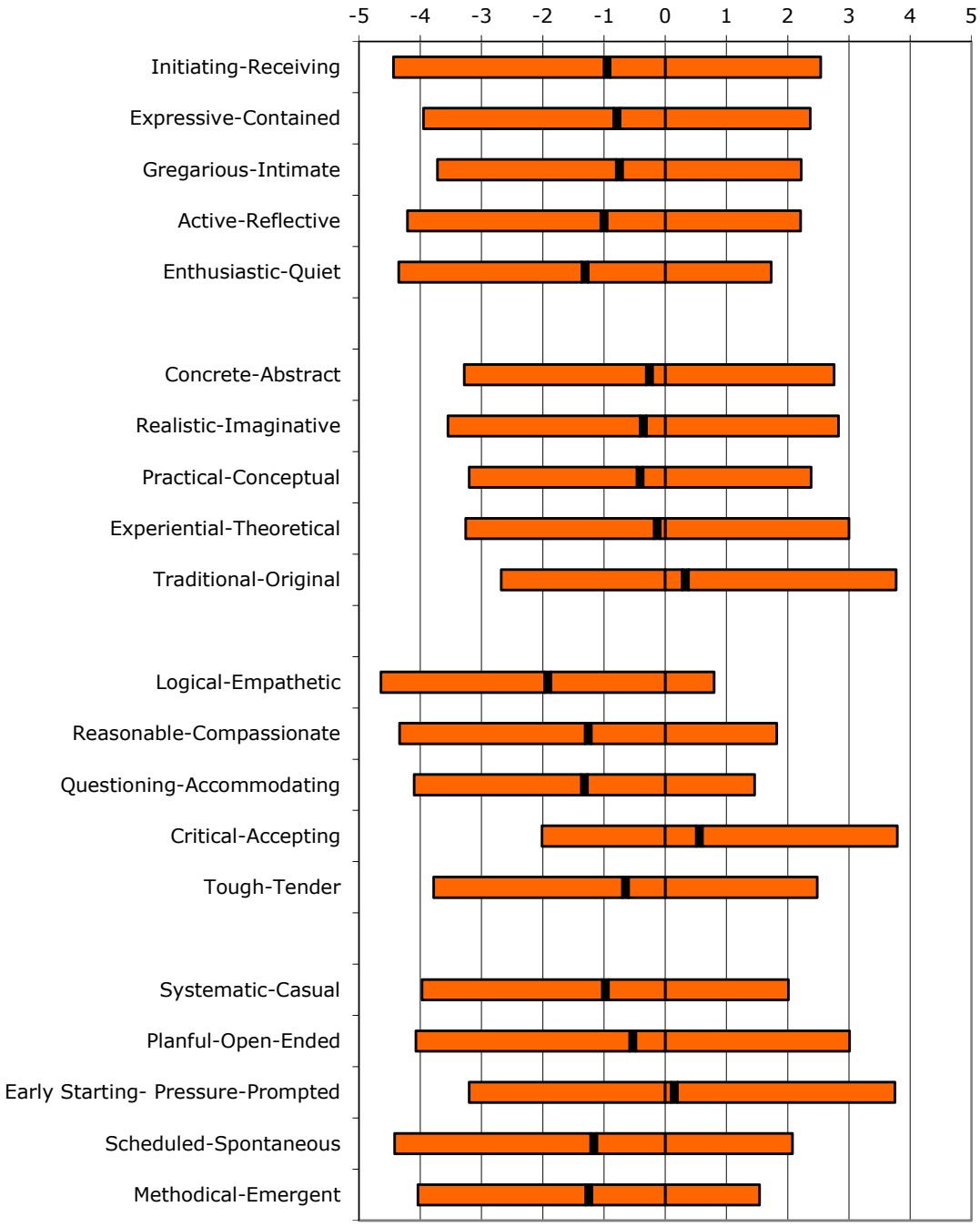
The means and standard deviations of the Step II facets are shown in Table 2.2, and are illustrated graphically in Figure 2.1. The mean score for a scale is calculated by adding together the scores for each individual in the sample and then dividing the sum by the number of individuals. Note that negative values indicate that mean scores are nearer the left-hand pole of the facet and positive values indicate that the mean scores are nearer the right-hand pole. The standard deviation (SD) is a statistical measure describing the degree to which the scores from the sample either bunch up close to, or are scattered widely around, the mean for the sample.

Table 2.2: Means and standard deviations of the facet scales

Step II facet scale	Mean[3]	SD
E–I facet scales		
Initiating–Receiving	–0.97	3.47
Expressive–Contained	–0.77	3.14
Gregarious–Intimate	–0.49	3.01
Active–Reflective	–0.92	3.22
Enthusiastic–Quiet	–1.12	3.08
S–N facet scales		
Concrete–Abstract	–0.56	2.93
Realistic–Imaginative	–0.71	3.14
Practical–Conceptual	–0.53	2.78
Experiential–Theoretical	–0.38	3.06
Traditional–Original	0.03	2.98
T–F facet scales		
Logical–Empathetic	–1.92	2.72
Reasonable–Compassionate	–1.27	3.06
Questioning–Accommodating	–1.15	2.80
Critical–Accepting	0.53	2.57
Tough–Tender	–0.63	3.08
J–P facet scales		
Systematic–Casual	–1.27	2.97
Planful–Open-Ended	–0.84	3.49
Early Starting–Pressure-Prompted	-0.10	3.34
Scheduled–Spontaneous	–1.47	3.15
Methodical–Emergent	–1.41	2.79

[3] Scale means: in Step II reports, scores are given from 5 on one pole, through 0, to 5 on the opposite pole. In compiling statistical information, however, one pole needs to be clearly distinguished from the other. Throughout this data supplement, therefore, a convention has been adopted of scores 5, 4, 3, 2 or 1 on the left-hand pole being assigned values of –5, –4, –3, –2 or –1, and scores on the right-hand pole being assigned positive scores. This does not, of course, imply any suggestion that positive scores are 'better' than negative scores.

Figure 2.1: Facet scale mean scores[4]



[4] For each facet scale, the central line shows the mean and the coloured bars to either side indicate the standard deviation.

In general, the means centre quite close to the midpoint of each scale. None of the scales has a mean greater than 2 points from the midpoint in either direction, and each has a standard deviation of between 2.5 and 3.5. However, a clear pattern does emerge. The mean scores tend towards the E, S, T and J direction (with the exceptions of Traditional–Original and Critical–Accepting), which is consistent with the most common four-letter type preference amongst this sample.

Facet scale intercorrelations

Correlations among the Step II facet scales are shown in Table 2.3. Facet scales within each dichotomy consistently correlate more highly (usually substantially so) with the other scales of the same dichotomy than they do with scales in the other three dichotomies. For example, every E–I facet scale correlates more highly with the other four E–I scales than with any of the S–N, T–F or J–P facet scales.

There is only one scale that correlates **significantly** more highly with scales in other dichotomies. The T–F scale Questioning–Accommodating correlates with many scales, across dichotomies. For example, it correlates at a level of 0.17 with Enthusiastic–Quiet on the E–I dichotomy. It correlates negatively at between –0.19 and –0.38 with all of the S–N scales. It also correlates negatively at –0.16 with both Early Starting–Pressure-Prompted and Scheduled–Spontaneous on the J–P dichotomy. By way of comparison, Questioning–Accommodating correlates at between 0.15 and 0.20 with the other four scales on the same dichotomy as itself (T–F). The negative correlations between Questioning–Accommodating and the S–N scales are consistent with findings in the US version of the instrument, and suggest that a questioning approach to differences of opinion seems to be related to a range of iNtuition (N) facets. These patterns are similar to those found when the European Step II instrument was developed.

Table 2.3: Intercorrelations of Step II facet scales

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Initiating–Receiving	**1.00**																			
2. Expressive–Contained	**0.51**	**1.00**																		
3. Gregarious–Intimate	**0.56**	**0.48**	**1.00**																	
4. Active–Reflective	**0.73**	**0.51**	**0.56**	**1.00**																
5. Enthusiastic–Quiet	**0.61**	**0.58**	**0.63**	**0.64**	**1.00**															
6. Concrete–Abstract	-0.08	-0.08	-0.05	-0.08	-0.13	**1.00**														
7. Realistic–Imaginative	-0.09	-0.11	-0.09	-0.10	-0.18	**0.66**	**1.00**													
8. Practical–Conceptual	-0.06	-0.03	-0.02	-0.05	-0.10	**0.60**	**0.62**	**1.00**												
9. Experiential–Theoretical	-0.03	-0.02	0.01	-0.03	-0.04	**0.64**	**0.54**	**0.53**	**1.00**											
10. Traditional–Original	-0.14	-0.09	-0.08	-0.15	-0.21	**0.62**	**0.57**	**0.54**	**0.45**	**1.00**										
11. Logical–Empathetic	-0.09	-0.20	-0.05	-0.08	-0.12	0.21	0.22	0.03	0.09	0.08	**1.00**									
12. Reasonable–Compassionate	-0.06	-0.14	-0.04	-0.03	-0.05	0.24	0.24	0.07	0.14	0.07	**0.67**	**1.00**								
13. Questioning–Accommodating	0.13	0.05	0.07	0.13	0.17	-0.25	-0.19	-0.28	-0.22	-0.38	**0.16**	**0.15**	**1.00**							
14. Critical–Accepting	-0.16	-0.18	-0.14	-0.12	-0.12	0.15	0.15	0.04	0.06	0.03	**0.41**	**0.49**	**0.16**	**1.00**						
15. Tough–Tender	0.00	-0.12	-0.01	0.02	0.01	0.19	0.18	0.03	0.11	0.00	**0.58**	**0.65**	**0.20**	**0.50**	**1.00**					
16. Systematic–Casual	-0.10	-0.13	-0.12	-0.12	-0.20	0.46	0.43	0.28	0.28	0.48	0.33	0.29	-0.10	0.19	0.25	**1.00**				
17. Planful–Open-Ended	-0.02	-0.01	-0.02	-0.04	-0.06	0.31	0.27	0.20	0.22	0.40	0.11	0.09	-0.14	0.05	0.06	**0.60**	**1.00**			
18. Early Starting–Pressure-Prompted	-0.07	-0.06	-0.08	-0.09	-0.15	0.26	0.24	0.19	0.20	0.36	0.04	0.04	-0.16	-0.02	-0.02	**0.47**	**0.47**	**1.00**		
19. Scheduled–Spontaneous	-0.04	-0.04	-0.05	-0.06	-0.11	0.37	0.32	0.24	0.26	0.46	0.14	0.11	-0.16	0.04	0.06	**0.66**	**0.69**	**0.57**	**1.00**	
20. Methodical–Emergent	0.03	-0.03	-0.03	0.01	-0.05	0.24	0.22	0.14	0.18	0.26	0.12	0.11	-0.06	0.02	0.07	**0.51**	**0.47**	**0.47**	**0.56**	**1.00**

Correlations of Step II facet scales with Step I scales

Correlations between Step II facet scales and the continuous scores from the MBTI Step I™ instrument are shown in Table 2.4.[5] Positive correlations between facet scales and the corresponding Step I dimension (e.g. between Initiating–Receiving and the E–I continuous score) indicate relationships in the expected direction, with higher facet scores tending to be associated with higher continuous scores, and vice versa. Negative correlations would indicate an inverse relationship between the facet scale scores and the corresponding Step I dimension.

Table 2.4: Correlations of Step II facet scales with Step I continuous scores

	Step I continuous score			
Step II facet scales	**E–I**	**S–N**	**T–F**	**J–P**
E–I facet scales				
Initiating–Receiving	**0.84**	–0.10	–0.10	–0.03
Expressive–Contained	**0.72**	–0.09	–0.22	–0.05
Gregarious–Intimate	**0.68**	–0.06	-0.07	–0.06
Active–Reflective	**0.81**	–0.11	–0.08	–0.06
Enthusiastic–Quiet	**0.82**	–0.17	–0.11	–0.11
S–N facet scales				
Concrete–Abstract	–0.10	**0.87**	0.21	0.39
Realistic–Imaginative	–0.12	**0.78**	0.20	0.34
Practical–Conceptual	–0.05	**0.70**	0.02	0.25
Experiential–Theoretical	–0.02	**0.72**	0.08	0.27
Traditional–Original	–0.16	**0.75**	0.03	0.47
T–F facet scales				
Logical–Empathetic	–0.15	0.15	**0.86**	0.17
Reasonable–Compassionate	–0.08	0.17	**0.78**	0.14
Questioning–Accommodating	0.14	–0.32	**0.25**	–0.15
Critical–Accepting	–0.17	0.10	**0.60**	0.05
Tough–Tender	–0.02	0.11	**0.76**	0.10
J–P facet scales				
Systematic–Casual	–0.15	0.46	0.32	**0.76**
Planful–Open-Ended	–0.02	0.33	0.09	**0.83**
Early Starting–Pressure-Prompted	–0.10	0.30	0.01	**0.71**
Scheduled–Spontaneous	–0.06	0.40	0.10	**0.84**
Methodical–Emergent	0.00	0.25	0.09	**0.68**

The E–I facet scales correlate at 0.68–0.84 with the E–I continuous scores from Step I; the S–N facet scales correlate at 0.70–0.87 with the S–N continuous scores; the T–F facet scales correlate at 0.25–0.86

[5] Continuous scores place an individual's score on each dimension onto a continuous scale with a mid-point of 100. To calculate continuous scores, Preference Clarity Index (PCI) scores for each dimension are either subtracted or added to 100, depending on which direction the overall preference is. PCI scores in the direction of E, S, T or J are subtracted from 100. PCI scores in the direction of I, N, F or P are added to 100.

with the T–F continuous scores; and the J–P facet scales correlate at 0.68–0.84 with the J–P continuous scores. The scale that showed the lowest correlation with its associated Step I continuous score was Questioning–Accommodating at 0.25. This scale is also lowest on internal consistency (see page 21), and has been previously found to have the lowest test-retest reliability (Quenk, Hammer and Majors, 2004), which would result in the true correlations being underestimated.

These correlations are very similar to those found during the development of the Step II questionnaire. This consistency, alongside the fact that the correlations between Step II facet scales and Step I continuous scores associated with their own dimension are substantially higher than correlations with the other three dimensions, provides compelling evidence for the theoretical hierarchical structure of the Step II facet scales in relation to the Step I scales.

Out-of-preference scores

It is known that the five facet scales relating to each type dimension do not describe the dimension in its entirety; there will not be a precise and exact overlap between, for example, an individual's score on E–I and their total score across the five facet scales that relate to this dimension. For example, it is not uncommon to see an Enthusiastic Introvert or an Intimate Extravert. Such apparent inconsistencies are usually known as 'out-of-preference scores' (OOPS).

Although it is usual to have a number of OOPS in any one profile, it should be unusual to find that OOPS make up more than half of the facet scores relating to any one type dichotomy. Table 2.5 demonstrates that the proportion of people for whom this happens ranges from 3.3% for the T–F block to 1.3% for the J–P block. It is therefore a very infrequent occurrence.

Table 2.5: Proportion of OOPS by type dichotomy[6]

	Proportion of 'reported type' OOPS					
	None	One	Two	Three	Four	Five
E–I	68.0%	23.0%	7.4%	1.5%	0.2%	0.0%
S–N	66.2%	23.8%	7.8%	2.0%	0.3%	0.0%
T–F	53.0%	33.0%	10.7%	2.9%	0.4%	0.0%
J–P	62.0%	28.4%	8.4%	1.2%	0.1%	0.0%

[6] In this table, an OOPS is defined as a score of 2, 3, 4 or 5 on the 'wrong' pole of the facet when compared with the reported type.

For any individual facet scale, an OOPS tends to occur in approximately 10% of profiles, with the exception of Questioning–Accommodating (20%) and Critical–Accepting (22%).

Reliability

The reliability of a test or questionnaire relates to how consistent and precise it is. Internal consistency reliability addresses the question of whether all the questions in a scale measure the same construct. For example, are the Step II facet scales consistent within themselves, and do they hold together well as scales? A common measure of internal consistency reliability is coefficient alpha (Cronbach, 1951). The alpha coefficients for the Step II facet scales are shown in Table 2.6.

Table 2.6: Internal consistency reliability

Step II facet scale	No. of items	Coefficient alpha
E–I facet scales		
Initiating–Receiving	8	0.86
Expressive–Contained	7	0.77
Gregarious–Intimate	7	0.68
Active–Reflective	8	0.73
Enthusiastic–Quiet	9	0.74
S–N facet scales		
Concrete–Abstract	9	0.75
Realistic–Imaginative	7	0.74
Practical–Conceptual	8	0.52
Experiential–Theoretical	8	0.78
Traditional–Original	8	0.74
T–F facet scales		
Logical–Empathetic	9	0.81
Reasonable–Compassionate	8	0.73
Questioning–Accommodating	7	0.46
Critical–Accepting	8	0.53
Tough–Tender	8	0.78
J–P facet scales		
Systematic–Casual	8	0.76
Planful–Open-Ended	6	0.81
Early Starting–Pressure-Prompted	6	0.73
Scheduled–Spontaneous	8	0.77
Methodical–Emergent	8	0.67
	Median	0.74

The internal consistency reliability of most scales is good, and the average (median) reliability is above 0.7. However, three scales (Practical–Conceptual, Questioning–Accommodating and Critical–Accepting) have lower alpha coefficients. This pattern is consistent with what was found during the development of the Step II

instrument, where these three facet scales were also found to show lower reliability than the others (Quenk, Hammer and Majors, 2004).

It should be noted, however, that coefficient alpha reliability statistics will give an underestimate of the reliability of the Step II facet scales, given the use of item response theory (IRT) methods in the actual scoring process.[7]

Group differences

This section looks at the way in which people from different groups responded to the European English version of the European Step II questionnaire.

Gender

The means and standard deviations of the Step II facet scales are shown separately for males and females in Table 2.7, along with the difference in mean scores. This is represented graphically in Figure 2.2. Statistically significant differences were found between male and female mean scores for 18 of the 20 facet scales, with some consistent patterns emerging. However in most cases, these differences were negligible.

- On the T–F facet scales, there were statistically significant gender differences for four of the five facet scales, with mean scores tending more towards the T pole for males than for females.

[7] Item response theory (IRT) is an approach to measurement that is concerned with modelling the relationship between item responses and the underlying characteristic assessed by the scale or test the item is designed to measure. IRT can be used to select items for a test and/or to score the items.

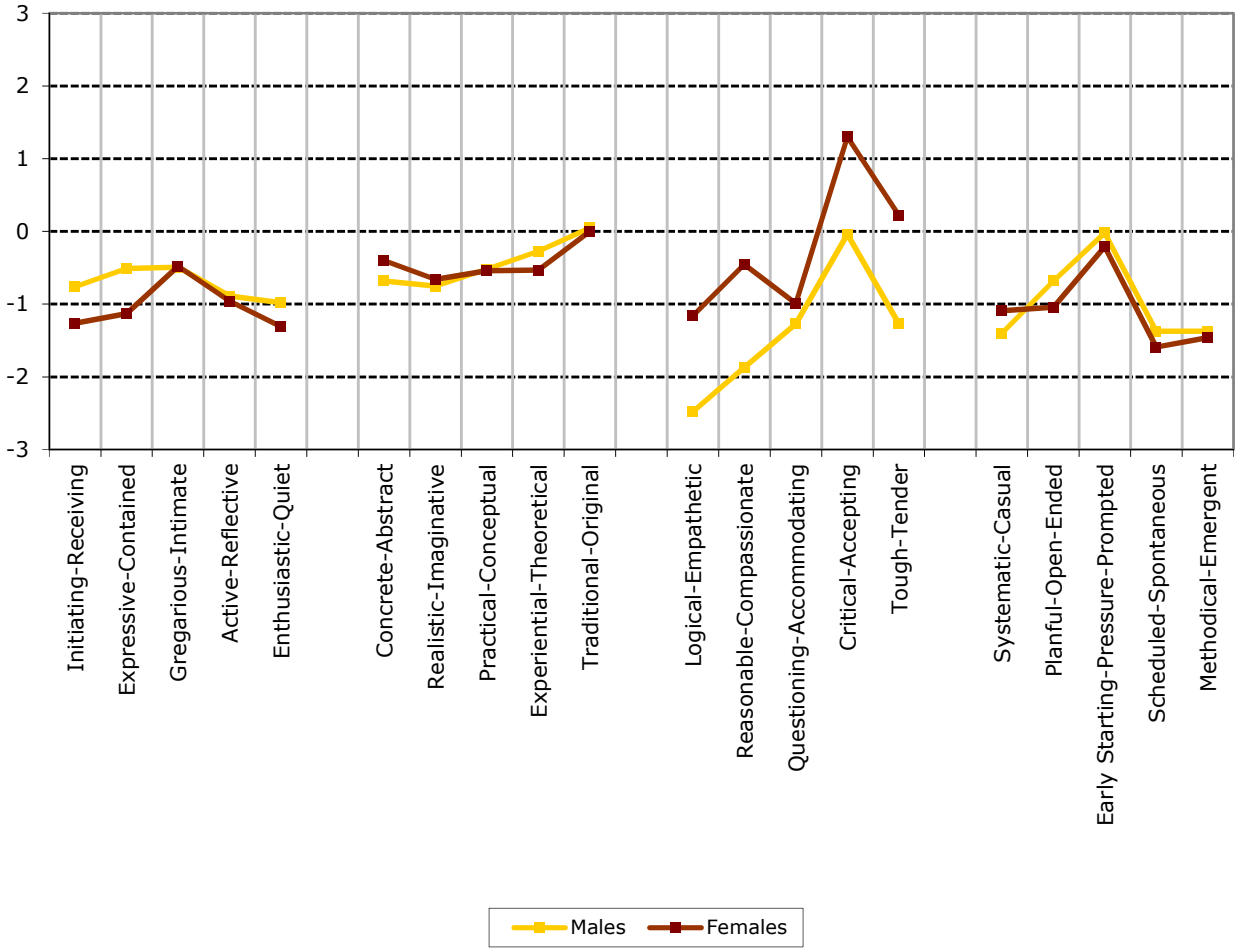
Table 2.7: Gender differences in facet scale scores

	Males (n=78,465)		Females (n=58,372)		Difference (M–F)[8]
Step II facet scale	**Mean**	**SD**	**Mean**	**SD**	
E–I facet scales					
Initiating–Receiving	-0.76	3.48	-1.26	3.44	0.50**
Expressive–Contained	-0.51	3.13	-1.13	3.12	0.63**
Gregarious–Intimate	-0.49	3.03	-0.48	2.98	0.00
Active–Reflective	-0.89	3.21	-0.96	3.23	0.07**
Enthusiastic–Quiet	-0.98	3.09	-1.31	3.05	0.33**
S–N facet scales					
Concrete–Abstract	-0.68	2.88	-0.40	2.98	-0.29**
Realistic–Imaginative	-0.75	3.15	-0.66	3.13	-0.08**
Practical–Conceptual	-0.52	2.72	-0.54	2.86	0.02
Experiential–Theoretical	-0.27	3.09	-0.53	3.02	0.26**
Traditional–Original	0.05	2.98	0.00	2.99	0.05**
T–F facet scales					
Logical–Empathetic	-2.48	2.38	-1.16	2.95	-1.32**
Reasonable–Compassionate	-1.87	2.91	-0.45	3.07	-1.42**
Questioning–Accommodating	-1.27	2.74	-0.99	2.87	-0.27**
Critical–Accepting	-0.04	2.45	1.30	2.53	-1.33**
Tough–Tender	-1.27	2.94	0.22	3.05	-1.49**
J–P facet scales					
Systematic–Casual	-1.40	2.96	-1.09	2.97	-0.32**
Planful–Open-Ended	-0.68	3.49	-1.04	3.48	0.36**
Early Starting–Pressure-Prompted	-0.02	3.32	-0.21	3.37	0.19**
Scheduled–Spontaneous	-1.37	3.14	-1.59	3.16	0.21**
Methodical–Emergent	-1.37	2.79	-1.46	2.78	0.09**

Difference significant at: *p<0.05, **p<0.01 (based on an independent samples t-test).

[8] A positive value indicates that male scores tend more towards I, N, F or P, and a negative value indicates that female scores tend more towards I, N, F or P.

Figure 2.2: Gender differences in facet scale scores



Age

There were found to be no meaningful links between age and facet scale scores. Although correlational analysis showed several facet scales to be significantly correlated with age, the significance levels were more the result of the very large sample size rather than being indicative of a meaningful relationship. The highest correlation was 0.12 (Critical–Accepting), and even this is too small to be considered to be meaningful.

For interpretation purposes, it is reasonable to conclude that there are no clear relationships between age and facet scale scores.

Ethnic origin

The UK sample contained people with a broad range of ethnic origins. While most groups were large enough to allow a full exploration of ethnic differences in facet scale scores, there were a couple of groups that were not large enough. To overcome this, Bangladeshi, Indian and Pakistani people were re-classified as 'South Asian', while Black-African, Black-Caribbean and Black-Other were re-classified as the group 'Black- African/Caribbean/Other' for the purposes of analysis. A

comparison of mean facet scales across all eight groups highlighted some interesting differences, and this is shown in Table 2.8 and Figure 2.3. The findings can be summarised as follows:

- The most noticeable pattern regarding the facets on the E–I dimension was that mean scores for the 'Black-African/Caribbean/Other' group tended further towards the I pole (or less towards the E pole) than the other groups on three of the facet scales.

- There were no noticeable patterns regarding the facet scales of the S–N or T–F dimensions.

- There were three noticeable groupings regarding the facet scales of the J–P dimension. Firstly, mean facet scores for the 'White British', 'White Irish' 'White-Other European' and 'White-Other' groups were similar and, as a grouping, tended less towards the J pole than the remaining groups. Secondly, mean facet scores for the 'South Asian', 'Asian-Other' and 'Black- African/Caribbean/Other' groups were similar and, as a grouping, tended more towards the J pole than the previous grouping. Thirdly, mean facet scores for the 'Chinese' group tended noticeably further towards the J pole than all of the other groups on four of the five facet scales.

Table 2.8: Mean facet scale scores by ethnic origin

Step II facet scale	White British (n=61,766)		White Irish (n=5005)		White -Other European (n=20,187)		White-Other (n=6,757)		South Asian (n=4,244)		Asian -Other (n=1,973)		Chinese (n=1,172)		Black - African/Caribbean/Other (n=3,295)		Sig.
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD					
E–I facet scales																	
Initiating–Receiving	-0.96	3.57	-1.47	3.40	-1.06	3.31	-1.05	3.47	-0.80	3.34	-0.23	3.24	-0.71	3.26	-0.77	3.30	**
Expressive–Contained	-0.79	3.23	-0.82	3.18	-1.12	2.99	-1.02	3.12	-0.46	2.99	-0.93	2.87	-0.65	2.89	0.15	3.01	**
Gregarious–Intimate	-0.73	3.03	-0.74	2.96	-0.11	2.93	-0.42	3.07	-0.38	2.97	-0.35	3.04	-0.31	3.04	0.52	2.96	**
Active–Reflective	-0.91	3.33	-1.26	3.17	-1.15	3.08	-0.98	3.19	-0.91	3.07	-0.22	2.95	-0.23	3.08	-0.65	3.04	**
Enthusiastic–Quiet	-1.13	3.14	-1.16	3.03	-1.38	2.99	-1.16	3.06	-1.21	2.99	-0.87	2.86	-0.84	2.92	-0.06	3.05	**
S–N facet scales																	
Concrete–Abstract	-0.43	3.08	-0.58	3.01	-0.64	2.66	-0.69	2.95	-0.96	2.63	-1.04	2.40	-0.95	2.60	-1.13	2.71	**
Realistic–Imaginative	-0.47	3.20	-0.63	3.19	-0.94	2.97	-0.75	3.11	-1.47	3.02	-1.43	2.92	-1.27	3.07	-1.22	3.02	**
Practical–Conceptual	-0.62	2.85	-0.36	2.89	-0.42	2.59	-0.42	2.77	-0.45	2.72	-0.42	2.65	-0.42	2.68	-0.35	2.69	**
Experiential–Theoretical	-0.12	3.14	-0.37	3.13	-0.64	2.94	-0.44	3.12	-0.93	2.77	-0.98	2.87	-0.83	2.87	-0.56	2.85	**
Traditional–Original	-0.02	3.03	-0.11	3.01	0.39	2.96	0.02	2.98	-0.21	2.82	-0.86	2.65	-0.44	2.82	-0.32	2.78	**
T–F facet scales																	
Logical–Empathetic	-1.77	2.78	-1.91	2.72	-2.09	2.63	-2.14	2.75	-2.05	2.61	-2.25	2.52	-2.01	2.68	-2.14	2.58	**
Reasonable–Compassionate	-0.82	3.14	-0.85	3.11	-2.16	2.71	-1.63	3.03	-1.58	3.01	-2.07	2.76	-1.49	2.98	-1.33	3.03	**
Questioning–Accommodating	-1.21	2.85	-1.18	2.78	-1.14	2.70	-1.13	2.77	-0.93	2.72	-0.19	2.70	-0.45	2.76	-1.27	2.71	**
Critical–Accepting	0.81	2.65	1.05	2.56	-0.12	2.24	0.36	2.58	0.50	2.45	-0.19	2.46	0.23	2.47	0.86	2.44	**
Tough–Tender	-0.36	3.20	-0.23	3.10	-1.17	2.82	-0.95	2.98	-0.87	2.85	-0.88	2.73	-0.64	2.91	-1.03	2.81	**
J–P facet scales																	
Systematic–Casual	-1.22	3.00	-1.21	2.96	-1.19	2.93	-1.13	2.99	-1.79	2.91	-1.89	2.73	-1.60	2.89	-1.65	2.99	**
Planful–Open-Ended	-0.85	3.56	-0.59	3.55	-0.56	3.41	-0.76	3.54	-1.47	3.30	-1.99	3.02	-1.60	3.18	-1.29	3.40	**
Early Starting–Pressure-Prompted	0.08	3.36	0.02	3.36	0.13	3.33	-0.19	3.33	-0.75	3.25	-1.35	3.02	-0.81	3.19	-0.92	3.30	**
Scheduled–Spontaneous	-1.41	3.27	-1.57	3.15	-1.34	3.03	-1.50	3.11	-1.91	2.90	-2.24	2.72	-1.78	2.80	-1.81	2.98	**
Methodical–Emergent	-1.29	2.85	-1.31	2.81	-1.33	2.74	-1.35	2.91	-2.17	2.56	-1.94	2.42	-1.67	2.59	-2.06	2.69	**

Significant at: *p<0.05, **p<0.01 (based on an independent samples t-test).

Occupational level

Research using the European English version of the MBTI Step I questionnaire has demonstrated that individuals in higher-level jobs in organisations are more likely to have preferences for iNtuition and (to some extent) Thinking than those in lower-level jobs.[9]

Table 2.9 shows the facet scale means and standard deviation for different occupational levels. The findings can be summarised as follows:

- While there was no clear overall pattern regarding the facets relating to the E-I dimension, mean facet scores tended to be further towards the E pole with each rise in occupational level.
- A clear pattern was found regarding the facets relating to the S–N dimension and the higher three occupational levels. Mean facet scores for the three higher occupational levels tended to be further towards the N pole (or less towards the S pole) than those for the lower three occupational level groups. Furthermore, mean scores for these three higher occupational levels tended further towards the N pole the higher the occupational level, so that mean scores for the 'Top level' group tended furthest towards the N pole.
- The most noticeable pattern regarding the facets relating to the T–F dimension was that mean facet scores for the 'Employee' group tended to be further towards the F pole (or less towards the T pole) than those for the higher occupational level groups on all five facet scales.
- A reasonably consistent pattern was found regarding the facets relating to the J–P dimension. Mean facet scores amongst higher occupational levels tended to be further towards the P pole (or less towards the J pole) than those for the lower occupational level groups, with facet scores tending less towards the J pole with each rise in occupational level. The only exception was the 'Employee' group, which tended less towards the J pole than some of the higher occupational level groups on three of the five facet scales.

[9] See the MBTI Step I European Data Supplement for details.

Table 2.9: Mean facet scale scores by occupational level

Step II facet scale	Top level (n=4,759)		Senior Executive (n=18,438)		Upper Middle Management (n=25,938)		Middle Management (n=25,282)		First Level Mgt/Supervisory (n=12,778)		Employee (n=17,178)		Sig.
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales													
Initiating–Receiving	-1.32	3.39	-1.27	3.42	-1.02	3.47	-0.89	3.48	-0.81	3.49	-0.89	3.50	**
Expressive–Contained	-1.01	3.13	-1.03	3.16	-0.88	3.17	-0.77	3.15	-0.71	3.13	-0.75	3.11	**
Gregarious–Intimate	-0.64	2.98	-0.69	3.01	-0.60	3.03	-0.50	3.02	-0.39	3.01	-0.38	3.00	**
Active–Reflective	-1.23	3.11	-1.25	3.14	-1.04	3.22	-0.84	3.27	-0.74	3.28	-0.78	3.27	**
Enthusiastic–Quiet	-1.53	2.98	-1.45	3.04	-1.24	3.08	-1.09	3.10	-0.93	3.10	-0.94	3.08	**
S–N facet scales													
Concrete–Abstract	0.13	3.00	-0.12	2.99	-0.54	2.94	-0.82	2.90	-1.01	2.82	-0.79	2.87	**
Realistic–Imaginative	0.18	3.19	-0.16	3.16	-0.67	3.12	-1.00	3.11	-1.22	3.04	-0.99	3.07	**
Practical–Conceptual	0.20	2.63	-0.10	2.64	-0.47	2.75	-0.74	2.79	-0.95	2.80	-0.90	2.81	**
Experiential–Theoretical	0.22	3.08	0.02	3.07	-0.29	3.08	-0.58	3.06	-0.73	3.02	-0.57	3.02	**
Traditional–Original	0.86	2.96	0.60	2.96	0.16	2.97	-0.20	2.96	-0.53	2.91	-0.46	2.94	**
T–F facet scales													
Logical–Empathetic	-2.15	2.56	-2.30	2.49	-2.25	2.53	-2.03	2.66	-1.67	2.83	-1.02	3.02	**
Reasonable–Compassionate	-1.26	3.07	-1.45	3.03	-1.54	2.99	-1.46	3.02	-1.15	3.09	-0.58	3.17	**
Questioning–Accommodating	-1.46	2.62	-1.42	2.63	-1.30	2.71	-1.09	2.78	-0.91	2.89	-0.67	3.01	**
Critical–Accepting	0.48	2.51	0.54	2.52	0.47	2.55	0.48	2.58	0.53	2.60	0.69	2.63	**
Tough–Tender	-1.01	3.01	-1.06	2.97	-0.99	3.00	-0.71	3.07	-0.34	3.11	0.15	3.17	**
J–P facet scales													
Systematic–Casual	-1.25	2.93	-1.36	2.90	-1.34	2.94	-1.36	2.99	-1.35	2.98	-0.99	3.04	**
Planful–Open-Ended	-0.88	3.51	-0.84	3.51	-0.76	3.52	-0.94	3.48	-0.99	3.44	-0.89	3.50	**
Early Starting–Pressure-Prompted	0.51	3.34	0.39	3.31	0.25	3.32	-0.12	3.33	-0.44	3.31	-0.73	3.29	**
Scheduled–Spontaneous	-1.22	3.29	-1.33	3.23	-1.35	3.20	-1.55	3.15	-1.71	3.03	-1.64	3.02	**
Methodical–Emergent	-1.26	2.84	-1.35	2.77	-1.33	2.80	-1.42	2.79	-1.50	2.79	-1.51	2.81	**

Significant at: *p<0.05, **p<0.01 (based on a one-way analysis of variance).

Education

Specific educational qualifications were not collected for the OPPassessment sample; however, the age at which individuals left full-time education was collected. No significant and meaningful correlations were found between the age at which people left full-time education and their facet scale scores.

Work area

Information regarding the area of work people engage in was collected for the group. Many different categories were used, but for the purposes of analysis the focus was on the five most commonly occurring. These were as follows:

- Finance
- HR, training, guidance
- Sales, customer service
- Health, social services
- Business services

Table 2.10 shows the facet scale means and standard deviation for these five work areas.

The findings can be summarised as follows:

- An interesting pattern was found regarding the facets relating to the E–I dimension. Although mean scores for all groups tended towards the E pole for each facet scale, scores were considerably further towards the E pole for the 'Sales, customer service' group than for the other groups on four of the five facet scales. Mean scores for the 'Business services' group fell somewhere in between.
- On facets relating to the S–N dimension, mean scores generally tended slightly towards the S pole for all groups, with the 'Finance' group tending further towards the S pole than all other groups on four of the five facet scales.
- The most noticeable pattern regarding the facets in the T–F dimension was that although mean scores for all groups tended towards the T pole on three out of the five facet scales, scores for all groups tended towards the F pole for the Critical–Accepting facet scale. This is a common finding with this scale. Mean scores for the 'HR, training, guidance' and 'Health, social services' groups tended more towards the F pole (or less towards the T pole) than the other groups on four of the five facet scales.

On facets relating to the J–P dimension, mean scores generally tended towards the J pole for all groups, with the 'HR, training, guidance' group tending less towards the J pole than all other groups on all five facet scales.

Table 2.10: Mean facet scale scores by work area

Step II facet scale	Finance (n=15,840)		HR, training, guidance (n=12,975)		Sales, customer service (n=9.956)		Health, social services (9,494)		Business services (n=9,263)		Sig.
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales											
Initiating–Receiving	-0.70	3.51	-1.23	3.45	-1.92	3.14	-1.00	3.52	-1.10	3.43	**
Expressive–Contained	-0.62	3.14	-1.29	3.17	-1.35	2.98	-0.77	3.21	-0.91	3.11	**
Gregarious–Intimate	-0.50	2.97	-0.57	3.10	-1.11	2.91	-0.30	3.01	-0.64	3.01	**
Active–Reflective	-0.82	3.26	-1.10	3.25	-1.82	3.02	-0.77	3.24	-1.09	3.19	**
Enthusiastic–Quiet	-1.08	3.07	-1.45	3.12	-2.00	2.79	-0.71	3.18	-1.35	3.01	**
S–N facet scales											
Concrete–Abstract	-1.08	2.78	0.02	3.15	-0.91	2.68	-0.21	3.08	-0.43	2.89	**
Realistic–Imaginative	-1.26	2.95	-0.30	3.32	-0.88	3.06	-0.62	3.16	-0.48	3.12	**
Practical–Conceptual	-0.86	2.70	-0.48	2.85	-0.73	2.57	-0.32	2.94	-0.37	2.69	**
Experiential–Theoretical	-0.82	3.09	-0.09	3.17	-0.79	2.88	0.07	3.10	-0.26	3.04	**
Traditional–Original	-0.33	2.94	0.31	3.13	-0.03	2.85	0.03	2.99	0.23	2.97	**
T–F facet scales											
Logical–Empathetic	-2.38	2.45	-1.05	3.12	-1.92	2.64	-1.58	2.86	-2.05	2.61	**
Reasonable–Compassionate	-1.88	2.90	-0.38	3.18	-1.66	2.91	-0.29	3.10	-1.48	2.99	**
Questioning–Accommodating	-1.16	2.81	-0.88	2.87	-1.15	2.69	-1.26	2.82	-1.17	2.77	**
Critical–Accepting	0.13	2.53	1.31	2.61	0.27	2.45	1.26	2.52	0.40	2.53	**
Tough–Tender	-1.10	2.98	0.16	3.14	-1.25	2.85	0.23	3.14	-0.83	2.98	**
J–P facet scales											
Systematic–Casual	-1.48	2.90	-0.58	3.18	-1.17	2.95	-1.46	2.94	-1.13	2.95	**
Planful–Open-Ended	-0.89	3.48	-0.70	3.66	-0.88	3.43	-0.87	3.53	-0.78	3.49	**
Early Starting–Pressure-Prompted	-0.13	3.36	0.18	3.49	-0.13	3.27	0.02	3.35	0.10	3.36	**
Scheduled–Spontaneous	-1.58	3.09	-1.03	3.45	-1.48	3.04	-1.49	3.20	-1.40	3.16	**
Methodical–Emergent	-1.45	2.76	-1.05	3.00	-1.44	2.80	-1.27	2.88	-1.43	2.75	**

Significant at: *p<0.05, **p<0.01 (based on a one-way analysis of variance).

Nationality

Information on nationality was available for the group. Seventy per cent of the group were British. Although a number of other European nationalities were represented in fairly large numbers, we would normally expect these individuals to complete the questionnaire in their own first language. Therefore, we have not explored national differences using the European English version of the instrument. The exception to this is individuals who described themselves as Irish. These formed approximately 4% of the group.

A comparison of mean facet scales across the British and Irish groups highlighted some interesting differences, as shown in Table 2.11 and Figure 2.4. It is worth noting that although the differences below were statistically significant, none of the differences in mean scores was more than 0.5 points. In real terms this is actually quite small.

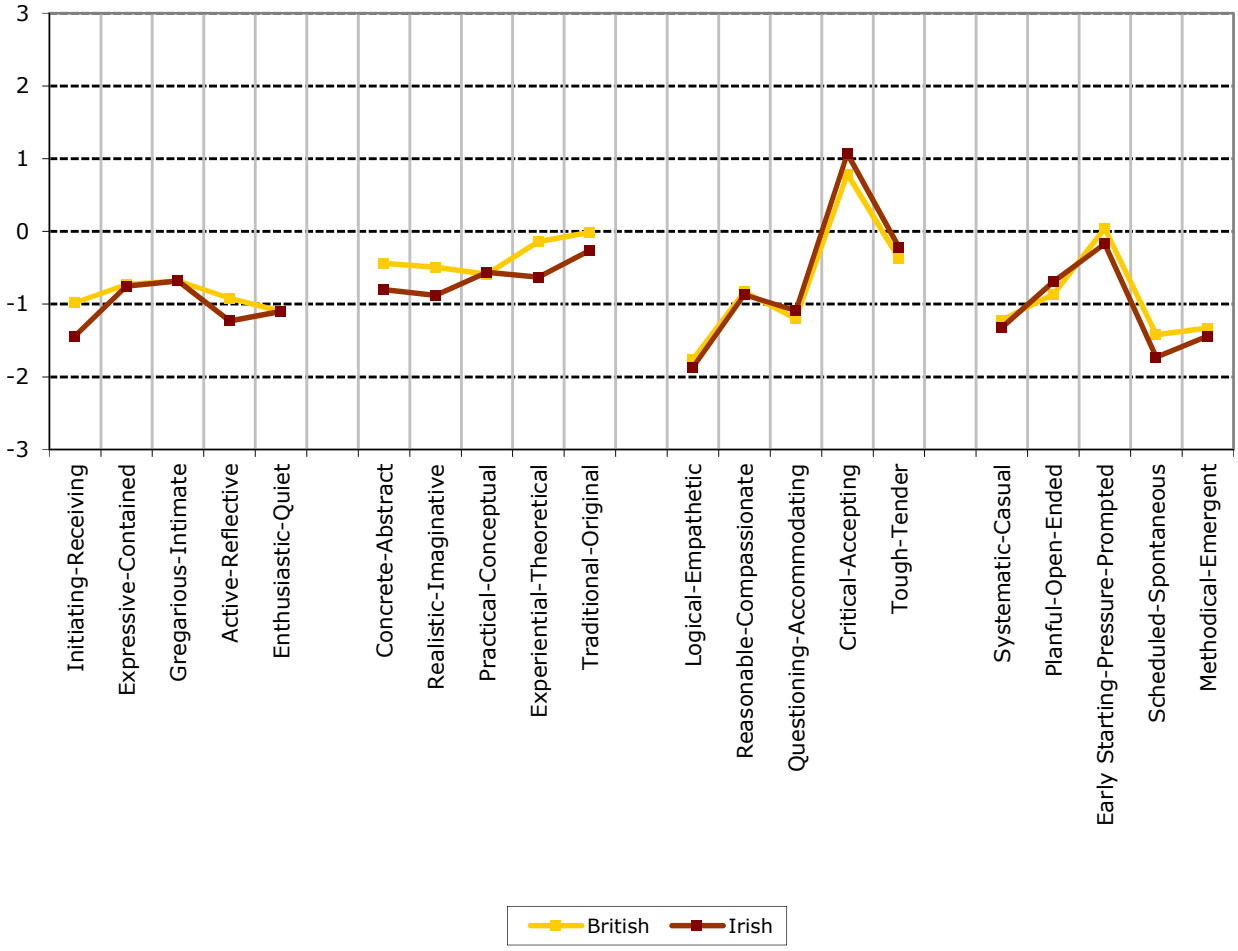
- On two of the five facet scales linked to the E–I dimension (Initiating–Receiving, Active–Reflective), the Irish group mean scores were significantly further towards the E pole than those for the British group.
- On four of the five facet scales linked to the S–N dimension, the Irish group mean scores were significantly further towards the S pole than those for the British group.
- On one of the five facet scales linked to the T–F dimension (Critical–Accepting), the British group mean scores were significantly further towards the T pole than those for the Irish group.
- On one of the five facet scales linked to the J–P dimension (Scheduled–Spontaneous), the Irish group mean scores were significantly further towards the J pole than those for the British group.

Table 2.11: Mean facet scale scores by nationality

Step II facet scale	British (n=72,647)		Irish (n=4,655)		Sig.
	Mean	SD	Mean	SD	
E–I facet scales					
Initiating–Receiving	-0.98	3.55	-1.44	3.37	**
Expressive–Contained	-0.73	3.22	-0.75	3.16	
Gregarious–Intimate	-0.68	3.03	-0.68	2.94	
Active–Reflective	-0.92	3.31	-1.23	3.16	**
Enthusiastic–Quiet	-1.10	3.13	-1.10	3.01	
S–N facet scales					
Concrete–Abstract	-0.44	3.07	-0.80	2.95	**
Realistic–Imaginative	-0.49	3.20	-0.88	3.13	**
Practical–Conceptual	-0.59	2.85	-0.56	2.89	
Experiential–Theoretical	-0.14	3.12	-0.63	3.08	**
Traditional–Original	-0.01	3.02	-0.26	2.99	**
T–F facet scales					
Logical–Empathetic	-1.76	2.77	-1.87	2.73	**
Reasonable–Compassionate	-0.83	3.13	-0.87	3.10	
Questioning–Accommodating	-1.20	2.85	-1.09	2.78	*
Critical–Accepting	0.79	2.64	1.07	2.57	**
Tough–Tender	-0.37	3.18	-0.22	3.11	**
J–P facet scales					
Systematic–Casual	-1.22	2.99	-1.32	2.95	*
Planful–Open-Ended	-0.86	3.55	-0.69	3.53	**
Early Starting–Pressure-Prompted	0.04	3.35	-0.17	3.33	**
Scheduled–Spontaneous	-1.42	3.25	-1.73	3.08	**
Methodical–Emergent	-1.33	2.83	-1.44	2.76	*

Significant at: *p<0.05, **p<0.01 (based on an independent samples t-test).

Figure 2.4: Mean facet scale scores by nationality



Employment status

Employment status information was available for the OPPassessment sample. The vast majority of the group worked either full-time or part-time, or were self-employed. Table 2.12 shows the mean facet scale scores for each group, with the data illustrated in graphical form in Figure 2.5.

The analyses showed statistically significant differences across the groups on all the facet scales on the S–N, T–F and J–P dimensions, and three of the five facets on the E–I dimension. The clearest patterns were as follows.

- The part-time group tended to score more towards the E pole on the E–I facet scales than the other two groups on two of the facet scales.
- The self-employed group tended to score more towards the N pole on all five of the S–N facet scales than the other two groups, and less towards the J pole on all five of the J–P facet scales.

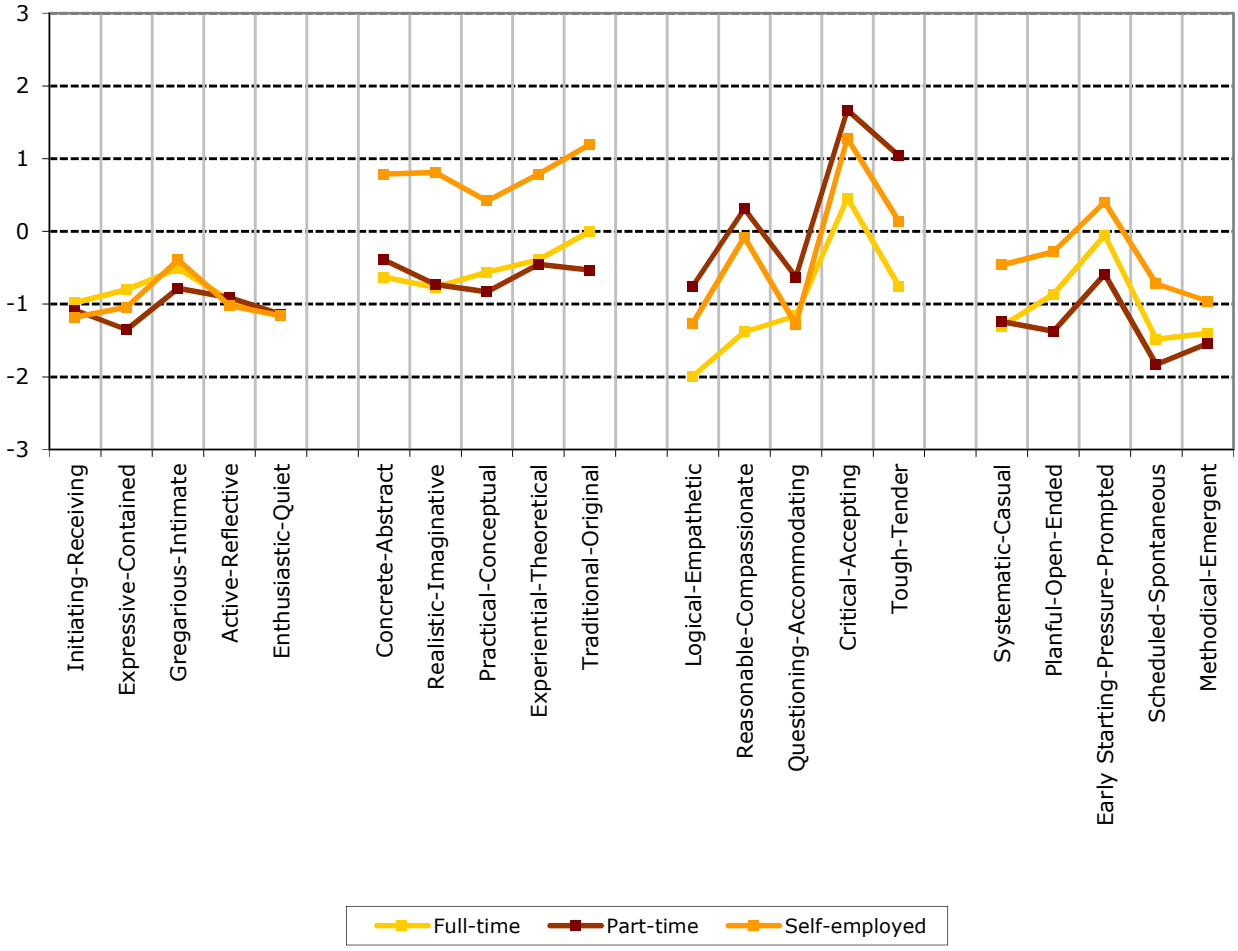
- The full-time group tended to score more towards the T pole on four of the five T–F facet scales than the other two groups, particularly the part-time group. This is likely to be at least partly a gender effect; 88% of part-time workers were female, compared with 42% of the total group and 39% of full-time workers.

Table 2.12: Mean facet scale scores by employment status

Step II facet scale	Full-time (n=101,960)		Part-time (n=4,826)		Self-employed (n=3,085)		Sig.
	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales							
Initiating–Receiving	-0.98	3.47	-1.08	3.49	-1.18	3.49	**
Expressive–Contained	-0.80	3.14	-1.35	3.12	-1.04	3.23	**
Gregarious–Intimate	-0.51	3.02	-0.78	2.97	-0.39	3.11	**
Active–Reflective	-0.94	3.23	-0.91	3.26	-1.02	3.09	
Enthusiastic–Quiet	-1.15	3.08	-1.14	3.11	-1.16	3.10	
S–N facet scales							
Concrete–Abstract	-0.63	2.92	-0.39	3.06	0.79	3.06	**
Realistic–Imaginative	-0.77	3.13	-0.73	3.16	0.81	3.19	**
Practical–Conceptual	-0.56	2.76	-0.83	2.96	0.42	2.74	**
Experiential–Theoretical	-0.39	3.07	-0.45	3.09	0.79	3.00	**
Traditional–Original	0.00	2.98	-0.53	3.02	1.20	3.00	**
T–F facet scales							
Logical–Empathetic	-1.99	2.69	-0.75	3.02	-1.26	2.97	**
Reasonable–Compassionate	-1.38	3.04	0.31	3.05	-0.08	3.21	**
Questioning–Accommodating	-1.16	2.78	-0.63	3.02	-1.28	2.77	**
Critical–Accepting	0.46	2.56	1.67	2.50	1.28	2.61	**
Tough–Tender	-0.75	3.06	1.04	3.07	0.14	3.14	**
J–P facet scales							
Systematic–Casual	-1.30	2.96	-1.24	3.00	-0.46	3.07	**
Planful–Open-Ended	-0.86	3.49	-1.37	3.45	-0.28	3.61	**
Early Starting–Pressure-Prompted	-0.05	3.34	-0.59	3.38	0.40	3.44	**
Scheduled–Spontaneous	-1.48	3.14	-1.83	3.15	-0.72	3.50	**
Methodical–Emergent	-1.40	2.79	-1.54	2.80	-0.96	2.96	**

Significant at: *p<0.05, **p<0.01 (based on a one-way analysis of variance).

Figure 2.5: Mean facet scale scores by employment status



Appendix 1: Sample description

Sample 1: Data from OPPassessment (representative European English-speaking professional and managerial sample)

This sample consists of 136,837 individuals who completed the MBTI Step II instrument in European English via the OPPassessment system between November 2005 and July 2016. Fifty-seven per cent of the respondents were male and 43% were female. Age ranged from 16 to 90 years, with a mean and median of 40.

Nationality was provided by 87% of respondents. Of these, 61% were British and 4% were Irish. Many other nationalities were represented, but each one formed less than 3% of the total group:

Nationality	Percentage
British	60.8%
Irish	3.9%
Other	35.3%

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

Employment status	Percentage
Full-time	92.0%
Part-time	4.3%
Self-employed	2.8%
Unemployed	1.0%
Retired	0.1%
Homemaker	0.1%

The majority of the group were of managerial level or above, with the largest single group being upper middle management (24%):

Occupational level	Percentage
Top level	4.4%
Senior executive	17.0%
Upper middle management	23.9%
Middle management	23.3%
First level management/supervisor	11.8%
Employee	15.8%
Other	3.8%

A range of work areas were represented:

Work area (job type)	Percentage
Finance	14.3%
HR, training, guidance	11.7%
Sales, customer service	9.0%
Health, social services, etc.	8.6%
Business services	8.4%
Science, engineering	7.5%
IT	6.2%
Education	3.3%
Research and development	3.0%
Admin or secretarial	2.9%
Military, police, prison, fire	2.5%
Leisure, personal service	0.6%
Land, sea or air transport	0.6%
Skilled operative	0.6%
Unskilled operative	0.1%
Other public sector	6.4%
Other private sector	5.7%
Other	8.5%

MBTI® Step II™ instrument

European Data Supplement

Danish

January 2018

2nd edition

OPP Ltd
Elsfield Hall
15–17 Elsfield Way
Oxford OX2 8EP
UK
t: +44 (0)1865 404 500

Contents

Introduction

Data collected for the European MBTI® Step II™ instrument were analysed to produce the findings in this supplement. This is the second data supplement produced for this instrument. A brief description of the sample is given below. Further details of the sample are provided in Appendix 1.

- The sample consisted of 15,183 individuals who completed the MBTI Step II questionnaire in Danish via the OPPassessment™ system between 2006 and mid-2016.[1] This sample is considered to be representative of the groups of people with whom the Danish MBTI Step II instrument has been and will be used for applications such as management development, coaching, counselling and teambuilding. As such, it is likely to represent a cross-section of the Danish-speaking professional and managerial population.

The results of the analyses are outlined below.

[1] OPPassessment allows personality questionnaires such as the MBTI instrument to be administered via email and/or completed online.

Type distribution

Type tables are a way of illustrating the proportion of each type within a particular group. Below is a type table taken from the sample described on the previous page.

For each of the 16 different types, the number of cases, the percentage of the total that this represents and the self-selection ratio (SSR) are shown. The SSR (Myers et al., 1998) is a way of demonstrating whether a given type appears more or less often in a particular group than would be expected given its frequency in a reference group. Ideally, the type distribution from a large representative sample of the Danish population would be used to calculate SSRs in this data supplement. However, such a sample does not currently exist. In its place, SSRs have been calculated using type data from the UK general population (Kendall, 1998), which can be justified by the fact that type distributions for comparable Danish and British groups, such as managers and professionals, are similar. Evidence (eg Hackston and Kendall, 2004; Quenk et al., 2004) does suggest that although type-related behaviours vary a good deal from country to country and from culture to culture, the frequencies of underlying MBTI types do not.

An SSR of greater than 1 indicates that a type is over-represented, and an SSR of less than 1 denotes that it is under-represented. Asterisks are used to denote whether the over- or under-representations are statistically significant, based on the results of chi-square analysis.[2]

[2] Chi-square analysis (often abbreviated to χ^2) is a technique used to explore whether observed frequency distributions differ significantly from other, predefined, distributions. In this case, the UK general population group is used as the reference group, and the chi-square analysis indicates whether the proportion of people of each type within a particular sample differs significantly from the proportion of people reporting the same type within the reference group.

OPPassessment data (representative Danish-speaking professional and managerial sample)

Table 3.1: Type table for OPPassessment data (reported type, n=15,183)

ISTJ	ISFJ	INFJ	INTJ
n=1703 11.2% SSR=0.82**	n=319 2.1% SSR=0.17**	n=97 0.6% SSR=0.35**	n=449 3.0% SSR=2.14**
ISTP	**ISFP**	**INFP**	**INTP**
n=474 3.1% SSR=0.48**	n=77 0.5% SSR=0.08**	n=128 0.8% SSR=0.25**	n=484 3.2% SSR=1.33
ESTP	**ESFP**	**ENFP**	**ENTP**
n=1193 7.9% SSR=1.36**	n=290 1.9% SSR=0.22**	n=862 5.7% SSR=0.90	n=1974 13.0% SSR=4.64**
ESTJ	**ESFJ**	**ENFJ**	**ENTJ**
n=3846 25.3% SSR=2.43**	n=918 6.0% SSR=0.48**	n=500 3.3% SSR=1.18	n=1869 12.3% SSR=4.24**

Type	n	%
E	11,452	75.4%**
I	3,731	24.6%**
S	8,820	58.1%**
N	6,363	41.9%**
T	11,992	79.0%**
F	3,191	21.0%**
J	9,701	63.9%**
P	5,482	36.1%**

*Difference significant at p<0.05, based on chi-square results.
**Difference significant at p<0.01, based on chi-square results.

The most common single type preference is ESTJ (25% of the total); this is a common finding with managerial groups in other countries. The SSR results suggest that, in comparison with the UK general population, those with preferences for NT are over-represented, and those with preferences for SF are under-represented. Again, this is a common finding with managerial groups.

Properties of the Step II facet scales

Facet scale score distribution

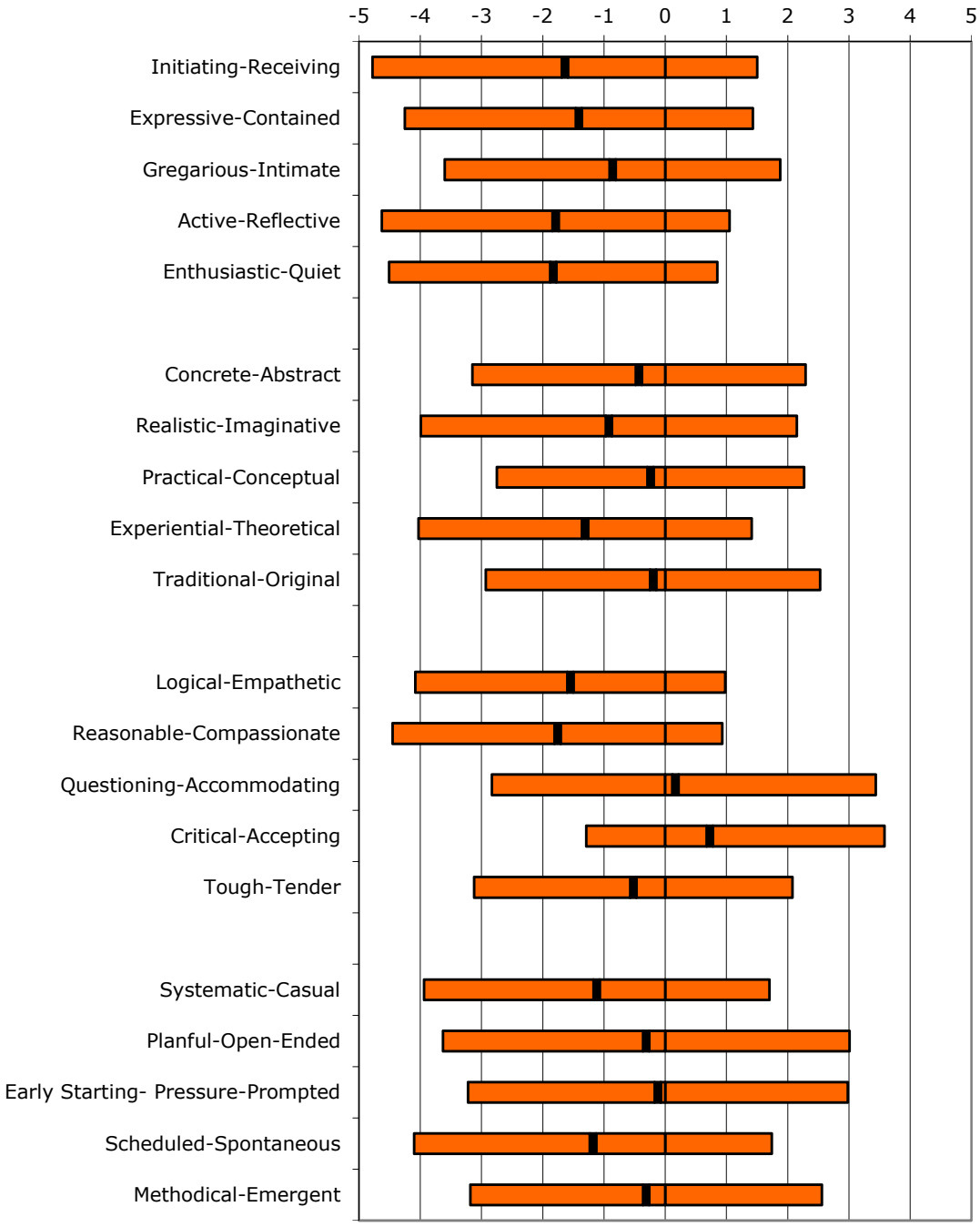
The means and standard deviations of the Step II facets are shown in Table 3.2, and are illustrated graphically in Figure 2.1. The mean score for a scale is calculated by adding together the scores of each individual in the sample and then dividing the sum by the number of individuals. Note that negative values indicate that mean scores are nearer the left-hand pole of the facet and positive values indicate that the mean scores are nearer the right-hand pole. The standard deviation (SD) is a statistical measure describing the degree to which the scores from the sample either bunch up close to, or are scattered widely around, the mean for the sample.

Table 3.2: Means and standard deviations of the facet scales

Step II facet scale	Mean[3]	SD
E–I facet scales		
Initiating–Receiving	–1.64	3.14
Expressive–Contained	–1.41	2.84
Gregarious–Intimate	–0.86	2.74
Active–Reflective	–1.79	2.84
Enthusiastic–Quiet	–1.83	2.68
S–N facet scales		
Concrete–Abstract	–0.43	2.72
Realistic–Imaginative	–0.92	3.07
Practical–Conceptual	–0.24	2.51
Experiential–Theoretical	–1.31	2.72
Traditional–Original	- 0.20	2.73
T–F facet scales		
Logical–Empathetic	–1.55	2.53
Reasonable–Compassionate	–1.76	2.69
Questioning–Accommodating	0.17	3.00
Critical–Accepting	0.73	2.02
Tough–Tender	–0.52	2.60
J–P facet scales		
Systematic–Casual	–1.12	2.82
Planful–Open-Ended	- 0.31	3.32
Early Starting–Pressure-Prompted	- 0.12	3.10
Scheduled–Spontaneous	–1.18	2.92
Methodical–Emergent	–0.31	2.87

[3] Scale means: in Step II reports, scores are given from 5 on one pole, through 0, to 5 on the opposite pole. In compiling statistical information, however, one pole needs to be clearly distinguished from the other. Throughout this data supplement, therefore, a convention has been adopted of scores 5, 4, 3, 2 or 1 on the left-hand pole being assigned values of –5, –4, –3, –2, or –1 and scores on the right-hand pole being assigned positive scores. This does not, of course, imply any suggestion that positive scores are 'better' than negative scores.

Figure 3.1: Facet scale mean scores[4]



[4] For each facet scale, the central line shows the mean, and the coloured bars on either side indicate the standard deviation.

In general, the means centre quite close to the midpoint of each scale. None of the scales has a mean greater than 2 points from the midpoint in either direction. However, a clear pattern does emerge. The mean scores tend towards the E, S, T and J direction (with the exception of Questioning–Accommodating and Critical–Accepting), which is consistent with the most common four-letter type preference amongst this sample. Amongst the group as a whole, there are more individuals with a preference for E, S, T and J.

Facet scale intercorrelations

Correlations among the Step II facet scales are shown in Table 3.3. Facet scales within each dichotomy consistently correlate more highly (usually substantially so) with the other scales of the same dichotomy than they do with scales in the other three dichotomies. For example, every E–I facet scale correlates more highly with the other four E–I scales than with any of the S–N, T–F or J–P facet scales.

There is only one scale that correlates **significantly** more highly with scales in other dichotomies. The T–F scale Questioning–Accommodating correlates with many scales, across dichotomies. For example, it correlates at a level of 0.22 with Enthusiastic–Quiet, and 0.20 with Initiating–Receiving and Active–Reflective on the E–I dichotomy. It correlates negatively at between –0.22 and –0.38 with all of the S–N scales of the S–N dichotomy. It also correlates negatively at a level of -0.23 with Early Starting–Pressure-Prompted, and -0.15 with Planful–Open Ended and Scheduled–Spontaneous on the J–P dichotomy. By way of comparison, it correlates at between 0.13 and 0.25 with the other four scales on the same dichotomy as itself (T–F). The negative correlations between Questioning–Accommodating and the S–N scales are consistent with findings in the US version of the Step II instrument, and suggest that a questioning approach to differences of opinion seems to be related to a range of iNtuition (N) facets. These patterns are similar to those found when the European Step II instrument was developed.

Table 3.3: Intercorrelations of Step II facet scales

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Initiating–Receiving	1.00																			
2. Expressive–Contained	0.53	1.00																		
3. Gregarious–Intimate	0.49	0.46	1.00																	
4. Active–Reflective	0.67	0.53	0.52	1.00																
5. Enthusiastic–Quiet	0.57	0.61	0.57	0.61	1.00															
6. Concrete–Abstract	-0.15	-0.17	-0.08	-0.12	-0.26	1.00														
7. Realistic–Imaginative	-0.14	-0.20	-0.14	-0.15	-0.31	0.65	1.00													
8. Practical–Conceptual	-0.13	-0.11	-0.09	-0.11	-0.24	0.58	0.57	1.00												
9. Experiential–Theoretical	-0.09	-0.10	-0.04	-0.07	-0.14	0.54	0.45	0.46	1.00											
10. Traditional–Original	-0.20	-0.18	-0.12	-0.18	-0.30	0.60	0.60	0.57	0.39	1.00										
11. Logical–Empathetic	-0.08	-0.23	-0.03	-0.07	-0.16	0.25	0.28	0.04	0.06	0.10	1.00									
12. Reasonable–Compassionate	-0.02	-0.14	-0.02	0.00	-0.09	0.17	0.19	-0.03	0.02	0.01	0.65	1.00								
13. Questioning–Accommodating	0.20	0.11	0.11	0.20	0.22	-0.26	-0.22	-0.31	-0.27	-0.38	0.13	0.21	1.00							
14. Critical–Accepting	-0.14	-0.20	-0.11	-0.11	-0.19	0.19	0.18	0.07	0.05	0.08	0.35	0.37	0.15	1.00						
15. Tough–Tender	0.02	-0.11	0.01	0.04	-0.06	0.18	0.19	0.02	0.04	0.01	0.47	0.55	0.25	0.40	1.00					
16. Systematic–Casual	-0.10	-0.16	-0.11	-0.12	-0.24	0.46	0.50	0.32	0.18	0.46	0.35	0.28	-0.10	0.22	0.25	1.00				
17. Planful–Open-Ended	-0.03	-0.02	-0.04	-0.05	-0.09	0.32	0.32	0.27	0.15	0.40	0.12	0.05	-0.15	0.04	0.03	0.56	1.00			
18. Early Starting–Pressure-Prompted	-1.12	-0.09	-0.13	-0.13	-0.20	0.27	0.27	0.27	0.19	0.35	0.02	-0.05	-0.23	-0.01	-0.07	0.37	0.45	1.00		
19. Scheduled–Spontaneous	-0.04	-0.06	-0.06	-0.07	-0.13	0.33	0.34	0.25	0.18	0.40	0.17	0.09	-0.15	0.04	0.06	0.59	0.64	0.53	1.00	
20. Methodical–Emergent	0.00	-0.03	-0.06	-0.04	-0.10	0.21	0.23	0.17	0.12	0.24	0.13	0.08	-0.09	0.04	0.06	0.46	0.47	0.44	0.58	1.00

Correlations of Step II facet scales with Step I scales

Correlations between Step II facet scales and the continuous scores from the MBTI Step I™ instrument are shown in Table 3.4.[5] Positive correlations between facet scales and the corresponding Step I dimension (e.g. between Initiating–Receiving and the E–I continuous score) indicate relationships in the expected direction, with higher facet scores tending to be associated with higher continuous scores, and vice versa. Negative correlations would indicate an inverse relationship between the facet scale scores and the corresponding Step I dimension.

Table 3.4: Correlations of Step II facet scales with Step I continuous scores

	Step I continuous score			
Step II facet scales	**E–I**	**S–N**	**T–F**	**J–P**
E–I facet scales				
Initiating–Receiving	**0.81**	-0.17	-0.08	-0.06
Expressive–Contained	**0.75**	-0.18	-0.24	-0.07
Gregarious–Intimate	**0.63**	-0.10	-0.05	-0.09
Active–Reflective	**0.80**	-0.15	-0.06	-0.09
Enthusiastic–Quiet	**0.81**	-0.28	-0.17	-0.17
S–N facet scales				
Concrete–Abstract	-0.19	**0.85**	0.23	0.41
Realistic–Imaginative	-0.22	**0.77**	0.25	0.42
Practical–Conceptual	-0.16	**0.68**	0.00	0.33
Experiential–Theoretical	-0.11	**0.64**	0.03	0.22
Traditional–Original	-0.24	**0.76**	0.05	0.48
T–F facet scales				
Logical–Empathetic	-0.16	0.20	**0.83**	0.18
Reasonable–Compassionate	-0.08	0.10	**0.77**	0.10
Questioning–Accommodating	0.21	-0.34	**0.28**	-0.20
Critical–Accepting	-0.18	0.14	**0.51**	0.07
Tough–Tender	-0.02	0.12	**0.70**	0.07
J–P facet scales				
Systematic–Casual	-0.17	0.48	0.35	**0.72**
Planful–Open-Ended	-0.05	0.36	0.07	**0.82**
Early Starting–Pressure-Prompted	-0.15	0.32	-0.04	**0.68**
Scheduled–Spontaneous	-0.08	0.38	0.11	**0.80**
Methodical–Emergent	-0.04	0.23	0.09	**0.68**

The E–I facet scales correlate at a level of 0.63 to 0.81 with the E–I continuous scores from Step I; the S–N facet scales correlate at 0.64 to 0.85 with the S–N continuous scores; the T–F facet scales correlate

[5] Continuous scores place an individual's score on each dimension onto a continuous scale with a midpoint of 100. To calculate continuous scores, Preference Clarity Index (PCI) scores for each dimension are either subtracted or added to 100, depending on which direction the overall preference is. PCI scores in the direction of E, S, T or J are subtracted from 100. PCI scores in the direction of I, N, F or P are added to 100.

at 0.28 to 0.83 with the T–F continuous scores; and the J–P facet scales correlate at 0.68 to 0.82 with the J–P continuous scores. The scale that showed the lowest correlation with its associated Step I continuous score was Questioning–Accommodating at 0.28. This scale is also amongst the lowest on internal consistency (see page 52), and has been previously found to have the lowest test-retest reliability (Quenk, Hammer and Majors, 2004), which would result in the true correlations being underestimated.

These correlations are very similar to those found during the development of the Step II questionnaire. This consistency, alongside the fact that the correlations between Step II facet scales and Step I continuous scores associated with their own dimension are substantially higher than correlations with the other three dimensions, provides compelling evidence for the theoretical hierarchical structure of the Step II facet scales in relation to the Step I scales.

Out-of-preference scores

It is known that the five facet scales relating to each type dimension do not describe the dimension in its entirety; there will not be a precise and exact overlap between, for example, an individual's score on E–I and their total score across the five facet scales that relate to this dimension. For example, it is not uncommon to see an Enthusiastic Introvert or an Intimate Extravert. Such apparent inconsistencies are usually known as 'out-of-preference scores' (OOPS).

Although it is usual to have a number of OOPS in any one profile, it should be unusual to find that OOPS make up more than half of the facet scores relating to any one type dichotomy. Table 3.5 demonstrates that the proportion of people for whom this happens ranges from 3.9% for the T–F block to 1.0% for the E–I block. It is therefore a very infrequent occurrence.

Table 3.5: Proportion of OOPS by type dichotomy[6]

	Proportion of 'reported type' OOPS					
	None	One	Two	Three	Four	Five
E–I	73.6%	20.0%	5.4%	0.9%	0.1%	0.0%
S–N	65.5%	26.2%	7.0%	1.2%	0.1%	0.0%
T–F	44.1%	36.7%	15.3%	3.5%	0.4%	0.0%
J–P	60.8%	29.6%	8.5%	1.1%	0.0%	0.0%

[6] In this table, an OOPS is defined as a score of 2, 3, 4 or 5 on the 'wrong' pole of the facet when compared with the reported type.

For any individual facet scale, an OOPS tends to occur in approximately 10% of profiles, with the exception of Questioning–Accommodating (34%) and Critical–Accepting (27%).

Reliability

The reliability of a test or questionnaire relates to how consistent and precise it is. Internal consistency reliability addresses the question of whether all the questions in a scale measure the same construct. For example, are the Step II facet scales consistent within themselves, and do they hold together well as scales? A common measure of internal consistency reliability is coefficient alpha (Cronbach, 1951). The alpha coefficients for the Step II facet scales are shown in Table 3.6.

Table 3.6: Internal consistency reliability

Step II facet scale	No. of items	Coefficient alpha
E–I facet scales		
Initiating–Receiving	8	0.83
Expressive–Contained	7	0.73
Gregarious–Intimate	7	0.63
Active–Reflective	8	0.70
Enthusiastic–Quiet	9	0.68
S–N facet scales		
Concrete–Abstract	9	0.73
Realistic–Imaginative	7	0.75
Practical–Conceptual	8	0.42
Experiential–Theoretical	8	0.71
Traditional–Original	8	0.70
T–F facet scales		
Logical–Empathetic	9	0.74
Reasonable–Compassionate	8	0.67
Questioning–Accommodating	7	0.53
Critical–Accepting	8	0.39
Tough–Tender	8	0.72
J–P facet scales		
Systematic–Casual	8	0.78
Planful–Open-Ended	6	0.76
Early Starting–Pressure-Prompted	6	0.72
Scheduled–Spontaneous	8	0.71
Methodical–Emergent	8	0.62
	Median	0.71

The internal consistency reliability of most scales is good, and the average (median) reliability is above 0.7. However, three scales (Practical–Conceptual, Questioning–Accommodating and Critical–Accepting) do have lower alpha coefficients. This pattern is consistent with what was found during the development of the Step II

instrument, where these three facet scales were also found to show lower reliability than the others (Quenk, Hammer and Majors, 2004).

It should be noted, however, that coefficient alpha reliability statistics will give an underestimate of the reliability of the Step II facet scales, given the use of item response theory (IRT) methods in the actual scoring process.[7]

Group differences

This section looks at the way in which people from different groups responded to the Danish version of the European Step II questionnaire.

Gender

- The means and standard deviations of the Step II facet scales are shown separately for males and females in Table 3.7, along with the difference in mean scores. This is represented graphically in Figure 3.2. Statistically significant differences were found between male and female mean scores for 17 of the 20 facet scales, although in most cases these were too small to be of any practical significance.
- On the E–I facet scales, all five mean scores tended slightly toward the E pole for both males and females. There were statistically significant gender differences on four of the five facets, with females tending further towards the E pole than males. There was no statistically significant difference between genders on the remaining scale.
- On the S–N facet scales, all five mean scores tended slightly toward the S pole for both males and females. There were statistically significant gender differences on all five facet scales, with male mean scores tending more towards the S pole than females.
- On the T–F facet scales, four of the five mean scores tended toward the T pole for males, whereas three did so for females. There were statistically significant differences on all five facet scales, with male mean scores tending more towards the T pole than females.
- On the J–P facet scales, three of the five mean scores tended toward the J pole for both males and females. The other two tended towards the J pole for females and the P pole for males, though for males it was very close to the midpoint. There were significant gender differences on three of the five scales, with male mean scores tending more towards the J pole than females on one scale and female mean scores tending more towards the J pole than males on the other two facets. There was no consistent pattern

[7] Item response theory (IRT) is an approach to measurement that is concerned with modelling the relationship between item responses and the underlying characteristic assessed by the scale or test the item is designed to measure. IRT can be used to select items for a test and/or to score the items.

across facet scales of either males or females tending further towards the J pole. There were no statistically significant differences on the other two scales.

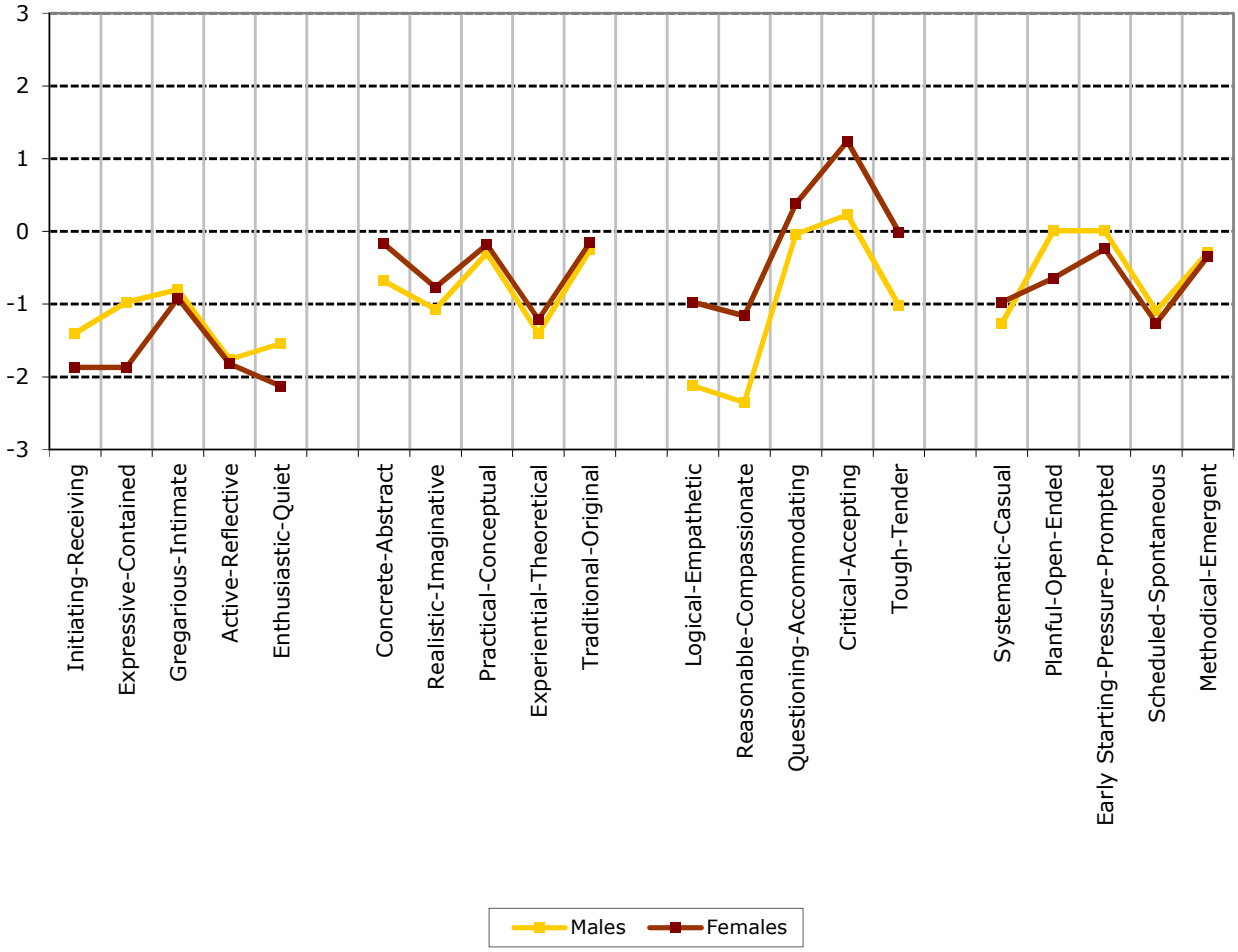
Table 3.7: Gender differences in facet scale scores

	Males (n=2,372)		Females (n=1,877)		Difference (M–F)[8]
Step II facet scale	**Mean**	**SD**	**Mean**	**SD**	
E–I facet scales					
Initiating–Receiving	-1.40	3.20	-1.87	3.07	0.47**
Expressive–Contained	-0.97	2.30	-1.87	2.70	0.90**
Gregarious–Intimate	-0.80	2.79	-0.92	2.69	0.12**
Active–Reflective	-1.76	2.82	-1.82	2.87	0.06
Enthusiastic–Quiet	-1.54	2.74	-2.13	2.58	0.59**
S–N facet scales					
Concrete–Abstract	-0.68	2.70	-0.17	2.71	–0.51**
Realistic–Imaginative	-1.07	3.11	-0.77	3.01	–0.30**
Practical–Conceptual	-0.30	2.50	-0.18	2.52	–0.13**
Experiential–Theoretical	-1.40	2.76	-1.21	2.68	–0.19**
Traditional–Original	-0.25	2.77	-0.15	2.68	–0.10*
T–F facet scales					
Logical–Empathetic	-2.12	2.33	-0.97	2.60	–1.14**
Reasonable–Compassionate	-2.35	2.56	-1.16	2.68	–1.19**
Questioning–Accommodating	-0.04	2.97	0.38	3.02	–0.42**
Critical–Accepting	0.23	2.01	1.24	1.91	–1.01**
Tough–Tender	-1.02	2.59	-0.01	2.51	–1.01**
J–P facet scales					
Systematic–Casual	-1.26	2.81	-0.97	2.83	–0.28**
Planful–Open-Ended	0.01	3.34	-0.64	3.27	0.64**
Early Starting–Pressure-Prompted	0.01	3.09	-0.24	3.09	0.25
Scheduled–Spontaneous	-1.09	2.94	-1.26	2.90	0.17**
Methodical–Emergent	-0.29	2.87	-0.34	2.87	0.05

Difference significant at: *p<0.05, **p<0.01 (based on an independent samples t-test).

[8] A positive value indicates that male scores tend more towards I, N, F or P, and a negative value indicates that female scores tend more towards I, N, F or P.

Figure 3.2: Gender differences in facet scale scores



Age

There were found to be no meaningful links between age and facet scale scores. Although correlational analysis showed several facet scales to be significantly correlated with age, the significance levels were more the result of the very large sample size rather than being indicative of a meaningful relationship. The highest correlations were 0.14 (Critical–Accepting) and 0.12 (Tough–Tender) and even these are too small to be considered to be meaningful.

For interpretation purposes, it is reasonable to conclude that there are no clear relationships between age and facet scale scores.

Ethnic origin

Ethnic origin information was not captured for people who completed the Danish language version of the questionnaire, so no analyses were conducted.

Occupational level

Research using the Danish version of the MBTI Step I questionnaire has demonstrated that individuals in higher-level jobs in organisations

are more likely to have preferences for Extraversion, iNtuition, Thinking and Perceiving than those in lower-level jobs.[9]

Table 3.8 shows the facet scale means and standard deviation for different occupational levels. The findings can be summarised as follows.

- A pattern was found regarding the facets relating to the E–I dimensions, with mean scores of those at more senior levels tending to be further towards the E pole than those for the lower occupational groups. Those at the top level were also significantly more Gregarious than those at all other levels.
- A pattern was found regarding the facets relating to the S–N dimensions, with mean scores of those at more senior levels tending to be further towards the N pole than those for the lower occupational level groups. Those at first level management/supervisory level tended more towards the S pole than all other levels on all facets.
- A pattern was found regarding the facets relating to the T–F dimension, with the employee group mean scores tending further towards the F pole (or less towards the T pole) than other groups. No clear patterns were found between groups above employee level. This could be a reflection of the fact that females formed a considerably higher proportion of the employee group than they did of other groups.
- There were no clear overall patterns regarding the facets in the J–P preference block.

[9] See the MBTI Step I European Data Supplement for details.

Table 3.8: Mean facet scale scores by occupational level

Step II facet scale	Top level (n=665)		Senior Executive (n=1,073)		Upper Middle Management (n=1,518)		Middle Management (n=3,271)		First Level Mgt/ Supervisory (n=565)		Employee (n=4,754)		Sig.
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales													
Initiating–Receiving	-2.44	2.83	-2.01	2.90	-2.03	2.94	-1.83	3.04	-1.10	3.27	-1.31	3.25	**
Expressive–Contained	-1.72	2.76	-1.60	2.79	-1.52	2.85	-1.45	2.78	-1.11	2.93	-1.42	2.88	**
Gregarious–Intimate	-1.22	2.69	-0.98	2.75	-0.97	2.69	-1.00	2.71	-0.81	2.79	-0.75	2.75	**
Active–Reflective	-2.58	2.55	-2.04	2.71	-2.16	2.71	-1.98	2.76	-1.47	2.97	-1.54	2.95	**
Enthusiastic–Quiet	-2.36	2.64	-2.08	2.61	-2.13	2.59	-1.95	2.65	-1.61	2.73	-1.69	2.71	**
S–N facet scales													
Concrete–Abstract	-0.01	2.77	-0.21	2.75	-0.29	2.73	-0.53	2.69	-0.70	2.52	-0.53	2.68	**
Realistic–Imaginative	-0.51	3.14	-0.82	3.04	-0.78	3.07	-1.04	3.02	-1.20	3.00	-0.96	3.06	**
Practical–Conceptual	0.09	2.28	0.05	2.30	-0.04	2.36	-0.29	2.39	-0.44	2.47	-0.34	2.65	**
Experiential–Theoretical	-1.18	2.70	-1.06	2.70	-1.07	2.69	-1.32	2.71	-1.63	2.62	-1.38	2.73	**
Traditional–Original	0.41	2.65	0.22	2.67	0.05	2.68	-0.25	2.67	-0.40	2.72	-0.40	2.71	**
T–F facet scales													
Logical–Empathetic	-1.72	2.45	-2.09	2.32	-1.97	2.38	-1.72	2.43	-1.87	2.41	-1.26	2.61	**
Reasonable–Compassionate	-2.37	2.59	-2.47	2.43	-2.27	2.51	-2.01	2.52	-2.02	2.64	-1.29	2.78	**
Questioning–Accommodating	-0.41	2.90	-0.50	2.90	-0.21	2.95	-0.09	2.90	0.28	2.95	0.60	3.05	**
Critical–Accepting	0.44	2.04	0.47	1.97	0.61	1.94	0.77	1.95	0.65	1.96	0.83	2.05	**
Tough–Tender	-1.18	2.43	-1.10	2.45	-0.89	2.52	-0.67	2.52	-0.75	2.37	-0.10	2.65	**
J–P facet scales													
Systematic–Casual	-1.06	2.67	-1.23	2.67	-1.26	2.75	-1.30	2.80	-1.23	2.74	-1.00	2.88	**
Planful–Open–Ended	0.26	3.36	-0.27	3.25	-0.24	3.29	-0.43	3.29	-0.21	3.31	-0.44	3.32	**
Early Starting–Pressure–Prompted	0.89	3.08	0.54	2.96	0.43	3.11	-0.03	3.04	-0.16	2.91	-0.52	3.07	**
Scheduled–Spontaneous	-0.94	2.93	-1.09	2.88	-1.16	2.92	-1.26	2.84	-1.21	2.78	-1.22	2.95	
Methodical–Emergent	-0.01	2.85	-0.28	2.85	-0.34	2.92	-0.31	2.85	-0.28	2.84	-0.31	2.88	

Significant at: *p<0.05, **p<0.01 (based on a one-way analysis of variance).

Education

Specific educational qualifications were not collected for the OPPassessment sample; however, the age at which individuals left full-time education was. No significant and meaningful correlations were found between the age at which people left full-time education and their facet scale scores.

Work area

Information regarding the area of work people engage in was collected for the group. Many different categories were used, but for the purposes of analysis the focus was on the five most commonly occurring. These were as follows:

- Health, social services
- Science, engineering
- IT
- HR, training, guidance
- Finance

Table 3.9 shows the facet scale means and standard deviation for these five work areas. The findings can be summarised as follows:

- A clear pattern was found regarding the facets relating to the E–I dimensions. Although mean scores for all groups tended towards the E pole for each facet scale, scores were further towards the E pole for the 'HR, training, guidance' groups than for the other four groups.
- Another clear pattern was found regarding the facets relating to the S–N dimensions. Mean scores for all groups tended towards the S pole (or the mid-point), except the 'HR, training, guidance' group, which tended towards the N pole on four of the five facets, the fifth facet being a mid-point score.
- The most noticeable pattern regarding the facets in the T–F preference block was that although mean scores for all groups tended towards the T pole (or the mid-point) on three out of the five facet scales, scores for all five work area groups tended towards the F pole for the Critical–Accepting facet scale. This is a common finding with this scale.
- The pattern regarding the facets in the J–P facet block was that mean scores for four of the groups tended towards the J pole (or the mid-point) on all five facet scales. With the 'HR, training, guidance' group, however, mean scores for 3 of the five facet scales tended towards the P pole and the remaining two facet scale

scores, while tending to the J pole, were significantly less so than the mean scores for the other four groups.

Table 3.9: Mean facet scale scores by work area

Step II facet scale	Health, social services (n=1,312)		Science, engineering (n=1,010)		IT (n=969)		HR, training, guidance (n=946)		Finance (n=837)		Sig.
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales											
Initiating–Receiving	-1.97	2.99	-1.25	3.17	-1.20	3.26	-2.25	2.88	-1.37	3.32	**
Expressive–Contained	-1.54	2.69	-1.13	2.92	-1.24	2.87	-2.24	2.71	-0.97	2.87	**
Gregarious–Intimate	-0.60	2.66	-0.96	2.73	-0.68	2.68	-1.21	2.83	-0.94	2.72	**
Active–Reflective	-1.87	2.80	-1.71	2.84	-1.69	2.88	-2.20	2.77	-1.69	2.92	**
Enthusiastic–Quiet	-1.84	2.60	-1.71	2.71	-1.75	2.75	-2.53	2.47	-1.74	2.72	**
S–N facet scales											
Concrete–Abstract	-0.07	2.66	-0.71	2.63	-0.34	2.61	0.94	2.90	-1.03	2.64	**
Realistic–Imaginative	-0.76	2.92	-1.13	3.05	-0.98	3.02	0.36	3.18	-1.57	2.90	**
Practical–Conceptual	-0.10	2.44	-0.13	2.43	0.02	2.47	0.47	2.37	-0.50	2.40	**
Experiential–Theoretical	-0.75	2.68	-1.30	2.68	-1.37	2.74	-0.04	2.89	-1.73	2.71	**
Traditional–Original	-0.15	2.57	-0.32	2.66	-0.03	2.71	0.90	2.77	-0.55	2.65	**
T–F facet scales											
Logical–Empathetic	-1.19	2.53	-2.35	2.18	-2.11	2.28	-0.83	2.94	-2.16	2.31	**
Reasonable–Compassionate	-1.42	2.68	-2.52	2.48	-2.37	2.47	-1.28	2.76	-2.35	2.56	**
Questioning–Accommodating	-0.28	3.07	0.17	2.96	-0.02	2.96	-0.19	3.07	0.24	3.07	**
Critical–Accepting	1.15	1.76	0.38	1.94	0.47	2.02	1.14	2.06	0.42	2.01	**
Tough–Tender	-0.16	2.46	-0.88	2.59	-0.82	2.50	-0.22	2.54	-0.88	2.54	**
J–P facet scales											
Systematic–Casual	-0.81	2.82	-1.59	2.69	-1.23	2.85	-0.30	3.02	-1.58	2.73	**
Planful–Open-Ended	-0.33	3.23	-0.44	3.31	-0.14	3.39	0.23	3.43	-0.42	3.35	**
Early Starting–Pressure-Prompted	0.04	3.09	-0.23	3.11	-0.07	3.02	0.70	3.22	-0.12	3.05	**
Scheduled–Spontaneous	-1.11	2.86	-1.31	2.98	-1.10	2.95	-0.26	3.25	-1.53	2.79	**
Methodical–Emergent	0.04	2.86	-0.67	2.91	-0.54	2.83	0.29	2.95	-0.36	2.94	**

Significant at: *p<0.05, **p<0.01 (based on a one-way analysis of variance).

Nationality

Information on nationality was available for the group. Ninety-seven per cent of the group were Danish, and no other nationality was represented in large numbers. Therefore, no analyses of national differences were conducted.

Employment status

Employment status information was available for the OPPassessment sample. The vast majority of the group worked either full-time or part-time, or were self-employed. Table 3.10 shows the mean facet scale scores for each group, with the data illustrated in graphical form in Figure 3.3.

The analyses showed statistically significant differences across the groups on several of the facet scales on the S–N, T–F and J–P dimensions, but none of the five facets on the E–I dimension. The clearest patterns were as follows:

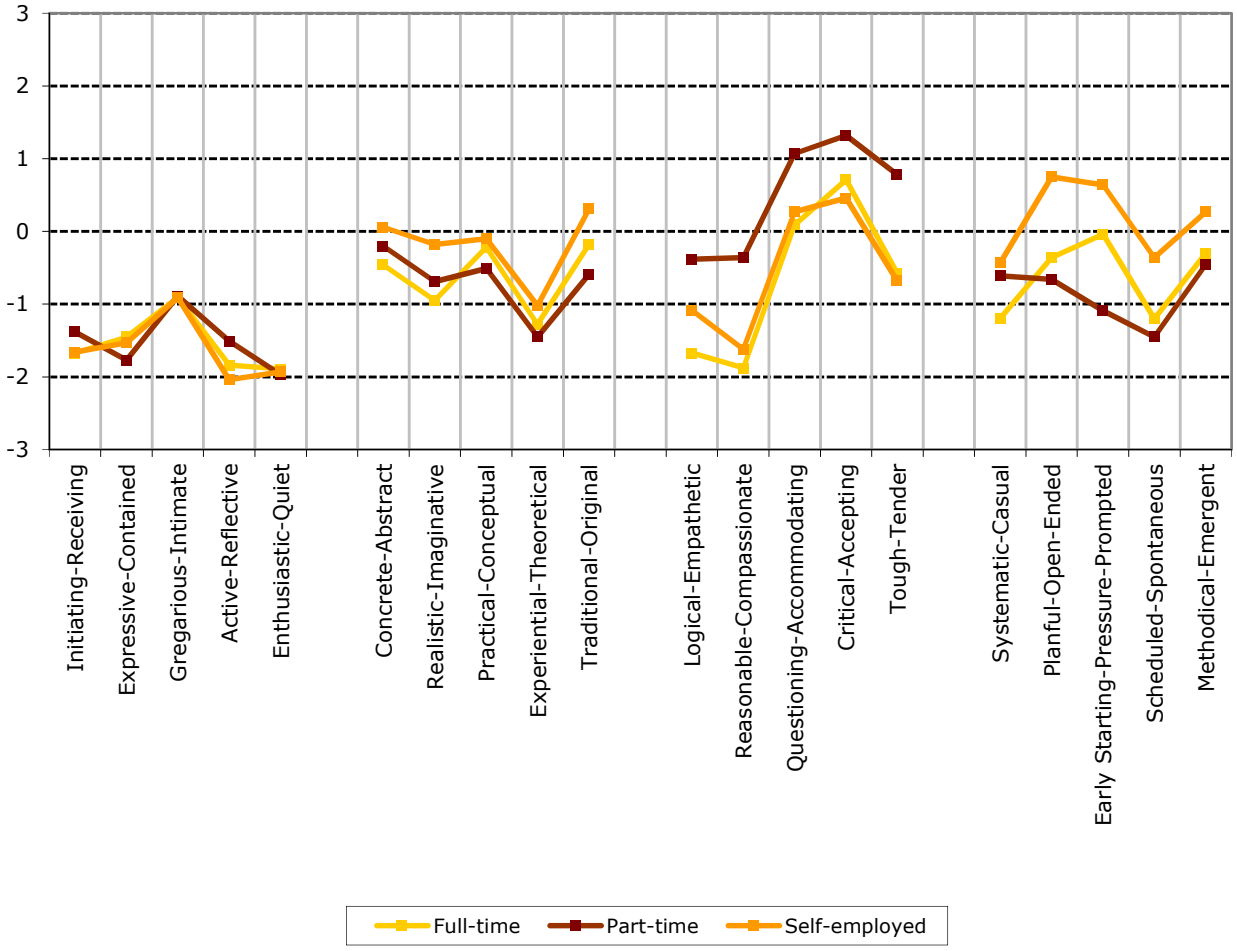
- The self-employed group tended to score more towards the N pole on the S–N facet scales and more towards the P pole on the J–P facet scales, than the other two groups.
- The full-time group tended to score more towards the T pole on the T–F facet scales than the other two groups, particularly the part-time group. This is likely to be at least partly a gender effect; 81% of part-time workers were female, compared with 49% of the total group and 48% of full-time workers.

Table 3.10: Mean facet scale scores by employment status

Step II facet scale	Full-time (n=11,707)		Part-time (n=496)		Self-employed (n=357)		Sig.
	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales							
Initiating–Receiving	-1.67	3.12	-1.38	3.12	-1.66	3.33	
Expressive–Contained	-1.45	2.84	-1.77	2.67	-1.53	2.97	
Gregarious–Intimate	-0.91	2.74	-0.89	2.60	-0.91	2.79	
Active–Reflective	-1.84	2.84	-1.51	2.95	-2.04	2.84	*
Enthusiastic–Quiet	-1.89	2.67	-1.97	2.48	-1.93	2.74	
S–N facet scales							
Concrete–Abstract	-0.46	2.70	-0.20	2.67	0.06	3.02	**
Realistic–Imaginative	-0.95	3.05	-0.69	3.00	-0.18	3.29	**
Practical–Conceptual	-0.22	2.49	-0.51	2.60	-0.10	2.50	*
Experiential–Theoretical	-1.28	2.72	-1.45	2.63	-1.02	2.86	
Traditional–Original	-0.18	2.71	-0.59	2.59	0.31	2.85	**
T–F facet scales							
Logical–Empathetic	-1.67	2.48	-0.38	2.69	-1.08	2.78	**
Reasonable–Compassionate	-1.88	2.64	-0.36	2.80	-1.62	2.80	**
Questioning–Accommodating	0.09	2.99	1.07	3.07	0.27	2.92	**
Critical–Accepting	0.71	1.20	1.32	1.95	0.46	2.20	**
Tough–Tender	-0.58	2.58	0.78	2.54	-0.68	2.51	**
J–P facet scales							
Systematic–Casual	-1.19	2.80	-0.61	2.92	-0.43	2.89	**
Planful–Open-Ended	-0.36	3.31	-0.66	3.29	0.75	3.31	**
Early Starting–Pressure-Prompted	-0.04	3.08	-1.09	3.01	0.64	3.11	**
Scheduled–Spontaneous	-1.20	2.90	-1.45	2.94	-0.36	3.11	**
Methodical–Emergent	-0.30	2.87	-0.45	2.86	0.27	2.93	**

Significant at: *p<0.05, **p<0.01 (based on a one-way analysis of variance).

Figure 3.3: Mean facet scale scores by employment status



Appendix 1: Sample description

Sample 1: Data from OPPassessment (representative Danish-speaking professional and managerial sample)

This sample consists of 15,183 individuals who completed the MBTI Step II instrument in Danish via the OPPassessment system between April 2006 and July 2016. 50% of the respondents were male and 50% were female. Of those who supplied their age, this ranged from 19 to 72 years, with a mean of 42 and a median of 42.

Nationality was disclosed by 91% of respondents. Of these, 97% were Danish. Many other nationalities were represented, but each one formed less than 1% of the total group:

Nationality	Percentage
Danish	97.1%
Other	2.9%

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

Employment status	Percentage
Full-time	92.1%
Part-time	3.9%
Self employed	2.8%
Unemployed	1.1%
Retired	0.1%
Homemaker	0.0%

The majority of the group were of managerial level or above, but with the largest single group being employee (38%):

Occupational level	Percentage
Top level	5.3%
Senior executive	8.6%
Upper middle management	12.2%
Middle management	26.3%
First level management/supervisor	4.5%
Employee	38.2%
Other	4.8%

A range of work areas were represented:

Work area (job type)	Percentage
Health, social services, etc.	10.5%
Science, engineering	8.1%
IT	7.7%
HR, training, guidance	7.5%
Finance	6.7%
Admin or secretarial	6.6%
Sales, customer service	5.6%
Research and development	4.8%
Education	3.4%
Skilled operative	3.0%
Land, sea or air transport	2.1%
Military, police, prison, fire	0.8%
Unskilled operative	0.8%
Leisure, personal service	0.4%
Other public sector	13.8%
Other private sector	9.1%
Other service area	3.4%
Other	5.6%

MBTI® Step II™ instrument

European Data Supplement

Dutch

January 2018

2nd edition

OPP Ltd
Elsfield Hall
15–17 Elsfield Way
Oxford OX2 8EP
UK
t: +44 (0)1865 404 500

Contents

Introduction

Data collected for the European MBTI® Step II™ instrument were analysed to produce the findings in this supplement. This is the second data supplement produced for this instrument. A brief description of the sample is given below. Further details of the sample are provided in Appendix 1.

- The sample consisted of 19,354 individuals who completed the MBTI Step II instrument in Dutch via the OPPassessment™ system between 2006 and mid-2016.[1] This sample is considered to be representative of the groups of people with whom the Dutch MBTI Step II instrument has been and will be used for applications such as management development, coaching, counselling and teambuilding. As such, it is likely to represent a cross-section of the Dutch-speaking professional and managerial population.

The results of the analyses are outlined below.

[1] OPPassessment allows personality questionnaires such as the MBTI instrument to be administered via email and/or completed online.

Type distribution

Type tables are a way of illustrating the proportion of each type within a particular group. Below is a type table taken from the sample described on the previous page.

For each of the 16 different types, the number of cases, the percentage of the total that this represents and the self-selection ratio (SSR) are shown. The SSR (Myers et al., 1998) is a way of demonstrating whether a given type appears more or less often in a particular group than would be expected given its frequency in a reference group. Ideally, the type distribution from a large representative sample of the Dutch and/or Belgian population would be used to calculate SSRs in this data supplement. However, such a sample does not currently exist. In its place, SSRs have been calculated using type data from the UK general population (Kendall, 1998), which can be justified by the fact that type distributions for comparable Dutch/Belgian and British groups, such as managers and professionals, are similar. Evidence (eg Hackston and Kendall, 2004; Quenk et al., 2004) does suggest that although type-related behaviours vary a good deal from country to country and from culture to culture, the frequencies of underlying MBTI types do not.

An SSR of greater than 1 indicates that a type is over-represented, and an SSR of less than 1 denotes that it is under-represented. Asterisks are used to denote whether the over- or under-representations are statistically significant, based on the results of chi-square analysis.[2]

[2] Chi-square analysis (often abbreviated to χ^2) is a technique used to explore whether observed frequency distributions differ significantly from other, predefined, distributions. In this case, the UK general population group is used as the reference group, and the chi-square analysis indicates whether the proportion of people of each type within a particular sample differs significantly from the proportion of people reporting the same type within the reference group.

OPPassessment data (representative Dutch-speaking professional and managerial sample)

Table 4.1: Type table for OPPassessment data (reported type, n=19,354)

ISTJ	ISFJ	INFJ	INTJ
n=2153 11.1% SSR=0.81**	n=738 3.8% SSR=0.30**	n=210 1.1% SSR=0.65*	n=556 2.9% SSR=2.07**
ISTP	**ISFP**	**INFP**	**INTP**
n=843 4.4% SSR=0.69**	n=328 1.7% SSR=0.28**	n=464 2.4% SSR=0.75	n=810 4.2% SSR=1.75**
ESTP	**ESFP**	**ENFP**	**ENTP**
n=1558 8.1% SSR=1.40**	n=793 4.1% SSR=0.47**	n=1494 7.7% SSR=1.22*	n=2101 10.9% SSR=3.89**
ESTJ	**ESFJ**	**ENFJ**	**ENTJ**
n=3704 19.1% SSR=1.84**	n=1426 7.4% SSR=0.59**	n=672 3.5% SSR=1.25	n=1504 7.8% SSR=2.69**

Type	n	%
E	13,252	68.5%**
I	6,102	31.5%**
S	11,543	59.6%**
N	7,811	40.4%**
T	13,229	68.4%**
F	6,125	31.6%**
J	10,963	56.6%
P	8,391	43.4%

*Difference significant at p<.05, based on chi-square results.
**Difference significant at p<.01, based on chi-square results.

The most common single type preference is ESTJ (19% of the total); this is a common finding with managerial groups in other countries. The SSR results suggest that, in comparison with the UK general population, those with preferences for NT are over-represented, and those with preferences for SF are under-represented. Again, this is a common finding with managerial groups.

Properties of the Step II facet scales

Facet scale score distribution

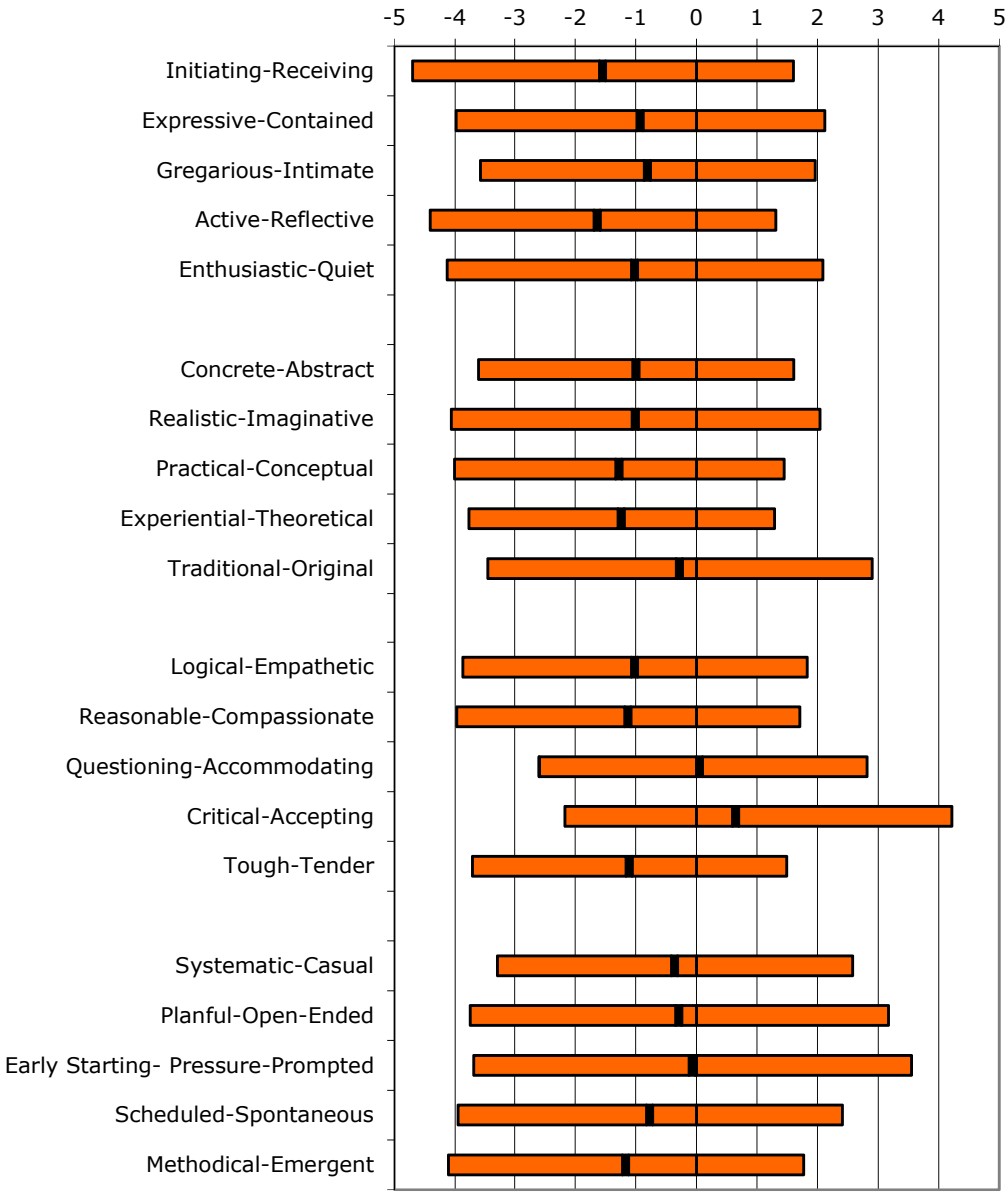
The means and standard deviations of the Step II facets are shown in Table 4.2, and illustrated graphically in Figure 4.1. The mean score for a scale is calculated by adding together the scores of each individual in the sample and then dividing the sum by the number of individuals. Note that negative values indicate that mean scores are nearer the left-hand pole of the facet, and positive values indicate that the mean scores are nearer the right-hand pole. The standard deviation (SD) is a statistical measure describing the degree to which the scores from the sample either bunch up close to, or are scattered widely around, the mean for the sample.

Table 4.2: Means and standard deviations of the facet scales

Step II facet scale	Mean[3]	SD
E–I facet scales		
Initiating–Receiving	–1.55	3.15
Expressive–Contained	–0.93	3.05
Gregarious–Intimate	–0.81	2.77
Active–Reflective	-1.64	2.95
Enthusiastic–Quiet	-1.02	3.11
S–N facet scales		
Concrete–Abstract	-1.00	2.61
Realistic–Imaginative	-1.01	3.05
Practical–Conceptual	-1.28	2.73
Experiential–Theoretical	-1.24	2.53
Traditional–Original	-0.28	3.18
T–F facet scales		
Logical–Empathetic	-1.02	2.85
Reasonable–Compassionate	-1.13	2.84
Questioning–Accommodating	0.04	2.63
Critical–Accepting	0.65	2.82
Tough–Tender	-1.11	2.60
J–P facet scales		
Systematic–Casual	-0.36	2.94
Planful–Open-Ended	-0.29	3.46
Early Starting–Pressure-Prompted	-0.07	3.62
Scheduled–Spontaneous	-0.77	3.18
Methodical–Emergent	-1.17	2.94

[3] Scale means: in Step II reports, scores are given from 5 on one pole, through 0, to 5 on the opposite pole. In compiling statistical information, however, one pole needs to be clearly distinguished from the other. Throughout this data supplement, therefore, a convention has been adopted of scores 5, 4, 3, 2 or 1 on the left-hand pole being assigned values of –5, –4, –3, –2, or –1, and scores on the right-hand pole being assigned positive scores. This does not, of course, imply any suggestion that positive scores are 'better' than negative scores.

Figure 4.1: Facet scale mean scores[4]



[4] For each facet scale, the central line shows the mean, and the coloured bars on either side indicate the standard deviation.

In general, the means centre quite close to the midpoint of each scale. None of the scales has a mean greater than 2 points from the midpoint in either direction. However, a clear pattern does emerge. The mean scores tend towards the E, S, T and J direction (with the exception of Questioning–Accommodating and Critical–Accepting), which is consistent with the most common four-letter type preference amongst this sample. Amongst the group as a whole, there are more individuals with a preference for E, S, T and J.

Facet scale intercorrelations

Correlations among the Step II facet scales are shown in Table 4.3. Facet scales within each dichotomy consistently correlate more highly (usually substantially so) with the other scales of the same dichotomy than they do with scales in the other three dichotomies. For example, every E–I facet scale correlates higher with the other four E–I scales than with any of the S–N, T–F or J–P facet scales.

There is only one scale that correlates **significantly** more highly with scales in other dichotomies. The T–F scale Questioning–Accommodating correlates with a number of scales across the S-N dichotomy. It correlates negatively at between –0.21 and –0.37 with all of the S–N scales except Realistic–Imaginative. By way of comparison, it correlates at between 0.21 and 0.30 with the other four scales on the same dichotomy as itself (T–F). The negative correlations between Questioning–Accommodating and the S–N scales are consistent with findings from the US version of the Step II instrument, and suggest that a questioning approach to differences of opinion seems to be related to a range of iNtuition (N) facets. These patterns are similar to those found when the European Step II instrument was developed.

Table 4.3: Intercorrelations of Step II facet scales

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Initiating–Receiving	**1**																			
2. Expressive–Contained	**0.51**	**1**																		
3. Gregarious–Intimate	**0.56**	**0.46**	**1**																	
4. Active–Reflective	**0.70**	**0.50**	**0.56**	**1**																
5. Enthusiastic–Quiet	**0.57**	**0.54**	**0.58**	**0.61**	**1**															
6. Concrete–Abstract	-0.11	-0.11	-0.06	-0.10	-0.21	**1**														
7. Realistic–Imaginative	-0.09	-0.12	-0.08	-0.11	-0.23	**0.66**	**1**													
8. Practical–Conceptual	-0.09	-0.06	-0.04	-0.09	-0.18	**0.63**	**0.60**	**1**												
9. Experiential–Theoretical	-0.03	-0.03	0.02	-0.02	-0.06	**0.56**	**0.46**	**0.52**	**1**											
10. Traditional–Original	-0.17	-0.12	-0.10	-0.19	-0.29	**0.62**	**0.57**	**0.57**	**0.39**	**1**										
11. Logical–Empathetic	-0.08	-0.23	-0.06	-0.04	-0.14	0.20	0.18	0.01	0.03	0.06	**1**									
12. Reasonable–Compassionate	-0.03	-0.16	-0.06	0.01	-0.07	0.15	0.13	-0.03	0.04	-0.01	**0.70**	**1**								
13. Questioning–Accommodating	0.13	0.05	0.09	0.17	0.19	-0.25	-0.20	-0.31	-0.21	-0.37	**0.21**	**0.26**	**1**							
14. Critical–Accepting	-0.18	-0.23	-0.17	-0.12	-0.16	0.14	0.13	0.01	0.03	0.02	**0.50**	**0.52**	**0.30**	**1**						
15. Tough–Tender	0.07	-0.10	0.02	0.11	0.07	0.08	0.04	-0.08	-0.00	-0.11	**0.57**	**0.60**	**0.30**	**0.47**	**1**					
16. Systematic–Casual	-0.15	-0.20	-0.17	-0.16	-0.28	0.45	0.41	0.27	0.20	0.49	0.38	0.31	-0.08	0.27	0.22	**1**				
17. Planful–Open-Ended	-0.04	-0.02	-0.03	-0.06	-0.11	0.36	0.30	0.25	0.19	0.47	0.13	0.08	-0.16	0.06	0.01	**0.59**	**1**			
18. Early Starting–Pressure-Prompted	-0.09	-0.06	-0.11	-0.11	-0.19	0.30	0.28	0.24	0.18	0.42	0.04	0.00	-0.21	-0.02	-0.07	**0.49**	**0.51**	**1**		
19. Scheduled–Spontaneous	-0.04	-0.02	-0.05	-0.06	-0.14	0.37	0.31	0.27	0.21	0.49	0.12	0.08	-0.19	0.03	0.00	**0.61**	**0.68**	**0.62**	**1**	
20. Methodical–Emergent	0.01	-0.03	-0.05	-0.02	-0.09	0.26	0.22	0.17	0.14	0.30	0.11	0.10	-0.09	0.04	0.03	**0.52**	**0.51**	**0.50**	**0.61**	**1**

Correlations of Step II facet scales with Step I scales

Correlations between Step II facet scales and the continuous scores from the MBTI Step I™ instrument are shown in Table 4.4.[5] Positive correlations between facet scales and the corresponding Step I dimension (e.g. between Initiating–Receiving and the E–I continuous score) indicate relationships in the expected direction, with higher facet scores tending to be associated with higher continuous scores, and vice versa. Negative correlations would indicate an inverse relationship between the facet scale scores and the corresponding Step I dimension.

Table 4.4: Correlations of Step II facet scales with Step I continuous scores

	Step I continuous score			
Step II facet scales	**E–I**	**S–N**	**T–F**	**J–P**
E–I facet scales				
Initiating–Receiving	**0.83**	–0.12	–0.09	–0.05
Expressive–Contained	**0.71**	–0.11	–0.25	–0.04
Gregarious–Intimate	**0.68**	–0.07	–0.09	–0.07
Active–Reflective	**0.82**	–0.13	–0.04	–0.08
Enthusiastic–Quiet	**0.79**	–0.24	–0.14	–0.17
S–N facet scales				
Concrete–Abstract	–0.13	**0.84**	0.20	0.43
Realistic–Imaginative	–0.14	**0.78**	0.17	0.37
Practical–Conceptual	–0.11	**0.72**	0.01	0.31
Experiential–Theoretical	–0.02	**0.64**	0.04	0.23
Traditional–Original	–0.20	**0.74**	0.03	0.55
T–F facet scales				
Logical–Empathetic	–0.13	0.14	**0.88**	0.15
Reasonable–Compassionate	–0.06	0.09	**0.78**	0.10
Questioning–Accommodating	0.18	–0.32	**0.33**	–0.20
Critical–Accepting	–0.18	0.09	**0.66**	0.06
Tough–Tender	0.07	0.00	**0.72**	0.01
J–P facet scales				
Systematic–Casual	–0.21	0.44	0.39	**0.73**
Planful–Open-Ended	–0.05	0.37	0.11	**0.84**
Early Starting–Pressure-Prompted	–0.13	0.33	0.02	**0.75**
Scheduled–Spontaneous	–0.06	0.39	0.10	**0.84**
Methodical–Emergent	0.02	0.26	0.11	**0.69**

The E–I facet scales correlate at a level of 0.68 to 0.83 with the E–I continuous scores from Step I; the S–N facet scales correlate at 0.64 to 0.84 with the S–N continuous scores; the T–F facet scales correlate at 0.33 to 0.88 with the T–F continuous scores; and the J–P facet

[5] Continuous scores place an individual's score on each dimension onto a continuous scale with a midpoint of 100. To calculate continuous scores, Preference Clarity Index (PCI) scores for each dimension are either subtracted or added to 100, depending on which direction the overall preference is. PCI scores in the direction of E, S, T or J are subtracted from 100. PCI scores in the direction of I, N, F or P are added to 100.

scales correlate at 0.69 to 0.84 with the J–P continuous scores. The scale that showed the lowest correlation with its associated Step I continuous score was Questioning–Accommodating at 0.33. This scale is also lowest on internal consistency (see page 72), and has been previously found to have the lowest test–retest reliability (Quenk, Hammer and Majors, 2004), which would result in the true correlations being underestimated.

These correlations are very similar to those found during the development of the Step II questionnaire. This consistency, alongside the fact that the correlations between Step II facet scales and Step I continuous scores associated with their own dimension are substantially higher than correlations with the other three dimensions, provides compelling evidence for the theoretical hierarchical structure of the Step II facet scales in relation to the Step I scales.

Out-of-preference scores

It is known that the five facet scales relating to each type dimension do not describe the dimension in its entirety; there will not be a precise and exact overlap between, for example, an individual's score on E–I and their total score across the five facet scales that relate to this dimension. For example, it is not uncommon to see an Enthusiastic Introvert or an Intimate Extravert. Such apparent inconsistencies are usually known as 'out-of-preference scores' (OOPS).

Although it is usual to have a number of OOPS in any one profile, it should be unusual to find that OOPS make up more than half of the facet scores relating to any one type dichotomy. Table 4.5 demonstrates that the proportion of individuals for whom this happens ranges from 2.7% for the S–N block to 1.0% for the E–I block. It is therefore a very infrequent occurrence.

Table 4.5: Proportion of OOPS by type dichotomy[6]

	Proportion of 'reported type' OOPS					
	None	One	Two	Three	Four	Five
E–I	70.3%	22.7%	6.0%	0.9%	0.1%	0.0%
S–N	65.5%	24.1%	7.7%	2.4%	0.3%	0.0%
T–F	56.5%	31.6%	9.6%	2.0%	0.2%	0.0%
J–P	60.7%	28.6%	9.0%	1.6%	0.1%	0.0%

For any individual facet scale, an OOPS tends to occur in approximately 10% of profiles, with the exception of Questioning–Accommodating (23%) and Critical-Accepting (17%).

[6] In this table, an OOPS is defined as a score of 2, 3, 4 or 5 on the 'wrong' pole of the facet when compared with the reported type.

Reliability

The reliability of a test or questionnaire relates to how consistent and precise it is. Internal consistency reliability addresses the question of whether all the questions in a scale measure the same construct. For example, are the Step II facet scales consistent within themselves, and do they hold together well as scales? A common measure of internal consistency reliability is coefficient alpha (Cronbach, 1951). The alpha coefficients for the Step II facet scales are shown in Table 4.6.

Table 4.6: Internal consistency reliability

Step II facet scale	No. of items	Coefficient alpha
E–I facet scales		
Initiating–Receiving	8	0.82
Expressive–Contained	7	0.77
Gregarious–Intimate	7	0.66
Active–Reflective	8	0.74
Enthusiastic–Quiet	9	0.75
S–N facet scales		
Concrete–Abstract	9	0.73
Realistic–Imaginative	7	0.73
Practical–Conceptual	8	0.54
Experiential–Theoretical	8	0.65
Traditional–Original	8	0.77
T–F facet scales		
Logical–Empathetic	9	0.81
Reasonable–Compassionate	8	0.69
Questioning–Accommodating	7	0.48
Critical–Accepting	8	0.57
Tough–Tender	8	0.73
J–P facet scales		
Systematic–Casual	8	0.74
Planful–Open-Ended	6	0.79
Early Starting–Pressure-Prompted	6	0.78
Scheduled–Spontaneous	8	0.76
Methodical–Emergent	8	0.64
	Median	0.73

The internal consistency reliability of most scales is good, and the average (median) reliability is above 0.7. However, three scales (Practical–Conceptual, Questioning–Accommodating and Critical–Accepting) do have lower alpha coefficients. This pattern is consistent with what was found during the development of the Step II instrument, where these three facet scales were also found to show lower reliability than the others (Quenk, Hammer and Majors, 2004).

It should be noted, however, that coefficient alpha reliability statistics will give an underestimate of the reliability of the Step II facet scales,

given the use of item response theory (IRT) methods in the actual scoring process.[7]

Group differences

This section looks at the way in which people from different groups responded to the Dutch version of the European Step II questionnaire.

Gender

The means and standard deviations of the Step II facet scales are shown separately for males and females in Table 4.7, along with the difference in mean scores. This is represented graphically in Figure 4.2. Statistically significant differences were found between male and female mean scores for 17 of the 20 facet scales, with some consistent patterns emerging.

- On the E–I facet scales, all five mean scores tended slightly toward the E pole for both males and females. There were significant gender differences for four of the five facet scales, with male mean scores tending more toward the E pole than females on two of these facet scales and female mean scores tending more towards the E pole than males on the other two facets. There was therefore no consistent pattern across facet scales of either males or females tending further towards the E pole. There was no statistically significant gender difference on the remaining facet scale.

- On the S–N facet scales, all five mean scores tended slightly toward the S pole for both males and females. There were statistically significant gender differences on four of the five scales, with female mean scores tending more towards the S pole than males. There was no statistically significant gender difference on the remaining facet scale.

- On the T–F facet scales, three mean scores tended toward the T pole and one towards the F pole for both males and females. Mean scores on the other facet (Questioning–Accommodating) tended towards the T pole for males and slightly towards the F pole for females. There were statistically significant gender differences on all five facets, with males mean scores significantly further toward the T pole than females.

- On the J–P facet scales, four of the five mean scores tended toward the J pole for both males and females, though only three of these were the same for both genders. There were statistically significant gender differences on four of the five facet scales, with female mean scores tending more towards the J pole than males on three

[7] Item response theory (IRT) is an approach to measurement that is concerned with modelling the relationship between item responses and the underlying characteristic assessed by the scale or test the item is designed to measure. IRT can be used to select items for a test and/or to score the items.

of these. There was no statistically significant gender difference on the remaining facet scale.

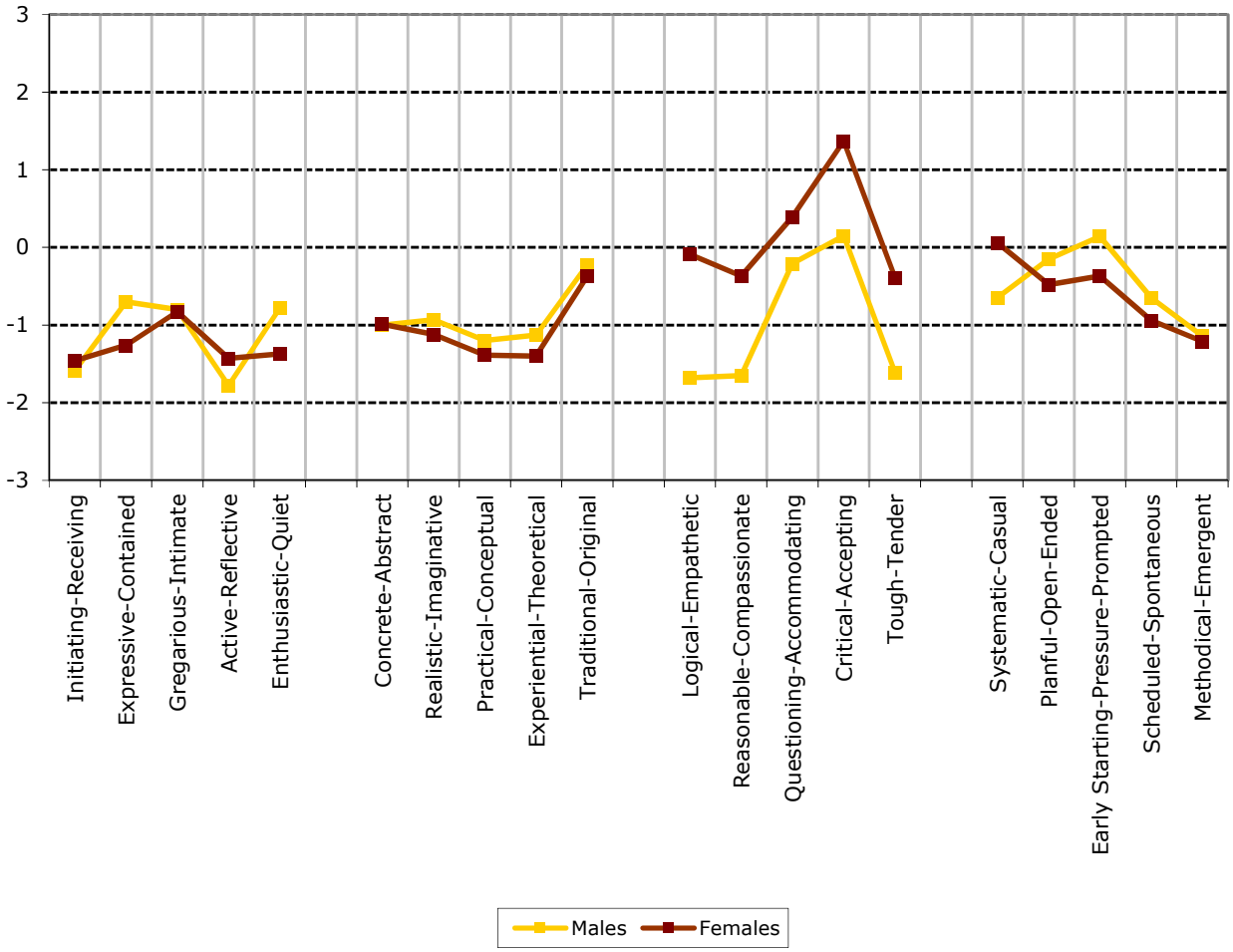
Table 4.7: Gender differences in facet scale scores

	Males (n=11,383)		Females (n=7,969)		Difference (M–F)[8]
Step II facet scale	**Mean**	**SD**	**Mean**	**SD**	
E–I facet scales					
Initiating–Receiving	-1.61	3.15	-1.46	3.16	-0.15**
Expressive–Contained	-0.70	3.08	-1.26	2.98	0.57**
Gregarious–Intimate	-0.80	2.82	-0.83	2.70	0.04
Active–Reflective	-1.78	2.94	-1.43	2.95	–0.35**
Enthusiastic–Quiet	-0.78	3.16	-1.37	3.00	0.59**
S–N facet scales					
Concrete–Abstract	-1.00	2.64	-0.99	2.57	–0.02
Realistic–Imaginative	-0.93	3.04	-1.12	3.05	0.19**
Practical–Conceptual	-1.20	2.70	-1.39	2.77	0.19**
Experiential–Theoretical	-1.13	2.62	-1.40	2.38	0.27**
Traditional–Original	-0.22	3.16	-0.37	3.21	0.14**
T–F facet scales					
Logical–Empathetic	-1.68	2.66	-0.09	2.86	–1.59**
Reasonable–Compassionate	-1.65	2.75	-0.37	2.80	–1.28**
Questioning–Accommodating	-0.21	2.60	0.39	2.65	–0.60**
Critical–Accepting	0.15	2.75	1.37	2.75	–1.22**
Tough–Tender	-1.61	2.52	-0.39	2.55	–1.22**
J–P facet scales					
Systematic–Casual	-0.65	2.90	0.06	2.96	–0.70**
Planful–Open-Ended	-0.15	3.44	-0.48	3.49	0.34**
Early Starting–Pressure-Prompted	0.14	3.58	-0.37	3.64	0.51**
Scheduled–Spontaneous	-0.65	3.16	-0.94	3.20	0.29**
Methodical–Emergent	-1.14	2.94	-1.21	2.93	0.08

*Difference significant at: p<0.05, **p<0.01 (based on an independent samples t-test).

[8] A positive value indicates that male scores tend more towards I, N, F or P, and a negative value indicates that female scores tend more towards I, N, F or P.

Figure 4.2: Gender differences in facet scale scores



Age

There were found to be no meaningful links between age and facet scale scores. Although correlational analysis showed several facet scales to be significantly correlated with age, the significance levels were more the result of the very large sample size rather than being indicative of a meaningful relationship. The highest correlations were 0.15 (Gregarious–Intimate) and 0.13 (Questioning–Accommodating), and even these are too small to be considered meaningful.

For interpretation purposes, it is reasonable to conclude that there are no clear relationships between age and facet scale scores.

Ethnic origin

Ethnic origin information was not captured for people who completed the Dutch language version of the questionnaire, so no analyses were conducted.

Occupational level

Research using the Dutch version of the MBTI Step I questionnaire has demonstrated that individuals in higher-level jobs in organisations are

more likely to have preferences for Extraversion, iNtuition and (to some extent) Thinking than those in lower-level jobs.[9]

Table 4.8 shows the facet scale means and standard deviation for different occupational levels. The findings can be summarised as follows:

- A pattern was found regarding the facets relating to the E–I dimension, with mean scores of those at more senior levels tending further towards the E pole than those for the lower occupational groups. Those at top level were significantly more Initiating, Expressive, Active and Enthusiastic than those at all other levels.

- A pattern was found regarding the facets relating to the S–N dimension, with mean scores of those at the top level considerably further towards the N pole (or less towards the S pole) than those for the lower occupational level groups. No clear patterns were found between groups below the top level.

- A consistent pattern was found regarding the facets relating to the T–F dimension. The employee group mean scores tended further towards the F pole (or less towards the T pole) than any other groups. No clear patterns were found between groups above employee level. This could be a reflection of the fact that females formed a considerably higher proportion of the employee group than they did of other groups.

- An interesting pattern was found regarding the facets relating to the J–P dimension. The mean scores of all groups except employee level followed a similar pattern, with the Early Starting–Pressure Prompted facet score tending considerably more towards the P pole than the other four facets. The mean scores for all facets of the top level group also tended more towards the P pole (or less towards the J pole) than the other four groups. With the employee group however, the Early Starting–Pressure Prompted facet score tended significantly more towards the J pole than those at all other levels.

[9] See the MBTI Step I European Data Supplement for details.

Table 4.8: Mean facet scale scores by occupational level

Step II facet scale	Top level (n=771)		Senior Executive (n=3,699)		Upper Middle Management (n=1,403)		Middle Management (n=2,191)		First Level Mgt/ Supervisory (n=1,044)		Employee (n=4,797)		Sig.
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales													
Initiating–Receiving	-2.39	2.86	-1.93	3.03	-1.84	3.04	-1.74	3.11	-1.26	3.22	-1.13	3.27	**
Expressive–Contained	-1.48	2.99	-1.12	3.05	-1.15	3.04	-0.97	3.07	-0.84	3.13	-0.78	3.05	**
Gregarious–Intimate	-1.11	2.76	-0.92	2.74	-0.80	2.77	-0.88	2.74	-1.02	2.73	-0.70	2.80	**
Active–Reflective	-2.51	2.62	-1.90	2.86	-2.03	2.85	-1.87	2.84	-1.52	3.07	-1.19	3.07	**
Enthusiastic–Quiet	-1.91	2.88	-1.20	3.08	-1.39	3.06	-1.12	3.08	-0.85	3.14	-0.71	3.15	**
S–N facet scales													
Concrete–Abstract	-0.09	2.68	-1.07	2.61	-0.78	2.57	-1.17	2.52	-1.34	2.53	-1.21	2.56	**
Realistic–Imaginative	0.07	3.06	-1.10	3.04	-0.81	2.98	-1.18	3.00	-1.31	2.96	-1.26	2.96	**
Practical–Conceptual	-0.23	2.68	-1.33	2.66	-0.93	2.67	-1.40	2.67	-1.62	2.63	-1.64	2.71	**
Experiential–Theoretical	-0.39	2.69	-1.29	2.48	-1.10	2.58	-1.39	2.48	-1.51	2.55	-1.42	2.42	**
Traditional–Original	1.10	3.08	-0.24	3.14	0.07	3.10	-0.37	3.19	-0.71	3.17	-0.76	3.13	**
T–F facet scales													
Logical–Empathetic	-1.32	2.66	-1.38	2.71	-1.57	2.65	-1.38	2.74	-1.63	2.77	-0.39	2.96	**
Reasonable–Compassionate	-1.46	2.73	-1.52	2.71	-1.74	2.68	-1.47	2.71	-1.70	2.76	-0.51	2.91	**
Questioning–Accommodating	-0.27	2.37	-0.03	2.44	-0.13	2.52	-0.04	2.57	-0.16	2.75	0.38	2.75	**
Critical–Accepting	0.67	2.84	0.61	2.77	0.47	2.78	0.44	2.74	0.09	2.77	0.88	2.87	**
Tough–Tender	-1.80	2.23	-1.46	2.47	-1.66	2.35	-1.30	2.53	-1.30	2.65	-0.52	2.77	**
J–P facet scales													
Systematic–Casual	-0.18	2.80	-0.67	2.85	-0.60	2.94	-0.53	2.97	-0.68	2.95	-0.15	2.99	**
Planful–Open-Ended	0.03	3.50	-0.48	3.44	-0.35	3.43	-0.31	3.47	-0.49	3.43	-0.33	3.49	**
Early Starting–Pressure-Prompted	0.93	3.48	0.10	3.59	0.46	3.55	0.11	3.59	0.03	3.64	-0.53	3.60	**
Scheduled–Spontaneous	-0.36	3.17	-0.97	3.10	-0.68	3.15	-0.81	3.15	-0.86	3.16	-0.84	3.17	**
Methodical–Emergent	-0.93	2.96	-1.40	2.85	-1.13	2.97	-1.29	2.92	-1.29	2.91	-1.08	2.97	**

Significant at: *p<0.05, **p<0.01 (based on a one-way analysis of variance).

Education

Specific educational qualifications were not collected for the OPPassessment sample; however, the age at which individuals left full-time education was. No significant and meaningful correlations were found between the age at which people left full-time education and their facet scale scores. All the correlations were less than 0.1.

Work area

Information regarding the area of work people engage in was collected for the group. Many different categories were used, but for the purposes of analysis the focus was on the five most commonly occurring. These were as follows:

- Finance
- HR, training, guidance
- Science, engineering
- Sales, customer service
- IT.

Table 4.9 shows the facet scale means and standard deviation for these five work areas. The findings can be summarised as follows:

- A pattern was found regarding the facets relating to the E–I dimensions. Although mean scores for all groups tended towards the E pole for each facet scale, scores were further towards the E pole for the 'Sales, customer service' group than for the 'Science and Engineering' and 'IT' groups. Mean scores for the 'HR, training, guidance' and 'Finance' groups fell in between.
- On the facets relating to the S–N dimension, mean scores for all groups generally tended towards the S pole except 'HR, training, guidance'. Four mean scores for this group tended towards the S pole and one (Traditional–Original) tended towards the N pole. The four scores for this group that did tend towards the S pole were less far towards this pole than the scores of the other groups.
- The most noticeable pattern regarding the facets in the T–F preference block was that although mean scores for all groups tended towards the T pole on three out of the five facet scales, scores for all of the work area groups tended towards the F pole or mid-point for the Questioning–Accommodating and Critical–Accepting facets scale. It was also observed that the mean scores across facet scales tended more towards the F pole (or less towards the T pole) for the 'HR, training, guidance' group than for the other groups.
- A pattern was also found regarding the facets for the J–P dimensions. The mean scores for the 'Science, engineering' group

were further towards the J pole than the scores of the other four groups, while the mean scores for the 'HR, training, guidance' group were further towards the P pole than the scores of the other four groups.

Table 4.9: Mean facet scale scores by work area

Step II facet scale	Finance (n=2,091)		HR, training, guidance (n=1,976)		Science, engineering (n=1,473)		Sales, customer service (n=1,421)		IT (n=1,283)		Sig.
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales											
Initiating–Receiving	-1.79	3.10	-1.74	3.04	-1.24	3.20	-2.44	2.84	-1.00	3.35	**
Expressive–Contained	-1.02	3.08	-1.52	3.04	-0.55	3.10	-1.47	2.88	-0.46	3.12	**
Gregarious–Intimate	-1.04	2.67	-0.85	2.83	-0.67	2.78	-1.50	2.63	-0.43	2.82	**
Active–Reflective	-1.94	2.88	-1.87	2.86	-1.34	3.03	-2.48	2.72	-1.13	3.13	**
Enthusiastic–Quiet	-1.13	3.12	-1.55	3.03	-0.39	3.09	-1.81	3.00	-0.44	3.22	**
S–N facet scales											
Concrete–Abstract	-1.40	2.54	-0.22	2.75	-1.51	2.42	-1.27	2.45	-0.89	2.59	**
Realistic–Imaginative	-1.57	2.86	-0.19	3.29	-1.61	2.76	-1.05	2.96	-1.03	3.02	**
Practical–Conceptual	-1.65	2.65	-0.78	2.85	-1.64	2.60	-1.53	2.52	-1.11	2.71	**
Experiential–Theoretical	-1.57	2.54	-0.71	2.67	-1.30	2.53	-1.67	2.25	-1.12	2.61	**
Traditional–Original	-0.54	3.11	0.50	3.27	-0.89	3.08	-0.42	3.13	-0.39	3.14	**
T–F facet scales											
Logical–Empathetic	-1.35	2.71	-0.11	2.92	-2.11	2.49	-0.98	2.85	-1.84	2.63	**
Reasonable–Compassionate	-1.44	2.72	-0.34	2.85	-1.93	2.64	-1.23	2.79	-1.93	2.65	**
Questioning–Accommodating	-0.02	2.58	0.27	2.67	0.07	2.55	-0.01	2.56	-0.19	2.62	**
Critical–Accepting	0.42	2.81	1.46	2.85	-0.17	2.66	0.78	2.82	-0.07	2.76	**
Tough–Tender	-1.51	2.48	-0.67	2.51	-1.67	2.56	-1.17	2.62	-1.47	2.56	**
J–P facet scales											
Systematic–Casual	-0.72	2.88	0.57	2.98	-1.29	2.75	-0.37	2.94	-0.67	2.90	**
Planful–Open-Ended	-0.36	3.48	0.25	3.68	-0.77	3.24	-0.56	3.44	-0.17	3.42	**
Early Starting–Pressure-Prompted	-0.06	3.64	0.61	3.74	-0.46	3.50	-0.09	3.56	0.15	3.61	**
Scheduled–Spontaneous	-0.85	3.15	-0.20	3.45	-1.28	2.98	-0.85	3.12	-0.67	3.16	**
Methodical–Emergent	-1.14	2.88	-0.51	3.10	-1.65	2.68	-1.36	3.01	-1.07	2.93	**

Significant at: *p<0.05, **p<0.01 (based on a one-way analysis of variance).

Nationality

Information on nationality was available for the group. Fifty per cent of the group were Belgian and 49% were Dutch.

A comparison of mean facet scales across the Dutch and Belgian groups highlighted some interesting differences, as shown in Table 4.10 and Figure 4.3, although in most cases these were too small to be of any practical significance.

- On all five facet scales linked to the E–I dimension, the Dutch group mean scores were significantly further towards the E pole than those for the Belgian group.
- On all five of the facet scales linked to the S–N dimension, the Belgian group mean scores were significantly further towards the S pole than those for the Dutch group.
- On four of the five facet scales linked to the T–F dimension, the Belgian group mean scores were significantly further towards the T pole than those for the Dutch group. The exception to this was the Tough–Tender facet scale, where the Dutch group mean score tended more towards the T pole.
- On all five of the facet scales linked to the J–P dimension, the Belgian group mean scores were significantly further towards the J pole than those for the Dutch group.

These differences are consistent with the patterns found when the MBTI Step I data were analysed, which showed the Dutch group to be significantly more likely to have preferences for Extraversion, iNtuition, Feeling and Perceiving than the Belgian group.[10]

[10] See the MBTI Step I European Data Supplement for details.

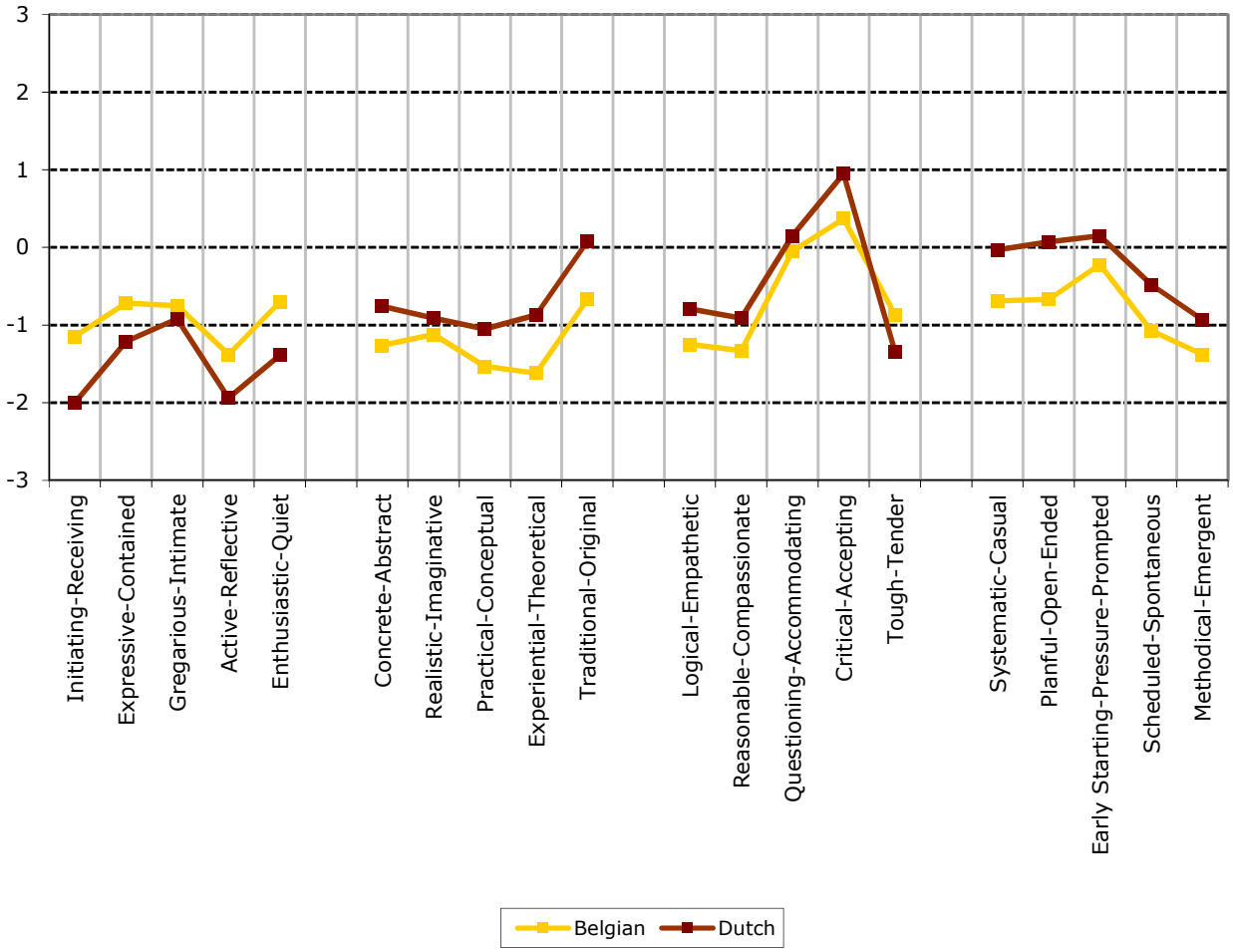
Table 4.10: Mean facet scale scores by nationality

Step II facet scale	Belgian (n=8,880)		Dutch (n=8,851)		Difference (B–D)[11]
	Mean	SD	Mean	SD	
E–I facet scales					
Initiating–Receiving	-1.15	3.24	-2.00	2.99	0.85**
Expressive–Contained	-0.72	3.11	-1.21	2.97	0.49**
Gregarious–Intimate	-0.75	2.82	-0.92	2.72	0.16**
Active–Reflective	-1.38	3.05	-1.94	2.82	0.55**
Enthusiastic–Quiet	-0.70	3.16	-1.39	3.02	0.70**
S–N facet scales					
Concrete–Abstract	-1.26	2.61	-0.76	2.58	–0.51**
Realistic–Imaginative	-1.12	3.04	-0.91	3.03	–0.22**
Practical–Conceptual	-1.53	2.67	-1.05	2.75	–0.48**
Experiential–Theoretical	-1.62	2.42	-0.87	2.57	–0.75**
Traditional–Original	-0.67	3.21	0.08	3.10	–0.75**
T–F facet scales					
Logical–Empathetic	-1.25	2.94	-0.79	2.74	–0.46**
Reasonable–Compassionate	-1.33	2.90	-0.91	2.76	–0.42**
Questioning–Accommodating	-0.05	2.71	0.15	2.55	–0.20**
Critical–Accepting	0.37	2.79	0.95	2.81	–0.58**
Tough–Tender	-0.87	2.78	-1.34	2.40	0.47**
J–P facet scales					
Systematic–Casual	-0.69	3.01	-0.03	2.85	–0.66**
Planful–Open-Ended	-0.67	3.48	0.07	3.43	–0.75**
Early Starting–Pressure-Prompted	-0.22	3.68	0.15	3.54	–0.37**
Scheduled–Spontaneous	-1.07	3.17	-0.48	3.15	–0.60**
Methodical–Emergent	-1.38	2.96	-0.93	2.90	–0.45**

Significant at: *p<0.05, **p<0.01 (based on an independent samples t-test).

[11] A positive value indicates that Belgian scores tend more towards I, N, F or P, and a negative value indicates that Dutch scores tend more towards I, N, F or P.

Figure 4.3: Mean facet scale scores by nationality



Employment status

Employment status information was available for the OPPassessment sample. The vast majority of the group worked either full-time or part-time, or were self-employed. Table 4.11 shows the mean facet scale scores for each group, with the data illustrated in graphical form in Figure 4.4.

The analyses showed statistically significant differences across the groups on all the facet scales on the E-I, S–N, T–F and J–P dimensions. The clearest patterns were as follows:

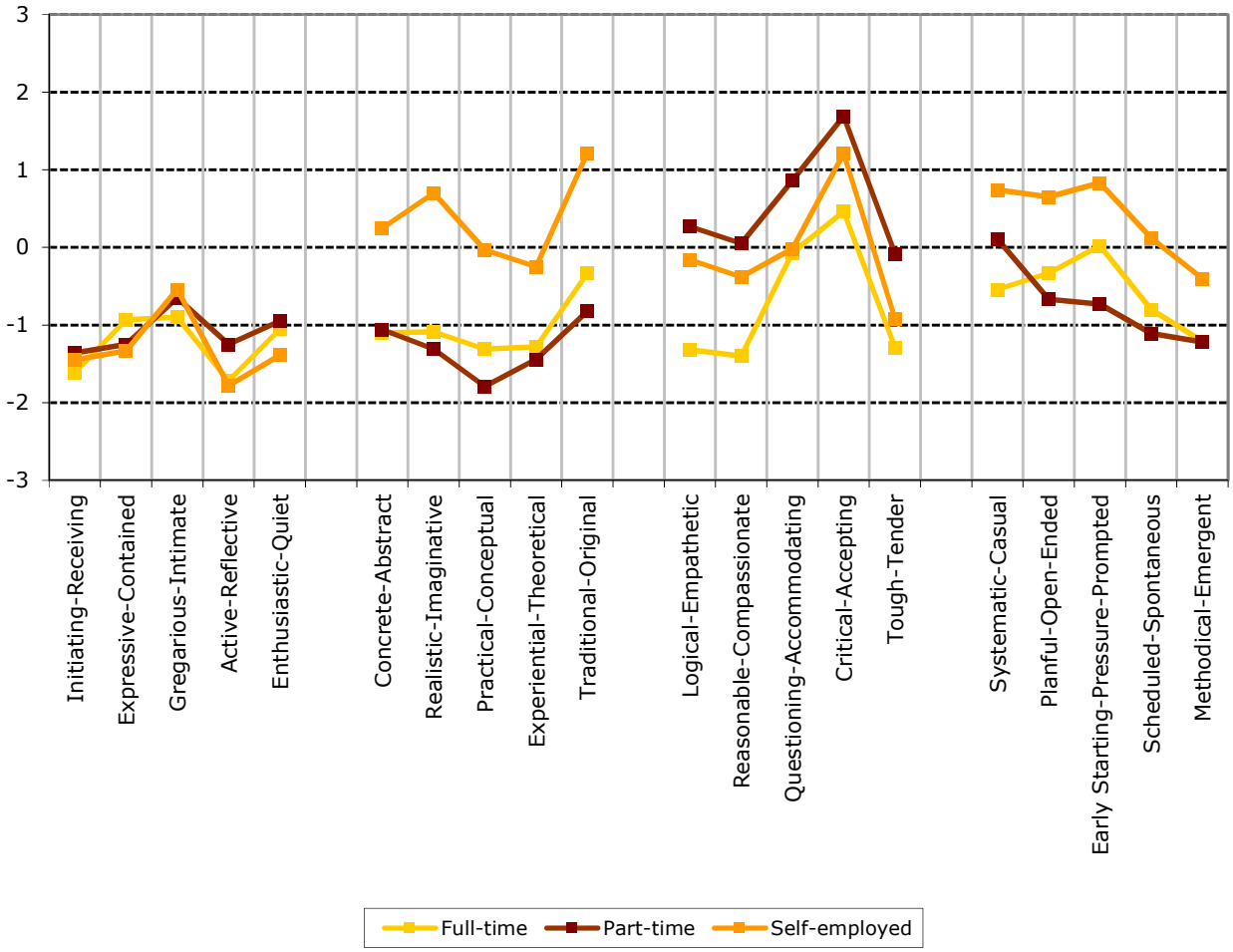
- The self-employed group tended to score more towards the N pole on the S–N facet scales than the other two groups, and less towards the J pole on the J–P facet scales.
- The full-time group tended to score more towards the T pole on the T–F facets than the other two groups, particularly the part-time group. This is likely to be at least partly a gender effect; 83% of part-time workers were female, compared with 41% of the total group and 34% of full-time workers.

Table 4.11: Mean facet scale scores by employment status

Step II facet scale	Full-time (n=12,769)		Part-time (n=2,012)		Self-employed (n=869)		Sig.
	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales							
Initiating–Receiving	-1.62	3.15	-1.36	3.13	-1.45	3.14	**
Expressive–Contained	-0.93	3.06	-1.25	2.98	-1.33	3.11	**
Gregarious–Intimate	-0.90	2.77	-0.65	2.62	-0.54	2.99	**
Active–Reflective	-1.72	2.96	-1.25	2.88	-1.78	2.84	**
Enthusiastic–Quiet	-1.05	3.13	-0.95	3.05	-1.39	3.02	**
S–N facet scales							
Concrete–Abstract	-1.10	2.59	-1.06	2.51	0.25	2.77	**
Realistic–Imaginative	-1.09	3.00	-1.31	3.04	0.70	3.21	**
Practical–Conceptual	-1.31	2.70	-1.79	2.71	-0.03	2.74	**
Experiential–Theoretical	-1.28	2.53	-1.44	2.31	-0.25	2.70	**
Traditional–Original	-0.33	3.18	-0.82	3.07	1.22	3.13	**
T–F facet scales							
Logical–Empathetic	-1.32	2.79	0.27	2.80	-0.16	2.89	**
Reasonable–Compassionate	-1.40	2.79	0.05	2.73	-0.38	2.90	**
Questioning–Accommodating	-0.07	2.62	0.86	2.52	-0.02	2.71	**
Critical–Accepting	0.46	2.80	1.69	2.66	1.21	2.92	**
Tough–Tender	-1.29	2.59	-0.08	2.58	-0.92	2.50	**
J–P facet scales							
Systematic–Casual	-0.54	2.91	0.11	3.00	0.74	2.97	**
Planful–Open-Ended	-0.33	3.47	-0.67	3.40	0.65	3.61	**
Early Starting–Pressure-Prompted	0.02	3.60	-0.73	3.56	0.83	3.67	**
Scheduled–Spontaneous	-0.80	3.15	-1.11	3.15	0.12	3.45	**
Methodical–Emergent	-1.21	2.93	-1.22	2.88	-0.40	3.10	**

Significant at: *p<0.05, **p<0.01 (based on a one-way analysis of variance).

Figure 4.4: Mean facet scale scores by employment status



Appendix 1: Sample description

Sample 1: Data from OPPassessment (representative Dutch-speaking professional and managerial sample)

This sample consists of 19,354 individuals who completed the MBTI Step II instrument in Dutch via the OPPassessment system between June 2006 and July 2016. 59% of the respondents were male and 41% were female. Age ranged from 16 to 78 years, with a mean of 40 and a median of 40.

Nationality was disclosed by 93% of respondents. Of these, 49.5% were Belgian and 49.3% were Dutch. Many other nationalities were represented, but each one formed less than 1% of the total group:

Nationality	Percentage
Belgian	49.5%
Dutch	49.3%
Other	1.2%

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

Employment status	Percentage
Full-time	80.9%
Part-time	12.7%
Self-employed	5.5%
Unemployed	0.7%
Homemaker	0.1%
Retired	0.1%

The majority of the group were of managerial level or above, but with the largest single group being employee (32%):

Occupational level	Percentage
Top level	5.1%
Senior executive	24.7%
Upper middle management	9.4%
Middle management	14.6%
First level management/supervisor	7.0%
Employee	32.0%
Other	7.2%

A range of work areas were represented:

Work area (job type)	Percentage
Finance	13.7%
HR, training, guidance	12.9%
Science, engineering	9.7%
Sales, customer service	9.3%
IT	8.4%
Business services	8.3%
Admin or secretarial	6.2%
Research and development	3.5%
Health, social services, etc.	3.0%
Education	1.5%
Land, sea or air transport	0.9%
Military, police, prison, fire	0.8%
Skilled operative	0.5%
Leisure, personal service	0.3%
Unskilled operative	0.1%
Other public sector	5.2%
Other private sector	3.0%
Other	12.8%

MBTI® Step II™ instrument

European Data Supplement

French

January 2018

2nd edition

OPP Ltd
Elsfield Hall
15–17 Elsfield Way
Oxford OX2 8EP
UK
t: +44 (0)1865 404 500

Contents

Introduction

Data collected for the European MBTI® Step II™ instrument were analysed to produce the findings in this supplement. This is the second data supplement produced for this instrument. A brief description of the sample is given below. Further details of the sample are provided in Appendix 1.

- The sample consisted of 13,610 individuals who completed the MBTI Step II instrument in French via the OPPassessment™ system between 2006 and mid-2016.[1] This sample is considered to be representative of the groups of people with whom the French MBTI Step II instrument has been and will be used for applications such as management development, coaching, counselling and teambuilding. As such, it is likely to represent a cross-section of the European French-speaking professional and managerial population.

The results of the analyses are outlined below.

[1] OPPassessment allows personality questionnaires such as the MBTI instrument to be administered via email and/or completed online.

Type distribution

Type tables are a way of illustrating the proportion of each type within a particular group. Below is a type table taken from the sample described on the previous page.

For each of the 16 different types, the number of cases, the percentage of the total that this represents and the self-selection ratio (SSR) are shown. The SSR (Myers et al., 1998) is a way of demonstrating whether a given type appears more or less often in a particular group than would be expected given its frequency in a reference group. Ideally, the type distribution from a large representative sample of the French population would be used to calculate SSRs in this data supplement. However, such a sample does not currently exist. In its place, SSRs have been calculated using type data from the UK general population (Kendall, 1998), which can be justified by the fact that type distributions for comparable French and British groups, such as managers and professionals, are similar. Evidence (eg Hackston and Kendall, 2004; Quenk et al., 2004) does suggest that although type-related behaviours vary a good deal from country to country and from culture to culture, the frequencies of underlying MBTI types do not.

An SSR of greater than 1 indicates that a type is over-represented, and an SSR of less than 1 denotes that it is under-represented. Asterisks are used to denote whether the over- or under-representations are statistically significant, based on the results of chi-square analysis.[2]

[2] Chi-square analysis (often abbreviated to χ^2) is a technique used to explore whether observed frequency distributions differ significantly from other, predefined, distributions. In this case, the UK general population group is used as the reference group, and the chi-square analysis indicates whether the proportion of people of each type within a particular sample differs significantly from the proportion of people reporting the same type within the reference group.

OPPassessment data (representative European French-speaking professional and managerial sample)

Table 5.1: Type table for OPPassessment data (reported type, N=13,610)

ISTJ	ISFJ	INFJ	INTJ
n=1706 12.5% SSR=0.91	n=617 4.5% SSR=0.35**	n=288 2.1% SSR=1.24	n=675 5.0% SSR=3.57**
ISTP	**ISFP**	**INFP**	**INTP**
n=537 3.9% SSR=0.61**	n=269 2.0% SSR=0.33**	n=485 3.6% SSR=1.13	n=649 4.8% SSR=2.00**
ESTP	**ESFP**	**ENFP**	**ENTP**
n=642 4.7% SSR=0.81	n=455 3.3% SSR=0.38**	n=941 6.9% SSR=1.10	n=989 7.3% SSR=2.61**
ESTJ	**ESFJ**	**ENFJ**	**ENTJ**
n=2522 18.5% SSR=1.78**	n=1006 7.4% SSR=0.59**	n=565 4.2% SSR=1.50**	n=1264 9.3% SSR=3.21**

Type	n	%
E	8,384	61.6%**
I	5,226	38.4%**
S	7,754	57.0%**
N	5,856	43.0%**
T	8,984	66.0%**
F	4,626	34.0%**
J	8,643	63.5%**
P	4,967	36.5%**

*Difference significant at p<0.05, based on chi-square results.
**Difference significant at p<0.01, based on chi-square results.

The most common single type preference is ESTJ (18.5% of the total); this is a common finding with managerial groups in other countries. The SSR results suggest that, in comparison with the UK general population, those with preferences for NT are over-represented and those with preferences for SF are under-represented. Again, this is a common finding with managerial groups.

Properties of the Step II facet scales

Facet scale score distribution

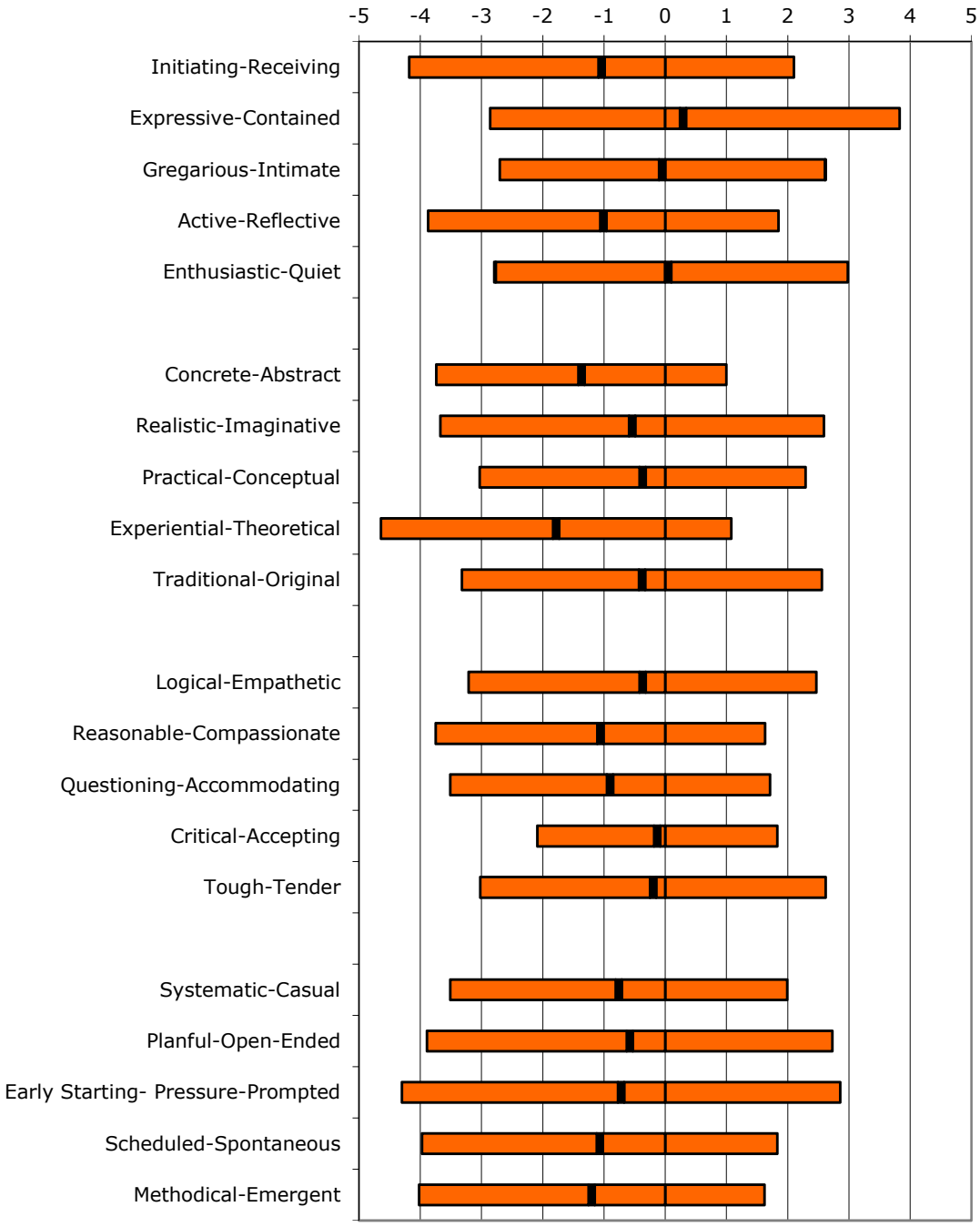
The means and standard deviations of the Step II facets are shown in Table 5.2 and illustrated graphically in Figure 5.1. The mean score for a scale is calculated by adding together the scores of each individual in the sample and then dividing the sum by the number of individuals. Note that negative values indicate that mean scores are nearer the left-hand pole of the facet and positive values indicate that the mean scores are nearer the right-hand pole. The standard deviation (SD) is a statistical measure describing the degree to which the scores from the sample either bunch up close to, or are scattered widely around, the mean for the sample.

Table 5.2: Means and standard deviations of the facet scales

Step II facet scale	Mean[3]	SD
E–I facet scales		
Initiating–Receiving	–1.04	3.14
Expressive–Contained	0.29	3.15
Gregarious–Intimate	–0.04	2.65
Active–Reflective	–1.01	2.86
Enthusiastic–Quiet	0.03	2.80
S–N facet scales		
Concrete–Abstract	–1.37	2.37
Realistic–Imaginative	–0.54	3.13
Practical–Conceptual	–0.37	2.66
Experiential–Theoretical	–1.78	2.86
Traditional–Original	–0.38	2.94
T–F facet scales		
Logical–Empathetic	–0.37	2.84
Reasonable–Compassionate	–1.06	2.69
Questioning–Accommodating	–0.90	2.61
Critical–Accepting	–0.13	1.96
Tough–Tender	–0.20	2.82
J–P facet scales		
Systematic–Casual	–0.76	2.75
Planful–Open-Ended	–0.58	3.31
Early Starting–Pressure-Prompted	–0.72	3.58
Scheduled–Spontaneous	–1.07	2.90
Methodical–Emergent	–1.20	2.82

[3] Scale means: in Step II reports, scores are given from 5 on one pole, through 0, to 5 on the opposite pole. In compiling statistical information, however, one pole needs to be clearly distinguished from the other. Throughout this data supplement, therefore, a convention has been adopted of scores 5, 4, 3, 2 or 1 on the left-hand pole being assigned values of –5, –4, –3, –2, or –1 and scores on the right-hand pole being assigned positive scores. This does not, of course, imply any suggestion that positive scores are 'better' than negative scores.

Figure 5.1: Facet scale mean scores[4]



[4] For each facet scale, the central line shows the mean and the coloured bars on either side indicate the standard deviation.

In general, the means centre quite close to the midpoint of each scale. None of the scales has a mean greater than 2 points from the midpoint in either direction, and each has a standard deviation of between 1.9 and 3.6. However, a clear pattern does emerge. The mean scores tend towards the E, S, T and J direction (with the exception of Expressive–Contained and Enthusiastic–Quiet), which is consistent with the most common four-letter type preference amongst this sample. Amongst the group as a whole, there are more individuals with a preference for E, S, T and J.

Facet scale intercorrelations

Correlations among the Step II facet scales are shown in Table 5.3. Facet scales within each dichotomy usually correlate more highly (often substantially so) with the other scales of the same dichotomy than they do with scales in the other three dichotomies. For example, every E–I facet scale correlates higher with the other four E–I scales than with any of the S–N, T–F or J–P facet scales.

There is only one scale that correlates **significantly** more highly with scales in other dichotomies. The T–F scale Questioning–Accommodating correlates with many scales, across dichotomies. For example, it correlates at a level of 0.15 with both Initiating–Receiving and Enthusiastic–Quiet, and 0.14 with Active–Reflective on the E-I dichotomy. It correlates negatively at a level between -0.14 and -0.30 with all of the scales on the S–N dichotomy. It also negatively correlates at a level of -0.12 with both Planful–Open-Ended and Scheduled–Spontaneous on the J–P dichotomy. By way of comparison, it correlates at between 0.11 and 0.23 with the other four scales on the same dichotomy as itself (T–F). The negative correlations between Questioning–Accommodating and the S–N scales are consistent with findings in the US version of the Step II instrument, and suggest that a questioning approach to differences of opinion seems to be related to a range of iNtuition (N) facets. These patterns are similar to those found when the European Step II instrument was developed.

Table 5.3: Intercorrelations of Step II facet scales

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Initiating–Receiving	1.00																			
2. Expressive–Contained	0.50	1.00																		
3. Gregarious–Intimate	0.55	0.47	1.00																	
4. Active–Reflective	0.72	0.50	0.55	1.00																
5. Enthusiastic–Quiet	0.55	0.54	0.60	0.58	1.00															
6. Concrete–Abstract	-0.02	-0.04	0.00	-0.02	-0.10	1.00														
7. Realistic–Imaginative	-0.05	-0.09	-0.06	-0.05	-0.17	0.63	1.00													
8. Practical–Conceptual	-0.05	-0.02	-0.03	-0.05	-0.12	0.59	0.58	1.00												
9. Experiential–Theoretical	0.03	0.00	0.06	0.04	0.02	0.57	0.44	0.45	1.00											
10. Traditional–Original	-0.12	-0.08	-0.08	-0.12	-0.23	0.56	0.57	0.50	0.32	1.00										
11. Logical–Empathetic	0.00	-0.17	-0.04	0.00	-0.12	0.20	0.26	0.09	0.03	0.16	1.00									
12. Reasonable–Compassionate	0.00	-0.13	-0.06	0.01	-0.06	0.21	0.22	0.07	0.07	0.14	0.64	1.00								
13. Questioning–Accommodating	0.15	0.01	0.07	0.14	0.15	-0.23	-0.17	-0.24	-0.14	-0.30	0.11	0.11	1.00							
14. Critical–Accepting	-0.08	-0.11	-0.11	-0.03	-0.02	0.00	0.03	-0.04	-0.05	-0.09	0.34	0.38	0.23	1.00						
15. Tough–Tender	0.07	-0.09	-0.02	0.06	0.02	0.12	0.13	0.00	0.03	-0.01	0.54	0.53	0.15	0.36	1.00					
16. Systematic–Casual	-0.05	-0.09	-0.08	-0.06	-0.19	0.40	0.41	0.29	0.17	0.48	0.32	0.31	-0.11	0.07	0.19	1.00				
17. Planful–Open-Ended	0.04	0.04	0.04	0.03	-0.05	0.31	0.29	0.22	0.14	0.44	0.16	0.16	-0.12	-0.01	0.08	0.55	1.00			
18. Early Starting–Pressure-Prompted	-0.04	-0.04	-0.03	-0.05	-0.15	0.28	0.27	0.21	0.16	0.38	0.10	0.08	-0.11	-0.08	0.00	0.51	0.50	1.00		
19. Scheduled–Spontaneous	0.02	-0.01	0.00	0.01	-0.09	0.36	0.33	0.25	0.17	0.48	0.21	0.19	-0.12	-0.04	0.10	0.60	0.67	0.56	1.00	
20. Methodical–Emergent	0.05	-0.02	0.00	0.04	-0.08	0.27	0.24	0.18	0.15	0.31	0.17	0.18	-0.06	-0.02	1.00	0.53	0.48	0.49	0.56	1.00

Correlations of Step II facet scales with Step I scales

Correlations between Step II facet scales and the continuous scores from the MBTI Step I™ instrument are shown in Table 5.4.[5] Positive correlations between facet scales and the corresponding Step I dimension (e.g. between Initiating–Receiving and the E–I continuous score) indicate relationships in the expected direction, with higher facet scores tending to be associated with higher continuous scores, and vice versa. Negative correlations would indicate an inverse relationship between the facet scale scores and the corresponding Step I dimension.

Table 5.4: Correlations of Step II facet scales with Step I continuous scores

	Step I continuous score			
Step II facet scales	**E–I**	**S–N**	**T–F**	**J–P**
E–I facet scales				
Initiating–Receiving	**0.82**	–0.05	–0.04	0.03
Expressive–Contained	**0.72**	–0.05	–0.20	–0.01
Gregarious–Intimate	**0.67**	–0.02	–0.12	0.00
Active–Reflective	**0.80**	–0.05	–0.03	0.01
Enthusiastic–Quiet	**0.77**	–0.14	–0.14	–0.11
S–N facet scales				
Concrete–Abstract	–0.04	**0.83**	0.20	0.39
Realistic–Imaginative	–0.09	**0.77**	0.25	0.36
Practical–Conceptual	–0.05	**0.68**	0.07	0.28
Experiential–Theoretical	0.05	**0.64**	0.02	0.20
Traditional–Original	–0.15	**0.70**	0.14	0.50
T–F facet scales				
Logical–Empathetic	–0.07	0.15	**0.85**	0.21
Reasonable–Compassionate	–0.04	0.15	**0.74**	0.21
Questioning–Accommodating	0.15	–0.27	**0.18**	–0.13
Critical–Accepting	–0.06	–0.06	**0.50**	–0.02
Tough–Tender	0.04	0.05	**0.68**	0.11
J–P facet scales				
Systematic–Casual	–0.10	0.42	0.35	**0.72**
Planful–Open-Ended	0.04	0.34	0.17	**0.82**
Early Starting–Pressure-Prompted	–0.07	0.31	0.10	**0.74**
Scheduled–Spontaneous	–0.01	0.38	0.21	**0.80**
Methodical–Emergent	0.02	0.28	0.17	**0.70**

The E–I facet scales correlate at a level of 0.67 to 0.82 with the E–I continuous scores from Step I; the S–N facet scales correlate at 0.64 to 0.83 with the S–N continuous scores; the T–F facet scales correlate at 0.18 to 0.85 with the T–F continuous scores; and the J–P facet

[5] Continuous scores place an individual's score on each dimension onto a continuous scale with a midpoint of 100. To calculate continuous scores, Preference Clarity Index (PCI) scores for each dimension are either subtracted or added to 100, depending on which direction the overall preference is. PCI scores in the direction of E, S, T or J are subtracted from 100. PCI scores in the direction of I, N, F or P are added to 100.

scales correlate at 0.70 to 0.82 with the J–P continuous scores. The scale that showed the lowest correlation with its associated Step I continuous score was Questioning–Accommodating at 0.18. This scale is also lowest on internal consistency (see page 109), and has been previously found to have the lowest test-retest reliability (Quenk, Hammer and Majors, 2004), which would result in the true correlations being underestimated.

These correlations are very similar to those found during the development of the Step II questionnaire. This consistency, alongside the fact that the correlations between Step II facet scales and Step I continuous scores associated with their own dimension are substantially higher than correlations with the other three dimensions, provides compelling evidence for the theoretical hierarchical structure of the Step II facet scales in relation to the Step I scales.

Out-of-preference scores

It is known that the five facet scales relating to each type dimension do not describe the dimension in its entirety; there will not be a precise and exact overlap between, for example, an individual's score on E–I and their total score across the five facet scales that relate to this dimension. For example, it is not uncommon to see an Enthusiastic Introvert or an Intimate Extravert. Such apparent inconsistencies are usually known as 'out-of-preference scores' (OOPS).

Although it is usual to have a number of OOPS in any one profile, it should be unusual to find that OOPS make up more than half of the facet scores relating to any one type dichotomy. Table 5.5 demonstrates that the proportion of people for whom this occurs ranges from 2.5% for the S–N block to 1.1% for the J–P block. It is therefore a very infrequent occurrence.

Table 5.5: Proportion of OOPS by type dichotomy[6]

	Proportion of 'reported type' OOPS					
	None	One	Two	Three	Four	Five
E–I	65.9%	24.5%	7.9%	1.6%	0.1%	0.0%
S–N	61.5%	27.1%	9.0%	2.0%	0.4%	0.1%
T–F	55.5%	33.5%	9.1%	1.8%	0.1%	0.0%
J–P	64.0%	27.4%	7.6%	1.0%	0.1%	0.0%

For any individual facet scale, an OOPS tends to occur in approximately 10% of profiles, with the exception of Expressive–

[6] In this table, an OOPS is defined as a score of 2, 3, 4 or 5 on the 'wrong' pole of the facet when compared with the reported type.

Contained (16%), Experiential–Theoretical (17%) and Questioning–Accommodating (24%).

Reliability

The reliability of a test or questionnaire relates to how consistent and precise it is. Internal consistency reliability addresses the question of whether all the questions in a scale measure the same construct. For example, are the Step II facet scales consistent within themselves, and do they hold together well as scales? A common measure of internal consistency reliability is coefficient alpha (Cronbach, 1951). The alpha coefficients for the Step II facet scales are shown in Table 5.6.

Table 5.6: Internal consistency reliability

Step II facet scale	No. of items	Coefficient alpha
E–I facet scales		
Initiating–Receiving	8	0.81
Expressive–Contained	7	0.77
Gregarious–Intimate	7	0.65
Active–Reflective	8	0.70
Enthusiastic–Quiet	9	0.69
S–N facet scales		
Concrete–Abstract	9	0.69
Realistic–Imaginative	7	0.73
Practical–Conceptual	8	0.56
Experiential–Theoretical	8	0.74
Traditional–Original	8	0.72
T–F facet scales		
Logical–Empathetic	9	0.79
Reasonable–Compassionate	8	0.62
Questioning–Accommodating	7	0.30
Critical–Accepting	8	0.38
Tough–Tender	8	0.69
J–P facet scales		
Systematic–Casual	8	0.72
Planful–Open-Ended	6	0.76
Early Starting–Pressure-Prompted	6	0.77
Scheduled–Spontaneous	8	0.72
Methodical–Emergent	8	0.64
	Median	0.70

The internal consistency reliability of most scales is good, and the average (median) reliability is 0.7. However, three scales (Practical–Conceptual, Questioning–Accommodating and Critical–Accepting) have lower alpha coefficients. This pattern is consistent with what was found during the development of the Step II instrument, where these three

facet scales were also found to show lower reliability than the others (Quenk, Hammer and Majors, 2004).

It should be noted, however, that coefficient alpha reliability statistics will give an underestimate of the reliability of the Step II facet scales, given the use of item response theory (IRT) methods in the actual scoring process.[7]

Group differences

This section looks at the way in which people from different groups responded to the French version of the European Step II questionnaire.

Gender

The means and standard deviations of the Step II facet scales are shown separately for males and females in Table 5.7, along with the difference in mean scores. This is represented graphically in Figure 5.2. Statistically significant differences were found between male and female mean scores for 18 of the 20 facet scales, with some consistent patterns emerging.

- On the E–I facet scales, all five mean scores tended slightly toward the E pole for females, while only three of the five did for males. There was no consistent gender difference across the dimension, with females tending further towards the E pole than males on two facet scales, and males tending more towards the E pole than females on one of the facet scales.

- On the S–N facet scales, all five mean scores tended slightly toward the S pole for both males and females. There were statistically significant gender differences on four of the five facet scales, with male mean scores noticeably tending more towards the S pole than females on three of these.

- On the T–F facet scales, all five mean scores tended slightly toward the T pole for males, whilst two did so for females. There were statistically significant gender differences on all five facet scales, with male mean scores noticeably tending more towards the T pole than females on four of these.

- On the J–P facet scales, all five mean scores tended slightly toward the J pole for both males and females. There were statistically significant gender differences on four of the five facet scales, with male mean scores noticeably tending more towards the J pole than females on three of these.

[7] Item response theory (IRT) is an approach to measurement that is concerned with modelling the relationship between item responses and the underlying characteristic assessed by the scale or test the item is designed to measure. IRT can be used to select items for a test and/or to score the items.

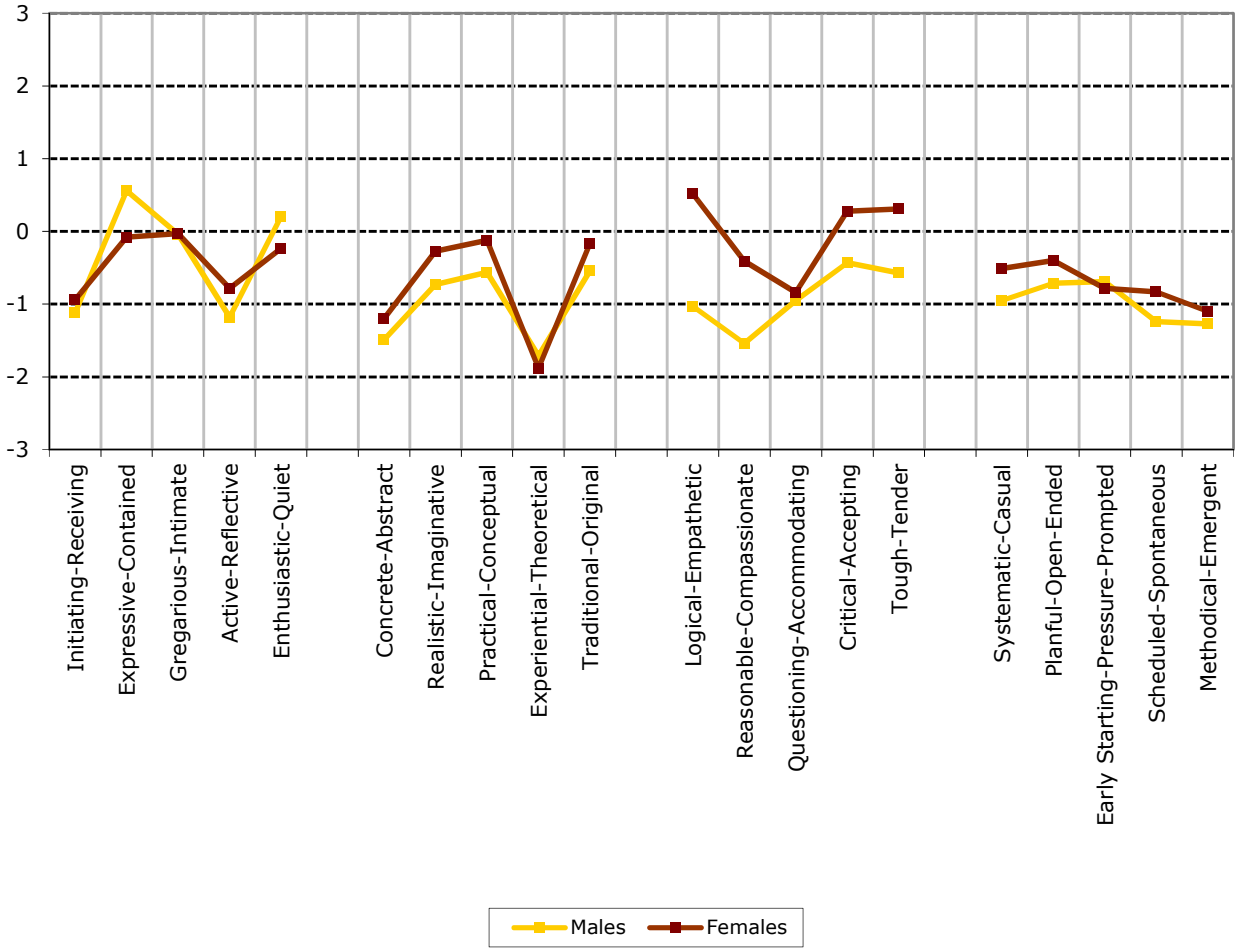
Table 5.7: Gender differences in facet scale scores

	Males (n=2,570)		Females (n=1,682)		Difference (M–F)[8]
Step II facet scale	**Mean**	**SD**	**Mean**	**SD**	
E–I facet scales					
Initiating–Receiving	-1.12	3.15	-0.94	3.13	–0.19**
Expressive–Contained	0.56	3.10	-0.08	3.18	0.64**
Gregarious–Intimate	-0.04	2.70	-0.03	2.56	0.00
Active–Reflective	-1.18	2.84	-0.78	2.87	–0.39**
Enthusiastic–Quiet	0.21	2.79	-0.23	2.81	0.44**
S–N facet scales					
Concrete–Abstract	-1.49	2.37	-1.20	2.37	–0.29**
Realistic–Imaginative	-0.73	3.10	-0.27	3.16	–0.47**
Practical–Conceptual	-0.56	2.65	-0.12	2.64	–0.44**
Experiential–Theoretical	-1.71	2.92	-1.88	2.78	0.17**
Traditional–Original	-0.54	2.90	-0.16	2.99	–0.38**
T–F facet scales					
Logical–Empathetic	-1.03	2.66	0.52	2.84	–1.55**
Reasonable–Compassionate	-1.54	2.61	-0.41	2.65	–1.12**
Questioning–Accommodating	-0.95	2.61	-0.84	2.61	–0.11*
Critical–Accepting	-0.43	1.87	0.28	2.00	–0.71**
Tough–Tender	-0.57	2.76	0.31	2.82	–0.88**
J–P facet scales					
Systematic–Casual	-0.95	2.75	-0.51	2.73	–0.44**
Planful–Open-Ended	-0.71	3.29	-0.40	3.34	–0.31**
Early Starting–Pressure-Prompted	-0.69	3.58	-0.78	3.59	0.09
Scheduled–Spontaneous	-1.24	2.88	-0.83	2.90	–0.41**
Methodical–Emergent	-1.27	2.81	-1.10	2.83	–0.17**

Difference significant at: *p<0.05, **p<0.01 (based on an independent samples t-test).

[8] A positive value indicates that male scores tend more towards I, N, F or P, and a negative value indicates that female scores tend more towards I, N, F or P.

Figure 5.2: Gender differences in facet scale scores



Age

There were found to be no meaningful links between age and facet scale scores. Although correlational analysis showed several facet scales to be significantly correlated with age, the significance levels were more the result of the very large sample size rather than being indicative of a meaningful relationship. The highest correlation was 0.07 (Concrete–Abstract and Experiential–Theoretical), and even this is too small to be considered meaningful.

For interpretation purposes, it is reasonable to conclude that there are no clear relationships between age and facet scale scores.

Ethnic origin

Ethnic origin information was not captured for people who completed the French language version of the questionnaire, so no analyses were conducted.

Occupational level

Research using the French version of the MBTI Step I questionnaire has demonstrated that individuals in higher-level jobs in organisations

are more likely to have preferences for iNtuition and (to some extent) Thinking than those in lower-level jobs.[9]

Table 5.8 shows the facet scale means and standard deviation for different occupational levels. The findings can be summarised as follows:

- A consistent pattern was found regarding the facets relating to the E–I dimension, in that mean facet scores amongst higher occupational levels tended to be further towards the E pole than those for the lower occupational level groups. As such, mean scores for the 'Employee' and 'First-level management/supervisory' levels tended most towards the I pole (or less towards the E pole) and mean scores for the 'Top level' group tended most towards the E pole.

- A similar, but more pronounced, pattern was found regarding the facets relating to the S–N dimension, in that mean facet scores amongst higher occupational levels tended to be further towards the N pole than those for the lower occupational level groups. As such, mean scores for the 'Employee' and 'First-level management/supervisory' levels tended most towards the S pole, and mean scores for the 'Top level' group tended most towards the N pole (or less towards the S pole).

- A reasonably consistent pattern was found regarding the facets relating to the T–F dimension. Mean facet scores amongst higher occupational levels tended to be further towards the T pole than those for the lower occupational level groups. The employee group mean scores were consistently considerably further towards the F pole (or less towards the T pole) than any other groups. This could be a reflection of the fact that females formed a considerably higher proportion of the employee group than they did of other groups.

- There were no clear overall patterns regarding the facets in the J–P preference block. The only consistent finding was that mean facet scores amongst higher occupational levels tended to be further towards the P pole of Early Starting–Pressure-Prompted than those for the lower occupational level groups. As such, mean scores for the 'Employee' and 'First-level management/supervisory' levels tended most towards the J pole, and mean scores for the 'Top level' group tended most towards the P pole.

[9] See the MBTI Step I European Data Supplement for details.

Table 5.8: Mean facet scale scores by occupational level

Step II facet scale	Top level (n=396)		Senior Executive (n=2,915)		Upper Middle Management (n=1,760)		Middle Management (n=2,419)		First Level Mgt/ Supervisory (n=680)		Employee (n=1,263)		Sig.
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales													
Initiating–Receiving	-1.46	3.03	-1.32	3.12	-1.09	3.16	-0.92	3.14	-0.95	3.15	-0.83	3.12	**
Expressive–Contained	-0.03	3.27	0.16	3.16	0.15	3.20	0.27	3.19	0.52	3.10	0.29	3.14	*
Gregarious–Intimate	-0.17	2.71	-0.17	2.62	-0.08	2.63	-0.08	2.64	0.08	2.62	0.13	2.61	*
Active–Reflective	-1.53	2.71	-1.27	2.74	-1.17	2.85	-0.88	2.96	-0.76	2.90	-0.70	2.85	**
Enthusiastic–Quiet	-0.46	2.88	-0.28	2.80	-0.16	2.83	0.19	2.83	0.46	2.69	0.32	2.72	**
S–N facet scales													
Concrete–Abstract	-0.90	2.51	-1.09	2.37	-1.24	2.46	-1.71	2.25	-1.74	2.28	-1.64	2.29	**
Realistic–Imaginative	0.48	3.26	-0.15	3.16	-0.36	3.11	-0.96	3.07	-1.25	3.05	-0.97	2.96	**
Practical–Conceptual	0.44	2.61	0.10	2.55	-0.14	2.68	-0.75	2.65	-1.19	2.63	-0.86	2.65	**
Experiential–Theoretical	-1.10	3.14	-1.42	2.98	-1.60	2.91	-2.13	2.75	-2.19	2.68	-2.23	2.58	**
Traditional–Original	0.37	2.82	-0.02	2.93	-0.20	2.95	-0.81	2.87	-1.02	2.92	-0.91	2.89	**
T–F facet scales													
Logical–Empathetic	-0.67	2.77	-0.74	2.71	-0.63	2.76	-0.43	2.85	-0.52	2.76	0.52	3.00	**
Reasonable–Compassionate	-1.40	2.66	-1.37	2.64	-1.25	2.68	-1.12	2.63	-1.03	2.62	-0.34	2.78	**
Questioning–Accommodating	-0.83	2.59	-0.99	2.53	-0.99	2.55	-0.88	2.59	-0.70	2.61	-0.58	2.68	**
Critical–Accepting	-0.31	1.83	-0.24	1.81	-0.26	1.85	-0.14	1.96	-0.03	2.18	0.32	2.19	**
Tough–Tender	-0.87	2.68	-0.68	2.66	-0.30	2.85	-0.02	2.78	-0.11	2.81	0.75	2.93	**
J–P facet scales													
Systematic–Casual	-0.70	2.69	-0.75	2.65	-0.64	2.73	-0.88	2.78	-0.99	2.72	-0.94	2.83	**
Planful–Open-Ended	-0.82	3.23	-0.80	3.30	-0.47	3.39	-0.61	3.32	-0.64	3.22	-0.58	3.38	*
Early Starting–Pressure-Prompted	0.24	3.58	-0.31	3.62	-0.47	3.61	-0.87	3.56	-1.07	3.58	-1.66	3.35	**
Scheduled–Spontaneous	-1.01	3.02	-1.14	2.89	-1.03	2.97	-1.20	2.87	-1.28	2.94	-1.11	2.77	*
Methodical–Emergent	-0.88	2.92	-1.15	2.75	-1.10	2.80	-1.31	2.81	-1.28	2.93	-1.38	2.90	**

Difference significant at: *p<0.05, **p<0.01 (based on a one-way analysis of variance).

Education

Specific educational qualifications were not collected for the OPPassessment sample; however, the age at which individuals left full-time education was. No significant and meaningful correlations were found between the age at which people left full-time education and their facet scale scores. All the correlations were less than 0.1, which, in real terms, is very small.

Work area

Information regarding the area of work people engage in was collected for the group. Many different categories were used, but for the purposes of analysis the focus was on the five most commonly occurring. These were as follows:

- HR, training, guidance
- Finance
- Science, engineering
- Sales, customer service
- IT.

Table 5.9 shows the facet scale means and standard deviation for these five work areas. The findings can be summarised as follows:

- There were no clear overall patterns regarding the facets in the E–I preference block. However, it is worth noting that mean scores for the 'Sales, customer service' group tended further towards the E pole than any of other groups on four of the five facet scales.
- The most noticeable pattern regarding the facets in the S–N preference block was that mean scores for the 'HR, training, guidance' group tended further towards the N pole (or less towards the S pole) than any of the other groups on all five facet scales.
- While there was no clear overall pattern regarding the facets in the T–F preference block, mean scores for the 'HR, training, guidance' group tended further towards the F pole (or less towards the T pole) than any of the other groups on four of the five facet scales. In contrast, mean scores for the 'Science, engineering' group tended further towards the T pole than any of the other groups on these same four facet scales. The scale which demonstrated no real difference between any of the groups was Questioning–Accommodating.
- The most noticeable pattern regarding the facets in the J–P preference block was that mean scores for the 'HR, training, guidance' group tended further towards the P pole (or less towards the J pole) than any of the other groups on four of the five facet scales.

Table 5.9: Mean facet scale scores by work area

Step II facet scale	HR, training, guidance (n=1,848)		Finance (n=1,522)		Science, engineering (n=1,243)		Sales, customer service (n=1,155)		IT (n=678)		Sig.
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales											
Initiating–Receiving	-1.09	3.19	-0.99	3.12	-0.95	3.12	-1.94	2.89	-0.63	3.29	**
Expressive–Contained	-0.17	3.22	0.56	3.08	0.42	3.20	-0.23	3.07	0.57	3.13	**
Gregarious–Intimate	0.00	2.71	-0.08	2.60	-0.25	2.62	-0.50	2.50	0.31	2.69	**
Active–Reflective	-0.93	2.86	-0.97	2.84	-0.96	2.84	-1.80	2.64	-0.65	3.02	**
Enthusiastic–Quiet	-0.14	2.89	0.15	2.74	0.19	2.75	-0.67	2.73	0.64	2.77	**
S–N facet scales											
Concrete–Abstract	-0.60	2.58	-1.83	2.25	-1.73	2.28	-1.70	2.07	-1.45	2.30	**
Realistic–Imaginative	0.34	3.25	-1.29	2.97	-0.98	2.93	-0.76	2.92	-0.88	3.06	**
Practical–Conceptual	0.33	2.67	-0.72	2.70	-0.82	2.66	-0.61	2.44	-0.58	2.68	**
Experiential–Theoretical	-0.98	3.05	-2.34	2.76	-1.55	2.92	-2.43	2.43	-1.81	2.87	**
Traditional–Original	0.42	3.04	-0.94	2.84	-0.77	2.86	-0.52	2.81	-0.69	2.89	**
T–F facet scales											
Logical–Empathetic	0.42	2.94	-0.88	2.76	-1.35	2.52	-0.21	2.81	-0.82	2.75	**
Reasonable–Compassionate	-0.20	2.77	-1.64	2.60	-1.75	2.55	-1.13	2.56	-1.46	2.75	**
Questioning–Accommodating	-0.97	2.59	-0.79	2.54	-0.93	2.64	-0.89	2.56	-0.83	2.49	
Critical–Accepting	0.21	1.97	-0.33	1.90	-0.55	1.79	-0.07	1.93	-0.32	2.06	**
Tough–Tender	0.24	2.84	-0.28	2.82	-0.62	2.69	-0.31	2.81	-0.07	2.92	**
J–P facet scales											
Systematic–Casual	0.02	2.80	-1.18	2.71	-1.14	2.69	-0.92	2.69	-1.03	2.86	**
Planful–Open-Ended	-0.8	3.44	-0.87	3.30	-0.79	3.28	-0.94	3.23	-0.64	3.36	**
Early Starting–Pressure-Prompted	-0.2	3.68	-0.92	3.61	-0.73	3.57	-0.84	3.47	-0.74	3.62	**
Scheduled–Spontaneous	-0.43	3.09	-1.28	2.85	-1.48	2.83	-1.30	2.82	-1.22	2.94	**
Methodical–Emergent	-0.58	2.92	-1.23	2.79	-1.31	2.81	-1.49	2.71	-1.46	2.72	**

Difference significant at: *p<0.05, **p<0.01 (based on a one-way analysis of variance).

Nationality

Information on nationality was not captured for individuals who completed the French language version of the Step II questionnaire, so no analyses were conducted.

Employment status

Employment status information was available for the OPPassessment sample. The vast majority of the group worked either full-time or part-time, or were self-employed. Table 5.11 shows the mean facet scale scores for each group, with the data illustrated in graphical form in Figure 5.4.

The analyses showed statistically significant differences across the groups on all the facet scales on the S–N and J–P dimensions, and four of the five facets on the E–I and T-F dimensions. The clearest patterns were as follows:

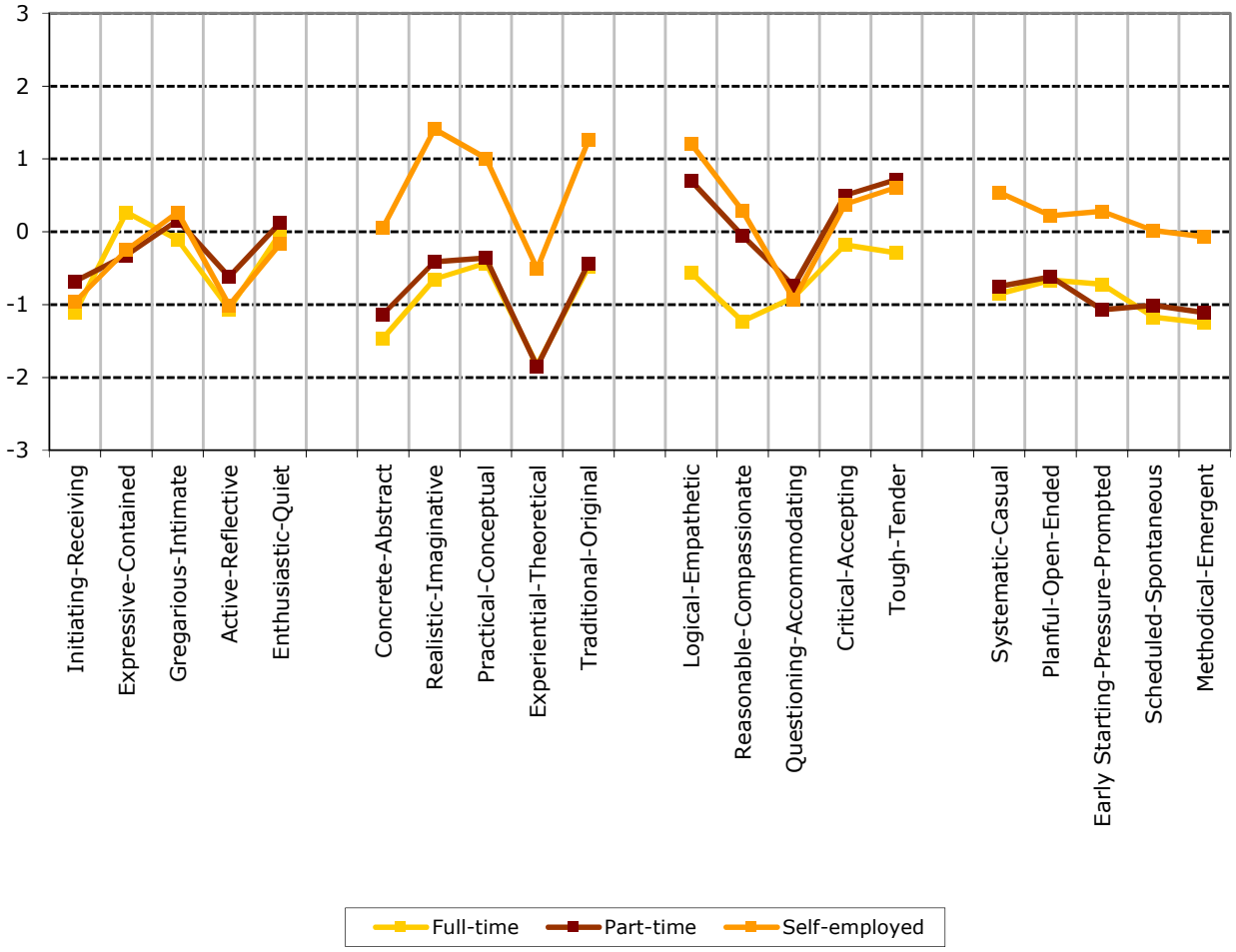
- There were no clear patterns regarding the facets in the E–I preference block.
- The self-employed group tended to score more towards the N pole on the S–N facet scales and more towards the P pole on the J–P facet scales than the other two groups.
- Apart from Questioning-Accommodating, the full-time group tended to score more towards the T pole on the T–F facets than the other two groups. This is likely to be at least partly a gender effect; 87% of part-time workers and 59% of self-employed people were female, compared with 42% of the total group and 38% of full-time workers.

Table 5.11: Mean facet scale scores by employment status

Step II facet scale	Full-time (n=8,924)		Part-time (n=548)		Self-employed (n=396)		Sig.
	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales							
Initiating–Receiving	-1.11	3.14	-0.68	3.15	-0.96	3.24	**
Expressive–Contained	0.27	3.15	-0.32	3.24	-0.25	3.46	**
Gregarious–Intimate	-0.11	2.63	0.16	2.54	0.27	2.88	**
Active–Reflective	-1.06	2.86	-0.62	2.89	-1.01	2.88	**
Enthusiastic–Quiet	-0.01	2.82	0.13	2.71	-0.16	2.89	
S–N facet scales							
Concrete–Abstract	-1.46	2.34	-1.13	2.48	0.06	2.59	**
Realistic–Imaginative	-0.65	3.11	-0.41	3.07	1.41	3.09	**
Practical–Conceptual	-0.43	2.65	-0.36	2.72	1.01	2.56	**
Experiential–Theoretical	-1.83	2.85	-1.85	2.81	-0.51	3.05	**
Traditional–Original	-0.48	2.92	-0.43	2.93	1.26	2.87	**
T–F facet scales							
Logical–Empathetic	-0.56	2.80	0.70	2.75	1.21	2.98	**
Reasonable–Compassionate	-1.23	2.65	-0.05	2.72	0.29	2.83	**
Questioning–Accommodating	-0.90	2.57	-0.74	2.70	-0.93	2.64	
Critical–Accepting	-0.18	1.93	0.50	2.06	0.38	2.04	**
Tough–Tender	-0.29	2.79	0.72	2.77	0.61	3.02	**
J–P facet scales							
Systematic–Casual	-0.85	2.72	-0.75	2.73	0.54	2.87	**
Planful–Open-Ended	-0.66	3.32	-0.62	3.30	0.22	3.39	**
Early Starting–Pressure-Prompted	-0.72	3.59	-1.07	3.55	0.28	3.65	**
Scheduled–Spontaneous	-1.17	2.88	-1.01	2.88	0.02	3.14	**
Methodical–Emergent	-1.25	2.80	-1.11	2.84	-0.07	3.12	**

Difference significant at: *p<0.05, **p<0.01 (based on a one-way analysis of variance).

Figure 5.4: Mean facet scale scores by employment status



Appendix 1: Sample description

Sample 1: Data from OPPassessment (representative European French-speaking professional and managerial sample)

This sample consists of 13,610 individuals who completed the MBTI Step II instrument in French via the OPPassessment system between June 2006 and July 2016. Fifty eight per cent of the respondents were male and 42% were female. Age ranged from 16 to 68 years, with a mean of 41 and a median of 42.

Nationality was not collected by OPPassessment.

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

Employment status	Percentage
Full-time	89.1%
Part-time	5.5%
Self-employed	4.0%
Unemployed	1.2%
Homemaker	0.2%
Retired	0.1%

The majority of the group were of managerial level or above, with the largest single group being senior executive (31%):

Occupational level	Percentage
Top level	4.0%
Senior executive	29.6%
Upper middle management	17.8%
Middle management	24.5%
First level management/supervisor	6.9%
Employee	12.8%
Other	4.4%

A range of work areas were represented:

Work area (job type)	Percentage
HR, training, guidance	18.6%
Finance	15.3%
Science, engineering	12.5%
Sales, customer service	11.6%
IT	6.8%
Business services	4.6%
Research and development	3.9%
Admin or secretarial	3.4%
Health, social services, etc.	2.7%
Education	1.0%
Land, sea or air transport	0.9%
Leisure, personal service	0.4%
Military, police, prison, fire	0.1%
Skilled operative	0.5%
Unskilled operative	0.1%
Other private sector	6.8%
Other public sector	1.8%
Other	8.9%

MBTI® Step II™ instrument

European Data Supplement

German

January 2018

2nd edition

OPP Ltd
Elsfield Hall
15–17 Elsfield Way
Oxford OX2 8EP
UK
t: +44 (0)1865 404 500

Contents

Introduction

Data collected for the European MBTI® Step II™ instrument were analysed to produce the findings in this supplement. This is the second data supplement produced for this instrument. A brief description of the sample is given below. Further details of the sample are provided in Appendix 1.

- The sample consisted of 15,228 individuals who completed the MBTI Step II questionnaire in German via the OPPassessment™ system between 2006 and mid-2016.[1] This sample is considered to be representative of the groups of people with whom the German MBTI Step II instrument has been and will be used for applications such as management development, coaching, counselling and teambuilding. As such, it is likely to represent a cross-section of the German-speaking professional and managerial population.

The results of the analyses are outlined below.

[1] OPPassessment allows personality questionnaires such as the MBTI instrument to be administered via email and/or completed online.

Type distribution

Type tables are a way of illustrating the proportion of each type within a particular group. Below is a type table taken from the sample described on the previous page.

For each of the 16 different types, the number of cases, the percentage of the total that this represents and the self-selection ratio (SSR) are shown. The SSR (Myers et al., 1998) is a way of demonstrating whether a given type appears more or less often in a particular group than would be expected given its frequency in a reference group. Ideally, the type distribution from a large representative sample of the German population would be used to calculate SSRs in this data supplement. However, such a sample does not currently exist. In its place, SSRs have been calculated using type data from the UK general population (Kendall, 1998), which can be justified by the fact that type distributions for comparable German and British groups, such as managers and professionals, are similar. Evidence (eg Hackston and Kendall, 2004; Quenk et al., 2004) does suggest that although type-related behaviours vary a good deal from country to country and from culture to culture, the frequencies of underlying MBTI types do not.

An SSR of greater than 1 indicates that a type is over-represented, and an SSR of less than 1 denotes that it is under-represented. Asterisks are used to denote whether the over- or under-representations are statistically significant, based on the results of chi-square analysis.[2]

[2] Chi-square analysis (often abbreviated to χ^2) is a technique used to explore whether observed frequency distributions differ significantly from other, predefined, distributions. In this case, the UK general population group is used as the reference group, and the chi-square analysis indicates whether the proportion of people of each type within a particular sample differs significantly from the proportion of people reporting the same type within the reference group.

OPPassessment data (representative German-speaking professional and managerial sample)

Table 6.1: Type table for OPPassessment data (reported type, n=15,228)

ISTJ	ISFJ	INFJ	INTJ
n=1,963 12.9% SSR=0.94	n=336 2.2% SSR=0.17**	n=194 1.3% SSR=0.76	n=821 5.4% SSR=3.86**
ISTP	**ISFP**	**INFP**	**INTP**
n=392 2.6% SSR=0.41**	n=90 0.6% SSR=0.10**	n=181 1.2% SSR=0.38**	n=515 3.4% SSR=1.42
ESTP	**ESFP**	**ENFP**	**ENTP**
n=707 4.6% SSR=0.79*	n=254 1.7% SSR=0.20**	n=756 5.0% SSR=0.79*	n=1,290 8.5% SSR=3.03**
ESTJ	**ESFJ**	**ENFJ**	**ENTJ**
n=3,663 24.1% SSR=2.32**	n=816 5.4% SSR=0.43**	n=704 4.6% SSR=1.64**	n=2,546 16.7% SSR=5.76**

Type	n	%
E	10,736	70.5%**
I	4,492	29.5%**
S	8,221	54.0%**
N	7,007	46.0%**
T	11,897	78.1%**
F	3,331	21.9%**
J	11,043	72.5%**
P	4,185	27.5%**

*Difference significant at p<0.05, based on chi-square results.
**Difference significant at p<0.01, based on chi-square results.

The most common single type preferences are ESTJ (24% of the total) and ENTJ (17%); this is a common finding with managerial groups in other countries. The SSR results suggest that, in comparison with the UK general population, those with preferences for NT are over-represented and those with preferences for SF are under-represented. Again, this is a common finding with managerial groups.

Properties of the Step II facet scales

Facet scale score distribution

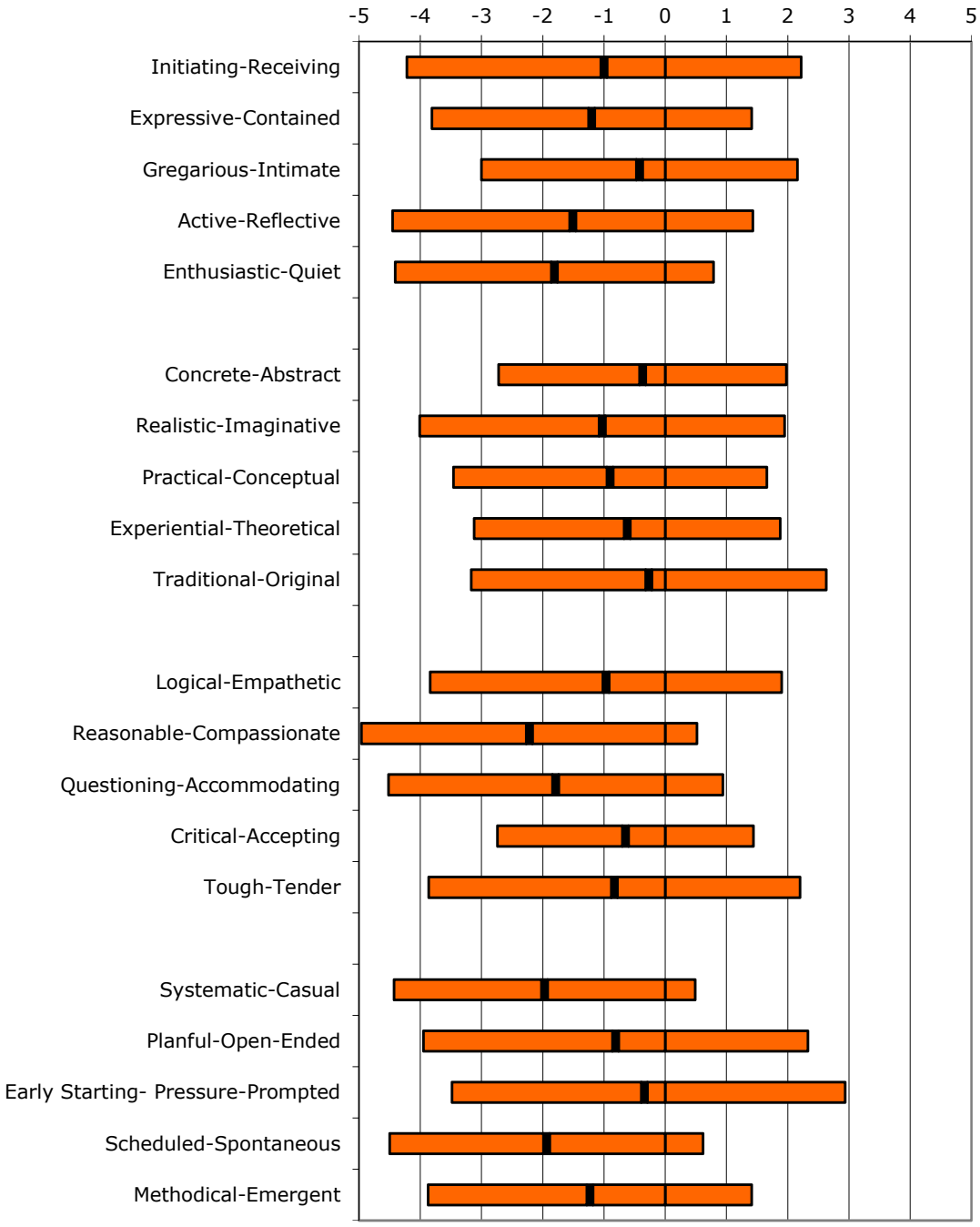
The means and standard deviations of the Step II facets are shown in Table 6.2, and illustrated graphically in Figure 6.1. The mean score for a scale is calculated by adding together the scores of each individual in the sample and then dividing the sum by the number of individuals. Note that negative values indicate that mean scores are nearer the left-hand pole of the facet and positive values indicate that the mean scores are nearer the right-hand pole. The standard deviation (SD) is a statistical measure describing the degree to which the scores from the sample either bunch up close to, or are scattered widely around, the mean for the sample.

Table 6.2: Means and standard deviations of the facet scales

Step II facet scale	Mean[3]	SD
E–I facet scales		
Initiating–Receiving	–1.00	3.22
Expressive–Contained	–1.20	2.61
Gregarious–Intimate	–0.42	2.58
Active–Reflective	–1.51	2.94
Enthusiastic–Quiet	–1.81	2.60
S–N facet scales		
Concrete–Abstract	–0.37	2.35
Realistic–Imaginative	–1.03	2.98
Practical–Conceptual	–0.90	2.56
Experiential–Theoretical	–0.62	2.50
Traditional–Original	–0.27	2.90
T–F facet scales		
Logical–Empathetic	–0.97	2.87
Reasonable–Compassionate	–2.22	2.74
Questioning–Accommodating	–1.79	2.73
Critical–Accepting	–0.65	2.09
Tough–Tender	–0.83	3.03
J–P facet scales		
Systematic–Casual	–1.97	2.46
Planful–Open-Ended	–0.81	3.14
Early Starting–Pressure-Prompted	–0.34	3.28
Scheduled–Spontaneous	–1.94	2.56
Methodical–Emergent	–1.23	2.64

[3] Scale means: in Step II reports, scores are given from 5 on one pole, through 0, to 5 on the opposite pole. In compiling statistical information, however, one pole needs to be clearly distinguished from the other. Throughout this data supplement, therefore, a convention has been adopted of scores 5, 4, 3, 2 or 1 on the left-hand pole being assigned values of –5, –4, –3, –2, or –1 and scores on the right-hand pole being assigned positive scores. This does not, of course, imply any suggestion that positive scores are 'better' than negative scores.

Figure 6.1: Facet scale mean scores[4]



[4] For each facet scale, the central line shows the mean and the coloured bars on either side indicate the standard deviation.

In general, the means centre quite close to the midpoint of each scale. Only one of the scales has a mean greater than 2 points from the midpoint in either direction, and this has a standard deviation of 2.74. However, a clear pattern does emerge. The mean scores all tend towards the E, S, T and J direction, which is consistent with the most common four-letter type preference amongst this sample. Amongst the group as a whole, there are more individuals with a preference for E, S, T and J.

Facet scale intercorrelations

Correlations among the Step II facet scales are shown in Table 6.3. Facet scales within each dichotomy usually correlate more highly (often substantially so) with the other scales of the same dichotomy than they do with scales in the other three dichotomies. For example, every E–I facet scale correlates more highly with the other four E–I scales than with any of the S–N, T–F or J–P facet scales.

There is only one scale that correlates more highly with scales in other dichotomies. The T–F scale Questioning–Accommodating correlates with many scales across the E–I and S–N dichotomies. For example, it correlates at a level of 0.23 with Enthusiastic–Quiet and at 0.21 with both Initiating–Receiving and Active–Reflective on the E–I dichotomy, whilst correlating negatively at between –0.16 and –0.35 with all of the S–N scales except Realistic–Imaginative on the S–N dichotomy. By way of comparison, it correlates at between 0.14 and 0.22 with the other four scales on the same dichotomy as itself (T–F). The negative correlations between Questioning–Accommodating and the S–N scales are consistent with findings from the US version of the Step II instrument, and suggest that a questioning approach to differences of opinion seems to be related to a range of iNtuition (N) facets. These patterns are similar to those found when the European Step II instrument was developed.

Table 6.3: Intercorrelations of Step II facet scales

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Initiating–Receiving	**1.00**																			
2. Expressive–Contained	**0.47**	**1.00**																		
3. Gregarious–Intimate	**0.49**	**0.43**	**1.00**																	
4. Active–Reflective	**0.68**	**0.50**	**0.50**	**1.00**																
5. Enthusiastic–Quiet	**0.58**	**0.55**	**0.57**	**0.64**	**1.00**															
6. Concrete–Abstract	-0.12	-0.10	-0.12	-0.12	-0.21	**1.00**														
7. Realistic–Imaginative	-0.10	-0.13	-0.13	-0.11	-0.21	**0.58**	**1.00**													
8. Practical–Conceptual	-0.12	-0.03	-0.05	-0.08	-0.14	**0.49**	**0.46**	**1.00**												
9. Experiential–Theoretical	-0.03	-0.03	-0.02	-0.01	-0.05	**0.47**	**0.40**	**0.45**	**1.00**											
10. Traditional–Original	-0.21	-0.12	-0.14	-0.21	-0.26	**0.56**	**0.49**	**0.49**	**0.32**	**1.00**										
11. Logical–Empathetic	-0.08	-0.20	-0.06	-0.08	-0.17	0.21	0.33	-0.03	0.05	0.07	**1.00**									
12. Reasonable–Compassionate	-0.01	-0.13	-0.03	-0.01	-0.08	0.12	0.23	-0.08	0.04	-0.02	**0.70**	**1.00**								
13. Questioning–Accommodating	0.21	0.09	0.10	0.21	0.23	-0.22	-0.10	-0.30	-0.16	-0.35	**0.14**	**0.21**	**1.00**							
14. Critical–Accepting	-0.06	-0.13	-0.09	-0.05	-008	0.07	0.18	-0.05	0.01	-0.07	**0.43**	**0.46**	**0.22**	**1.00**						
15. Tough–Tender	0.05	-0.09	-0.02	0.05	-0.04	0.15	0.26	-0.03	0.09	-0.05	**0.57**	**0.60**	**0.19**	**0.49**	**1.00**					
16. Systematic–Casual	-0.08	-0.14	-0.13	-0.10	-0.18	0.39	0.44	0.15	0.14	0.37	0.40	0.32	-0.01	0.23	0.29	**1.00**				
17. Planful–Open-Ended	-0.02	-0.02	-0.03	-0.03	-0.06	0.29	0.29	0.14	0.12	0.34	0.19	0.13	-0.07	0.08	0.12	**0.58**	**1.00**			
18. Early Starting–Pressure–Prompted	-0.07	-0.06	-0.08	-0.09	-0.14	0.22	0.21	0.16	0.11	0.31	0.05	0.00	-0.12	-0.05	-0.02	**0.39**	**0.38**	**1.00**		
19. Scheduled–Spontaneous	0.00	-0.05	-0.04	-0.02	-0.07	0.29	0.31	0.16	0.16	0.32	0.21	0.17	-0.02	0.08	0.15	**0.60**	**0.59**	**0.46**	**1.00**	
20. Methodical–Emergent	0.04	-0.02	-0.03	0.02	-0.04	0.20	0.21	0.09	0.09	0.20	0.13	0.12	0.01	0.05	0.09	**0.47**	**0.42**	**0.41**	**0.51**	**1.00**

Correlations of Step II facet scales with Step I scales

Correlations between Step II facet scales and the continuous scores from the MBTI Step I™ instrument are shown in Table 6.4.[5] Positive correlations between facet scales and the corresponding Step I dimension (eg between Initiating–Receiving and the E–I continuous score) indicate relationships in the expected direction, with higher facet scores tending to be associated with higher continuous scores, and vice versa. Negative correlations would indicate an inverse relationship between the facet scale scores and the corresponding Step I dimension.

Table 6.4: Correlations of Step II facet scales with Step I continuous scores

	Step I continuous score			
Step II facet scales	**E–I**	**S–N**	**T–F**	**J–P**
E–I facet scales				
Initiating–Receiving	**0.81**	–0.14	–0.07	–0.01
Expressive–Contained	**0.70**	–0.11	–0.22	–0.05
Gregarious–Intimate	**0.62**	–0.12	–0.08	–0.05
Active–Reflective	**0.80**	–0.13	–0.08	–0.03
Enthusiastic–Quiet	**0.81**	–0.21	–0.17	–0.09
S–N facet scales				
Concrete–Abstract	–0.16	**0.80**	0.20	0.36
Realistic–Imaginative	–0.15	**0.73**	0.33	0.36
Practical–Conceptual	–0.10	**0.61**	–0.05	0.20
Experiential–Theoretical	–0.02	**0.58**	0.05	0.18
Traditional–Original	–0.24	**0.69**	0.02	0.40
T–F facet scales				
Logical–Empathetic	–0.15	0.19	**0.88**	0.21
Reasonable–Compassionate	–0.06	0.09	**0.76**	0.16
Questioning–Accommodating	0.23	–0.26	**0.23**	–0.07
Critical–Accepting	–0.08	0.06	**0.57**	0.07
Tough–Tender	0.02	0.13	**0.71**	0.14
J–P facet scales				
Systematic–Casual	–0.14	0.38	0.40	**0.70**
Planful–Open-Ended	–0.04	0.29	0.19	**0.80**
Early Starting–Pressure-Prompted	–0.12	0.24	0.03	**0.67**
Scheduled–Spontaneous	–0.04	0.31	0.21	**0.79**
Methodical–Emergent	0.00	0.20	0.14	**0.64**

The E–I facet scales correlate at a level of 0.62 to 0.81 with the E–I continuous scores from Step I; the S–N facet scales correlate at 0.61 to 0.80 with the S–N continuous scores; the T–F facet scales correlate at 0.23 to 0.88 with the T–F continuous scores; and the J–P facet

[5] Continuous scores place an individual's score on each dimension onto a continuous scale with a midpoint of 100. To calculate continuous scores, Preference Clarity Index (PCI) scores for each dimension are either subtracted or added to 100, depending on which direction the overall preference is. PCI scores in the direction of E, S, T or J are subtracted from 100. PCI scores in the direction of I, N, F or P are added to 100.

scales correlate at 0.67 to 0.80 with the J–P continuous scores. The scale that showed the lowest correlation with its associated Step I continuous score was Questioning–Accommodating at 0.23. This scale is also one of the lowest on internal consistency (see page 135), and has been previously found to have the lowest test–retest reliability (Quenk, Hammer and Majors, 2004), which would result in the true correlations being underestimated.

These correlations are very similar to those found during the development of the Step II questionnaire. This consistency, alongside the fact that the correlations between Step II facet scales and Step I continuous scores associated with their own dimension are substantially higher than correlations with the other three dimensions, provides compelling evidence for the theoretical hierarchical structure of the Step II facet scales in relation to the Step I scales.

Out-of-preference scores

It is known that the five facet scales relating to each type dimension do not describe the dimension in its entirety; there will not be a precise and exact overlap between, for example, an individual's score on E–I and their total score across the five facet scales that relate to this dimension. For example, it is not uncommon to see an Enthusiastic Introvert or an Intimate Extravert. Such apparent inconsistencies are usually known as 'out-of-preference scores' (OOPS).

Although it is usual to have a number of OOPS in any one profile, it should be unusual to find that OOPS make up more than half of the facet scores relating to any one type dichotomy. Table 6.5 demonstrates that the proportion of individuals for whom this occurs ranges from 2.8% for the S–N block to 1.1% for the E–I block. It is therefore a very infrequent occurrence.

Table 6.5: Proportion of OOPS by type dichotomy[6]

	Proportion of 'reported type' OOPS					
	None	One	Two	Three	Four	Five
E–I	70.7%	22.0%	6.2%	1.0%	0.1%	0.0%
S–N	61.0%	26.7%	9.4%	2.4%	0.4%	0.0%
T–F	61.2%	28.0%	8.6%	1.9%	0.3%	0.0%
J–P	59.4%	31.1%	8.3%	1.1%	0.1%	0.0%

For any individual facet scale, an OOPS tends to occur in approximately 10% of profiles, with the exceptions of

[6] In this table, an OOPS is defined as a score of 2, 3, 4 or 5 on the 'wrong' pole of the facet when compared with the reported type.

Experiential–Theoretical (17%), Questioning–Accommodating (18%) and Early Starting–Pressure-Prompted (17%).

Reliability

The reliability of a test or questionnaire relates to how consistent and precise it is. Internal consistency reliability addresses the question of whether all the questions in a scale measure the same construct. For example, are the Step II facet scales consistent within themselves, and do they hold together well as scales? A common measure of internal consistency reliability is coefficient alpha (Cronbach, 1951). The alpha coefficients for the Step II facet scales are shown in Table 6.6.

Table 6.6: Internal consistency reliability

Step II facet scale	No. of items	Coefficient alpha
E–I facet scales		
Initiating–Receiving	8	0.80
Expressive–Contained	7	0.67
Gregarious–Intimate	7	0.59
Active–Reflective	8	0.70
Enthusiastic–Quiet	9	0.66
S–N facet scales		
Concrete–Abstract	9	0.65
Realistic–Imaginative	7	0.70
Practical–Conceptual	8	0.52
Experiential–Theoretical	8	0.67
Traditional–Original	8	0.70
T–F facet scales		
Logical–Empathetic	9	0.78
Reasonable–Compassionate	8	0.69
Questioning–Accommodating	7	0.40
Critical–Accepting	8	0.38
Tough–Tender	8	0.76
J–P facet scales		
Systematic–Casual	8	0.68
Planful–Open-Ended	6	0.74
Early Starting–Pressure-Prompted	6	0.68
Scheduled–Spontaneous	8	0.62
Methodical–Emergent	8	0.60
	Median	0.67

The internal consistency reliability of most scales is good, and the average (median) reliability is close to 0.7. However, three scales (Practical–Conceptual, Questioning–Accommodating and Critical–Accepting) do have lower alpha coefficients. This pattern is consistent

with what was found during the development of the Step II instrument, where these three facet scales were also found to show lower reliability than the others (Quenk, Hammer and Majors, 2004).

It should be noted, however, that coefficient alpha reliability statistics will give an underestimate of the reliability of the Step II facet scales, given the use of item response theory (IRT) methods in the actual scoring process.[7]

Group differences

This section looks at the way in which people from different groups responded to the German version of the European Step II questionnaire.

Gender

The means and standard deviations of the Step II facet scales are shown separately for males and females in Table 6.7, along with the difference in mean scores. This is represented graphically in Figure 6.2. Statistically significant differences were found between male and female mean scores for 18 of the 20 facet scales, with some consistent patterns emerging.

- On the E–I facet scales, all five mean scores tended slightly toward the E pole for both males and females. There were statistically significant gender differences on all five facet scales, with females tending further towards the E pole than males on three facet scales, and males tending more towards the E pole than females on the other two facet scales.
- On the S–N facet scales, all five mean scores tended slightly toward the S pole for both males and females. There were statistically significant gender differences on four of the five facet scales with male mean scores tending more towards the S pole than females on two of these and female mean scores tending more towards the towards the S pole on the other two. There was no statistically significant difference between genders on the remaining scale.
- On the T–F facet scales, all five mean scores tended slightly toward the T pole for males, whereas three did so for females. There were statistically significant gender differences on all five facet scales, with male mean scores tending more towards the T pole than females.
- On the J–P facet scales, all five mean scores tended slightly toward the J pole for both males and females. There were statistically

[7] Item response theory (IRT) is an approach to measurement that is concerned with modelling the relationship between item responses and the underlying characteristic assessed by the scale or test the item is designed to measure. IRT can be used to select items for a test and/or to score the items.

significant gender differences on four of the five facet scales, with male mean scores tending more towards the J pole than females.

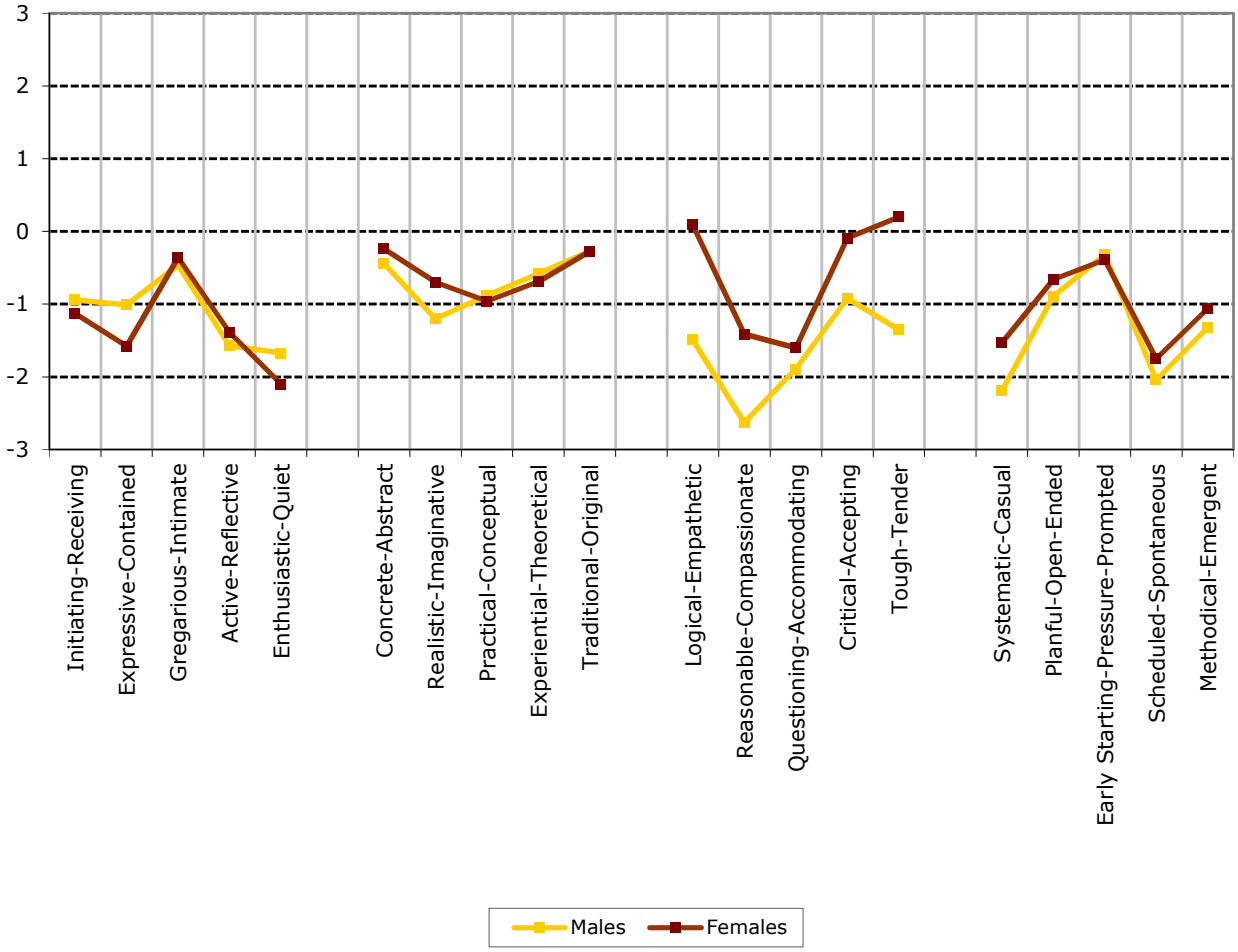
Table 6.7: Gender differences in facet scale scores

	Males (n=10,179)		Females (n=5,049)		Difference (M–F)[8]
Step II facet scale	**Mean**	**SD**	**Mean**	**SD**	
E–I facet scales					
Initiating–Receiving	-0.94	3.22	-1.13	3.23	0.18**
Expressive–Contained	-1.01	2.60	-1.58	2.58	0.57**
Gregarious–Intimate	-0.46	2.58	-0.36	2.58	–0.10*
Active–Reflective	-1.57	2.91	-1.39	3.00	–0.18**
Enthusiastic–Quiet	-1.67	2.64	-2.10	2.49	0.43**
S–N facet scales					
Concrete–Abstract	-0.44	2.39	-0.24	2.26	–0.21**
Realistic–Imaginative	-1.20	2.97	-0.70	2.98	–0.50**
Practical–Conceptual	-0.88	2.50	-0.96	2.66	0.09*
Experiential–Theoretical	-0.58	2.56	-0.69	2.38	0.11**
Traditional–Original	-0.27	2.89	-0.27	2.93	0.01
T–F facet scales					
Logical–Empathetic	-1.49	2.69	0.09	2.93	–1.58**
Reasonable–Compassionate	-2.63	2.59	-1.41	2.83	–1.22**
Questioning–Accommodating	-1.89	2.69	-1.60	2.81	–0.29**
Critical–Accepting	-0.92	2.05	-0.09	2.06	–0.84**
Tough–Tender	-1.35	2.94	0.20	2.94	–1.55**
J–P facet scales					
Systematic–Casual	-2.19	2.37	-1.53	2.56	–0.67**
Planful–Open-Ended	-0.89	3.09	-0.66	3.23	–0.23**
Early Starting–Pressure-Prompted	-0.32	3.26	-0.39	3.31	0.07
Scheduled–Spontaneous	-2.04	2.51	-1.75	2.63	–0.29**
Methodical–Emergent	-1.32	2.61	-1.06	2.69	–0.26**

Difference significant at: *p<0.05, **p<0.01 (based on an independent samples t-test).

[8] A positive value indicates that male scores tend more towards I, N, F or P, and a negative value indicates that female scores tend more towards I, N, F or P.

Figure 6.2: Gender differences in facet scale scores



Age

There were found to be no meaningful links between age and facet scale scores. Although correlational analysis showed several facet scales to be significantly correlated with age, the significance levels were more the result of the very large sample size rather than being indicative of a meaningful relationship. The highest correlations were 0.10 (Gregarious–Intimate) and 0.11 (Enthusiastic-Quiet), and even these are too small to be considered meaningful.

For interpretation purposes, it is reasonable to conclude that there are no clear relationships between age and facet scale scores.

Ethnic origin

Ethnic origin information was not captured for those who completed the German language version of the Step II questionnaire, so no analyses were conducted.

Occupational level

Research using the German version of the MBTI Step I questionnaire has demonstrated that individuals in higher-level jobs in organisations

are more likely to have preferences for Extraversion and Thinking than those in lower-level jobs.[9]

Table 6.8 shows the facet scale means and standard deviation for different occupational levels. The findings can be summarised as follows:

- No clear pattern was found regarding the facets relating to the E–I dimension.
- A pattern was found regarding the facets relating to the S–N dimension. Mean facet scores amongst people at the top level tended to be further towards the N pole than those for the lower occupational level groups.
- A reasonably consistent pattern was found regarding the facets relating to the T–F dimension. Mean facet scores amongst the employee group mean scores were consistently further towards the F pole (or less towards the T pole) than any other groups. This could be a reflection of the fact that females formed a considerably higher proportion of the employee group than they did of other groups.
- There were no clear overall patterns regarding the facets in the J–P preference block.

[9] See the MBTI Step I European Data Supplement for details.

Table 6.8: Mean facet scale scores by occupational level

Step II facet scale	Top level (n=533)		Senior Executive (n=2,084)		Upper Middle Management (n=1,309)		Middle Management (n=2,136)		First Level Mgt/Supervisory (n=1,242)		Employee (n=3,233)		Sig.
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales													
Initiating–Receiving	-1.35	3.16	-1.25	3.18	-1.25	3.16	-1.05	3.23	-0.68	3.28	-0.81	3.22	**
Expressive–Contained	-1.61	2.61	-1.28	2.58	-1.43	2.56	-1.24	2.61	-1.09	2.64	-1.23	2.63	**
Gregarious–Intimate	-0.38	2.74	-0.35	2.58	-0.56	2.57	-0.50	2.57	-0.39	2.65	-0.46	2.58	
Active–Reflective	-1.84	2.89	-1.65	2.89	-1.83	2.83	-1.55	2.93	-1.39	3.03	-1.36	3.02	**
Enthusiastic–Quiet	-2.09	2.61	-1.91	2.55	-2.03	2.60	-1.88	2.58	-1.65	2.68	-1.74	2.63	**
S–N facet scales													
Concrete–Abstract	0.15	2.44	-0.54	2.36	-0.27	2.28	-0.34	2.38	-0.43	2.35	-0.47	2.35	**
Realistic–Imaginative	-0.01	3.16	-1.26	2.97	-1.01	2.97	-1.06	2.91	-1.17	2.97	-1.15	2.96	**
Practical–Conceptual	-0.13	2.40	-0.95	2.53	-0.69	2.44	-0.77	2.57	-1.05	2.54	-1.24	2.58	**
Experiential–Theoretical	0.03	2.65	-0.69	2.48	-0.60	2.49	-0.56	2.54	-0.65	2.57	-0.72	2.48	**
Traditional–Original	0.55	2.91	-0.27	2.95	0.08	2.89	-0.12	2.92	-0.49	2.88	-0.64	2.84	**
T–F facet scales													
Logical–Empathetic	-0.67	3.02	-1.05	2.83	-1.26	2.67	-1.28	2.77	-1.19	2.77	-0.58	2.98	**
Reasonable–Compassionate	-2.09	2.88	-2.45	2.65	-2.70	2.51	-2.60	2.55	-2.32	2.68	-1.76	2.87	**
Questioning–Accommodating	-2.02	2.56	-1.88	2.66	-2.06	2.57	-1.96	2.66	-1.70	2.69	-1.50	2.91	**
Critical–Accepting	-0.92	2.07	-0.80	2.06	-0.95	1.97	-0.86	2.02	-0.61	2.11	-0.34	2.21	**
Tough–Tender	-0.84	3.11	-1.18	2.95	-1.23	2.91	-1.12	2.96	-0.85	2.96	-0.33	3.15	**
J–P facet scales													
Systematic–Casual	-1.77	2.60	-2.20	2.38	-2.11	2.41	-2.04	2.39	-1.97	2.38	-1.79	2.57	**
Planful–Open-Ended	-0.97	3.16	-1.07	3.11	-0.82	3.21	-0.85	3.15	-0.71	3.09	-0.76	3.18	**
Early Starting–Pressure-Prompted	0.26	3.21	-0.26	3.27	0.16	3.29	-0.07	3.27	-0.18	3.32	-0.61	3.28	**
Scheduled–Spontaneous	-1.71	2.69	-2.11	2.55	-1.94	2.55	-1.98	2.53	-1.86	2.49	-1.86	2.59	**
Methodical–Emergent	-0.82	2.65	-1.38	2.69	-1.22	2.66	-1.22	2.59	-1.04	2.66	-1.18	2.72	**

Significant at: *p<0.05, **p<0.01 (based on a one-way analysis of variance).

Education

Specific educational qualifications were not collected for the OPPassessment sample; however, the age at which individuals left full-time education was. Although correlational analysis showed several facet scales to be significantly correlated with age, all the correlations were less than 0.1. The significance levels were more the result of the very large sample size rather than being indicative of a meaningful relationship.

Work area

Information regarding the area of work people engage in was collected for the group. Many different categories were used, but for the purposes of analysis the focus was on the five most commonly occurring. These were as follows:

- Finance
- Business services
- HR, training, guidance
- Sales, customer service
- Science, engineering.

Table 6.9 shows the facet scale means and standard deviations for these five work areas. The findings can be summarised as follows:

- A clear pattern was observed regarding the facets in the E–I preference block in that mean scores for the 'Sales, customer service' group tended further towards the E pole than any of the other groups on four of the five facet scales.
- The most noticeable pattern regarding the facets in the S–N, T–F and J–P preference blocks was that whereas mean scores for all groups tended towards the S, T and J poles on the facet scales (or were close to the midpoint), the scores for the 'HR, training, guidance' group were less far towards the S, T and J poles than for the other groups, and actually tended towards the F pole on two facets. Also of note was that mean scores for the 'Sales, customer service' group tended further towards the T pole than any of the other groups on four of the five facet scales.

Table 6.9: Mean facet scale scores by work area

Step II facet scale	Finance (n=2,511)		Business services (n=1,351)		HR, training, guidance (n=1,299)		Sales, customer service (n=967)		Science, engineering (n=938)		Sig.
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales											
Initiating–Receiving	-0.95	3.28	-1.08	3.19	-1.25	3.18	-1.84	2.97	-0.63	3.18	**
Expressive–Contained	-1.18	2.62	-1.33	2.56	-1.80	2.58	-1.46	2.53	-0.95	2.63	**
Gregarious–Intimate	-0.51	2.52	-0.50	2.67	-0.48	2.62	-0.92	2.40	-0.32	2.70	**
Active–Reflective	-1.58	3.01	-1.65	2.92	-1.63	2.91	-2.24	2.65	-1.31	2.90	**
Enthusiastic–Quiet	-1.86	2.63	-1.94	2.62	-2.14	2.56	-2.31	2.39	-1.49	2.62	**
S–N facet scales											
Concrete–Abstract	-0.61	2.30	-0.30	2.37	0.13	2.43	-0.57	2.30	-0.75	2.31	**
Realistic–Imaginative	-1.46	2.80	-0.85	3.03	-0.11	3.15	-1.32	2.94	-1.50	2.84	**
Practical–Conceptual	-1.17	2.56	-0.69	2.64	-0.75	2.73	-1.30	2.46	-0.83	2.50	**
Experiential–Theoretical	-0.92	2.52	-0.54	2.58	-0.13	2.65	-0.98	2.35	-0.52	2.46	**
Traditional–Original	-0.50	2.88	-0.11	2.98	0.03	2.96	-0.56	2.85	-0.31	2.96	**
T–F facet scales											
Logical–Empathetic	-1.27	2.74	-1.01	2.84	0.44	2.95	-0.85	2.79	-1.84	2.57	**
Reasonable–Compassionate	-2.49	2.61	-2.27	2.70	-1.02	2.82	-2.39	2.67	-2.99	2.43	**
Questioning–Accommodating	-1.85	2.75	-1.86	2.79	-1.51	2.77	-1.77	2.67	-1.83	2.62	**
Critical–Accepting	-0.77	2.00	-0.68	2.07	0.03	2.12	-0.79	1.96	-1.09	2.12	**
Tough–Tender	-1.08	2.93	-0.88	3.09	0.40	3.00	-1.17	2.94	-1.47	2.95	**
J–P facet scales											
Systematic–Casual	-2.11	2.39	-1.85	2.48	-1.09	2.75	-2.08	2.35	-2.56	2.25	**
Planful–Open-Ended	-1.05	3.16	-0.70	3.19	-0.30	3.40	-1.11	3.05	-1.14	2.96	**
Early Starting–Pressure-Prompted	-0.27	3.32	-0.22	3.26	0.06	3.42	-0.32	3.31	-0.33	3.22	*
Scheduled–Spontaneous	-2.00	2.51	-1.92	2.59	-1.16	2.86	-2.26	2.43	-2.34	2.47	**
Methodical–Emergent	-1.27	2.61	-1.23	2.61	-0.60	2.81	-1.46	2.62	-1.52	2.59	**

Significant at: *p<0.05, **p<0.01 (based on a one-way analysis of variance).

Nationality

Information on nationality was available for the group. Seventy two per cent of the group were German, 20% were Swiss and 3% were Austrian.

A comparison of mean facet scales across the three main nationality groups highlighted a number of significant differences, as shown in Table 6.10 and Figure 6.3, although in most cases these were too small to be of any practical significance.

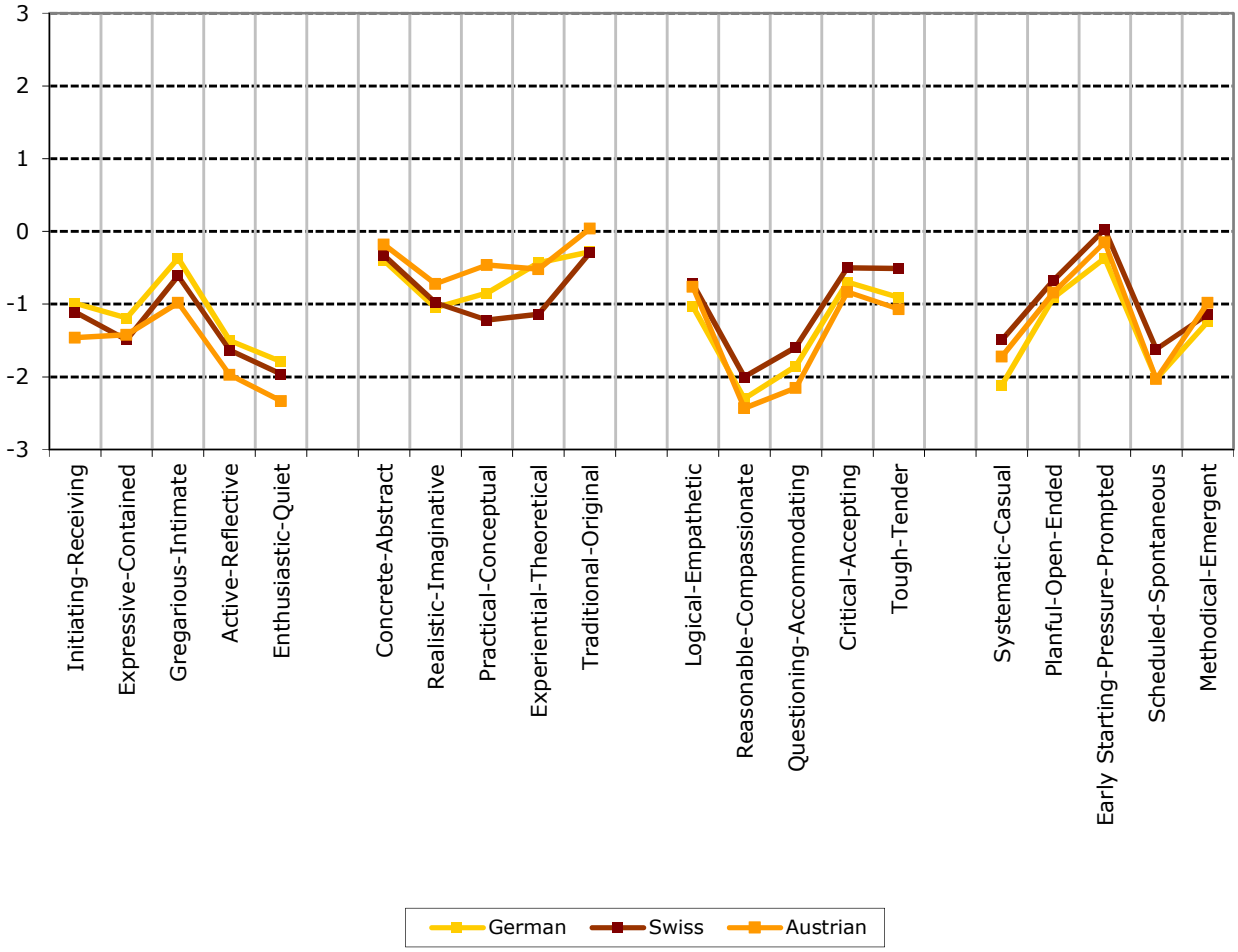
- The largest difference was found on the Experiential–Theoretical facet of the S–N preference block, where the mean score for the Swiss group was further towards the Experiential (S) pole than the German and Austrian groups.
- On the facet scales linked to the T–F dimension, the Swiss group mean scores tended further towards the F pole than those for the German and Austrian groups.

Table 6.10: Mean facet scale scores by nationality

Step II facet scale	German (n=8,921)		Swiss (n=2,534)		Austrian (n=387)		Sig.
	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales							
Initiating–Receiving	-0.99	3.23	-1.11	3.22	-1.46	3.19	**
Expressive–Contained	-1.19	2.61	-1.49	2.57	-1.42	2.59	**
Gregarious–Intimate	-0.37	2.59	-0.61	2.60	-0.98	2.42	**
Active–Reflective	-1.50	2.96	-1.63	2.91	-1.97	2.83	**
Enthusiastic–Quiet	-1.79	2.62	-1.96	2.58	-2.33	2.37	**
S–N facet scales							
Concrete–Abstract	-0.40	2.37	-0.33	2.30	-0.18	2.13	
Realistic–Imaginative	-1.05	3.00	-0.98	2.92	-0.72	2.93	
Practical–Conceptual	-0.85	2.56	-1.22	2.52	-0.46	2.43	**
Experiential–Theoretical	-0.43	2.53	-1.14	2.38	-0.52	2.46	**
Traditional–Original	-0.28	2.92	-0.29	2.89	0.04	2.65	
T–F facet scales							
Logical–Empathetic	-1.03	2.87	-0.72	2.89	-0.76	2.95	**
Reasonable–Compassionate	-2.30	2.73	-2.00	2.76	-2.43	2.67	**
Questioning–Accommodating	-1.85	2.72	-1.59	2.78	-2.15	2.57	**
Critical–Accepting	-0.70	2.12	-0.50	2.01	-0.83	1.88	*
Tough–Tender	-0.91	3.02	-0.51	3.11	-1.07	2.92	**
J–P facet scales							
Systematic–Casual	-2.11	2.42	-1.49	2.56	-1.72	2.41	**
Planful–Open-Ended	-0.91	3.11	-0.67	3.22	-0.84	3.07	*
Early Starting–Pressure-Prompted	-0.37	3.28	0.03	3.33	-0.15	3.18	**
Scheduled–Spontaneous	-2.03	2.56	-1.62	2.58	-2.03	2.31	**
Methodical–Emergent	-1.24	2.66	-1.14	2.70	-0.98	2.75	

Significant at: *p<0.05, **p<0.01 (based on a one-way analysis of variance).

Figure 6.3: Mean facet scale scores by nationality



Employment status

Employment status information was available for the OPPassessment sample. The vast majority of the group worked either full-time or part-time, or were self-employed. Table 6.11 shows the mean facet scale scores for each group, with the data illustrated in graphical form in Figure 6.4.

The analyses revealed statistically significant differences across the groups on all the facet scales on the S–N, T-F and J–P dimensions and three facets on the E–I dimension. The clearest patterns were as follows:

- There were no clear patterns regarding the facets in the E–I preference block.
- The self-employed group tended to score more towards the N pole on the S–N facet scales than the other two groups, and less towards the J pole on the J–P facet scales.
- The full-time group tended to score more towards the T pole on the T–F facets than the other two groups. This is likely to be at least

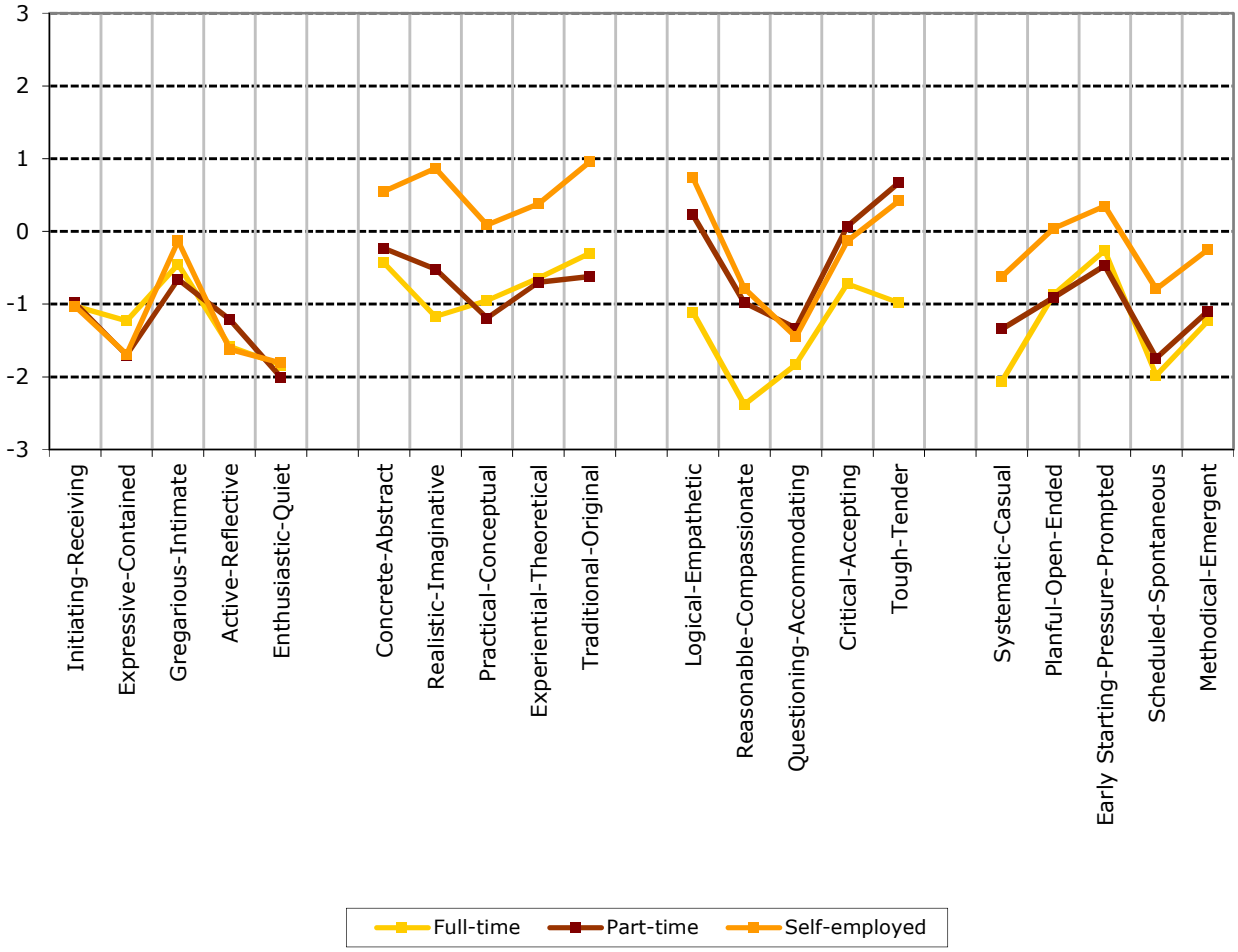
partly a gender effect; 84% of part-time workers and 49% of self-employed people were female, compared with 32% of the total group and 28% of full-time workers.

Table 6.11: Mean facet scale scores by employment status

Step II facet scale	Full–time (n=10,200)		Part–time (n=666)		Self–employed (n=469)		Sig.
	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales							
Initiating–Receiving	-1.03	3.22	-0.98	3.24	-1.03	3.25	
Expressive–Contained	-1.23	2.60	-1.70	2.51	-1.69	2.77	**
Gregarious–Intimate	-0.45	2.59	-0.66	2.54	-0.13	2.76	**
Active–Reflective	-1.58	2.94	-1.21	2.99	-1.62	2.83	**
Enthusiastic–Quiet	-1.84	2.61	-2.01	2.44	-1.81	2.69	
S–N facet scales							
Concrete–Abstract	-0.43	2.36	-0.23	2.35	0.55	2.33	**
Realistic–Imaginative	-1.17	2.96	-0.52	3.02	0.87	2.93	**
Practical–Conceptual	-0.95	2.55	-1.19	2.57	0.09	2.59	**
Experiential–Theoretical	-0.64	2.52	-0.70	2.36	0.38	2.57	**
Traditional–Original	-0.30	2.91	-0.62	2.95	0.96	2.89	**
T–F facet scales							
Logical–Empathetic	-1.11	2.83	0.23	2.91	0.74	3.02	**
Reasonable–Compassionate	-2.38	2.67	-0.98	2.93	-0.78	2.92	**
Questioning–Accommodating	-1.83	2.71	-1.33	2.92	-1.45	2.73	**
Critical–Accepting	-0.72	2.09	0.07	1.98	-0.13	2.20	**
Tough–Tender	-0.98	3.01	0.67	3.02	0.42	3.07	**
J–P facet scales							
Systematic–Casual	-2.06	2.44	-1.34	2.59	-0.62	2.61	**
Planful–Open-Ended	-0.87	3.14	-0.91	3.23	0.04	3.29	**
Early Starting–Pressure-Prompted	-0.26	3.28	-0.47	3.32	0.35	3.38	**
Scheduled–Spontaneous	-1.98	2.55	-1.74	2.57	-0.79	2.83	**
Methodical–Emergent	-1.23	2.67	-1.10	2.63	-0.25	2.85	**

Significant at: *p<0.05, **p<0.01 (based on a one-way analysis of variance).

Figure 6.4: Mean facet scale scores by employment status



Appendix 1: Sample description

Sample 1: Data from OPPassessment (representative German-speaking professional and managerial sample)

This sample consists of 15,228 individuals who completed the MBTI Step II instrument in German via the OPPassessment system between August 2006 and July 2016. Sixty-seven per cent of the respondents were male and 33% were female. Age ranged from 16 to 70 years, with a mean and median of 40.

Nationality was disclosed by 81% of respondents. Of these, 72% were German, 20% were Swiss and 3% were Austrian. Many other nationalities were represented, but each one formed less than 1% of the total group:

Nationality	Percentage
German	71.9%
Swiss	20.4%
Austrian	3.1%
Other	4.6%

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

Employment status	Percentage
Full-time	89.5%
Part-time	5.8%
Self-employed	4.1%
Unemployed	0.3%
Homemaker	0.2%
Retired	0.1%

The majority of the group were of managerial level or above, although the largest single group was employee (30%):

Occupational level	Percentage
Top level	4.9%
Senior executive	19.1%
Upper middle management	12.0%
Middle management	19.6%
First level management/supervisor	11.4%
Employee	29.6%
Other	3.5%

A range of work areas were represented:

Work area (job type)	Percentage
Finance	22.9%
Business services	12.3%
HR, training, guidance	11.8%
Sales, customer service	8.8%
Science, engineering	8.5%
IT	7.7%
Research and development	5.5%
Health, social services, etc.	3.1%
Admin or secretarial	2.4%
Education	1.8%
Land, sea or air transport	0.6%
Skilled operative	0.4%
Leisure, personal service	0.2%
Unskilled operative	0.1%
Other private sector	1.5%
Other public sector	0.9%
Other	11.5%

MBTI® Step II™ instrument

European Data Supplement

Italian

January 2018

2nd edition

unlocking potential

OPP Ltd
Elsfield Hall
15–17 Elsfield Way
Oxford OX2 8EP
UK
t: +44 (0)1865 404 500

Contents

Introduction

Data collected for the European MBTI® Step II™ instrument were analysed to produce the findings in this supplement. This is the second data supplement produced for this instrument. A brief description of the sample is given below. Further details of the sample are provided in Appendix 1.

- The sample consisted of 886 individuals who completed the MBTI Step II questionnaire in Italian via the OPPassessment™ system between 2006 and mid-2016.[1] This sample is considered to be representative of the groups of people with whom the Italian MBTI Step II instrument has been and will be used for applications such as management development, coaching, counselling and teambuilding. As such, it is likely to represent a cross-section of the Italian-speaking professional and managerial population.

The results of the analyses are outlined below.

[1] OPPassessment allows personality questionnaires such as the MBTI instrument to be administered via email and/or completed online.

Type distribution

Type tables are a way of illustrating the proportion of each type within a particular group. Below is a type table taken from the sample described on the previous page.

For each of the 16 different types, the number of cases, the percentage of the total that this represents and the self-selection ratio (SSR) are shown. The SSR (Myers et al., 1998) is a way of demonstrating whether a given type appears more or less often in a particular group than would be expected given its frequency in a reference group. Ideally, the type distribution from a large representative sample of the Italian population would be used to calculate SSRs in this data supplement. However, such a sample does not currently exist. In its place, SSRs have been calculated using type data from the UK general population (Kendall, 1998), which can be justified by the fact that type distributions for comparable Italian and British groups, such as managers and professionals, are similar. Evidence (eg Hackston and Kendall, 2004; Quenk et al, 2004) does suggest that although type-related behaviours vary a good deal from country to country and from culture to culture, the frequencies of underlying MBTI types do not.

An SSR of greater than 1 indicates that a type is over-represented, and an SSR of less than 1 denotes that it is under-represented. Asterisks are used to denote whether the over- or under-representations are statistically significant, based on the results of chi-square analysis.[2]

[2] Chi-square analysis (often abbreviated to χ^2) is a technique used to explore whether observed frequency distributions differ significantly from other, predefined, distributions. In this case, the UK general population group is used as the reference group, and the chi-square analysis indicates whether the proportion of people of each type within a particular sample differs significantly from the proportion of people reporting the same type within the reference group.

OPPassessment data (representative Italian-speaking professional and managerial sample)

Table 7.1: Type table for OPPassessment data (reported type, n=886)

ISTJ	ISFJ	INFJ	INTJ
n=114 12.9% SSR=0.94	n=22 2.5% SSR=0.20**	n=22 2.5% SSR=1.47	n=63 7.1% SSR=5.07**
ISTP	**ISFP**	**INFP**	**INTP**
n=22 2.5% SSR=0.39**	n=9 1.0% SSR=0.16**	n=21 2.4% SSR=0.75	n=34 3.8% SSR=1.58
ESTP	**ESFP**	**ENFP**	**ENTP**
n=25 2.8% SSR=0.48**	n=16 1.8% SSR=0.21**	n=45 5.1% SSR=0.81	n=67 7.6% SSR=2.71**
ESTJ	**ESFJ**	**ENFJ**	**ENTJ**
n=200 22.6% SSR=2.17**	n=52 5.9% SSR=0.47**	n=48 5.4% SSR=1.93**	n=126 14.2% SSR=4.90**

Type	n	%
E	579	65.3%**
I	307	34.7%**
S	460	51.9%**
N	426	48.1%**
T	651	73.5%**
F	235	26.5%**
J	647	73.0%**
P	239	27.0%**

*Difference significant at p<0.05, based on chi-square results.
**Difference significant at p<0.01, based on chi-square results.

The most common single type preference is ESTJ (23% of the total); this is a common finding with managerial groups in other countries. The SSR results suggest that, in comparison with the UK general population, those with preferences for NT are over-represented, and those with preferences for SF are under-represented. Again, this is a common finding with managerial groups.

Properties of the Step II facet scales

Facet scale score distribution

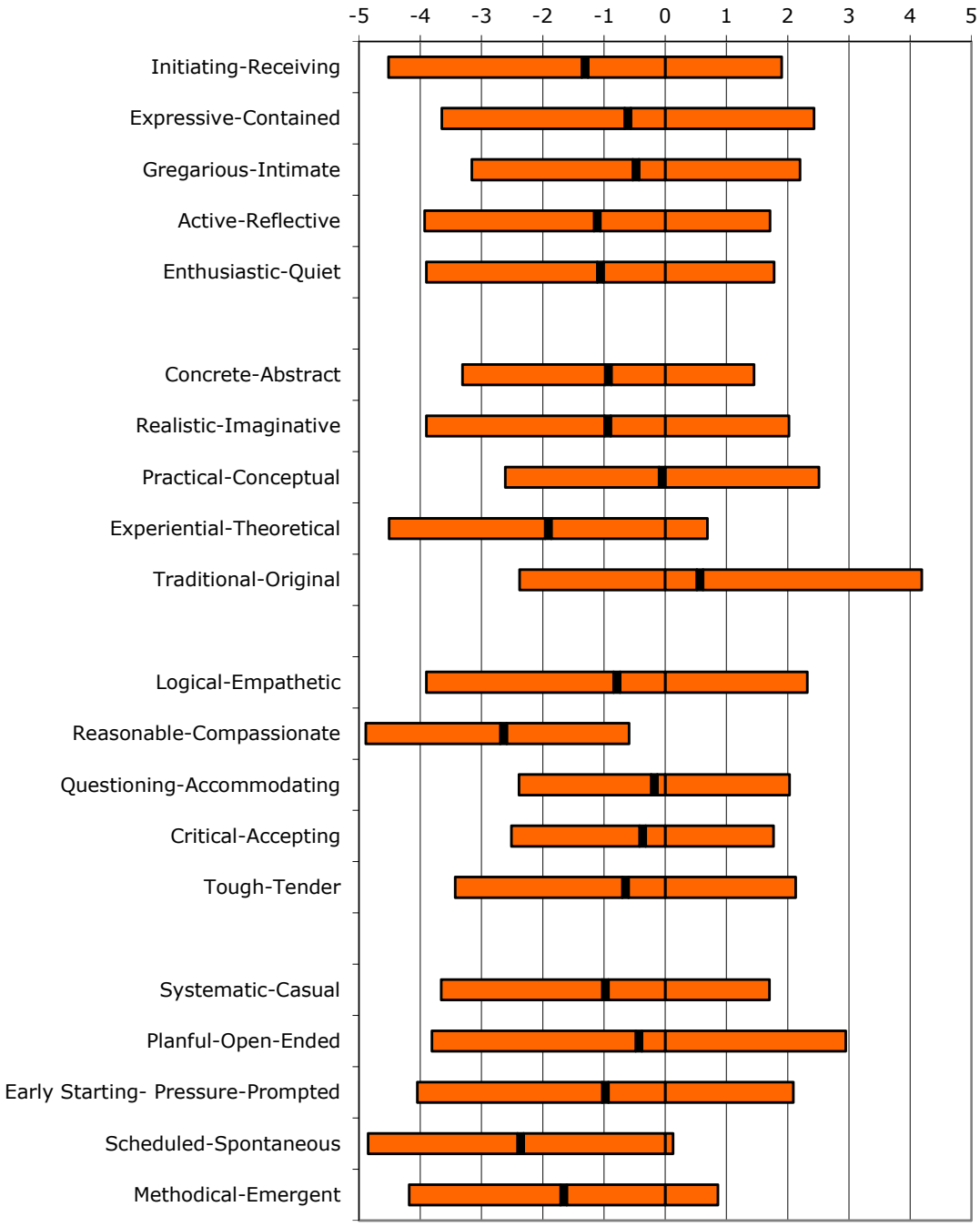
The means and standard deviations of the Step II facets are shown in Table 7.2 and illustrated graphically in Figure 7.1. The mean score for a scale is calculated by adding together the scores of each individual in the sample and then dividing the sum by the number of individuals. Note that negative values indicate that mean scores are nearer the left-hand pole of the facet and positive values indicate that the mean scores are nearer the right-hand pole. The standard deviation (SD) is a statistical measure describing the degree to which the scores from the sample either bunch up close to, or are scattered widely around, the mean for the sample.

Table 7.2: Means and standard deviations of the facet scales

Step II facet scale	Mean[3]	SD
E–I facet scales		
Initiating–Receiving	–1.31	3.21
Expressive–Contained	–0.61	3.04
Gregarious–Intimate	–0.48	2.68
Active–Reflective	–1.11	2.82
Enthusiastic–Quiet	–1.06	2.84
S–N facet scales		
Concrete–Abstract	–0.93	2.38
Realistic–Imaginative	–0.94	2.96
Practical–Conceptual	-0.05	2.56
Experiential–Theoretical	–1.91	2.60
Traditional–Original	0.57	2.95
T–F facet scales		
Logical–Empathetic	–0.79	3.11
Reasonable–Compassionate	–2.77	2.44
Questioning–Accommodating	–0.18	2.21
Critical–Accepting	–0.37	2.14
Tough–Tender	–0.65	2.78
J–P facet scales		
Systematic–Casual	–0.98	2.68
Planful–Open-Ended	–0.43	3.38
Early Starting–Pressure-Prompted	–0.98	3.07
Scheduled–Spontaneous	–2.36	2.49
Methodical–Emergent	–1.66	2.52

[3] Scale means: in Step II reports, scores are given from 5 on one pole, through 0, to 5 on the opposite pole. In compiling statistical information, however, one pole needs to be clearly distinguished from the other. Throughout this data supplement, therefore, a convention has been adopted of scores 5, 4, 3, 2 or 1 on the left-hand pole being assigned values of –5, –4, –3, –2, or –1 and scores on the right-hand pole being assigned positive scores. This does not, of course, imply any suggestion that positive scores are 'better' than negative scores.

Figure 7.1: Facet scale mean scores[4]



[4] For each facet scale, the central line shows the mean and the coloured bars either side indicate the standard deviation.

In general, the means centre quite close to the midpoint of each scale. Only two of the scales have means greater than 2 points from the midpoint in either direction, and both have a standard deviation of between 2.4 and 2.5. However, a clear pattern does emerge. With one exception (Traditional–Original), the mean scores tend towards the E, S, T and J direction, which is consistent with the most common four-letter type preference amongst this sample. Amongst the group as a whole, there are more individuals with a preference for E, S, T and J.

Facet scale intercorrelations

Correlations among the Step II facet scales are shown in Table 7.3. Facet scales within each dichotomy usually correlate more highly (often substantially so) with the other scales of the same dichotomy than they do with scales in the other three dichotomies. For example, every E–I facet scale correlates more highly with the other four E–I scales than with any of the S–N, T–F or J–P facet scales.

There is only one scale that correlates **significantly** more highly with scales in other dichotomies. The T–F scale Questioning–Accommodating correlates with many scales, across dichotomies. For example, it correlates at a level of –0.12 with Active–Reflective, at 0.10 with Enthusiastic–Quiet, and at 0.09 with Initiating–Receiving on the E–I dichotomy. It correlates at between –0.13 and –0.23 with all of the S–N scales except Realistic–Imaginative on the S–N dichotomy. It also correlates at a level of -0.11 with Planful–Open-Ended, at -0.12 with Scheduled–Spontaneous and at -0.13 with Early Starting–Pressure-Prompted on the J–P dichotomy. By way of comparison, it correlates at between 0.07 and 0.20 with the other four scales on the same dichotomy as itself (T–F). The negative correlations between Questioning–Accommodating and the S–N scales are consistent with findings with the US version of the Step II instrument, and suggest that a questioning approach to differences of opinion seems to be related to a range of iNtuition (N) facets. These patterns are similar to those found when the European Step II instrument was developed.

Table 7.3: Intercorrelations of Step II facet scales

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Initiating–Receiving	**1.00**																			
2. Expressive–Contained	**0.48**	**1.00**																		
3. Gregarious–Intimate	**0.54**	**0.47**	**1.00**																	
4. Active–Reflective	**0.69**	**0.43**	**0.51**	**1.00**																
5. Enthusiastic–Quiet	**0.61**	**0.58**	**0.58**	**0.57**	**1.00**															
6. Concrete–Abstract	0.00	-0.08	-0.02	0.01	-0.11	**1.00**														
7. Realistic–Imaginative	-0.04	-0.13	-0.10	-0.06	-0.19	**0.57**	**1.00**													
8. Practical–Conceptual	-0.06	-0.09	-0.05	-0.06	-0.14	**0.60**	**0.51**	**1.00**												
9. Experiential–Theoretical	0.14	0.02	0.12	0.13	0.11	**0.55**	**0.34**	**0.44**	**1.00**											
10. Traditional–Original	-0.16	-0.15	-0.12	-0.14	-0.24	**0.53**	**0.53**	**0.53**	**0.30**	**1.00**										
11. Logical–Empathetic	-0.05	-0.20	-0.05	0.00	-0.12	0.22	0.39	0.15	0.04	0.20	**1.00**									
12. Reasonable–Compassionate	-0.01	-0.16	-0.03	0.03	-0.06	0.29	0.38	0.15	0.16	0.22	**0.59**	**1.00**								
13. Questioning–Accommodating	0.09	0.04	0.05	0.12	0.10	-0.13	-0.06	-0.17	-0.13	-0.23	**0.09**	**0.07**	**1.00**							
14. Critical–Accepting	-0.08	-0.16	-0.12	-0.07	-0.09	0.18	0.29	0.11	0.01	0.19	**0.46**	**0.44**	**0.20**	**1.00**						
15. Tough–Tender	0.02	-0.13	-0.03	0.06	-0.10	0.24	0.34	0.12	0.13	0.13	**0.59**	**0.58**	**0.14**	**0.42**	**1.00**					
16. Systematic–Casual	-0.04	-0.17	-0.08	-0.05	-0.15	0.41	0.48	0.30	0.18	0.46	0.50	0.44	-0.04	0.34	0.39	**1.00**				
17. Planful–Open-Ended	0.02	-0.05	-0.05	-0.04	-0.05	0.31	0.30	0.25	0.17	0.43	0.20	0.21	-0.11	0.13	0.13	**0.59**	**1.00**			
18. Early Starting–Pressure-Prompted	-0.07	-0.06	-0.11	-0.08	-0.13	0.23	0.23	0.19	0.12	0.32	0.10	0.14	-0.13	0.02	0.06	**0.39**	**0.47**	**1.00**		
19. Scheduled–Spontaneous	-0.02	-0.13	-0.09	-0.06	-0.10	0.28	0.30	0.22	0.18	0.40	0.20	0.27	-0.12	0.11	0.14	**0.54**	**0.60**	**0.50**	**1.00**	
20. Methodical–Emergent	0.04	-0.04	-0.04	0.05	-0.05	0.23	0.25	0.18	0.15	0.22	0.15	0.22	-0.03	0.13	0.15	**0.45**	**0.48**	**0.44**	**0.53**	**1.00**

Correlations of Step II facet scales with Step I scales

Correlations between Step II facet scales and the continuous scores from the MBTI Step I™ instrument are shown in Table 7.4.[5] Positive correlations between facet scales and the corresponding Step I dimension (eg between Initiating–Receiving and the E–I continuous score) indicate relationships in the expected direction, with higher facet scores tending to be associated with higher continuous scores, and vice versa. Negative correlations would indicate an inverse relationship between the facet scale scores and the corresponding Step I dimension.

Table 7.4: Correlations of Step II facet scales with Step I continuous scores

	Step I continuous score			
Step II facet scales	**E–I**	**S–N**	**T–F**	**J–P**
E–I facet scales				
Initiating–Receiving	**0.83**	-0.03	-0.06	0.02
Expressive–Contained	**0.71**	-0.08	-0.23	-0.08
Gregarious–Intimate	**0.66**	-0.03	-0.10	-0.07
Active–Reflective	**0.76**	-0.03	-0.01	-0.01
Enthusiastic–Quiet	**0.81**	-0.13	-0.15	-0.07
S–N facet scales				
Concrete–Abstract	-0.06	**0.82**	0.26	0.37
Realistic–Imaginative	-0.13	**0.68**	0.42	0.38
Practical–Conceptual	-0.11	**0.67**	0.14	0.28
Experiential–Theoretical	0.14	**0.62**	0.07	0.21
Traditional–Original	-0.21	**0.69**	0.20	0.46
T–F facet scales				
Logical–Empathetic	-0.12	0.19	**0.87**	0.25
Reasonable–Compassionate	-0.06	0.26	**0.72**	0.31
Questioning–Accommodating	0.12	-0.20	**0.18**	-0.12
Critical–Accepting	-0.13	0.15	**0.60**	0.16
Tough–Tender	-0.04	0.19	**0.74**	0.21
J–P facet scales				
Systematic–Casual	-0.13	0,41	0.51	**0.69**
Planful–Open-Ended	-0.04	0.36	0.21	**0.85**
Early Starting–Pressure-Prompted	-0.12	0.27	0.10	**0.67**
Scheduled–Spontaneous	-0.09	0.35	0.21	**0.75**
Methodical–Emergent	0.02	0.25	0.20	**0.66**

The E–I facet scales correlate at 0.66 to 0.83 with the E–I continuous scores from Step I; the S–N facet scales correlate at 0.62 to 0.82 with the S–N continuous scores; the T–F facet scales correlate at 0.18 to 0.87 with the T–F continuous scores; and the J–P facet scales correlate

[5] Continuous scores place an individual's score on each dimension onto a continuous scale with a midpoint of 100. To calculate continuous scores, Preference Clarity Index (PCI) scores for each dimension are either subtracted or added to 100, depending on which direction the overall preference is. PCI scores in the direction of E, S, T or J are subtracted from 100. PCI scores in the direction of I, N, F or P are added to 100.

at 0.66 to 0.85 with the J–P continuous scores. The scale that showed the lowest correlation with its associated Step I continuous score was Questioning–Accommodating at 0.18. This scale is also lowest on internal consistency (see page 161), and has been previously found to have the lowest test-retest reliability (Quenk, Hammer and Majors, 2004), which would result in the true correlations being underestimated.

These correlations are very similar to those found during the development of the Step II questionnaire. This consistency, alongside the fact that the correlations between Step II facet scales and Step I continuous scores associated with their own dimension are substantially higher than correlations with the other three dimensions, provides compelling evidence for the theoretical hierarchical structure of the Step II facet scales in relation to the Step I scales.

Out-of-preference scores

It is known that the five facet scales relating to each type dimension do not describe the dimension in its entirety; there will not be a precise and exact overlap between, for example, an individual's score on E–I and their total score across the five facet scales that relate to this dimension. For example, it is not uncommon to see an Enthusiastic Introvert or an Intimate Extravert. Such apparent inconsistencies are usually known as 'out-of-preference scores' (OOPS).

Although it is usual to have a number of OOPS in any one profile, it should be unusual to find that OOPS make up more than half of the facet scores relating to any one type dichotomy. Table 7.5 demonstrates that the proportion of individuals for whom this occurs ranges from 3.9% for the S–N block to 1.0% for the E–I block. It is therefore a very infrequent occurrence.

Table 7.5: Proportion of OOPS by type dichotomy[6]

	Proportion of 'reported type' OOPS					
	None	One	Two	Three	Four	Five
E–I	67.6%	24.5%	6.9%	1.0%	0.0%	0.0%
S–N	52.5%	30.7%	13.0%	3.2%	0.7%	0.0%
T–F	56.5%	31.9%	9.4%	1.9%	0.2%	0.0%
J–P	61.1%	27.1%	9.8%	2.0%	0.0%	0.0%

For any individual facet scale, an OOPS tends to occur in approximately 10% of profiles, with the exception of Experiential–Theoretical (24%) and Questioning–Accommodating (20%).

[6] In this table, an OOPS is defined as a score of 2, 3, 4 or 5 on the 'wrong' pole of the facet when compared with the reported type.

Reliability

The reliability of a test or questionnaire relates to how consistent and precise it is. Internal consistency reliability addresses the question of whether all the questions in a scale measure the same construct. For example, are the Step II facet scales consistent within themselves, and do they hold together well as scales? A common measure of internal consistency reliability is coefficient alpha (Cronbach, 1951). The alpha coefficients for the Step II facet scales are shown in Table 7.6.

Table 7.6: Internal consistency reliability

Step II facet scale	No. of items	Coefficient alpha
E–I facet scales		
Initiating–Receiving	8	0.82
Expressive–Contained	7	0.74
Gregarious–Intimate	7	0.64
Active–Reflective	8	0.64
Enthusiastic–Quiet	9	0.67
S–N facet scales		
Concrete–Abstract	9	0.67
Realistic–Imaginative	7	0.69
Practical–Conceptual	8	0.50
Experiential–Theoretical	8	0.74
Traditional–Original	8	0.71
T–F facet scales		
Logical–Empathetic	9	0.84
Reasonable–Compassionate	8	0.65
Questioning–Accommodating	7	0.34
Critical–Accepting	8	0.45
Tough–Tender	8	0.73
J–P facet scales		
Systematic–Casual	8	0.72
Planful–Open-Ended	6	0.79
Early Starting–Pressure-Prompted	6	0.71
Scheduled–Spontaneous	8	0.61
Methodical–Emergent	8	0.59
	Median	0.67

The internal consistency reliability of most scales is good, and the average (median) reliability is close to 0.7. However, three scales (Practical–Conceptual, Questioning–Accommodating and Critical–Accepting) do have lower alpha coefficients. This pattern is consistent with what was found during the development of the Step II instrument, where these facet scales were also found to show lower reliability than the others (Quenk, Hammer and Majors, 2004).

It should be noted, however, that coefficient alpha reliability statistics will give an underestimate of the reliability of the Step II facet scales,

given the use of item response theory (IRT) methods in the actual scoring process.[7]

Group differences

This section looks at the way in which people from different groups responded to the Italian version of the European Step II questionnaire. Unfortunately, the total number of respondents is too small for the full range of analyses to be conducted. Further research will be conducted when additional data become available.

Gender

The means and standard deviations of the Step II facet scales are shown separately for males and females in Table 7.7, along with the difference in mean scores. This is represented graphically in Figure 7.2. Statistically significant differences were found between male and female mean scores for thirteen of the 20 facet scales, with a consistent pattern emerging.

- On the E–I facet scales, all five mean scores tended slightly toward the E pole for both males and females. Female mean scores tended further towards the E pole than males on three of the facet scales, with there being a significant different on two of these, and there was no significant difference between genders on the remaining facet scale.
- On the S–N facet scales, three mean scores tended slightly toward the S pole for both males and females, one mean score tended slightly toward the N pole for both males and females, and one mean score (Practical–Conceptual) tended slightly toward the S pole for males and the N pole for females. There were statistically significant gender differences on all five facet scales, with male mean scores tending more towards the S pole than females on four out of five scales.
- On the T–F facet scales, all five mean scores tended slightly toward the T pole for males, whereas two did so for females. There were statistically significant gender differences on all the five facet scales, with male mean scores tending more towards the T pole than females.
- On the J–P facet scales, all five mean scores tended toward the J pole for both males and females. There was a statistically significant gender difference on the Systematic–Casual facet scale, with males mean scores tending more towards the J pole than females. There were no statistically significant gender differences on any of the four facet scales.

[7] Item response theory (IRT) is an approach to measurement that is concerned with modelling the relationship between item responses and the underlying characteristic assessed by the scale or test the item is designed to measure. IRT can be used to select items for a test and/or to score the items.

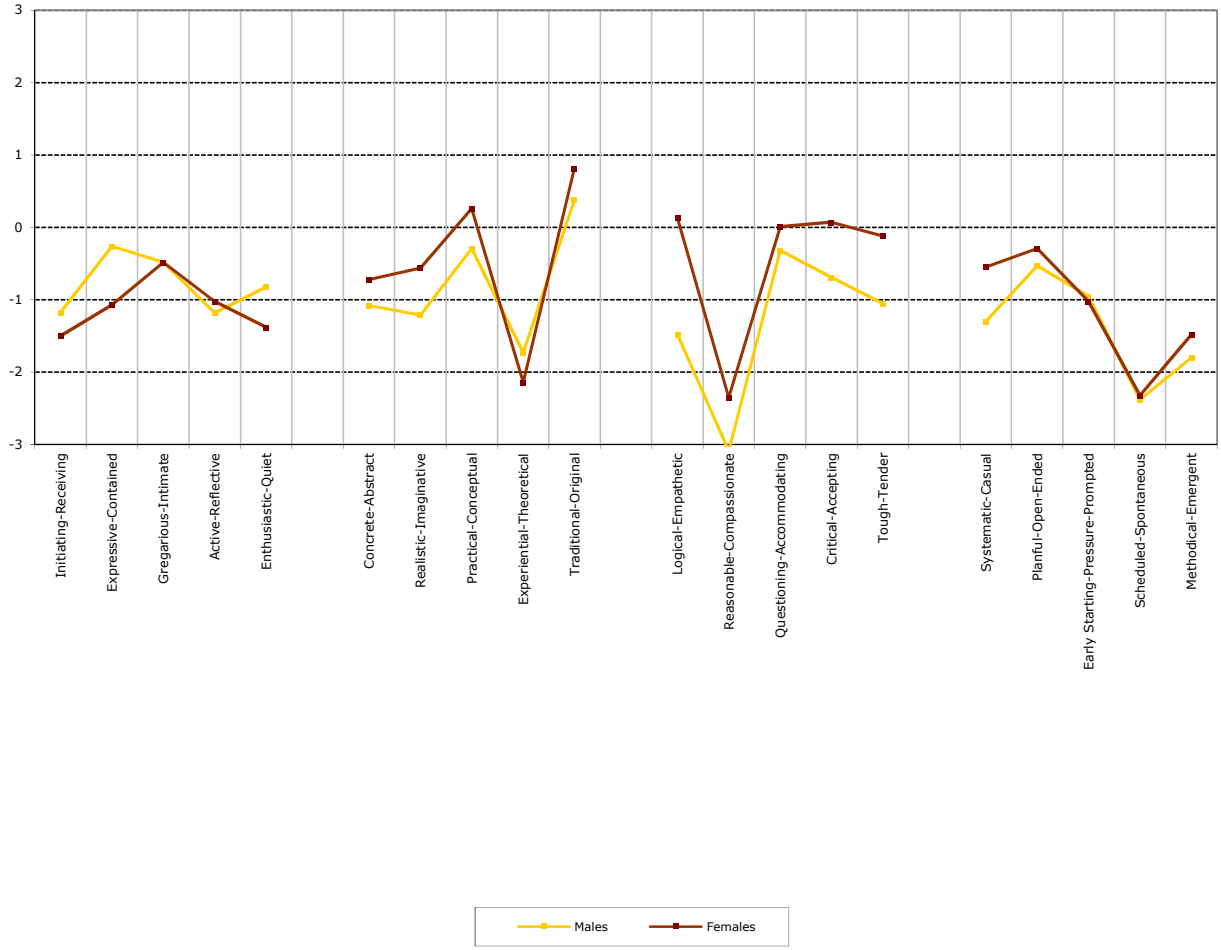
Table 7.7: Gender differences in facet scale scores

	Males (n=507)		Females (n=379)		Difference (M–F)[8]
Step II facet scale	**Mean**	**SD**	**Mean**	**SD**	
E–I facet scales					
Initiating–Receiving	-1.17	3.25	-1.50	3.14	0.33
Expressive–Contained	-0.26	3.05	-1.07	2.96	0.82**
Gregarious–Intimate	-0.48	2.78	-0.48	2.53	0.00
Active–Reflective	-1.18	2.83	-1.02	2.80	-0.16
Enthusiastic–Quiet	-0.82	2.87	-1.38	2.78	0.56**
S–N facet scales					
Concrete–Abstract	-1.08	2.33	-0.72	2.43	–0.37*
Realistic–Imaginative	-1.21	2.90	-0.56	3.01	–0.65**
Practical–Conceptual	-0.29	2.49	0.26	2.63	–0.55**
Experiential–Theoretical	-1.73	2.69	-2.14	2.47	0.41*
Traditional–Original	0.38	2.93	0.82	2.97	–0.44*
T–F facet scales					
Logical–Empathetic	-1.48	2.89	0.13	3.16	–1.61**
Reasonable–Compassionate	-3.09	2.23	-2.35	2.63	–0.75**
Questioning–Accommodating	-0.32	2.23	0.01	2.17	–0.33*
Critical–Accepting	-0.69	2.01	0.07	2.24	–0.76**
Tough–Tender	-1.05	2.73	-0.12	2.76	–0.93**
J–P facet scales					
Systematic–Casual	-1.30	2.67	-0.55	2.63	–0.75**
Planful–Open-Ended	-0.53	3.41	-0.29	3.33	–0.24
Early Starting–Pressure-Prompted	-0.95	3.08	-1.02	3.06	0.07
Scheduled–Spontaneous	-2.38	2.49	-2.32	2.48	–0.06
Methodical–Emergent	-1.80	2.49	-1.48	2.54	–0.32

Difference significant at: *p<0.05, **p<0.01 (based on an independent samples t-test).

[8] A positive value indicates that male scores tend more towards I, N, F or P, and a negative value indicates that female scores tend more towards I, N, F or P.

Figure 7.2: Gender differences in facet scale scores



Age

Correlational analysis showed a number of significant links between age and facet scores, with older people tending more towards the Abstract, Imaginative, Theoretical, Original, Accepting and Casual facets. This is not a pattern that has been found with other language versions.

Ethnic origin

Ethnic origin information was not captured for individuals who completed the Italian language version of the questionnaire, so no analyses were conducted.

Occupational level

Research using the Italian version of the MBTI Step I questionnaire has demonstrated that individuals in employee-level jobs in organisations

are more likely to have preferences for Feeling than those in higher-level jobs.[9]

Although occupational-level data were captured for the Italian sample, the number of people in each of the categories was too small to allow a full analysis. Therefore, individuals were split into three categories, 'Employee', 'Supervisory/management level' and 'Senior executive/top level'. These categories were used for the analysis. Table 7.8 shows the facet scale means and standard deviations for these groups, with the data represented graphically in Figure 7.3. The findings can be summarised as follows:

- A pattern was found regarding the facets relating to the E–I dimension. Mean scores for the 'Employee' level tended less towards the E pole on two of the facets than the other two groups, with a significant difference on the Enthusiastic–Quiet scale. The mean score for the 'Senior executive/top level' group also tended more towards the E pole than the 'Supervisory/management level' on this facet.
- A pattern observed regarding the facets relating to the S–N dimension was that mean scores for three of the facets tended more towards the N pole (or less towards the S pole) for the 'Senior executive/top level' group than the other two groups.
- Regarding the facets relating to the T–F dimension, mean scores for the 'Employee' group tended less towards the T pole on two of the facets than the other two groups.
- Regarding the facets relating to the J–P dimension, mean scores for the 'Employee' group tended less towards the J pole on two of the facets than the other two groups.

These findings should be treated with caution owing to the small sample size on which they are based.

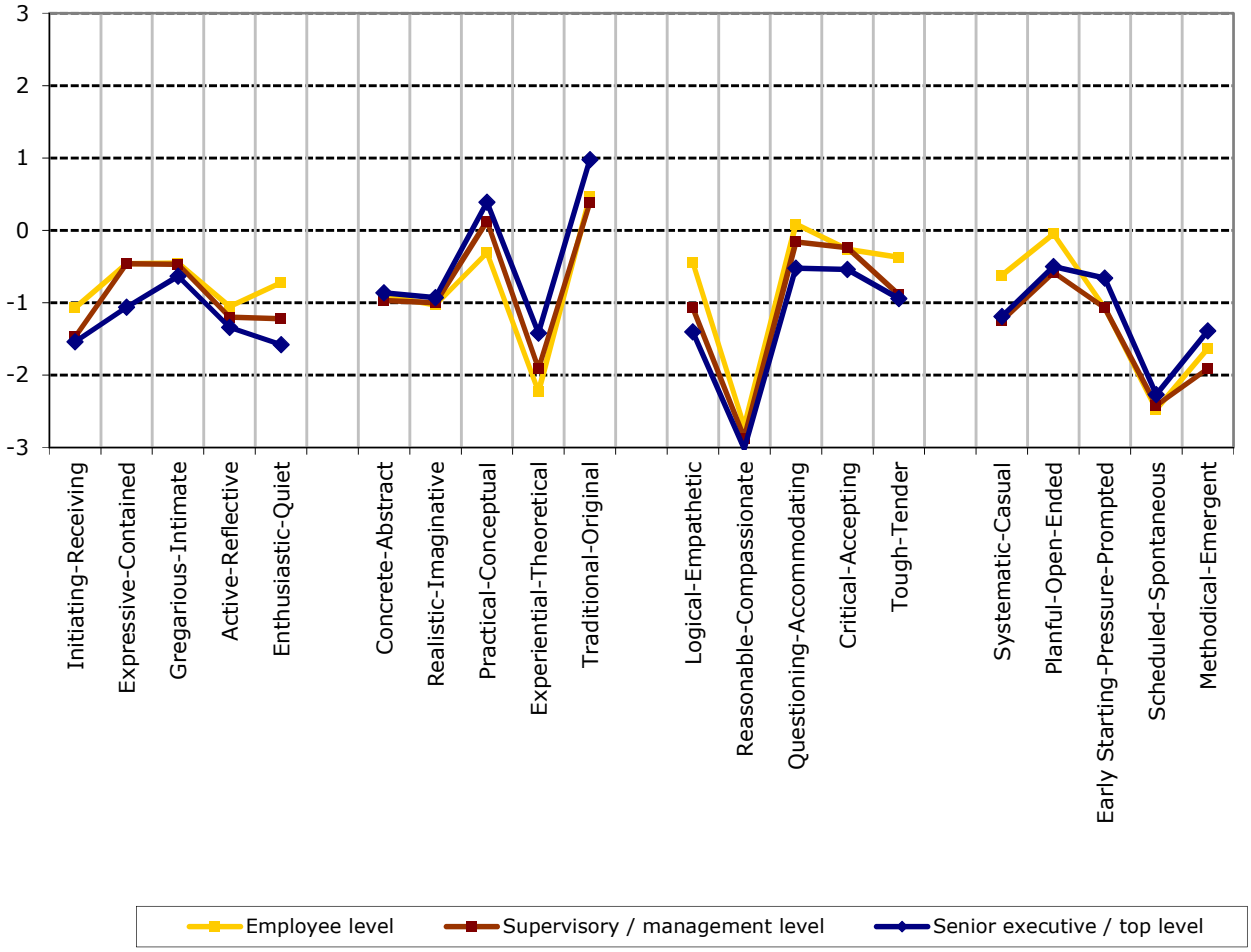
[9] See the MBTI Step I European Data Supplement for details.

Table 7.8: Occupational level differences in facet scale scores

	Employee (n=226)		Supervisory/ management level (n=178)		Senior executive/top level (161)		Sig
Step II facet scale	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	
E–I facet scales							
Initiating–Receiving	-1.07	3.16	-1.47	3.09	-1.54	3.23	
Expressive–Contained	-0.46	2.93	-0.46	3.10	-1.06	3.01	
Gregarious–Intimate	-0.45	2.40	-0.47	2.72	-0.63	2.69	
Active–Reflective	-1.05	2.87	-1.20	2.81	-1.34	2.77	
Enthusiastic–Quiet	-0.72	2.78	-1.22	2.86	-1.58	2.71	**
S–N facet scales							
Concrete–Abstract	-0.90	2.63	-0.97	2.42	-0.86	2.12	
Realistic–Imaginative	-1.03	2.96	-1.00	2.93	-0.93	3.00	
Practical–Conceptual	-0.31	2.74	0.13	2.47	0.39	2.41	*
Experiential–Theoretical	-2.23	2.45	-1.91	2.49	-1.42	2.74	**
Traditional–Original	0.47	3.01	0.39	3.01	0.98	2.78	
T–F facet scales							
Logical–Empathetic	-0.44	3.18	-1.07	2.86	-1.40	2.94	**
Reasonable–Compassionate	-2.71	2.40	-2.87	2.29	-3.02	2.47	
Questioning–Accommodating	0.09	2.35	-0.16	2.05	-0.52	2.11	*
Critical–Accepting	-0.26	2.05	-0.24	1.84	-0.54	2.28	
Tough–Tender	-0.37	2.89	-0.89	2.62	-0.94	2.64	
J–P facet scales							
Systematic–Casual	-0.62	2.66	-1.24	2.64	-1.19	2.71	*
Planful–Open-Ended	-0.05	3.37	-0.58	3.27	-0.50	3.35	
Early Starting–Pressure-Prompted	-1.07	3.21	-1.07	2.93	-0.66	3.07	
Scheduled–Spontaneous	-2.47	2.44	-2.42	2.41	-2.27	2.38	
Methodical–Emergent	-1.63	2.47	-1.91	2.64	-1.39	2.50	

Significant at: *p<0.05, **p<0.01 (based on a one-way analysis of variance).

Figure 7.3: Occupational-level differences in facet scale scores



Education

Specific educational qualifications were not collected for the OPPassessment sample; however, the age at which individuals left full-time education was. Only one significant correlation was found between the age at which people left full-time education and their facet scale scores, this being a correlation of -0.11 with Experiential–Theoretical. Further data will need to be gathered before it is possible to tell whether this correlation is meaningful.

Work area

Information regarding the area of work people engage in was collected for the group. For most categories, the sample size was too small for analysis, so the focus was on the three work areas with a reasonable sample size. These were as follows:

- HR, training, guidance
- Sales, customer service
- Finance

Table 6.9 shows the facet scale means and standard deviations for these three work areas. The findings can be summarised as follows:

A pattern observed regarding the facets of the E–I scale was that mean scores for the 'Sales, customer service' group tended more towards the E pole for three of the facets than the other two groups.

The most noticeable pattern regarding the facets in the S–N, T–F and J–P scales was that the 'HR, training, guidance' group tended more towards the N, F and P poles (or less towards the S, T and J poles) than the other two groups.

Table 6.9: Mean facet scale scores by work area

Step II facet scale	HR, training, guidance (n=108)		Sales, customer service (n=99)		Finance (n=96)		Sig.
	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales							
Initiating–Receiving	-1.19	3.35	-2.23	2.79	-1.24	3.31	*
Expressive–Contained	-1.31	3.06	-0.69	3.07	-0.05	3.04	*
Gregarious–Intimate	-0.46	2.96	-0.78	2.48	-0.53	2.76	
Active–Reflective	-0.82	3.13	-1.90	2.47	-1.17	2.78	*
Enthusiastic–Quiet	-1.26	2.91	-1.79	2.70	-0.67	3.05	*
S–N facet scales							
Concrete–Abstract	-0.13	2.61	-0.68	2.29	-1.15	2.44	*
Realistic–Imaginative	-0.02	3.23	-0.95	2.97	-0.92	2.96	*
Practical–Conceptual	0.79	2.75	-0.05	2.27	-0.05	2.82	*
Experiential–Theoretical	-1.25	2.85	-2.25	2.49	-1.86	2.74	*
Traditional–Original	1.72	2.78	0.80	3.13	0.29	3.06	**
T–F facet scales							
Logical–Empathetic	0.28	3.24	-0.71	3.19	-1.34	2.80	**
Reasonable–Compassionate	-1.86	3.10	-3.02	2.35	-3.09	2.07	**
Questioning–Accommodating	-0.16	2.30	-0.15	1.97	-0.21	2.43	
Critical–Accepting	0.76	2.24	-0.43	2.17	-0.67	2.14	**
Tough–Tender	0.07	2.94	-0.68	2.66	-1.06	2.68	*
J–P facet scales							
Systematic–Casual	0.24	2.86	-0.93	2.72	-1.21	2.45	**
Planful–Open-Ended	0.85	3.56	-0.46	3.47	-0.58	3.57	**
Early Starting–Pressure-Prompted	-0.37	3.31	-1.04	2.90	-0.40	3.48	
Scheduled–Spontaneous	-1.54	3.10	-2.37	2.48	-2.55	2.30	*
Methodical–Emergent	-0.96	3.06	-1.98	2.53	-1.62	2.61	*

Nationality

Information on nationality was gathered for 94% of the group. Of these, 94% were Italian. There were insufficient people of other nationalities for it to be possible to conduct any analyses across groups.

Employment status

Employment status information was gathered for 79% of the group. Of these, 91% worked full-time. There were insufficient people who did not work full-time for it to be possible to conduct any analyses across groups.

Appendix 1: Sample description

Sample 1: Data from OPPassessment (representative Italian-speaking professional and managerial sample)

This sample consists of 886 individuals who completed the MBTI Step II questionnaire in Italian via the OPPassessment system between September 2006 and June 2016. Fifty-seven per cent of the respondents were male and 43% were female. Age ranged from 20 to 70 years, with a mean of 39 and a median of 39.

Nationality was disclosed by 94% of respondents. Of these, 93.5% were Italian:

Nationality	Percentage
Italian	93.5%
Other	6.5%

Seventy-nine per cent of respondents stated their employment status. Of these, the majority of the group were in full-time employment:

Employment status	Percentage
Full-time	91.0%
Self-employed	1.9%
Part-time	5.8%
Unemployed	1.3%

And 74% of respondents stated their occupational level. The majority of the group were of managerial level or above, but the largest single group was employee (34.5%):

Occupational level	Percentage
Top level	1.1%
Senior executive	23.5%
Upper middle management	8.1%
Middle management	6.7%
First level management/supervisor	12.4%
Employee	34.5%
Other	13.7%

Seventy-seven per cent of respondents stated their work area. Amongst these, a range of work areas were represented:

Work area (job type)	Percentage
HR, training, guidance	15.8%
Sales, customer service	14.5%
Finance	14.1%
Research and development	5.1%
Business services	5.0%
Science, engineering	5.0%
Admin or secretarial	4.2%
IT	3.7%
Land, sea or air transport	1.8%
Education	0.4%
Other private sector	12.4%
Other	11.7%

MBTI® Step II™ instrument

European Data Supplement

Norwegian

January 2018

2nd edition

OPP Ltd
Elsfield Hall
15–17 Elsfield Way
Oxford OX2 8EP
UK
t: +44 (0)1965 404 500

Contents

Introduction

Data collected for the European MBTI® Step II™ instrument were analysed to produce the findings in this supplement. This is the second data supplement produced for this instrument. A brief description of the sample is given below. Further details of the sample are provided in Appendix 1.

- The sample consisted of 1395 individuals who completed the MBTI Step II questionnaire in Norwegian via the OPPassessment™ system between 2006 and mid-2016.[1] This sample is considered to be representative of the groups of people with whom the Norwegian MBTI Step II instrument has been and will be used for applications such as management development, coaching, counselling and teambuilding. As such, it is likely to represent a cross-section of the Norwegian-speaking professional and managerial population.

The results of the analyses are outlined below.

[1] OPPassessment allows personality questionnaires such as the MBTI instrument to be administered via email and/or completed online.

Type distribution

Type tables are a way of illustrating the proportion of each type within a particular group. Below is a type table taken from the sample described on the previous page.

For each of the 16 different types, the number of cases, the percentage of the total that this represents and the self-selection ratio (SSR) are shown. The SSR (Myers et al., 1998) is a way of demonstrating whether a given type appears more or less often in a particular group than would be expected given its frequency in a reference group. Ideally, the type distribution from a large representative sample of the Norwegian population would be used to calculate SSRs in this data supplement. However, such a sample does not currently exist. In its place, SSRs have been calculated using type data from the UK general population (Kendall, 1998), which can be justified by the fact that type distributions for comparable Norwegian and British groups, such as managers and professionals, are similar. Evidence (eg Hackston and Kendall, 2004; Quenk et al., 2004) does suggest that although type-related behaviours vary a good deal from country to country and from culture to culture, the frequencies of underlying MBTI types do not.

An SSR of greater than 1 indicates that a type is over-represented, and an SSR of less than 1 denotes that it is under-represented. Asterisks are used to denote whether the over- or under-representations are statistically significant, based on the results of chi-square analysis.[2]

[2] Chi-square analysis (often abbreviated to χ^2) is a technique used to explore whether observed frequency distributions differ significantly from other, predefined, distributions. In this case, the UK general population group is used as the reference group, and the chi-square analysis indicates whether the proportion of people of each type within a particular sample differs significantly from the proportion of people reporting the same type within the reference group.

OPPassessment data (representative Norwegian-speaking professional and managerial sample)

Table 8.1: Type table for OPPassessment data (reported type, n=1395)

ISTJ	ISFJ	INFJ	INTJ
n=147 10.5% SSR=0.77**	n=21 1.5% SSR=0.12**	n=8 0.6% SSR=0.35**	n=28 2.0% SSR=1.43
ISTP	**ISFP**	**INFP**	**INTP**
n=43 3.1% SSR=0.48**	n=6 0.4% SSR=0.07**	n=7 0.5% SSR=0.16**	n=39 2.8% SSR=1.67
ESTP	**ESFP**	**ENFP**	**ENTP**
n=132 9.5% SSR=1.64**	n=13 0.9% SSR=0.10**	n=41 2.9% SSR=0.46**	n=161 11.5% SSR=4.11**
ESTJ	**ESFJ**	**ENFJ**	**ENTJ**
n=456 32.7% SSR=3.14**	n=63 4.5% SSR=0.36**	n=35 2.5% SSR=0.89	n=195 14.0% SSR=4.83**

Type	n	%
E	1096	78.6%**
I	299	21.4%**
S	881	63.2%**
N	514	36.8%**
T	1201	86.1%**
F	194	13.9%**
J	953	68.3%**
P	442	31.7%**

*Difference significant at p<0.05, based on chi-square results.
**Difference significant at p<0.01, based on chi-square results.

The most common single type preference is ESTJ (33% of the total); this is a common finding with managerial groups in other countries. The SSR results suggest that, in comparison with the UK general population, those with preferences for NT are over-represented and those with preferences for SF are under-represented. Again, this is a common finding with managerial groups. In this group, those with preferences for NF are also under-represented.

Properties of the Step II facet scales

Facet scale score distribution

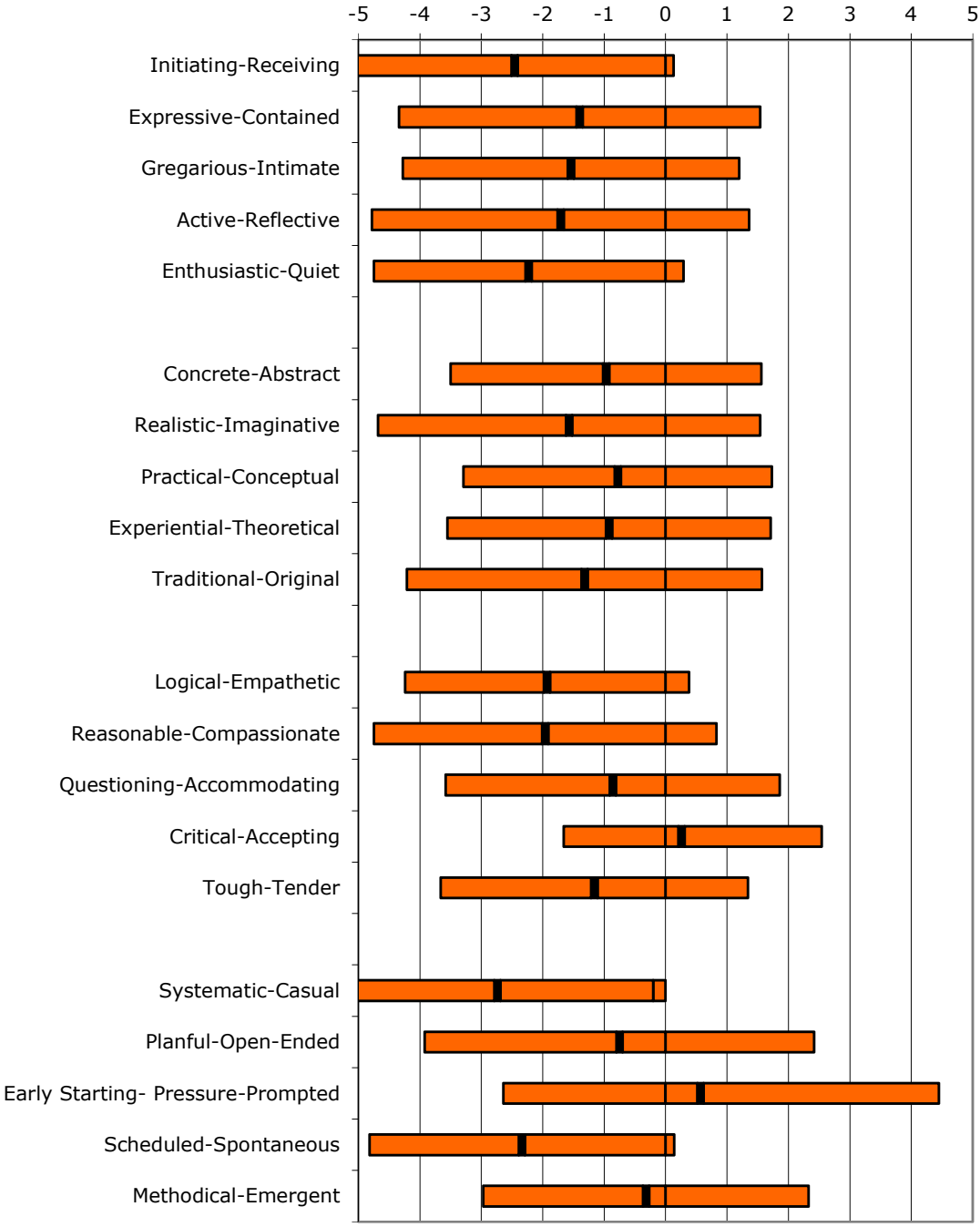
The means and standard deviations of the Step II facets are shown in Table 8.2 and illustrated graphically in Figure 8.1. The mean score for a scale is calculated by adding together the scores of each individual in the sample and then dividing the sum by the number of individuals. Note that negative values indicate that mean scores are nearer the left-hand pole of the facet and positive values indicate that the mean scores are nearer the right-hand pole. The standard deviation (SD) is a statistical measure describing the degree to which the scores from the sample either bunch up close to, or are scattered widely around, the mean for the sample.

Table 8.2: Means and standard deviations of the facet scales

Step II facet scale	Mean[3]	SD
E–I facet scales		
Initiating–Receiving	–2.46	2.59
Expressive–Contained	–1.40	2.94
Gregarious–Intimate	–1.54	2.74
Active–Reflective	–1.71	3.07
Enthusiastic–Quiet	–2.23	2.52
S–N facet scales		
Concrete–Abstract	–0.97	2.53
Realistic–Imaginative	–1.57	3.11
Practical–Conceptual	–0.78	2.51
Experiential–Theoretical	–0.92	2.63
Traditional–Original	–1.32	2.89
T–F facet scales		
Logical–Empathetic	–1.93	2.31
Reasonable–Compassionate	–1.96	2.79
Questioning–Accommodating	–0.86	2.72
Critical–Accepting	0.26	1.92
Tough–Tender	–1.16	2.50
J–P facet scales		
Systematic–Casual	–2.54	2.34
Planful–Open-Ended	-0.75	3.17
Early Starting–Pressure-Prompted	0.57	3.21
Scheduled–Spontaneous	–2.34	2.48
Methodical–Emergent	-0.32	2.65

[3] Scale means: in Step II reports, scores are given from 5 on one pole, through 0, to 5 on the opposite pole. In compiling statistical information, however, one pole needs to be clearly distinguished from the other. Throughout this data supplement, therefore, a convention has been adopted of scores 5, 4, 3, 2 or 1 on the left-hand pole being assigned values of –5, –4, –3, –2, or –1, and scores on the right-hand pole being assigned positive scores. This does not, of course, imply any suggestion that positive scores are 'better' than negative scores.

Figure 8.1: Facet scale mean scores[4]



[4] For each facet scale, the central line shows the mean and the coloured bars either side indicate the standard deviation.

In general, the means centre quite close to the midpoint of each scale. Only four of the scales have a mean greater than 2 points from the midpoint point in either direction, and each of these has a standard deviation of between 2.3 and 2.6. However, a clear pattern does emerge. The mean scores generally tend towards the E, S, T and J direction, which is consistent with the most common four-letter type preference amongst this sample. Amongst the group as a whole, there are more individuals with a preference for E, S, T and J.

Facet scale intercorrelations

Correlations among the Step II facet scales are shown in Table 8.3. Facet scales within each dichotomy usually correlate more highly (often substantially so) with the other scales of the same dichotomy than they do with scales in the other three dichotomies. For example, every E–I facet scale correlates more highly with the other four E–I scales than with any of the S–N, T–F or J–P facet scales.

There is only one scale that correlates **significantly** more highly with scales in other dichotomies. The T–F scale Questioning–Accommodating correlates with all scales, across dichotomies. For example, it correlates at between 0.18 and 0.30 with all of the E–I scales of the E–I dichotomy, between –0.24 and –0.40 with all of the S–N scales of the S–N dichotomy, and between –0.06 and –0.21 with all of the J–P scales of the J–P dichotomy. By way of comparison, it correlates at between 0.01 and 0.15 with the other four scales on the same dichotomy as itself (T–F). The negative correlations between Questioning–Accommodating and the S–N scales are consistent with findings with the US version of the Step II instrument, and suggest that a questioning approach to differences of opinion seems to be related to a range of iNtuition (N) facets. These patterns are similar to those found when the European Step II instrument was developed.

Table 8.3: Intercorrelations of Step II facet scales

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Initiating–Receiving	**1.00**																			
2. Expressive–Contained	**0.52**	**1.00**																		
3. Gregarious–Intimate	**0.53**	**0.44**	**1.00**																	
4. Active–Reflective	**0.70**	**0.51**	**0.53**	**1.00**																
5. Enthusiastic–Quiet	**0.61**	**0.58**	**0.61**	**0.63**	**1.00**															
6. Concrete–Abstract	-0.14	-0.15	-0.05	-0.10	-0.17	**1.00**														
7. Realistic–Imaginative	-0.12	-0.16	-0.09	-0.11	-0.18	**0.63**	**1.00**													
8. Practical–Conceptual	-0.09	-0.06	-0.04	-0.08	-0.13	**0.58**	**0.58**	**1.00**												
9. Experiential–Theoretical	-0.09	-0.11	-0.02	-0.07	-0.10	**0.53**	**0.42**	**0.49**	**1.00**											
10. Traditional–Original	-0.17	-0.16	-0.07	-0.17	-0.22	**0.58**	**0.55**	**0.57**	**0.40**	**1.00**										
11. Logical–Empathetic	-0.07	-0.20	0.02	-0.02	-0.05	0.13	0.24	-0.01	-0.01	0.01	**1.00**									
12. Reasonable–Compassionate	0.00	-0.11	0.02	0.05	0.01	0.06	0.17	-0.08	-0.04	-0.08	**0.67**	**1.00**								
13. Questioning–Accommodating	0.24	0.18	0.18	0.26	0.30	-0.32	-0.26	-0.36	-0.24	-0.40	**0.08**	**0.15**	**1.00**							
14. Critical–Accepting	-0.22	-0.19	-0.10	-0.15	-0.15	0.13	0.18	0.05	0.06	0.04	**0.37**	**0.38**	**0.01**	**1.00**						
15. Tough–Tender	0.01	-0.11	0.07	0.07	0.02	0.16	0.23	0.07	0.08	-0.00	**0.52**	**0.56**	**0.08**	**0.40**	**1.00**					
16. Systematic–Casual	-0.09	-0.18	-0.06	-0.11	-0.16	0.43	0.47	0.30	0.22	0.39	0.29	0.21	-0.18	0.16	0.21	**1.00**				
17. Planful–Open-Ended	0.00	0.02	-0.02	-0.02	0.00	0.27	0.24	0.20	0.15	0.31	0.03	0.01	-0.15	0.01	0.01	**0.48**	**1.00**			
18. Early Starting–Pressure-Prompted	-0.13	-0.08	-0.09	-0.13	-0.18	0.28	0.25	0.25	0.16	0.34	-0.04	-0.09	-0.21	0.00	0.00	**0.33**	**0.41**	**1.00**		
19. Scheduled–Spontaneous	-0.02	-0.06	-0.05	-0.04	-0.09	0.29	0.29	0.21	0.18	0.30	0.12	0.08	-0.13	0.03	0.06	**0.51**	**0.53**	**0.39**	**1.00**	
20. Methodical–Emergent	0.06	0.02	0.01	0.04	0.00	0.19	0.17	0.13	0.09	0.18	0.04	0.02	-0.06	-0.04	0.02	**0.42**	**0.46**	**0.39**	**0.48**	**1.00**

Correlations of Step II facet scales with Step I scales

Correlations between Step II facet scales and the continuous scores from the MBTI Step I™ instrument are shown in Table 8.4.[5] Positive correlations between facet scales and the corresponding Step I dimension (e.g. between Initiating–Receiving and the E–I continuous score) indicate relationships in the expected direction, with higher facet scores tending to be associated with higher continuous scores, and vice versa. Negative correlations would indicate an inverse relationship between the facet scale scores and the corresponding Step I dimension.

Table 8.4: Correlations of Step II facet scales with Step I continuous scores

	Step I continuous score			
Step II facet scales	**E–I**	**S–N**	**T–F**	**J–P**
E–I facet scales				
Initiating–Receiving	**0.82**	-0.15	–0.08	0.00
Expressive–Contained	**0.71**	–0.17	–0.21	-0.03
Gregarious–Intimate	**0.65**	-0.06	0.01	-0.01
Active–Reflective	**0.80**	–0.13	–0.01	-0.04
Enthusiastic–Quiet	**0.83**	–0.20	–0.05	–0.05
S–N facet scales				
Concrete–Abstract	–0.16	**0.84**	0.15	0.37
Realistic–Imaginative	–0.16	**0.75**	0.25	0.35
Practical–Conceptual	–0.10	**0.69**	–0.02	0.30
Experiential–Theoretical	–0.10	**0.63**	0.02	0.22
Traditional–Original	–0.20	**0.76**	–0.02	0.41
T–F facet scales				
Logical–Empathetic	–0.11	0.11	**0.87**	0.07
Reasonable–Compassionate	–0.01	0.01	**0.75**	0.02
Questioning–Accommodating	0.29	–0.38	**0.14**	–0.21
Critical–Accepting	–0.21	0.12	**0.51**	0.02
Tough–Tender	0.01	0.11	**0.68**	0.04
J–P facet scales				
Systematic–Casual	–0.15	0.46	0.27	**0.64**
Planful–Open-Ended	0.00	0.30	0.02	**0.82**
Early Starting–Pressure-Prompted	–0.15	0.33	–0.05	**0.67**
Scheduled–Spontaneous	-0.07	0.32	0.10	**0.70**
Methodical–Emergent	0.04	0.20	0.04	**0.65**

The E–I facet scales correlate at 0.65 to 0.83 with the E–I continuous scores from Step I; the S–N facet scales correlate at 0.63 to 0.84 with the S–N continuous scores; the T–F facet scales correlate at 0.14 to 0.87 with the T–F continuous scores; and the J–P facet scales correlate

[5] Continuous scores place an individual's score on each dimension onto a continuous scale with a midpoint of 100. To calculate continuous scores, Preference Clarity Index (PCI) scores for each dimension are either subtracted or added to 100, depending on which direction the overall preference is. PCI scores in the direction of E, S, T or J are subtracted from 100. PCI scores in the direction of I, N, F or P are added to 100.

at 0.64 to 0.82 with the J–P continuous scores. The scale that showed the lowest correlation with its associated Step I continuous score was Questioning–Accommodating at 0.14. This scale is also lowest on internal consistency (see page 185), and has been previously found to have the lowest test-retest reliability (Quenk, Hammer and Majors, 2004), which would result in the true correlations being underestimated.

These correlations are very similar to those found during the development of the Step II questionnaire. This consistency, alongside the fact that the correlations between Step II facet scales and Step I continuous scores associated with their own dimension are substantially higher than correlations with the other three dimensions, provides compelling evidence for the theoretical hierarchical structure of the Step II facet scales in relation to the Step I scales.

Out-of-preference scores

It is known that the five facet scales relating to each type dimension do not describe the dimension in its entirety; there will not be a precise and exact overlap between, for example, an individual's score on E–I and their total score across the five facet scales that relate to this dimension. For example, it is not uncommon to see an Enthusiastic Introvert or an Intimate Extravert. Such apparent inconsistencies are usually known as 'out-of-preference scores' (OOPS).

Although it is usual to have a number of OOPS in any one profile, it should be unusual to find that OOPS make up more than half of the facet scores relating to any one type dichotomy. Table 8.5 demonstrates that the proportion of people for whom this occurs ranges from 2.2% for the S–N block to 0.8% for the E–I block. It is therefore a very infrequent occurrence.

Table 8.5: Proportion of OOPS by type dichotomy[6]

	Proportion of 'reported type' OOPS					
	None	One	Two	Three	Four	Five
E–I	72.7%	21.6%	4.9%	0.8%	0.0%	0.0%
S–N	67.2%	23.1%	7.6%	1.9%	0.3%	0.0%
T–F	57.5%	33.5%	7.2%	1.6%	0.1%	0.1%
J–P	49.0%	35.7%	13.9%	1.4%	0.0%	0.0%

For any individual facet scale, an OOPS tends to occur in approximately 10% of profiles, with the exception of Questioning–Accommodating (23%) and Early Starting–Pressure-Prompted (22%).

[6] In this table, an OOPS is defined as a score of 2, 3, 4 or 5 on the 'wrong' pole of the facet when compared with the reported type.

Reliability

The reliability of a test or questionnaire relates to how consistent and precise it is. Internal consistency reliability addresses the question of whether all the questions in a scale measure the same construct. For example, are the Step II facet scales consistent within themselves, and do they hold together well as scales? A common measure of internal consistency reliability is coefficient alpha (Cronbach, 1951). The alpha coefficients for the Step II facet scales are shown in Table 8.6.

Table 8.6: Internal consistency reliability

Step II facet scale	No. of items	Coefficient alpha
E–I facet scales		
Initiating–Receiving	8	0.84
Expressive–Contained	7	0.76
Gregarious–Intimate	7	0.65
Active–Reflective	8	0.72
Enthusiastic–Quiet	9	0.72
S–N facet scales		
Concrete–Abstract	9	0.73
Realistic–Imaginative	7	0.73
Practical–Conceptual	8	0.51
Experiential–Theoretical	8	0.75
Traditional–Original	8	0.73
T–F facet scales		
Logical–Empathetic	9	0.80
Reasonable–Compassionate	8	0.71
Questioning–Accommodating	7	0.45
Critical–Accepting	8	0.49
Tough–Tender	8	0.76
J–P facet scales		
Systematic–Casual	8	0.75
Planful–Open-Ended	6	0.79
Early Starting–Pressure-Prompted	6	0.73
Scheduled–Spontaneous	8	0.74
Methodical–Emergent	8	0.65
	Median	0.73

The internal consistency reliability of most scales is good, and the average (median) reliability is above 0.7. However, three scales (Practical–Conceptual, Questioning–Accommodating and Critical–Accepting) do have lower alpha coefficients. This pattern is consistent with what was found during the development of the Step II instrument, where these three facet scales were also found to show lower reliability than the others (Quenk, Hammer and Majors, 2004).

It should be noted, however, that coefficient alpha reliability statistics will give an underestimate of the reliability of the Step II facet scales,

given the use of item response theory (IRT) methods in the actual scoring process.[7]

Group differences

This section looks at the way in which people from different groups responded to the Norwegian version of the European Step II questionnaire. Unfortunately, the majority of the respondents did not fill in any demographic data, so the variables available for analysis were very limited. Further research will be conducted when additional data become available.

Gender

The means and standard deviations of the Step II facet scales are shown separately for males and females in Table 8.7, along with the difference in mean scores. This is represented graphically in Figure 8.2. Statistically significant differences were found between male and female mean scores for 11 of the 20 facet scales, with some consistent patterns emerging.

- On the E–I facet scales, all five mean scores tended slightly toward the E pole for both males and females. Female mean scores tended further towards the E pole than males on four of the facet scales, and there was no significant difference between genders on the remaining facet scale.
- On the S–N facet scales, all five mean scores tended slightly toward the S pole for both males and females. There were statistically significant gender differences on three of the five facet scales, with male mean scores tending more towards the S pole than females.
- On the T–F facet scales, four of the five mean scores tended slightly toward the T pole for both males and females. On the Critical–Accepting facet, male mean scores tended towards the T pole whereas female mean scores tended towards the F pole. There were statistically significant gender differences on four of the five facet scales (including Critical–Accepting), with male mean scores tending more towards the T pole than females.
- On the J–P facet scales, four of the five mean scores tended toward the J pole for both males and females. There were no statistically significant gender differences on any of the five facet scales.

[7] Item response theory (IRT) is an approach to measurement that is concerned with modelling the relationship between item responses and the underlying characteristic assessed by the scale or test the item is designed to measure. IRT can be used to select items for a test and/or to score the items.

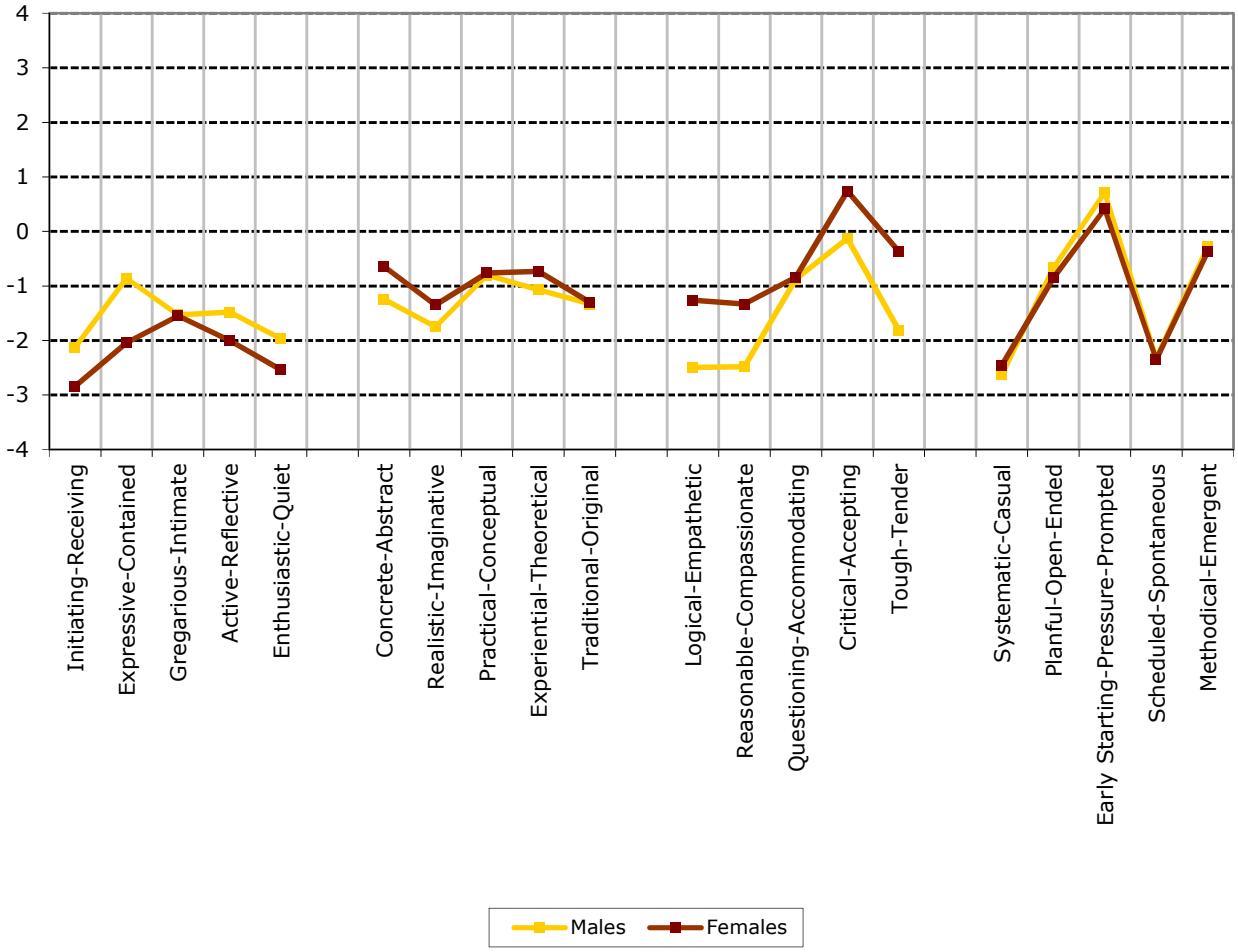
Table 8.7: Gender differences in facet scale scores

	Males (n=194)		Females (n=119)		Difference (M–F)[8]
Step II facet scale	**Mean**	**SD**	**Mean**	**SD**	
E–I facet scales					
Initiating–Receiving	-2.13	2.68	-2.84	2.43	0.71**
Expressive–Contained	-0.86	3.02	-2.04	2.70	1.19**
Gregarious–Intimate	-1.53	2.78	-1.55	2.70	0.02
Active–Reflective	-1.48	3.14	-2.00	2.98	0.52**
Enthusiastic–Quiet	-1.97	2.64	-2.54	2.34	0.57**
S–N facet scales					
Concrete–Abstract	-1.24	2.50	-0.64	2.52	−0.60**
Realistic–Imaginative	-1.75	3.05	-1.35	3.17	−0.40*
Practical–Conceptual	-0.80	2.53	-0.76	2.50	−0.05
Experiential–Theoretical	-1.07	2.61	-0.73	2.63	−0.34*
Traditional–Original	-1.33	2.86	-1.30	2.93	−0.03
T–F facet scales					
Logical–Empathetic	-2.49	2.04	-1.26	2.44	−1.23**
Reasonable–Compassionate	-2.48	2.60	-1.33	2.90	−1.15**
Questioning–Accommodating	-0.87	2.68	-0.84	2.78	−0.02
Critical–Accepting	-0.13	1.88	0.74	1.85	−0.87**
Tough–Tender	-1.81	2.30	-0.37	2.50	−1.44**
J–P facet scales					
Systematic–Casual	-2.62	2.39	-2.45	2.28	−0.17
Planful–Open-Ended	-0.67	3.23	-0.85	3.09	0.18
Early Starting–Pressure-Prompted	0.71	3.16	0.41	3.26	0.30
Scheduled–Spontaneous	-2.33	2.46	-2.35	2.50	0.02
Methodical–Emergent	-0.28	2.73	-0.37	2.56	0.09

Difference significant at: *p<.05, **p<0.01 (based on an independent samples t-test).

[8] A positive value indicates that male scores tend more towards I, N, F or P, and a negative value indicates that female scores tend more towards I, N, F or P.

Figure 8.2: Gender differences in facet scale scores



Age

Although age data was captured for the Norwegian sample, the number of people who provided this data was too small to allow a full analysis. These analyses will be conducted once additional data become available.

Ethnic origin

Ethnic origin information was not captured for individuals who completed the Norwegian language version of the Step II questionnaire, so no analyses were conducted.

Occupational level

Research using the Norwegian version of the MBTI Step I questionnaire has demonstrated that individuals in higher-level jobs in organisations are more likely to have preferences for Thinking than those in lower-level jobs.[9]

[9] See the MBTI Step I European Data Supplement for details.

Although occupational level data were captured for the Norwegian sample, the number of people in each of the categories was too small to allow a full analysis. These analyses will be conducted once additional data become available.

Education

Specific educational qualifications were not collected for the OPPassessment sample. The age at which individuals left full-time education was collected, however the number of people who provided this data was too small to allow a full analysis. These analyses will be conducted once additional data become available.

Work area

Information regarding the area of work people engage in was collected for the group. Many different categories were used, but the numbers of people in each category were too small for the purposes of analysis. There were no more than 48 people in any one group. These analyses will be conducted once additional data become available.

Nationality

Information on nationality was gathered for 17% of the group. Of these, 97% were Norwegian. There were insufficient people of other nationalities for it to be possible to conduct any analyses across groups.

Employment status

Employment status information was gathered for 16% of the group. Of these, 98% worked full-time. There were insufficient people who did not work full-time for it to be possible to conduct any analyses across groups.

Appendix 1: Sample description

Sample 1: Data from OPPassessment (representative Norwegian-speaking professional and managerial sample)

This sample consists of 1395 individuals who completed the MBTI Step II questionnaire in Norwegian via the OPPassessment system between September 2006 2003 and July 2016. Fifty-five per cent of the respondents were male and 45% were female. Age ranged from 28 to 55 years, with a mean of 42 and median of 41.

Nationality was disclosed by 17% of respondents. Of these, 97% were Norwegian. A few other nationalities were represented, but each one formed less than 1% of the total group:

Nationality	Percentage
Norwegian	96.6%
Other	1.7%

Sixteen per cent of the respondents stated their employment status. Of these, the majority of the group were in full-time employment:

Employment status	Percentage
Full-time	97.7%
Self-employed	1.4%
Part-time	0.9%

Fifteen per cent of respondents stated their occupational level. The majority of the group were of managerial level or above, with the largest single group being upper middle management (32%):

Occupational level	Percentage
Top level	18.7%
Senior executive	0.0%
Upper middle management	32.1%
Middle management	28.2%
First level management/supervisor	4.3%
Employee	14.8%
Other	1.9%

Fifteen per cent of respondents stated their work area. Amongst these, a range of work areas were represented:

Work area (job type)	Percentage
Finance	22.4%
HR, training, guidance	10.3%
Sales, customer service	9.8%
Admin or secretarial	8.4%
IT	7.9%
Education	5.6%
Science, engineering	4.7%
Skilled operative	1.9%
Business services	1.4%
Health, social services, etc.	1.4%
Land, sea or air transport	0.5%
Research and development	0.5%
Unskilled operative	0.5%
Other private sector	15.0%
Other public sector	3.7%
Other	6.1%

MBTI® Step II™ instrument

European Data Supplement

Polish

January 2018

2nd edition

OPP Ltd
Elsfield Hall
15–17 Elsfield Way
Oxford OX2 8EP
UK
t: +44 (0)19865 404 500

Contents

Introduction

The Polish language version of the MBTI® Step II™ questionnaire was launched in 2012.

Data collected for the European MBTI Step II instrument were analysed to produce the findings in this supplement. This is the second data supplement produced for this instrument. A brief description of the sample is given below. Further details of the sample are provided in Appendix 1.

- The sample consisted of 688 individuals who completed the MBTI Step II questionnaire in Polish via the OPPassessment™ system between 2012 and mid-2016.[1] This sample is considered to be representative of the groups of people with whom the Polish MBTI Step II instrument has been and will be used for applications such as management development, coaching, counselling and teambuilding. As such, it is likely to represent a cross-section of the Polish-speaking professional and managerial population.

The results of the analyses are shown below.

[1] OPPassessment allows personality questionnaires such as the MBTI instrument to be administered via email and/or completed online.

Type distribution

Type tables are a way of illustrating the proportion of each type within a particular group. Below is a type table for the Polish sample described on the previous page.

For each of the 16 different types, the number of cases, the percentage of the total that this represents and the self-selection ratio (SSR) are shown. The SSR (Myers et al., 1998) is a way of demonstrating whether a given type appears more or less often in a particular group than would be expected given its frequency in a reference group. An SSR of greater than 1 indicates that a type is over-represented, and an SSR of less than 1 denotes that it is under-represented. Asterisks are used to denote whether the over- or under-representations are statistically significant, based on the results of chi-square analysis[2].

Ideally, the type distribution from a large representative sample of the Polish population would be used to calculate SSRs in this data supplement. However, such a sample does not currently exist. In its place, SSRs have been calculated using type data from the UK general population (Kendall, 1998). Evidence (eg Hackston and Kendall, 2004; Quenk et al., 2004; Kirby, Kendall and Barger, 2007) does suggest that although type-related behaviours vary a good deal from country to country and from culture to culture, the frequencies of underlying MBTI types do not.

[2] Chi-square analysis (often abbreviated to χ^2) is a technique used to explore whether observed frequency distributions differ significantly from other, predefined, distributions. In this case, the UK general population group is used as the reference group, and the chi-square analysis indicates whether the proportion of people of each type within a particular sample differs significantly from the proportion of people reporting the same type within the reference group.

Questionnaire development sample

Table 9.1: Type table for questionnaire development data Reported type (n=688)

ISTJ	ISFJ	INFJ	INTJ
n=112 16.3% SSR=1.19	n=6 0.9% SSR=0.07**	n=9 1.3% SSR=0.76	n=42 6.1% SSR=4.36**
ISTP	**ISFP**	**INFP**	**INTP**
n=14 2.0% SSR=0.31**	n=5 0.7% SSR=0.11**	n=10 1.5% SSR=0.47*	n=26 3.8% SSR=1.58
ESTP	**ESFP**	**ENFP**	**ENTP**
n=27 3.9% SSR=0.67	n=10 1.5% SSR=0.17**	n=58 8.4% SSR=1.33	n=59 8.6% SSR=3.07**
ESTJ	**ESFJ**	**ENFJ**	**ENTJ**
n=130 18.9% SSR=1.82**	n=28 4.1% SSR=0.33**	n=35 5.1% SSR=1.82**	n=117 17.0% SSR=5.86**

Type	n	%
E	464	67.4%**
I	224	32.6%**
S	332	48.3%**
N	356	51.7%**
T	527	76.6%**
F	161	23.4%**
J	479	69.6%**
P	209	30.4%**

*Difference significant at p<0.05, based on chi-square results.
**Difference significant at p<0.01, based on chi-square results.

The most common single type preference is ESTJ (19% of the total); this is a common finding with managerial groups in other countries. The SSR results suggest that, in comparison with the UK general population, those with preferences for NT are over-represented, and those with preferences for SF are under-represented. Again, this is a common finding with managerial groups.

Properties of the Step II facet scales

Facet scale score distribution

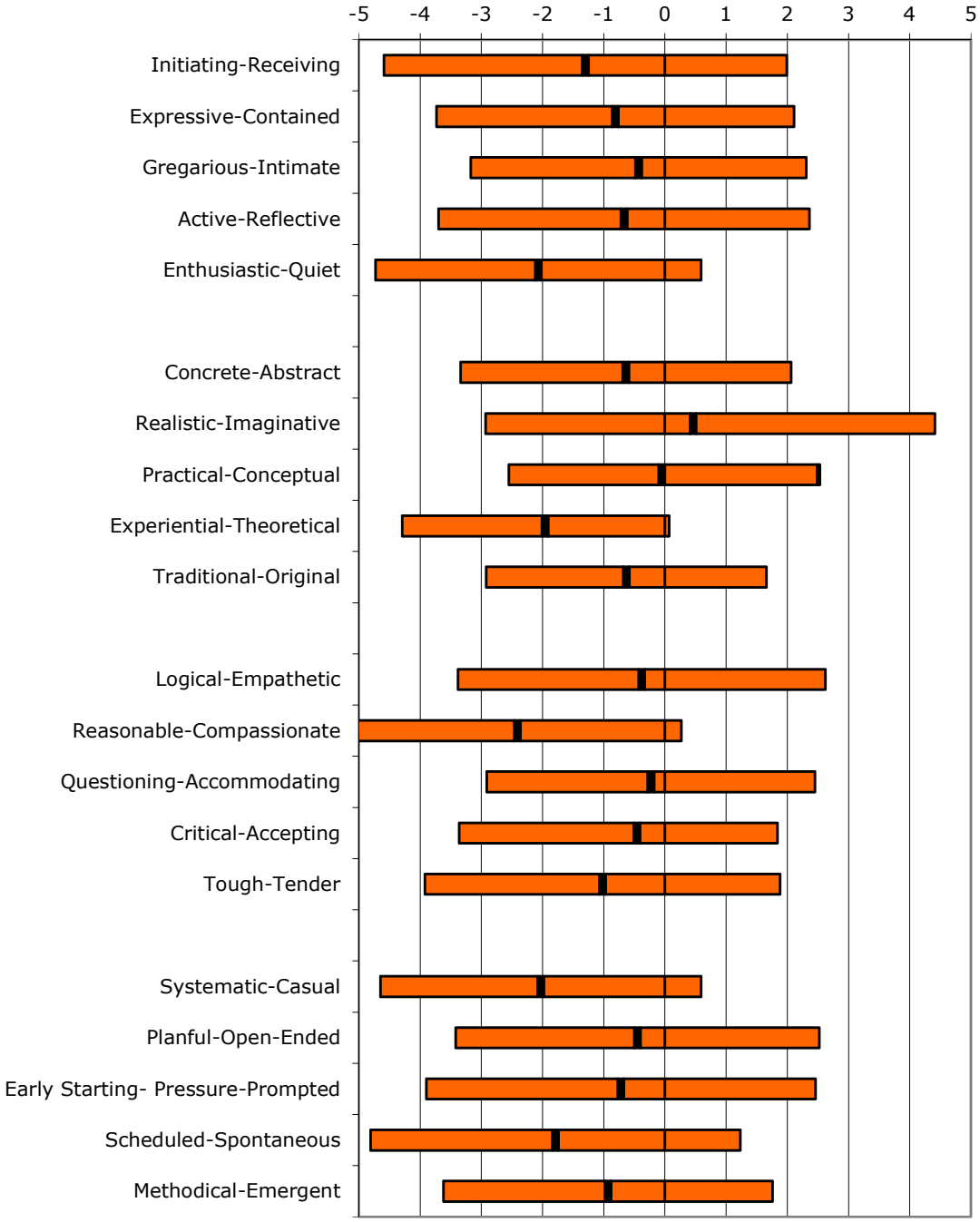
The means and standard deviations of the Step II facets are shown in Table 9.2, and are illustrated graphically in Figure 9.1. The mean score for a scale is calculated by adding together the scores for each individual in the sample and then dividing the sum by the number of individuals. Note that negative values indicate that mean scores are nearer the left-hand pole of the facet and positive values indicate that the mean scores are nearer the right-hand pole. The standard deviation (SD) is a statistical measure describing the degree to which the scores from the sample either bunch up close to or are scattered widely around the mean for the sample.

Table 9.2: Means and standard deviations of the facet scales

Step II facet scale	Mean[3]	SD
E–I facet scales		
Initiating–Receiving	–1.30	3.29
Expressive–Contained	-0.81	2.92
Gregarious–Intimate	-0.43	2.74
Active–Reflective	-0.67	3.03
Enthusiastic–Quiet	-2.07	2.66
S–N facet scales		
Concrete–Abstract	-0.64	2.70
Realistic–Imaginative	0.46	3.39
Practical–Conceptual	-0.01	2.50
Experiential–Theoretical	-1.81	2.75
Traditional–Original	-0.05	2.68
T–F facet scales		
Logical–Empathetic	-1.26	2.79
Reasonable–Compassionate	-2.49	2.58
Questioning–Accommodating	-0.58	2.74
Critical–Accepting	-0.19	2.22
Tough–Tender	-0.92	2.94
J–P facet scales		
Systematic–Casual	-1.95	2.75
Planful–Open-Ended	-0.95	3.26
Early Starting–Pressure-Prompted	-0.52	3.42
Scheduled–Spontaneous	-2.07	3.07
Methodical–Emergent	-1.17	2.81

[3] Scale means: in Step II reports, scores are given from 5 on one pole, through 0, to 5 on the opposite pole. In compiling statistical information, however, one pole needs to be clearly distinguished from the other. Throughout this data supplement, therefore, a convention has been adopted of scores 5, 4, 3, 2 or 1 on the left-hand pole being assigned values of –5, –4, –3, –2 or –1 and scores on the right-hand pole being assigned positive scores. This does not, of course, imply any suggestion that positive scores are 'better' than negative scores.

Figure 9.1: Facet scale mean scores[4]



[4] For each facet scale, the central line shows the mean and the coloured bars to either side indicate the standard deviation.

In general, the means centre quite close to the midpoint of each scale. Only three of the scales have a mean greater than 2 points from the midpoint in either direction, and these have a standard deviation of between 2.5 and 3.1. However, a clear pattern does emerge. The mean scores tend towards the E, S, T and J direction (with the exception of Realistic–Imaginative), which is consistent with the most common four-letter type preference amongst this sample. Amongst the group as a whole, there are more individuals with a preference for E, S, T and J.

Facet scale intercorrelations

Correlations among the Step II facet scales are shown in Table 9.3. Facet scales within each dichotomy consistently correlate more highly (usually substantially so) with the other scales of the same dichotomy than they do with scales in the other three dichotomies. For example, every E–I facet scale correlates more highly with the other four E–I scales than with any of the S–N, T–F or J–P facet scales.

There is only one scale that correlates **significantly** more highly with scales in other dichotomies. The T–F scale Questioning–Accommodating correlates with many scales, across dichotomies. For example, it correlates at a level of 0.27 with Enthusiastic–Quiet, 0.19 with Active–Reflective, 0.18 with Initiating–Receiving and 0.12 with Gregarious–Intimate on the E–I dichotomy. It correlates negatively at a level of between -0.17 and -0.36 with all the facets on the S–N dichotomy. It also correlates negatively at -0.17 with Early Starting-Pressure Prompted and at -0.13 with Systematic-Casual on the J–P dichotomy. By way of comparison, Questioning–Accommodating correlates at between 0.11 and 0.27 with the other four scales on the same dichotomy as itself (T–F). The negative correlations between Questioning–Accommodating and the S–N scales are consistent with findings in the US version of the instrument, and suggests that a questioning approach to differences of opinion seems to be related to a range of iNtuition (N) facets. These patterns are similar to those found when the European Step II instrument was developed.

Table 9.3: Intercorrelations of Step II facet scales

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Initiating–Receiving	**1.00**																			
2. Expressive–Contained	**0.52**	**1.00**																		
3. Gregarious–Intimate	**0.62**	**0.54**	**1.00**																	
4. Active–Reflective	**0.71**	**0.54**	**0.61**	**1.00**																
5. Enthusiastic–Quiet	**0.62**	**0.61**	**0.64**	**0.63**	**1.00**															
6. Concrete–Abstract	-0.15	-0.17	-0.12	-0.18	-0.25	**1.00**														
7. Realistic–Imaginative	-0.20	-0.22	-0.21	-0.26	-0.35	**0.67**	**1.00**													
8. Practical–Conceptual	-0.11	-0.05	-0.10	-0.13	-0.20	**0.63**	**0.60**	**1.00**												
9. Experiential–Theoretical	-0.09	-0.12	-0.03	-0.10	-0.11	**0.59**	**0.43**	**0.41**	**1.00**											
10. Traditional–Original	-0.20	-0.18	-0.20	-0.23	-0.33	**0.65**	**0.58**	**0.56**	**0.40**	**1.00**										
11. Logical–Empathetic	-0.11	-0.25	-0.10	-0.15	-0.22	0.36	0.42	0.14	0.29	0.23	**1.00**									
12. Reasonable–Compassionate	-0.09	-0.20	-0.09	-0.13	-0.17	0.29	0.30	0.09	0.28	0.19	**0.68**	**1.00**								
13. Questioning–Accommodating	0.18	0.08	0.12	0.19	0.27	-0.26	-0.20	-0.32	-0.17	-0.36	**0.11**	**0.19**	**1.00**							
14. Critical–Accepting	-0.18	-0.22	-0.15	-0.20	-0.27	0.28	0.34	0.15	0.20	0.21	**0.52**	**0.45**	**0.17**	**1.00**						
15. Tough–Tender	0.00	-0.13	-0.03	-0.02	-0.07	0.25	0.27	0.13	0.18	0.12	**0.58**	**0.59**	**0.27**	**0.46**	**1.00**					
16. Systematic–Casual	-0.14	-0.21	-0.15	-0.19	-0.28	0.47	0.46	0.27	0.41	0.53	0.44	0.38	-0.13	0.27	0.24	**1.00**				
17. Planful–Open-Ended	-0.07	-0.09	-0.10	-0.11	-0.16	0.38	0.35	0.29	0.29	0.44	0.28	0.21	-0.11	0.23	0.18	**0.68**	**1.00**			
18. Early Starting–Pressure-Prompted	-0.08	-0.08	-0.13	-0.11	-0.17	0.27	0.26	0.22	0.23	0.38	0.16	0.15	-0.17	0.11	0.09	**0.56**	**0.53**	**1.00**		
19. Scheduled–Spontaneous	-0.09	-0.12	-0.12	-0.12	-0.19	0.38	0.36	0.25	0.32	0.46	0.32	0.31	-0.10	0.20	0.20	**0.71**	**0.69**	**0.60**	**1.00**	
20. Methodical–Emergent	-0.03	-0.11	-0.07	-0.03	-0.15	0.30	0.28	0.17	0.28	0.33	0.24	0.25	-0.08	0.13	0.13	**0.61**	**0.57**	**0.55**	**0.66**	**1.00**

Correlations of Step II facet scales with Step I scales

Correlations between Step II facet scales and the continuous scores from the MBTI Step I™ instrument are shown in Table 9.4.[5] Positive correlations between facet scales and the corresponding Step I dimension (e.g. between Initiating–Receiving and the E–I continuous score) indicate relationships in the expected direction, with higher facet scores tending to be associated with higher continuous scores, and vice versa. Negative correlations would indicate an inverse relationship between the facet scale scores and the corresponding Step I dimension.

Table 9.4: Correlations of Step II facet scales with Step I continuous scores

	Step I continuous score			
Step II facet scales	**E–I**	**S–N**	**T–F**	**J–P**
E–I facet scales				
Initiating–Receiving	**0.83**	-0.19	-0.13	-0.07
Expressive–Contained	**0.74**	-0.17	-0.27	-0.12
Gregarious–Intimate	**0.73**	-0.17	0.13	-0.11
Active–Reflective	**0.80**	-0.21	-0.17	-0.10
Enthusiastic–Quiet	**0.83**	-0.31	-0.25	-0.18
S–N facet scales				
Concrete–Abstract	-0.21	**0.88**	0.38	0.41
Realistic–Imaginative	-0.29	**0.79**	0.42	0.38
Practical–Conceptual	-0.13	**0.68**	0.18	0.28
Experiential–Theoretical	-0.11	**0.64**	0.29	0.34
Traditional–Original	-0.27	**0.75**	0.24	0.47
T–F facet scales				
Logical–Empathetic	-0.21	0.37	**0.86**	0.30
Reasonable–Compassionate	-0.18	0.28	**0.78**	0.28
Questioning–Accommodating	0.21	-0.31	**0.21**	-0.11
Critical–Accepting	-0.24	0.30	**0.63**	0.21
Tough–Tender	-0.06	0.24	**0.75**	0.20
J–P facet scales				
Systematic–Casual	-0.22	0.51	0.39	**0.79**
Planful–Open-Ended	-0.10	0.40	0.27	**0.85**
Early Starting–Pressure-Prompted	-0.12	0.29	0.16	**0.76**
Scheduled–Spontaneous	-0.14	0.42	0.31	**0.84**
Methodical–Emergent	-0.08	0.31	0.23	**0.73**

The E–I facet scales correlate at 0.73 to 0.83 with the E–I continuous scores from Step I; the S–N facet scales correlate at 0.64 to 0.88 with the S–N continuous scores; the T–F facet scales correlate at 0.21 to

[5] Continuous scores place an individual's score on each dimension onto a continuous scale with a mid-point of 100. To calculate continuous scores, Preference Clarity Index (PCI) scores for each dimension are either subtracted or added to 100, depending on which direction the overall preference is. PCI scores in the direction of E, S, T or J are subtracted from 100. PCI scores in the direction of I, N, F or P are added to 100.

0.86 with the T–F continuous scores; and the J–P facet scales correlate at 0.73 to 0.85 with the J–P continuous scores. The scale that showed the lowest correlation with its associated Step I continuous score was Questioning–Accommodating at 0.21. This scale is also lowest on internal consistency (see page 204), and has been previously found to have the lowest test-retest reliability (Quenk, Hammer and Majors, 2004), which would result in the true correlations being underestimated.

These correlations are very similar to those found during the development of the Step II questionnaire. This consistency, alongside the fact that the correlations between Step II facet scales and Step I continuous scores associated with their own dimension are substantially higher than correlations with the other three dimensions, provides compelling evidence for the theoretical hierarchical structure of the Step II facet scales in relation to the Step I scales.

Out-of-preference scores

It is known that the five facet scales relating to each type dimension do not describe the dimension in its entirety; there will not be a precise and exact overlap between, for example, an individual's score on E–I and their total score across the five facet scales that relate to this dimension. For example, it is not uncommon to see an Enthusiastic Introvert or an Intimate Extravert. Such apparent inconsistencies are usually known as 'out-of-preference scores' (OOPS).

Although it is usual to have a number of OOPS in any one profile, it should be unusual to find that OOPS make up more than half of the facet scores relating to any one type dichotomy. Table 9.5 demonstrates that the proportion of people for whom this happens ranges from 2.3% for the S–N block to 1.0% for the E–I block. It is therefore a very infrequent occurrence.

Table 9.5: Proportion of OOPS by type dichotomy[6]

	Proportion of 'reported type' OOPS					
	None	One	Two	Three	Four	Five
E–I	72.7%	20.8%	5.5%	0.9%	0.1%	0.0%
S–N	56.8%	32.7%	8.1%	1.9%	0.4%	0.0%
T–F	56.5%	33.9%	7.6%	2.0%	0.0%	0.0%
J–P	67.6%	23.8%	7.4%	1.0%	0.1%	0.0%

For any individual facet scale, an OOPS tends to occur in approximately 10% of profiles. Two exceptions to this are found;

[6] In this table, an OOPS is defined as a score of 2, 3, 4 or 5 on the 'wrong' pole of the facet when compared with the reported type.

Experiential–Theoretical (25%) and Questioning–Accommodating (22%).

Reliability

The reliability of a test or questionnaire relates to how consistent and precise it is. Internal consistency reliability addresses the question of whether all the questions in a scale measure the same construct. For example, are the Step II facet scales consistent within themselves, and do they hold together well as scales? A common measure of internal consistency reliability is coefficient alpha (Cronbach, 1951). The alpha coefficients for the Step II facet scales are shown in Table 9.6.

Table 9.6: Internal consistency reliability

Step II facet scale	No. of items	Coefficient alpha
E–I facet scales		
Initiating–Receiving	8	0.84
Expressive–Contained	7	0.75
Gregarious–Intimate	7	0.69
Active–Reflective	8	0.72
Enthusiastic–Quiet	9	0.71
S–N facet scales		
Concrete–Abstract	9	0.73
Realistic–Imaginative	7	0.77
Practical–Conceptual	8	0.44
Experiential–Theoretical	8	0.74
Traditional–Original	8	0.69
T–F facet scales		
Logical–Empathetic	9	0.80
Reasonable–Compassionate	8	0.71
Questioning–Accommodating	7	0.48
Critical–Accepting	8	0.53
Tough–Tender	8	0.75
J–P facet scales		
Systematic–Casual	8	0.75
Planful–Open-Ended	6	0.77
Early Starting–Pressure-Prompted	6	0.75
Scheduled–Spontaneous	8	0.77
Methodical–Emergent	8	0.66
	Median	0.73

The internal consistency reliability of most scales is good, and the average (median) reliability achieves a value above 0.7. However, three scales (Practical–Conceptual, Questioning–Accommodating and Critical–Accepting) have notably lower alpha coefficients. This pattern is consistent with what was found during the development of the

Step II instrument, where these three facet scales were also found to show lower reliability than the others (Quenk, Hammer and Majors, 2004).

It should be noted, however, that coefficient alpha reliability statistics will give an underestimate of the reliability of the Step II facet scales, given the use of item response theory (IRT) methods in the actual scoring process.[7]

Group differences

This section looks at the way in which people from different groups responded to the Polish version of the European Step II questionnaire.

Gender

The means and standard deviations of the Step II facet scales are shown separately for males and females in Table 9.7, along with the difference in mean scores. This is represented graphically in Figure 9.2. Statistically significant differences were found between male and female mean scores for five of the 20 facet scales, with some consistent patterns emerging.

- On the E–I facet scales, the Gregarious–Intimate mean score was significantly further toward the E pole for males than for females.
- On the T–F facet scales, mean scores for four of the facets were significantly further toward the T pole for males than for females.
- There were no significant gender differences on the S–N or J–P facet scales.

[7] Item response theory (IRT) is an approach to measurement that is concerned with modelling the relationship between item responses and the underlying characteristic assessed by the scale or test the item is designed to measure. IRT can be used to select items for a test and/or to score the items.

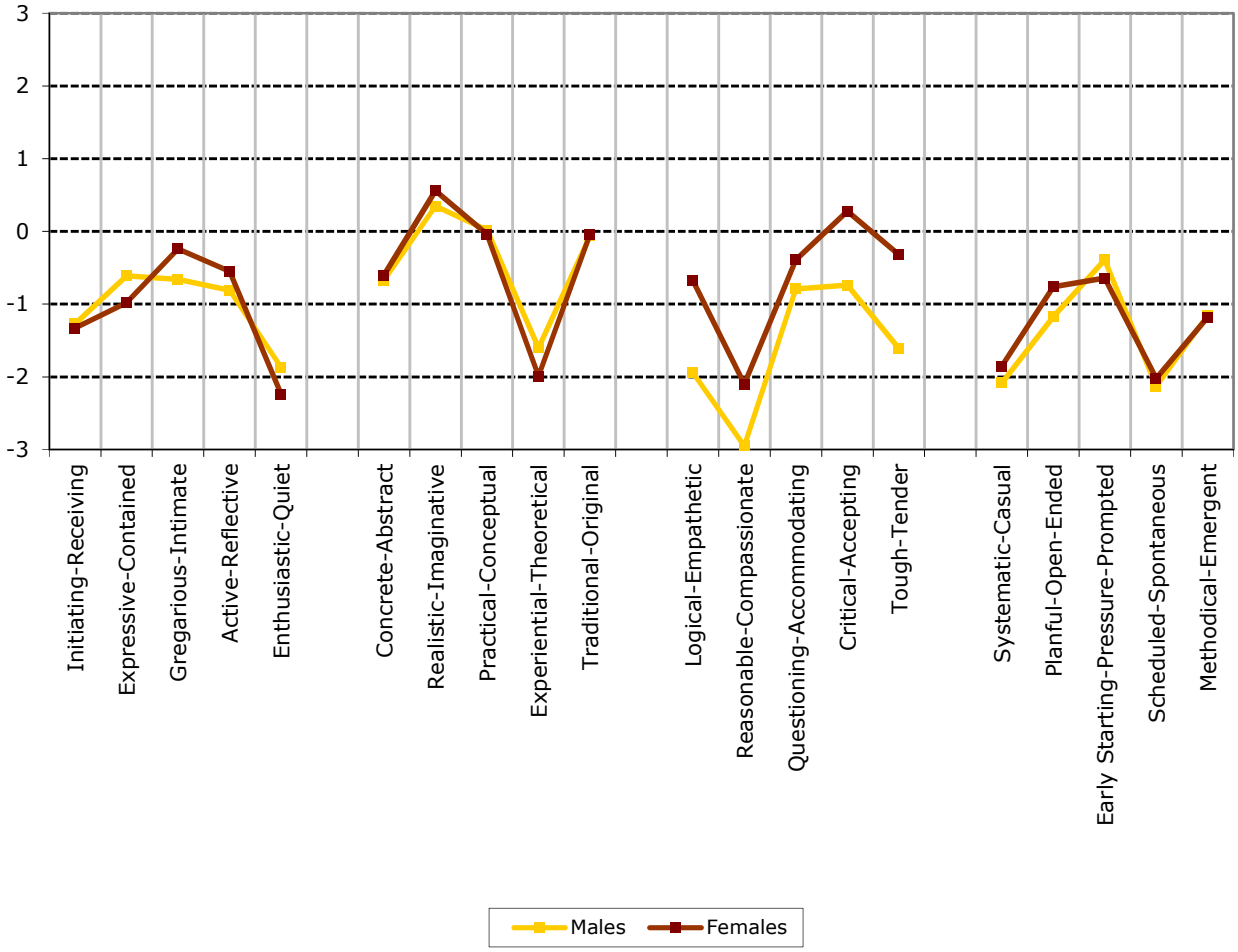
Table 9.7: Gender differences in facet scale scores

	Males (n=319)		Females (n=369)		Difference (M–F)[8]
Step II facet scale	**Mean**	**SD**	**Mean**	**SD**	
E–I facet scales					
Initiating–Receiving	-1.27	3.37	-1.33	3.21	0.06
Expressive–Contained	-0.61	2.98	-0.98	2.86	0.37
Gregarious–Intimate	-0.66	2.90	-0.24	2.57	-0.42*
Active–Reflective	-0.81	3.01	-0.55	3.04	-0.26
Enthusiastic–Quiet	-1.87	2.82	-2.24	2.50	0.37
S–N facet scales					
Concrete–Abstract	-0.67	2.65	-0.60	2.75	-0.07
Realistic–Imaginative	0.35	3.41	0.56	3.36	-0.21
Practical–Conceptual	0.02	2.52	-0.04	2.49	0.05
Experiential–Theoretical	-1.60	2.86	-1.99	2.65	0.39
Traditional–Original	-0.06	2.64	-0.04	2.72	-0.03
T–F facet scales					
Logical–Empathetic	-1.94	2.53	-0.67	2.87	-1.27**
Reasonable–Compassionate	-2.95	2.39	-2.10	2.67	-0.86**
Questioning–Accommodating	-0.79	2.66	-0.39	2.79	-0.39
Critical–Accepting	-0.74	2.10	0.28	2.22	-1.02**
Tough–Tender	-1.61	2.86	-0.32	2.88	-1.28**
J–P facet scales					
Systematic–Casual	-2.08	2.74	-1.85	2.76	-0.23
Planful–Open-Ended	-1.17	3.18	-0.76	3.33	-0.41
Early Starting–Pressure-Prompted	-0.39	3.37	-0.64	3.46	0.26
Scheduled–Spontaneous	-2.13	3.09	-2.02	3.05	-0.11
Methodical–Emergent	-1.15	2.87	-1.18	2.76	0.03

Difference significant at: *p<0.05, **p<0.01 (based on an independent samples t-test).

[8] A positive value indicates that male scores tend more towards I, N, F or P, and a negative value indicates that female scores tend more towards I, N, F or P.

Figure 9.2: Gender differences in facet scale scores



Age

There were found to be no meaningful links between age and facet scale scores. Correlational analysis showed only two facet scales to be significantly correlated with age, these correlations being 0.15 (Questioning–Accommodating) and 0.11 (Critical–Accepting), and even these are too small to be considered meaningful.

For interpretation purposes, it is reasonable to conclude that there are no clear relationships between age and facet scale scores.

Ethnic origin

Ethnic origin information was not captured for individuals who completed the Polish language version of the Step II questionnaire, so no analyses were conducted.

Occupational level

Research using the Polish version of the MBTI Step I questionnaire has demonstrated that individuals in employee-level jobs in organisations are more likely to have preferences for Feeling than those in higher-level jobs.[9] Although occupational-level data were captured for the Polish sample, the number of people in each of the categories was too small to allow a full analysis. Therefore, individuals were split into three categories, 'Employee', 'Supervisory/management level' and 'Senior executive/top level'. These categories were used for the analysis. Table 9.8 shows the facet scale means and standard deviations for these groups, with the data represented graphically in Figure 9.3. There were no significant differences in the findings.

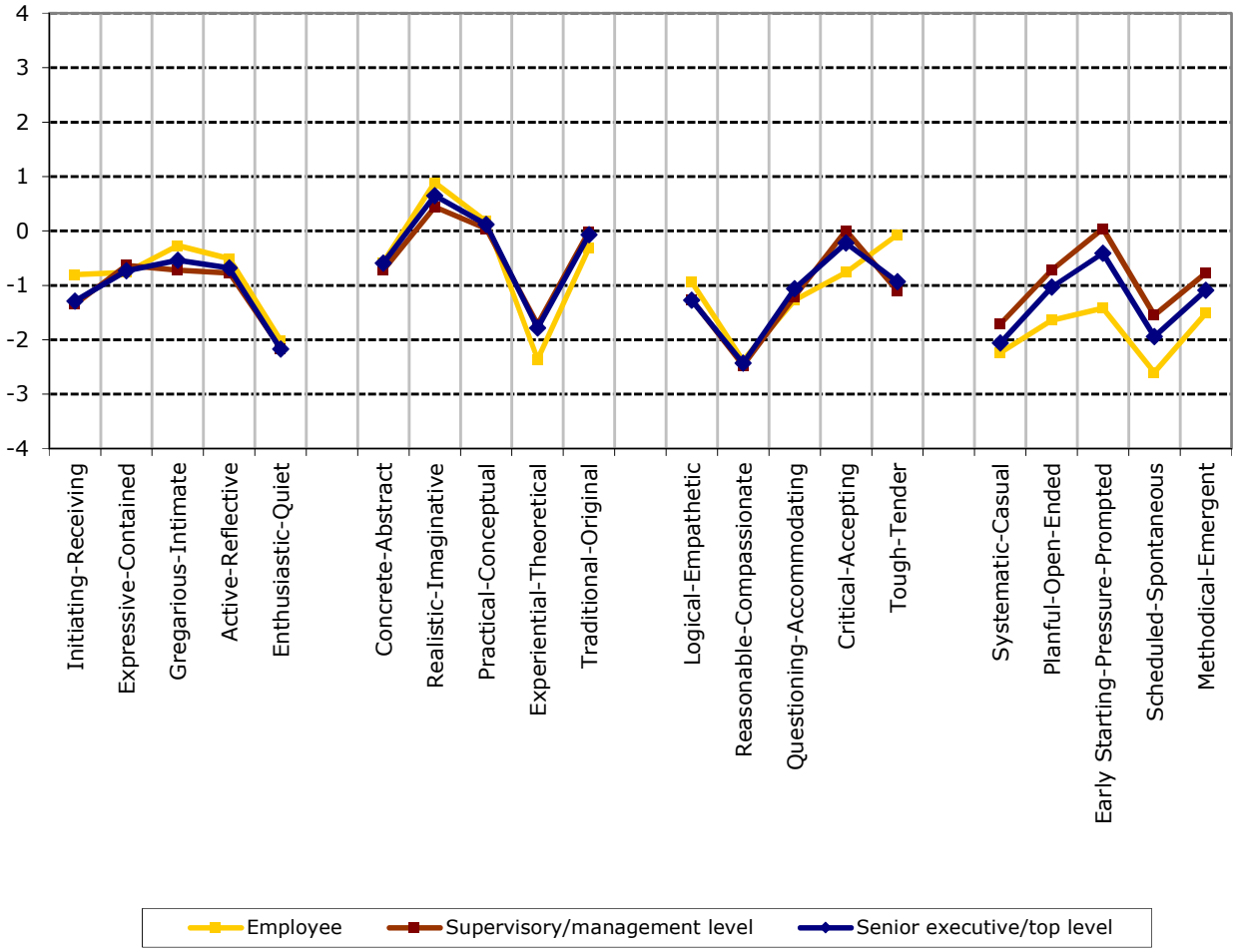
Table 9.8: Occupational level differences in facet scale scores

	Employee (n=113)		Supervisory/ management level (n=173)		Senior executive/top level (144)		Sig
Step II facet scale	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	
E–I facet scales							
Initiating–Receiving	-0.80	3.50	-1.35	3.30	-1.29	3.35	
Expressive–Contained	-0.76	3.01	-0.63	3.10	-0.73	2.98	
Gregarious–Intimate	-0.27	2.56	-0.72	2.89	-0.54	2.78	
Active–Reflective	-0.51	3.47	-0.77	3.30	-0.68	3.20	
Enthusiastic–Quiet	-2.02	2.30	-2.16	2.66	-2.17	2.55	
S–N facet scales							
Concrete–Abstract	-0.58	3.09	-0.72	2.59	-0.59	2.78	
Realistic–Imaginative	0.89	3.38	0.44	3.42	0.65	3.47	
Practical–Conceptual	0.18	2.87	0.04	2.31	0.12	2.54	
Experiential–Theoretical	-2.36	2.81	-1.72	2.92	-1.78	2.92	
Traditional–Original	-0.31	2.86	-0.01	2.78	-0.07	2.83	
T–F facet scales							
Logical–Empathetic	-0.93	3.09	-1.26	2.82	-1.27	2.86	
Reasonable–Compassionate	-2.38	2.71	-2.47	2.68	-2.43	2.70	
Questioning–Accommodating	-1.27	2.53	-1.21	2.83	-1.06	2.73	
Critical–Accepting	-0.76	1.89	0.01	2.32	-0.22	2.23	
Tough–Tender	-0.07	2.94	-1.10	2.93	-0.93	2.93	
J–P facet scales							
Systematic–Casual	-2.24	2.49	-1.71	2.90	-2.06	2.76	
Planful–Open-Ended	-1.64	3.10	-0.72	3.40	-1.03	3.36	
Early Starting–Pressure-Prompted	-1.42	3.37	0.04	3.44	-0.41	3.39	
Scheduled–Spontaneous	-2.60	2.74	-1.55	3.03	-1.94	3.03	
Methodical–Emergent	-1.51	2.66	-0.77	2.94	-1.09	2.87	

Significant at: *p<0.05, **p<0.01 (based on a one-way analysis of variance).

[9] See the MBTI Step I European Data Supplement for details.

Figure 9.3: Occupational level differences in facet scale scores



Education

Specific educational qualifications were not collected for the OPPassessment sample; however, the age at which individuals left full-time education was. There were found to be no meaningful links between the age at which people left full-time education and their facet scale scores. Correlational analysis found only two facet scales to be significantly correlated with the age people left full-time education. These two correlations were -0.08 (Enthusiastic–Quiet) and 0.11 (Reasonable–Compassionate) and both correlations were less than +/- 0.2 so too small to be considered meaningful.

Work area

Information regarding the area of work people engage in was collected for the group. For most categories were used, the sample size was too small for analysis, so the focus was on the three work areas with a reasonable sample size. These were as follows:

- Sales, customer service
- HR, training, guidance
- Finance

Table 9.9 (next page) shows the facet scale means and standard deviations for these three work areas. The findings can be summarised as follows:

- A pattern observed regarding the facets of the S–N scale was that mean scores for the 'HR, training, guidance' group tended more towards the N pole (or less to the S pole) for four of the facets than the other two groups.
- An interesting pattern was found regarding the facets of the T–F scale. Mean scores for the 'Finance' group tended more towards the T pole than the other two groups for four of the five facets, while the 'HR, training, guidance' group tended less towards the T pole than the other two groups for the same for facets.

Table 9.9: Mean facet scale scores by work area

Step II facet scale	Sales, customer service (n=117)		HR, training, guidance (n=88)		Finance (n=59)		Sig.
	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales							
Initiating–Receiving	-1.59	3.18	-1.24	3.41	-0.61	3.34	
Expressive–Contained	-0.94	3.02	-1.05	2.97	-0.25	2.91	
Gregarious–Intimate	-0.77	2.91	-0.26	2.82	-0.56	2.44	
Active–Reflective	-0.97	3.24	-0.58	3.18	-0.29	3.05	
Enthusiastic–Quiet	-2.57	2.56	-2.25	2.59	-1.66	2.37	
S–N facet scales							
Concrete–Abstract	-0.94	2.42	-0.15	3.05	-0.69	2.74	
Realistic–Imaginative	0.59	3.34	1.20	3.54	0.14	3.47	
Practical–Conceptual	-0.35	2.23	0.31	2.67	0.36	2.85	
Experiential–Theoretical	-2.07	2.61	-1.23	3.11	-2.29	2.59	*
Traditional–Original	-0.23	2.81	0.45	2.97	-0.37	2.56	
T–F facet scales							
Logical–Empathetic	-1.07	2.86	-0.60	3.07	-2.17	2.43	**
Reasonable–Compassionate	-2.44	2.59	-1.55	3.12	-3.22	2.16	**
Questioning–Accommodating	-0.96	2.65	-0.81	2.86	-1.02	2.84	
Critical–Accepting	-0.07	2.10	0.31	2.37	-0.95	2.19	**
Tough–Tender	-0.84	2.98	-0.34	3.20	-1.66	2.60	*
J–P facet scales							
Systematic–Casual	-2.03	2.90	-1.76	2.85	-2.36	2.35	
Planful–Open-Ended	-1.43	3.23	-0.78	3.48	-0.85	3.39	
Early Starting–Pressure-Prompted	-0.62	3.31	-0.39	3.60	-0.37	3.10	
Scheduled–Spontaneous	-2.09	2.89	-1.94	3.00	-1.92	3.30	
Methodical–Emergent	-1.18	2.93	-1.22	2.76	-0.95	2.69	

Significant at: *p<0.05, **p<0.01 (based on a one-way analysis of variance).

Nationality

Information on nationality was not captured for individuals who completed the Polish language version of the Step II questionnaire, so no analyses were conducted.

Employment status

The majority of people who disclosed their employment status worked full-time. There were insufficient people who did not work full-time for it to be possible to conduct any analyses across groups.

Appendix 1: Sample description

Sample 1: Data from OPPassessment (representative Polish-speaking professional and managerial sample)

This sample consists of 688 individuals who completed the MBTI Step II questionnaire in Polish via the OPPassessment system between September 2012 and July 2016. Of these individuals, 54% were female and 46% were male. Age ranged from 20 to 66 years, with a mean of 35 and a median of 34.

Nationality was not collected by OPPassessment.

Sixty-nine percent of people provided additional demographic information. The majority of those who disclosed their employment status were in full-time employment:

Employment status	Percentage
Full-time	93.4%
Self-employed	3.2%
Part-time	2.7%
Unemployed	0.4%
Homemaker	0.2%

Of the 68% who disclosed their occupational level, many were of managerial level or above, although the largest single group was employee level (24%):

Occupational level	Percentage
Top level	9.9%
Senior executive	21.0%
Upper middle management	8.4%
Middle management	10.5%
First-level management/supervisor	9.9%
Employee	24.2%
Other	7.9%

Sixty-nine per cent of respondents stated their work area. Amongst these, a range of work areas were represented:

Work area (job type)	Percentage
Sales, customer service	24.6%
HR, training, guidance	18.5%
Finance	12.4%
Research and development	7.8%
Business services	4.6%
IT	4.2%
Science, engineering	3.6%
Admin or secretarial	2.1%
Education	1.3%
Health, social services, etc	1.1%
Land, sea or air transport	0.4%
Leisure, personal service	0.2%
Other public sector	0.2%
Other private sector	8.2%
Other	10.9%

MBTI® Step II™ instrument

European Data Supplement

Russian

January 2018

2nd edition

OPP Ltd
Elsfield Hall
15–17 Elsfield Way
Oxford OX2 8EP
UK
t: +44 (0)1865 404 500

Contents

Introduction

Data collected for the European MBTI® Step II™ instrument were analysed to produce the findings in this supplement. This is the second data supplement produced for this instrument. A brief description of the sample is given below. Further details of the sample are provided in Appendix 1.

- The sample consisted of 1,281 individuals who completed the MBTI Step II questionnaire in Russian via the OPPassessment™ system between 2013 and mid-2016.[1] This sample is considered to be representative of the groups of people with whom the Russian MBTI Step II instrument has been and will be used for applications such as management development, coaching, counselling and teambuilding. As such, it is likely to represent a cross-section of the Russian-speaking professional and managerial population.

The results of the analyses are outlined below.

[1] OPPassessment allows personality questionnaires such as the MBTI instrument to be administered via email and/or completed online.

Type distribution

Type tables are a way of illustrating the proportion of each type within a particular group. Below is a type table for the Russian sample described on the previous page.

For each of the 16 different types, the number of cases, the percentage of the total that this represents and the self-selection ratio (SSR) are shown. The SSR (Myers et al., 1998) is a way of demonstrating whether a given type appears more or less often in a particular group than would be expected given its frequency in a reference group. An SSR of greater than 1 indicates that a type is over-represented, and an SSR of less than 1 denotes that it is under-represented. Asterisks are used to denote whether the over- or under-representations are statistically significant, based on the results of chi-square analysis[2].

Ideally, the type distribution from a large representative sample of the Russian population would be used to calculate SSRs in this data supplement. However, such a sample does not currently exist. In its place, SSRs have been calculated using type data from the UK general population (Kendall, 1998). Evidence (e.g. Hackston and Kendall, 2004; Quenk et al., 2004; Kirby, Kendall and Barger, 2007) does suggest that although type-related behaviours vary a good deal from country to country and from culture to culture, the frequencies of underlying MBTI types do not.

[2] Chi-square analysis (often abbreviated to χ^2) is a technique used to explore whether observed frequency distributions differ significantly from other, predefined, distributions. In this case, the UK general population group is used as the reference group, and the chi-square analysis indicates whether the proportion of people of each type within a particular sample differs significantly from the proportion of people reporting the same type within the reference group.

Questionnaire development sample

Table 10.1: Type table for OPPassessment data (reported type, n=1,281)

ISTJ	ISFJ	INFJ	INTJ
n=318 24.8% SSR=1.81**	n=16 1.2% SSR=0.09**	n=18 1.4% SSR=0.82	n=91 7.1% SSR=5.07**
ISTP	**ISFP**	**INFP**	**INTP**
n=50 3.9% SSR=0.61**	n=2 0.2% SSR=0.03**	n=18 1.4% SSR=0.44**	n=32 2.5% SSR=1.04
ESTP	**ESFP**	**ENFP**	**ENTP**
n=35 2.7% SSR=0.47**	n=10 0.8% SSR=0.09**	n=19 1.5% SSR=0.24**	n=65 5.1% SSR=1.82**
ESTJ	**ESFJ**	**ENFJ**	**ENTJ**
n=370 28.9% SSR=2.78**	n=27 2.1% SSR=0.17**	n=42 3.3% SSR=1.18	n=168 13.1% SSR=4.52**

Type	n	%
E	736	57.5%**
I	545	42.5%**
S	828	64.6%**
N	453	35.4%**
T	1129	88.1%**
F	152	11.9%**
J	1050	82.0%**
P	231	18.0%**

*Difference significant at p<0.05, based on chi-square results.
**Difference significant at p<0.01, based on chi-square results.

Looking at reported type, the most frequent type preference is ESTJ; this is a common finding with managerial groups in other countries. The SSR results suggest that, in comparison with the UK general population, those with preferences for NT are over-represented, and those with preferences for SF are under-represented. Again, this is a common finding with managerial groups.

Properties of the Step II facet scales

Facet scale score distribution

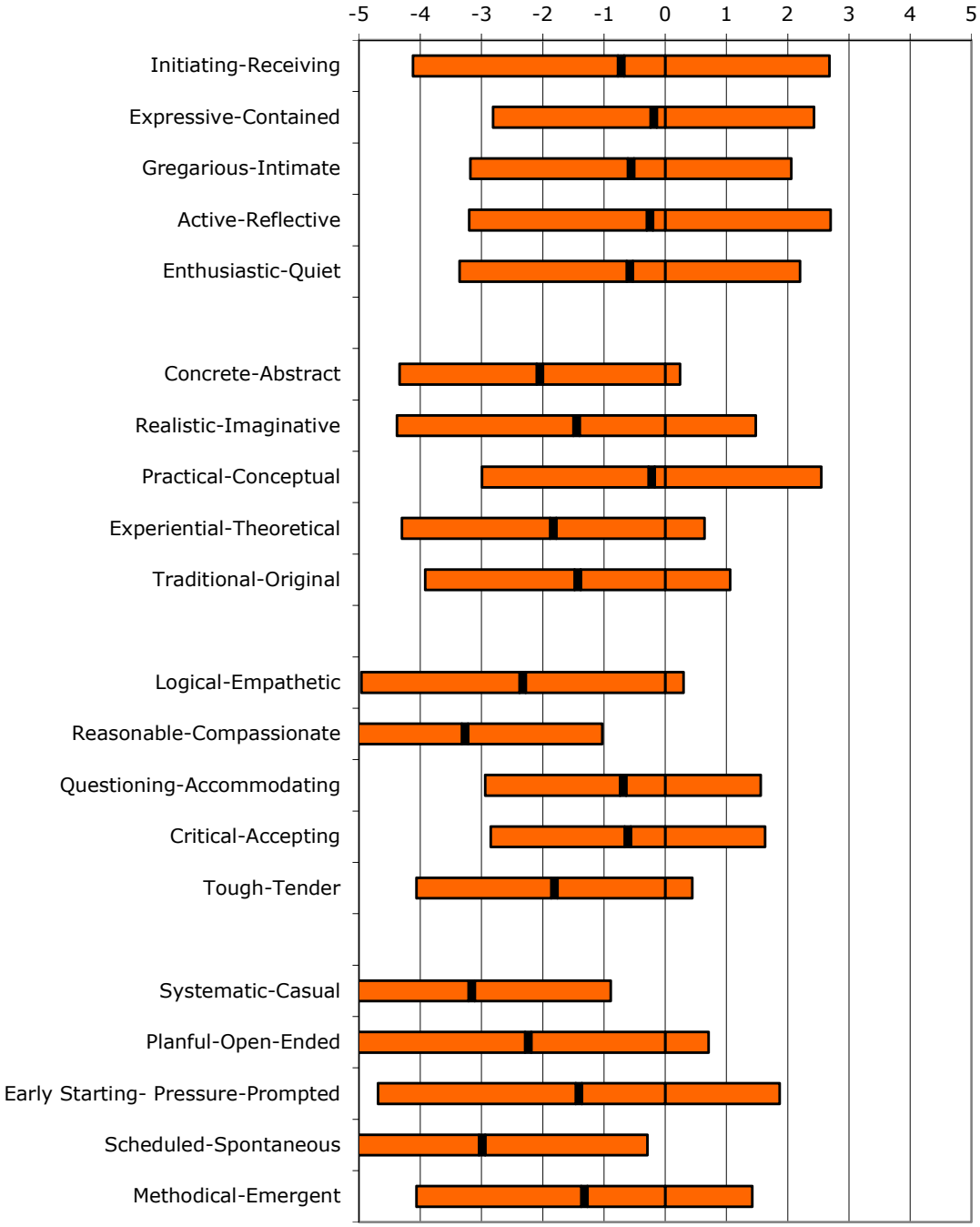
The means and standard deviations of the Step II facets are shown in Table 10.2, and are illustrated graphically in Figure 10.1. The mean score for a scale is calculated by adding together the scores for each individual in the sample and then dividing the sum by the number of individuals. Note that negative values indicate that mean scores are nearer the left-hand pole of the facet, and positive values indicate that the mean scores are nearer the right-hand pole. The standard deviation (SD) is a statistical measure describing the degree to which the scores from the sample either bunch up close to or are scattered widely around the mean for the sample.

Table 10.2: Means and standard deviations of the facet scales

Step II facet scale	Mean[3]	SD
E–I facet scales		
Initiating–Receiving	-0.72	3.40
Expressive–Contained	-0.19	2.62
Gregarious–Intimate	-0.56	2.62
Active–Reflective	-0.25	2.95
Enthusiastic–Quiet	-0.58	2.78
S–N facet scales		
Concrete–Abstract	-2.05	2.29
Realistic–Imaginative	-1.45	2.93
Practical–Conceptual	-0.22	2.77
Experiential–Theoretical	-1.83	2.47
Traditional–Original	-1.43	2.49
T–F facet scales		
Logical–Empathetic	-2.33	2.63
Reasonable–Compassionate	-3.27	2.24
Questioning–Accommodating	-0.69	2.25
Critical–Accepting	-0.51	2.14
Tough–Tender	-1.81	2.25
J–P facet scales		
Systematic–Casual	-3.11	2.22
Planful–Open-Ended	-2.24	2.95
Early Starting–Pressure-Prompted	-1.41	3.28
Scheduled–Spontaneous	-2.99	2.70
Methodical–Emergent	-1.32	2.74

[3] Scale means: in Step II reports, scores are given from 5 on one pole, through 0, to 5 on the opposite pole. In compiling statistical information, however, one pole needs to be clearly distinguished from the other. Throughout this data supplement, therefore, a convention has been adopted of scores 5, 4, 3, 2 or 1 on the left-hand pole being assigned values of –5, –4, –3, –2 or –1 and scores on the right-hand pole being assigned positive scores. This does not, of course, imply any suggestion that positive scores are 'better' than negative scores.

Figure 10.1: Facet scale mean scores[4]



In general the means centre quite close to the midpoint of each scale, with standard deviations between 2 and 3.5; however six scales have a mean greater than 2 points from the midpoint, in the S, T or J direction. Overall, a clear pattern emerges. The mean scores tend

[4] For each facet scale, the central line shows the mean and the coloured bars to either side indicate the standard deviation.

towards the E, S, T and J direction, which is consistent with the most common four-letter type preference amongst this sample. Amongst the group as a whole, there are more individuals with a preference for E, S, T and J.

Facet scale intercorrelations

Correlations among the Step II facet scales are shown in Table 10.3. Facet scales within each dichotomy usually correlate more highly (often substantially so) with the other scales of the same dichotomy than they do with scales in the other three dichotomies. For example, every E–I facet scale correlates more highly with the other four E–I scales than with any of the S–N, T–F or J–P facet scales.

There is only one scale that correlates more highly with scales in other dichotomies. The T–F scale Questioning–Accommodating correlates at a level of -0.28 with Traditional-Original, -0.26 with Practical–Conceptual and -0.19 with Concrete–Abstract on the S–N dichotomy. By way of comparison, it correlates at between 0.13 and 0.15 with the other four scales on the same dichotomy as itself (T–F). The negative correlations between Questioning–Accommodating and the S–N scales are consistent with findings from the US version of the Step II instrument, and suggest that a questioning approach to differences of opinion seems to be related to a range of iNtuition (N) facets. These patterns are similar to those found when the European Step II instrument was developed.

Table 10.3: Intercorrelations of Step II facet scales

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Initiating–Receiving	**1.00**																			
2. Expressive–Contained	**0.43**	**1.00**																		
3. Gregarious–Intimate	**0.56**	**0.46**	**1.00**																	
4. Active–Reflective	**0.64**	**0.46**	**0.52**	**1.00**																
5. Enthusiastic–Quiet	**0.53**	**0.51**	**0.60**	**0.58**	**1.00**															
6. Concrete–Abstract	-0.08	-0.09	-0.08	-0.05	-0.18	**1.00**														
7. Realistic–Imaginative	-0.08	-0.13	-0.11	-0.11	-0.24	**0.61**	**1.00**													
8. Practical–Conceptual	-0.08	-0.08	-0.07	-0.08	-0.17	**0.56**	**0.55**	**1.00**												
9. Experiential–Theoretical	-0.03	-0.04	0.00	-0.02	-0.06	**0.60**	**0.48**	**0.44**	**1.00**											
10. Traditional–Original	-0.13	-0.08	-0.13	-0.15	-0.28	**0.58**	**0.54**	**0.55**	**0.37**	**1.00**										
11. Logical–Empathetic	-0.03	-0.20	-0.07	0.01	-0.19	0.30	0.46	0.19	0.19	0.22	**1.00**									
12. Reasonable–Compassionate	0.01	-0.14	-0.01	0.06	-0.10	0.28	0.35	0.16	0.20	0.17	**0.65**	**1.00**								
13. Questioning–Accommodating	0.13	0.01	0.08	0.13	0.12	-0.19	-0.11	-0.26	-0.13	-0.28	**0.15**	**0.14**	**1.00**							
14. Critical–Accepting	-0.11	-0.16	-0.15	-0.07	-0.17	0.20	0.29	0.16	0.11	0.10	**0.43**	**0.41**	**0.13**	**1.00**						
15. Tough–Tender	-0.02	-0.16	-0.06	0.03	-0.15	0.25	0.32	0.15	0.17	0.13	**0.57**	**0.53**	**0.15**	**0.42**	**1.00**					
16. Systematic–Casual	0.04	-0.07	-0.01	-0.01	-0.13	0.48	0.48	0.34	0.30	0.44	0.34	0.29	-0.05	0.18	0.29	**1.00**				
17. Planful–Open-Ended	0.09	0.01	0.04	0.04	-0.04	0.30	0.26	0.21	0.21	0.27	0.19	0.16	-0.06	0.04	0.12	**0.63**	**1.00**			
18. Early Starting–Pressure-Prompted	0.02	-0.06	-0.02	-0.05	-0.10	0.22	0.19	0.16	0.17	0.24	0.12	0.09	-0.06	-0.02	0.09	**0.48**	**0.50**	**1.00**		
19. Scheduled–Spontaneous	0.11	-0.03	0.04	0.07	-0.03	0.29	0.29	0.19	0.21	0.24	0.22	0.19	-0.01	0.03	0.17	**0.65**	**0.65**	**0.53**	**1.00**	
20. Methodical–Emergent	0.11	-0.05	-0.01	0.04	-0.10	0.20	0.19	0.09	0.12	0.16	0.15	0.13	0.01	0.04	0.14	**0.44**	**0.50**	**0.45**	**0.56**	**1.00**

Correlations of Step II facet scales with Step I scales

Correlations between Step II facet scales and the continuous scores from the MBTI Step I™ instrument are shown in Table 10.4.[5] Positive correlations between facet scales and the corresponding Step I dimension (e.g. between Initiating–Receiving and the E–I continuous score) indicate relationships in the expected direction, with higher facet scores tending to be associated with higher continuous scores, and vice versa. Negative correlations would indicate an inverse relationship between the facet scale scores and the corresponding Step I dimension.

Table 10.4: Correlations of Step II facet scales with Step I continuous scores

	Step I continuous score			
Step II facet scales	**E–I**	**S–N**	**T–F**	**J–P**
E–I facet scales				
Initiating–Receiving	**0.80**	-0.11	-0.03	0.09
Expressive–Contained	**0.67**	-0.10	-0.21	-0.04
Gregarious–Intimate	**0.68**	-0.10	-0.08	0.02
Active–Reflective	**0.79**	-0.12	0.01	0.03
Enthusiastic–Quiet	**0.78**	-0.26	-0.18	-0.08
S–N facet scales				
Concrete–Abstract	-0.11	**0.81**	0.30	0.33
Realistic–Imaginative	-0.16	**0.76**	0.44	0.30
Practical–Conceptual	-0.10	**0.67**	0.20	0.22
Experiential–Theoretical	-0.04	**0.66**	0.18	0.24
Traditional–Original	-0.20	**0.73**	0.19	0.30
T–F facet scales				
Logical–Empathetic	-0.12	0.32	**0.88**	0.23
Reasonable–Compassionate	-0.04	0.26	**0.75**	0.19
Questioning–Accommodating	0.13	-0.26	**0.22**	-0.04
Critical–Accepting	-0.15	0.20	**0.57**	0.05
Tough–Tender	-0.07	0.22	**0.72**	0.19
J–P facet scales				
Systematic–Casual	-0.04	0.48	0.34	**0.74**
Planful–Open-Ended	0.04	0.29	0.18	**0.83**
Early Starting–Pressure-Prompted	-0.05	0.24	0.09	**0.72**
Scheduled–Spontaneous	0.05	0.28	0.20	**0.81**
Methodical–Emergent	0.02	0.18	0.15	**0.68**

The E–I facet scales correlate at 0.67–0.80 with the E–I continuous scores from Step I; the S–N facet scales correlate at 0.66–0.81 with the S–N continuous scores; the T–F facet scales correlate at 0.22–0.88

[5] Continuous scores place an individual's score on each dimension onto a continuous scale with a mid-point of 100. To calculate continuous scores, Preference Clarity Index (PCI) scores for each dimension are either subtracted or added to 100, depending on which direction the overall preference is. PCI scores in the direction of E, S, T or J are subtracted from 100. PCI scores in the direction of I, N, F or P are added to 100.

with the T–F continuous scores; and the J–P facet scales correlate at 0.68–0.83 with the J–P continuous scores. The scale that showed the lowest correlation with its associated Step I continuous score was Questioning–Accommodating at 0.22. This scale is also lowest on internal consistency (see page 227, and has been previously found to have the lowest test-retest reliability (Quenk, Hammer and Majors, 2004), which would result in the true correlations being underestimated.

These correlations are very similar to those found during the development of the Step II questionnaire. This consistency, alongside the fact that the correlations between Step II facet scales and Step I continuous scores associated with their own dimension tend to be higher than correlations with the other three dimensions, provides compelling evidence for the theoretical hierarchical structure of the Step II facet scales in relation to the Step I scales.

Out-of-preference scores

It is known that the five facet scales relating to each type dimension do not describe the dimension in its entirety; there will not be a precise and exact overlap between, for example, an individual's score on E–I and their total score across the five facet scales that relate to this dimension. For example, it is not uncommon to see an Enthusiastic Introvert or an Intimate Extravert. Such apparent inconsistencies are usually known as 'out-of-preference scores' (OOPS).

Although it is usual to have a number of OOPS in any one profile, it should be unusual to find that OOPS make up more than half of the facet scores relating to any one type dichotomy. Table 10.5 demonstrates that the proportion of people for whom this happens ranges from 2.8% for the S–N block to 0.7% for the T–F block. It is therefore a very infrequent occurrence.

Table 10.5: Proportion of OOPS by type dichotomy[6]

	Proportion of 'reported type' OOPS					
	None	One	Two	Three	Four	Five
E–I	69.5%	23.8%	5.8%	0.8%	0.2%	0.0%
S–N	63.2%	24.3%	9.7%	2.1%	0.5%	0.2%
T–F	70.2%	24.9%	4.2%	0.6%	0.1%	0.0%
J–P	67.1%	25.2%	6.6%	0.9%	0.1%	0.0%

[6] In this table, an OOPS is defined as a score of 2, 3, 4 or 5 on the 'wrong' pole of the facet when compared with the reported type.

For any individual facet scale, an OOPS tends to occur in approximately 10% of profiles. The exception to this is Experiential–Theoretical, where OOPS occur in approximately 17% of cases.

Reliability

The reliability of a test or questionnaire relates to how consistent and precise it is. Internal consistency reliability addresses the question of whether all the questions in a scale measure the same construct. For example, are the Step II facet scales consistent within themselves, and do they hold together well as scales? A common measure of internal consistency reliability is coefficient alpha (Cronbach, 1951). The alpha coefficients for the Step II facet scales are shown in Table 10.6.

Table 10.6: Internal consistency reliability

Step II facet scale	No. of items	Coefficient alpha
E–I facet scales		
Initiating–Receiving	8	0.85
Expressive–Contained	7	0.66
Gregarious–Intimate	7	0.61
Active–Reflective	8	0.71
Enthusiastic–Quiet	9	0.68
S–N facet scales		
Concrete–Abstract	9	0.67
Realistic–Imaginative	7	0.70
Practical–Conceptual	8	0.50
Experiential–Theoretical	8	0.70
Traditional–Original	8	0.66
T–F facet scales		
Logical–Empathetic	9	0.80
Reasonable–Compassionate	8	0.67
Questioning–Accommodating	7	0.33
Critical–Accepting	8	0.42
Tough–Tender	8	0.67
J–P facet scales		
Systematic–Casual	8	0.72
Planful–Open-Ended	6	0.79
Early Starting–Pressure-Prompted	6	0.71
Scheduled–Spontaneous	8	0.75
Methodical–Emergent	8	0.62
	Median	0.67

The internal consistency reliability of most scales is good, and the average (median) reliability achieves a value of 0.67. However, three scales (Practical–Conceptual, Questioning–Accommodating and Critical–Accepting) have notably lower alpha coefficients.

This pattern is fairly consistent with what was found during the development of the Step II instrument, where three of these facet scales (Practical–Conceptual, Questioning–Accommodating and Critical–Accepting) were found to show lower reliability than the others (Quenk, Hammer and Majors, 2004).

It should be noted, however, that coefficient alpha reliability statistics will give an underestimate of the reliability of the Step II facet scales, given the use of item response theory (IRT) methods in the actual scoring process.[7]

Group differences

This section looks at the way in which people from different groups responded to the Russian version of the European Step II questionnaire.

Gender

The means and standard deviations of the Step II facet scales are shown separately for males and females in Table 10.7, along with the difference in mean scores. This is represented graphically in Figure 10.2. Statistically significant differences were found between male and female mean scores for 15 of the 20 facet scales, with some consistent patterns emerging.

- On the E–I facet scales, all five mean scores tended slightly towards the E pole (or at the mid-point) for both males and females. There were statistically significant gender differences for two of the facet scales, with mean scores tending further towards the E pole for females than for males.
- On the S–N facet scales, all five mean scores tended towards the S pole (or at the mid-point) for both males and females. There were statistically significant gender differences for four of the facet scales, with males tending further towards the S pole than females on all four scales.
- On the T–F facet scales, all five mean scores tended towards the T pole (or at the mid-point) for both males and females. There were statistically significant gender differences for all five facet scales, with mean scores tending more towards the T pole for males than females on all five scales.
- On the J–P facet scales, all five mean scores tended further towards the J pole for both males and females. There were statistically significant gender differences for four of the facet scales, with male mean scores tending more towards the J pole than females.

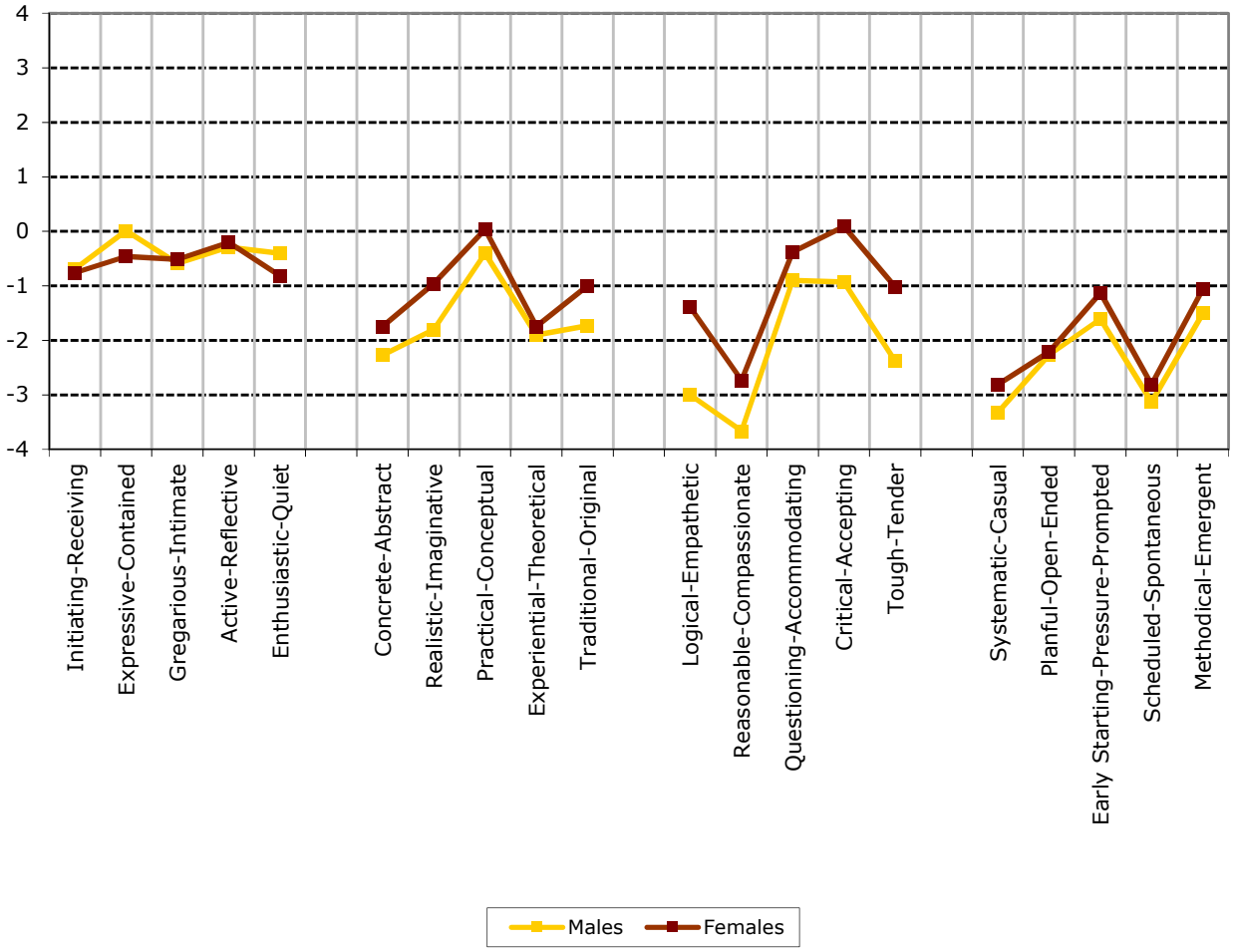
[7] Item response theory (IRT) is an approach to measurement that is concerned with modelling the relationship between item responses and the underlying characteristic assessed by the scale or test the item is designed to measure. IRT can be used to select items for a test and/or to score the items.

Table 10.7: Gender differences in facet scale scores

	Males (n=748)		Females (n=533)		Difference (M–F)
Step II facet scale	**Mean**	**SD**	**Mean**	**SD**	
E–I facet scales					
Initiating–Receiving	-0.69	3.43	-0.76	3.36	0.07
Expressive–Contained	0.01	2.62	-0.46	2.58	0.47**
Gregarious–Intimate	-0.59	2.70	-0.51	2.51	-0.08
Active–Reflective	-0.29	2.98	-0.20	2.90	-0.10
Enthusiastic–Quiet	-0.40	2.81	-0.82	2.73	0.43**
S–N facet scales					
Concrete–Abstract	-2.27	2.16	-1.74	2.42	-0.53**
Realistic–Imaginative	-1.80	2.78	-0.96	3.07	-0.85**
Practical–Conceptual	-0.40	2.77	0.04	2.76	-0.44**
Experiential–Theoretical	-1.90	2.49	-1.74	2.45	-0.16
Traditional–Original	-1.73	2.38	-1.00	2.58	-0.72**
T–F facet scales					
Logical–Empathetic	-3.00	2.20	-1.39	2.88	-1.62**
Reasonable–Compassionate	-3.66	1.98	-2.73	2.47	-0.93**
Questioning–Accommodating	-0.90	2.21	-0.38	2.26	-0.52**
Critical–Accepting	-0.93	1.99	0.09	2.20	-1.03**
Tough–Tender	-2.38	2.12	-1.02	2.19	-1.36**
J–P facet scales					
Systematic–Casual	-3.32	2.09	-2.81	2.35	-0.50**
Planful–Open-Ended	-2.26	3.01	-2.21	2.88	-0.04
Early Starting–Pressure-Prompted	-1.61	3.22	-1.13	3.36	-0.48*
Scheduled–Spontaneous	-3.13	2.67	-2.81	2.74	-0.32*
Methodical–Emergent	-1.50	2.70	-1.06	2.79	-0.44**

Difference significant at: *p<0.05, **p<0.01 (based on an independent samples t-test).

Figure 10.2: Gender differences in facet scale scores



Age

There were found to be no meaningful links between age and facet scale scores. Correlational analysis showed only four facet scales to be significantly correlated with age, the largest of these correlations being -0.15 (Traditional–Original) and even this is too small to be considered meaningful.

For interpretation purposes, it is reasonable to conclude that there are no clear relationships between age and facet scale scores.

Ethnic origin

Ethnic origin information was not captured for individuals who completed the Russian language version of the Step II questionnaire, so no analyses were conducted.

Occupational level

Although occupational level data were captured for the Russian sample, the number of people in some of the categories was too small to allow a full analysis. Therefore, the categories were re-coded into four broader groups, namely 'employee/supervisor/1st level

management' (containing 146 people), 'middle management level' (463 people), 'upper middle management level' (158 people) and 'senior executive/top level' (120 people). These categories were used for the analysis. Table 10.8 shows the facet scale means and standard deviations for these groups, with the data represented graphically in Figure 10.3. The findings can be summarised as follows:

No clear pattern was found regarding the facets relating to the E–I dimension or the S–N dimension.

A consistent pattern was found regarding the facets relating to the T–F dimension. Mean facet scores for the 'Employee/supervisor/1st level management' group tended least towards the T pole, with mean scores for the other three groups tending further towards the T pole with each progression of occupational level, so that the 'Senior executive/top level' group tended furthest towards the T pole.

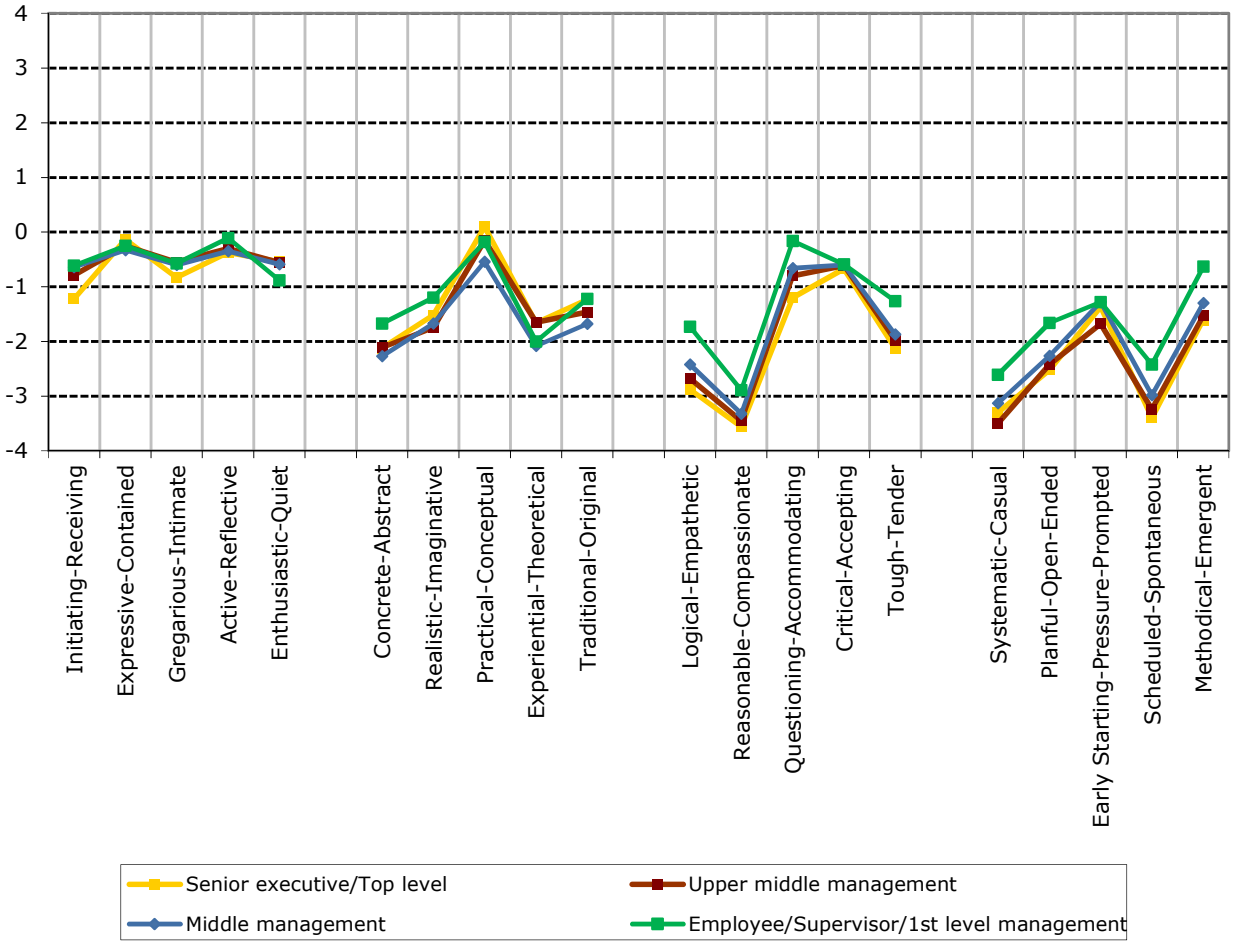
A pattern found with the facets relating to the J–P dimension was that mean facet scores for the 'Employee/supervisor/1st level management' group tended less towards the J pole than the other three level groups.

Table 10.8: Occupational level differences in facet scale scores

	Employee/supervisor/1st level management (n=146)		Middle management level (n=463)		Upper middle management level (n=158)		Senior executive/top level (120)	
Step II facet scale	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
E–I facet scales								
Initiating–Receiving	-0.61	3.39	-0.64	3.43	-0.79	3.22	-1.22	3.16
Expressive–Contained	-0.25	2.81	-0.33	2.62	-0.25	2.54	-0.13	2.36
Gregarious–Intimate	-0.57	2.71	-0.60	2.64	-0.55	2.59	-0.83	2.56
Active–Reflective	-0.11	3.12	-0.35	3.06	-0.30	2.88	-0.37	2.87
Enthusiastic–Quiet	-0.88	2.94	-0.59	2.75	-0.56	2.81	-0.54	2.67
S–N facet scales								
Concrete–Abstract	-1.67	2.64	-2.27	2.15	-2.11	2.28	-2.12	2.34
Realistic–Imaginative	-1.20	2.97	-1.67	2.88	-1.75	2.93	-1.52	3.00
Practical–Conceptual	-0.18	2.66	-0.54	2.90	-0.15	2.78	0.10	2.55
Experiential–Theoretical	-2.00	2.48	-2.08	2.36	-1.65	2.64	-1.66	2.36
Traditional–Original	-1.22	2.65	-1.68	2.44	-1.46	2.62	-1.23	2.38
T–F facet scales								
Logical–Empathetic	-1.73	3.06	-2.42	2.55	-2.68	2.59	-2.88	2.14
Reasonable–Compassionate	-2.89	2.61	-3.32	2.15	-3.45	2.32	-3.56	1.95
Questioning–Accommodating	-0.16	2.27	-0.66	2.23	-0.80	2.15	-1.20	1.20
Critical–Accepting	-0.59	2.42	-0.60	2.08	-0.61	2.01	-0.67	1.84
Tough–Tender	-1.26	2.57	-1.87	2.23	-1.98	2.32	-2.12	2.06
J–P facet scales								
Systematic–Casual	-2.61	2.44	-3.13	2.23	-3.50	2.00	-3.30	2.25
Planful–Open-Ended	-1.66	3.16	-2.26	2.96	-2.42	3.05	-2.52	2.85
Early Starting–Pressure-Prompted	-1.28	3.36	-1.28	3.29	-1.68	3.37	-1.37	3.14
Scheduled–Spontaneous	-2.42	2.83	-2.98	2.71	-3.24	2.69	-3.40	2.54
Methodical–Emergent	-0.63	2.95	-1.30	2.65	-1.53	2.82	-1.61	2.59

Significant at: *p<0.05, **p<0.01 (based on a one-way analysis of variance).

Figure 10.3: Occupational level differences in facet scale scores



Education

Specific educational qualifications were not collected for the sample; however, the age at which individuals left full-time education was. Correlational analysis showed only one facet scale to be significantly correlated with age, and this was less than 0.1, so too small to be indicative of any meaningful relationship.

Work area

Information regarding the area of work people engage in was collected for the group. For most categories the sample size was too small for analysis, so the focus was on the three work areas with a reasonable sample size. These were as follows:

- Sales, customer service
- HR, training, guidance
- Finance

Table 10.9 shows the facet scale means and standard deviations for these three work areas. The findings can be summarised as follows:

- An interesting pattern was found regarding the facets relating to the E–I dimensions. Although mean scores for all three groups tended towards the E pole for each facet scale, scores were further towards the E pole for the 'Sales, customer service' group than for the other two groups for three of the facet scales. In contrast, mean scores for the 'HR, training, guidance' group were less towards the E pole than the other two groups on four of the five facet scales
- The most noticeable pattern regarding the facets relating to the T–F dimension was that, while mean scores for all three groups tended towards the T pole for each facet scale, scores for the 'HR, training, guidance' group were less towards the T pole than the other two groups on all five facet scales.
- There were no clear overall patterns regarding the facets relating to the S–N dimensions or the J–P dimensions.

Table 10.9: Mean facet scale scores by work area

Step II facet scale	Sales, customer service (n=319)		HR, training, guidance (n=136)		Finance (n=119)		Sig.
	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales							
Initiating–Receiving	-1.03	3.27	-0.40	3.45	-0.82	3.14	
Expressive–Contained	-0.63	2.61	-0.29	2.62	-0.13	2.39	
Gregarious–Intimate	-0.98	2.62	-0.32	2.67	-0.63	2.23	*
Active–Reflective	-0.62	2.91	-0.04	2.95	-0.41	2.94	
Enthusiastic–Quiet	-1.12	2.78	-0.29	2.86	-0.40	2.64	**
S–N facet scales							
Concrete–Abstract	-2.26	2.14	-2.01	2.39	-2.15	2.06	
Realistic–Imaginative	-1.59	2.85	-1.51	2.99	-1.61	3.03	
Practical–Conceptual	-0.38	2.76	-0.32	2.93	-0.11	2.85	
Experiential–Theoretical	-2.07	2.22	-1.74	2.63	-1.99	2.44	
Traditional–Original	-1.54	2.42	-1.49	2.57	-1.41	2.22	
T–F facet scales							
Logical–Empathetic	-2.45	2.62	-1.72	2.75	-2.45	2.39	*
Reasonable–Compassionate	-3.46	2.14	-2.78	2.39	-3.37	2.05	**
Questioning–Accommodating	-0.78	2.05	-0.32	2.23	-0.51	2.25	
Critical–Accepting	-0.54	2.14	-0.07	1.96	-0.62	1.99	*
Tough–Tender	-2.02	2.32	-1.28	2.38	-1.53	2.16	**
J–P facet scales							
Systematic–Casual	-3.29	2.19	-2.96	2.25	-3.24	2.20	
Planful–Open-Ended	-2.46	2.95	-2.21	2.94	-2.40	2.85	
Early Starting–Pressure-Prompted	-1.60	3.21	-1.45	3.43	-1.62	3.09	
Scheduled–Spontaneous	-3.29	2.63	-2.84	2.78	-3.04	2.54	
Methodical–Emergent	-1.32	2.66	-1.45	2.72	-1.42	2.71	

Nationality

Information on nationality was not collected by OPPassessment, therefore it was not possible to conduct any analysis.

Employment status

The majority of people who disclosed their employment status worked full time. There were insufficient people who did not work full time for it to be possible to conduct any analyses across groups.

Appendix 1: Sample description

Sample 1: Data from OPPassessment (representative Russian-speaking professional and managerial sample)

This sample consists of 1281 individuals who completed the MBTI Step II questionnaire in Russian via the OPPassessment system between January 2013 and July 2016. Of these individuals, 58% were male and 42% were female. Age ranged from 22 to 61 years, with a mean of 36 and a median of 35.

Nationality was not collected by OPPassessment.

The majority of those who disclosed their employment status were in full-time employment:

Employment status	Percentage
Full-time	76.4%
Part-time	1.0%
Self-employed	0.4%
Unemployed	0.3%
Not disclosed	21.9%

Of those who disclosed their occupational level, many were of managerial level or above, although the largest single group was middle management level (36.1%):

Occupational level	Percentage
Top level	7.8%
Senior executive	1.6%
Upper middle management	12.3%
Middle management	36.1%
First-level management/supervisor	5.5%
Employee	5.9%
Other	3.8%
Not disclosed	26.9%

A range of work areas were represented:

Work area (job type)	Percentage
Sales, customer service	24.9%
HR, training, guidance	10.6%
Finance	9.3%
IT	3.4%
Business services	2.4%
Science, engineering	2.2%
Skilled operative	1.1%
Research and development	0.9%
Admin or secretarial	0.7%
Education	0.7%
Land, sea or air transport	0.5%
Health, social services, etc	0.2%
Leisure, personal service	0.1%
Military, police, prison, fire	0.1%
Other private sector	1.4%
Other public sector	0.2%
Other	16.5%
Not disclosed	24.8%

MBTI® Step II™ instrument

European Data Supplement

Spanish

January 2018

2nd edition

unlocking potential

OPP Ltd
Elsfield Hall
15–17 Elsfield Way
Oxford OX2 8EP
UK
t: +44 (0)1854 404 500

Contents

Introduction

Data collected for the European MBTI® Step II™ instrument were analysed to produce the findings in this supplement. This is the second data supplement produced for this instrument. A brief description of the sample is given below. Further details of the sample are provided in Appendix 1.

- The sample consisted of 2486 individuals who completed the MBTI Step II instrument in Spanish via the OPPassessment™ system between 2006 and mid-2016.[1] This sample is considered to be representative of the groups of people with whom the Spanish MBTI Step II instrument has been and will be used for applications such as management development, coaching, counselling and teambuilding. As such, it is likely to represent a cross-section of the Spanish-speaking professional and managerial population.

The results of the analyses are outlined below.

[1] OPPassessment allows personality questionnaires such as the MBTI instrument to be administered via email and/or completed online.

Type distribution

Type tables are a way of illustrating the proportion of each type within a particular group. Below is a type table taken from the sample described on the previous page.

For each of the 16 different types, the number of cases, the percentage of the total that this represents and the self-selection ratio (SSR) are shown. The SSR (Myers et al., 1998) is a way of demonstrating whether a given type appears more or less often in a particular group than would be expected given its frequency in a reference group. Ideally, the type distribution from a large representative sample of the Spanish population would be used to calculate SSRs in this data supplement. However, such a sample does not currently exist. In its place, SSRs have been calculated using type data from the UK general population (Kendall, 1998), which can be justified by the fact that type distributions for comparable Spanish and British groups, such as managers and professionals, are similar. Evidence (eg Hackston and Kendall, 2004; Quenk et al., 2004) does suggest that although type-related behaviours vary a good deal from country to country and from culture to culture, the frequencies of underlying MBTI types do not.

An SSR of greater than 1 indicates that a type is over-represented, and an SSR of less than 1 denotes that it is under-represented. Asterisks are used to denote whether the over- or under-representations are statistically significant, based on the results of chi-square analysis.[2]

[2] Chi-square analysis (often abbreviated to χ^2) is a technique used to explore whether observed frequency distributions differ significantly from other, predefined, distributions. In this case, the UK general population group is used as the reference group, and the chi-square analysis indicates whether the proportion of people of each type within a particular sample differs significantly from the proportion of people reporting the same type within the reference group.

OPPassessment data (representative Spanish-speaking professional and managerial sample)

Table 11.1: Type table for OPPassessment data (reported type, n=2,486)

ISTJ	ISFJ	INFJ	INTJ
n=354 14.2% SSR=1.04	n=27 1.1% SSR=0.08**	n=9 0.4% SSR=0.24**	n=122 4.9% SSR=3.50**
ISTP	**ISFP**	**INFP**	**INTP**
n=58 2.3% SSR=0.36**	n=6 0.2% SSR=0.03**	n=27 1.1% SSR=0.34**	n=77 3.1% SSR=1.29
ESTP	**ESFP**	**ENFP**	**ENTP**
n=142 5.7% SSR=0.98	n=20 0.8% SSR=0.09**	n=68 2.7% SSR=0.43**	n=193 7.8% SSR=2.79**
ESTJ	**ESFJ**	**ENFJ**	**ENTJ**
n=900 36.2% SSR=3.48**	n=50 2.0% SSR=0.16**	n=56 2.3% SSR=0.82	n=377 15.2% SSR=5.24**

Type	n	%
E	1,806	72.6%**
I	680	27.4%**
S	1,557	62.6%**
N	929	37.4%**
T	2,223	89.4%**
F	263	10.6%**
J	1,895	76.2%**
P	591	23.8%**

*Difference significant at p<0.05, based on chi-square results.
**Difference significant at p<0.01, based on chi-square results.

The most common single type preference is ESTJ (36% of the total); this is a common finding with managerial groups in other countries, although the actual percentage observed amongst this Spanish sample is higher than is usually seen. The SSR results suggest that, in comparison with the UK general population, those with preferences for NT are over-represented and those with preferences for SF are under-represented. Again, this is a common finding with managerial groups. In this group, those with preferences for NF are also under-represented. It is interesting to note that three types (ESTJ, ENTJ and ISTJ) account for 66% of the sample.

Properties of the Step II facet scales

Facet scale score distribution

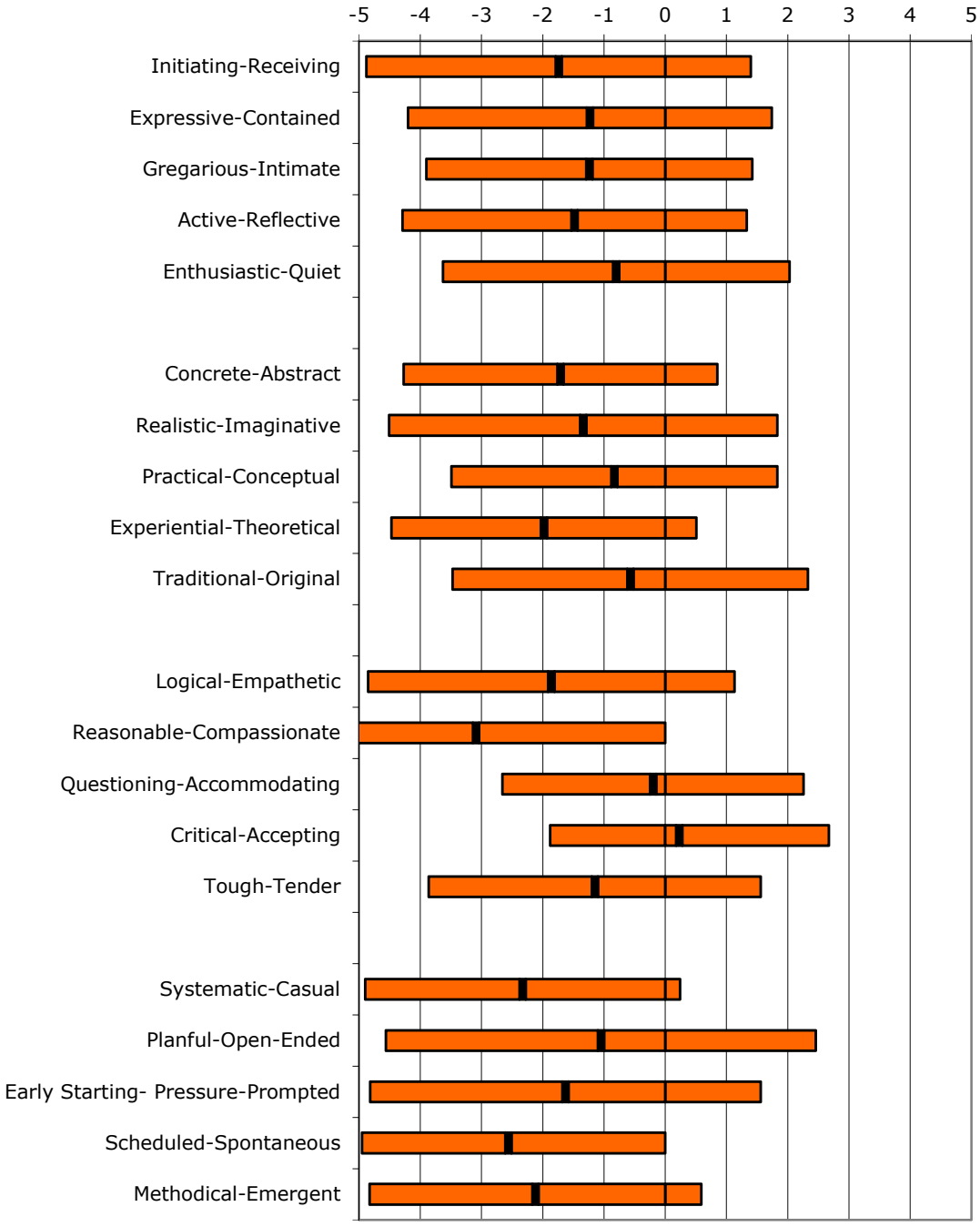
The means and standard deviations of the Step II facets are shown in Table 11.2 and illustrated graphically in Figure 11.1. The mean score for a scale is calculated by adding together the scores of each individual in the sample and then dividing the sum by the number of individuals. Note that negative values indicate that mean scores are nearer the left-hand pole of the facet and positive values indicate that the mean scores are nearer the right-hand pole. The standard deviation (SD) is a statistical measure describing the degree to which the scores from the sample either bunch up close to, or are scattered widely around, the mean for the sample.

Table 11.2: Means and standard deviations of the facet scales

Step II facet scale	Mean[3]	SD
E–I facet scales		
Initiating–Receiving	–1.74	3.14
Expressive–Contained	–1.23	2.97
Gregarious–Intimate	–1.24	2.66
Active–Reflective	–1.48	2.81
Enthusiastic–Quiet	–0.80	2.83
S–N facet scales		
Concrete–Abstract	–1.71	2.56
Realistic–Imaginative	–1.34	3.17
Practical–Conceptual	–0.83	2.66
Experiential–Theoretical	–1.98	2.49
Traditional–Original	–0.57	2.90
T–F facet scales		
Logical–Empathetic	–1.86	2.99
Reasonable–Compassionate	–3.09	2.40
Questioning–Accommodating	–0.20	2.46
Critical–Accepting	0.23	2.11
Tough–Tender	–1.15	2.71
J–P facet scales		
Systematic–Casual	–2.33	2.57
Planful–Open-Ended	–1.05	3.51
Early Starting–Pressure-Prompted	–1.63	3.19
Scheduled–Spontaneous	–2.56	2.39
Methodical–Emergent	–2.12	2.71

[3] Scale means: in Step II reports, scores are given from 5 on one pole, through 0, to 5 on the opposite pole. In compiling statistical information, however, one pole needs to be clearly distinguished from the other. Throughout this data supplement, therefore, a convention has been adopted of scores 5, 4, 3, 2 or 1 on the left-hand pole being assigned values of –5, –4, –3, –2, or –1 and scores on the right-hand pole being assigned positive scores. This does not, of course, imply any suggestion that positive scores are 'better' than negative scores.

Figure 11.1: Facet scale mean scores[4]



In general, the means centre quite close to the midpoint of each scale. Four of the scales have a mean greater than 2 points from the midpoint in either direction, and these have a standard deviation of

[4] For each facet scale, the central line shows the mean and the coloured bars on either side indicate the standard deviation.

between 2.3 and 2.8. However, a clear pattern does emerge. The mean scores tend towards the E, S, T and J direction (with the exception of Critical–Accepting), which is consistent with the most common four-letter type preference amongst this sample.

Facet scale intercorrelations

Correlations among the Step II facet scales are shown in Table 11.3. Facet scales within each dichotomy usually correlate more highly (often substantially so) with the other scales of the same dichotomy than they do with scales in the other three dichotomies. For example, every E–I facet scale correlates more highly with the other four E–I scales than with any of the S–N, T–F or J–P facet scales.

There is only one scale that correlates **significantly** more highly with scales in other dichotomies. The T–F scale Questioning–Accommodating correlates with many scales, across dichotomies. For example, it correlates at a level of between 0.11 and 0.17 with all of the E–I facet scales of the E–I dichotomy except for Gregarious–Intimate, and correlates negatively at between –0.13 and –0.32 with all of the S–N scales of the S–N dichotomy. It also correlates negatively at between -0.09 and -0.11 with all of the J–P facet scales of the J–P dichotomy except for Methodical–Emergent. By way of comparison, it correlates at between 0.05 and 0.15 with the other four scales on the same dichotomy as itself (T–F). The negative correlations between Questioning–Accommodating and the S–N scales are consistent with findings with the US version of the Step II instrument, and suggest that a questioning approach to differences of opinion seems to be related to a range of iNtuition (N) facets. These patterns are similar to those found when the European Step II instrument was developed.

Table 11.3: Intercorrelations of Step II facet scales

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Initiating–Receiving	**1.00**																			
2. Expressive–Contained	**0.52**	**1.00**																		
3. Gregarious–Intimate	**0.58**	**0.50**	**1.00**																	
4. Active–Reflective	**0.67**	**0.50**	**0.56**	**1.00**																
5. Enthusiastic–Quiet	**0.59**	**0.57**	**0.61**	**0.61**	**1.00**															
6. Concrete–Abstract	-0.01	-0.05	-0.01	-0.01	-0.10	**1.00**														
7. Realistic–Imaginative	-0.03	-0.06	-0.06	-0.03	-0.16	**0.65**	**1.00**													
8. Practical–Conceptual	-0.02	-0.01	-0.01	0.00	-0.09	**0.57**	**0.59**	**1.00**												
11. Experiential–Theoretical	0.08	0.03	0.08	0.10	0.08	**0.49**	**0.35**	**0.41**	**1.00**											
10. Traditional–Original	-0.09	-0.06	-0.08	-0.11	-0.23	**0.55**	**0.58**	**0.52**	**0.27**	**1.00**										
11. Logical–Empathetic	-0.04	-0.13	-0.04	-0.04	-0.10	0.29	0.31	0.15	0.08	0.16	**1.00**									
12. Reasonable–Compassionate	0.00	-0.10	-0.01	0.00	-0.05	0.27	0.28	0.15	0.12	0.17	**0.63**	**1.00**								
13. Questioning–Accommodating	0.12	0.11	0.04	0.11	0.17	-0.17	-0.13	-0.15	-0.13	-0.32	**0.05**	**0.05**	**1.00**							
14. Critical–Accepting	-0.15	-0.15	-0.13	-0.11	-0.10	0.13	0.10	0.05	0.00	0.01	**0.36**	**0.30**	**0.15**	**1.00**						
15. Tough–Tender	0.06	-0.07	0.02	0.06	0.04	0.28	0.24	0.15	0.13	0.08	**0.54**	**0.51**	**0.12**	**0.38**	**1.00**					
16. Systematic–Casual	-0.04	-0.10	-0.05	-0.06	-0.16	0.42	0.43	0.25	0.16	0.47	0.37	0.33	-0.11	0.14	0.27	**1.00**				
17. Planful–Open-Ended	0.06	0.03	0.04	0.04	-0.02	0.33	0.31	0.20	0.16	0.41	0.21	0.18	-0.09	0.07	0.15	**0.61**	**1.00**			
18. Early Starting–Pressure-Prompted	0.02	-0.03	-0.10	0.01	-0.09	0.28	0.28	0.17	0.16	0.35	0.15	0.13	-0.10	-0.02	0.10	**0.52**	**0.50**	**1.00**		
111. Scheduled–Spontaneous	0.02	0.00	0.03	-0.01	-0.06	0.32	0.29	0.18	0.14	0.41	0.21	0.18	-0.11	0.02	0.11	**0.61**	**0.62**	**0.53**	**1.00**	
20. Methodical–Emergent	0.08	-0.01	0.02	0.03	-0.04	0.24	0.24	0.14	0.09	0.29	0.22	0.19	-0.03	0.01	0.14	**0.58**	**0.54**	**0.52**	**0.54**	**1.00**

Correlations of Step II facet scales with Step I scales

Correlation between Step II facet scales and the continuous scores from the MBTI Step I™ instrument are shown in Table 11.4.[5] Positive correlations between facet scales and the corresponding Step I dimension (eg between Initiating–Receiving and the E–I continuous score) indicate relationships in the expected direction, with higher facet scores tending to be associated with higher continuous scores, and vice versa. Negative correlations would indicate an inverse relationship between the facet scale scores and the corresponding Step I dimension.

Table 11.4: Correlations of Step II facet scales with Step I continuous scores

	Step I continuous score			
Step II facet scales	**E–I**	**S–N**	**T–F**	**J–P**
E–I facet scales				
Initiating–Receiving	**0.82**	–0.02	0.04	0.06
Expressive–Contained	**0.72**	–0.03	–0.16	0.01
Gregarious–Intimate	**0.69**	–0.03	-0.07	0.03
Active–Reflective	**0.80**	–0.02	-0.04	0.02
Enthusiastic–Quiet	**0.80**	–0.12	–0.11	–0.06
S–N facet scales				
Concrete–Abstract	–0.03	**0.83**	0.34	0.38
Realistic–Imaginative	–0.07	**0.78**	0.34	0.36
Practical–Conceptual	–0.03	**0.68**	0.17	0.23
Experiential–Theoretical	0.09	**0.55**	0.12	0.19
Traditional–Original	–0.13	**0.70**	0.17	0.46
T–F facet scales				
Logical–Empathetic	–0.09	0.21	**0.87**	0.24
Reasonable–Compassionate	-0.03	0.21	**0.73**	0.22
Questioning–Accommodating	0.16	–0.25	**0.11**	–0.12
Critical–Accepting	–0.15	0.06	**0.48**	0.04
Tough–Tender	0.04	0.18	**0.70**	0.16
J–P facet scales				
Systematic–Casual	–0.08	0.42	0.41	**0.75**
Planful–Open-Ended	0.06	0.33	0.24	**0.84**
Early Starting–Pressure-Prompted	–0.02	0.29	0.18	**0.72**
Scheduled–Spontaneous	0.00	0.33	0.23	**0.78**
Methodical–Emergent	0.05	0.23	0.25	**0.71**

The E–I facet scales correlate at 0.69 to 0.82 with the E–I continuous scores from Step I; the S–N facet scales correlate at 0.68 to 0.83 with the S–N continuous scores; the T–F facet scales correlate at 0.11 to

[5] Continuous scores place an individual's score on each dimension onto a continuous scale with a midpoint of 100. To calculate continuous scores, Preference Clarity Index (PCI) scores for each dimension are either subtracted or added to 100, depending on which direction the overall preference is. PCI scores in the direction of E, S, T or J are subtracted from 100. PCI scores in the direction of I, N, F or P are added to 100.

0.87 with the T–F continuous scores; and the J–P facet scales correlate at 0.71 to 0.84 with the J–P continuous scores. The scale that showed the lowest correlation with its associated Step I continuous score was Questioning–Accommodating at 0.11. This scale is also lowest on internal consistency (see page 251), and has been previously found to have the lowest test-retest reliability (Quenk, Hammer and Majors, 2004), which would result in the true correlations being underestimated.

These correlations are very similar to those found during the development of the Step II questionnaire. This consistency, alongside the fact that the correlations between Step II facet scales and Step I continuous scores associated with their own dimension are substantially higher than correlations with the other three dimensions, provides compelling evidence for the theoretical hierarchical structure of the Step II facet scales in relation to the Step I scales.

Out-of-preference scores

It is known that the five facet scales relating to each type dimension do not describe the dimension in its entirety; there will not be a precise and exact overlap between, for example, an individual's score on E–I and their total score across the five facet scales that relate to this dimension. For example, it is not uncommon to see an Enthusiastic Introvert or an Intimate Extravert. Such apparent inconsistencies are usually known as 'out-of-preference scores' (OOPS).

Although it is usual to have a number of OOPS in any one profile, it should be unusual to find that OOPS make up more than half of the facet scores relating to any one type dichotomy. Table 11.5 demonstrates that the proportion of individuals for whom this occurs ranges from 4.1% for the T–F block to 0.7% for the E–I block. It is therefore a very infrequent occurrence.

Table 11.5: Proportion of OOPS by type dichotomy[6]

	Proportion of 'reported type' OOPS					
	None	One	Two	Three	Four	Five
E–I	73.7%	19.3%	6.3%	0.7%	0.0%	0.0%
S–N	62.2%	25.6%	8.7%	2.6%	0.8%	0.0%
T–F	49.4%	33.6%	12.9%	3.5%	0.6%	0.0%
J–P	67.1%	23.7%	7.5%	1.6%	0.2%	0.0%

For any individual facet scale, an OOPS tends to occur in approximately 10% of profiles, with the exception of Experiential–

[6] In this table, an OOPS is defined as a score of 2, 3, 4 or 5 on the 'wrong' pole of the facet when compared with the reported type.

Theoretical (19%), Questioning–Accommodating (26%) and Critical–Accepting (22%).

Reliability

The reliability of a test or questionnaire relates to how consistent and precise it is. Internal consistency reliability addresses the question of whether all the questions in a scale measure the same construct. For example, are the Step II facet scales consistent within themselves, and do they hold together well as scales? A common measure of internal consistency reliability is coefficient alpha (Cronbach, 1951). The alpha coefficients for the Step II facet scales are shown in Table 11.6.

Table 11.6: Internal consistency reliability

Step II facet scale	No. of items	Coefficient alpha
E–I facet scales		
Initiating–Receiving	8	0.82
Expressive–Contained	7	0.76
Gregarious–Intimate	7	0.68
Active–Reflective	8	0.69
Enthusiastic–Quiet	9	0.68
S–N facet scales		
Concrete–Abstract	9	0.71
Realistic–Imaginative	7	0.75
Practical–Conceptual	8	0.51
Experiential–Theoretical	8	0.70
Traditional–Original	8	0.71
T–F facet scales		
Logical–Empathetic	9	0.84
Reasonable–Compassionate	8	0.70
Questioning–Accommodating	7	0.32
Critical–Accepting	8	0.41
Tough–Tender	8	0.72
J–P facet scales		
Systematic–Casual	8	0.75
Planful–Open-Ended	6	0.81
Early Starting–Pressure-Prompted	6	0.73
Scheduled–Spontaneous	8	0.62
Methodical–Emergent	8	0.68
	Median	0.70

The internal consistency reliability of most scales is good, and the average (median) reliability is 0.7. However, three scales (Practical–Conceptual, Questioning–Accommodating and Critical–Accepting) do have lower alpha coefficients. This pattern is consistent with what was found during the development of the instrument, where these three

facet scales were also found to show lower reliability than the others (Quenk, Hammer and Majors, 2004).

It should be noted, however, that coefficient alpha reliability statistics will give an underestimate of the reliability of the Step II facet scales, given the use of item response theory (IRT) methods in the actual scoring process.[7]

Group differences

This section looks at the way in which people from different groups responded to the Spanish version of the European Step II questionnaire. Unfortunately, the total number of respondents is too small for the full range of analyses to be conducted. Further research will be conducted when additional data become available.

Gender

The means and standard deviations of the Step II facet scales are shown separately for males and females in Table 11.7, along with the difference in mean scores. This is represented graphically in Figure 11.2. Statistically significant differences were found between male and female mean scores for sixteen of the 20 facet scales, with one consistent pattern emerging.

- On the E–I facet scales, all five mean scores tended slightly toward the E pole for both males and females. There were statistically significant gender differences on three of the five facet scales, with females tending further towards the E pole than males. There was no statistically significant difference between genders on the remaining two scales.

- On the S–N facet scales, all five mean scores tended slightly toward the S pole for both males and females. There were statistically significant gender differences on three of the five facet scales with male mean scores tending more towards the S pole than females on two of these scales, and female mean score tending more towards the S pole on the other. There was no statistically significant difference between genders on the remaining two scales.

- On the T–F facet scales, all five mean scores tended slightly toward the T pole for males, whereas four did so for females. There were statistically significant gender differences on all five facet scales, with male mean scores tending more towards the T pole than females.

[7] Item response theory (IRT) is an approach to measurement that is concerned with modelling the relationship between item responses and the underlying characteristic assessed by the scale or test the item is designed to measure. IRT can be used to select items for a test and/or to score the items.

- On the J–P facet scales, all five mean scores tended slightly toward the J pole for both males and females. There were statistically significant gender differences on all five facet scales, with male mean scores tending more towards the J pole than females.

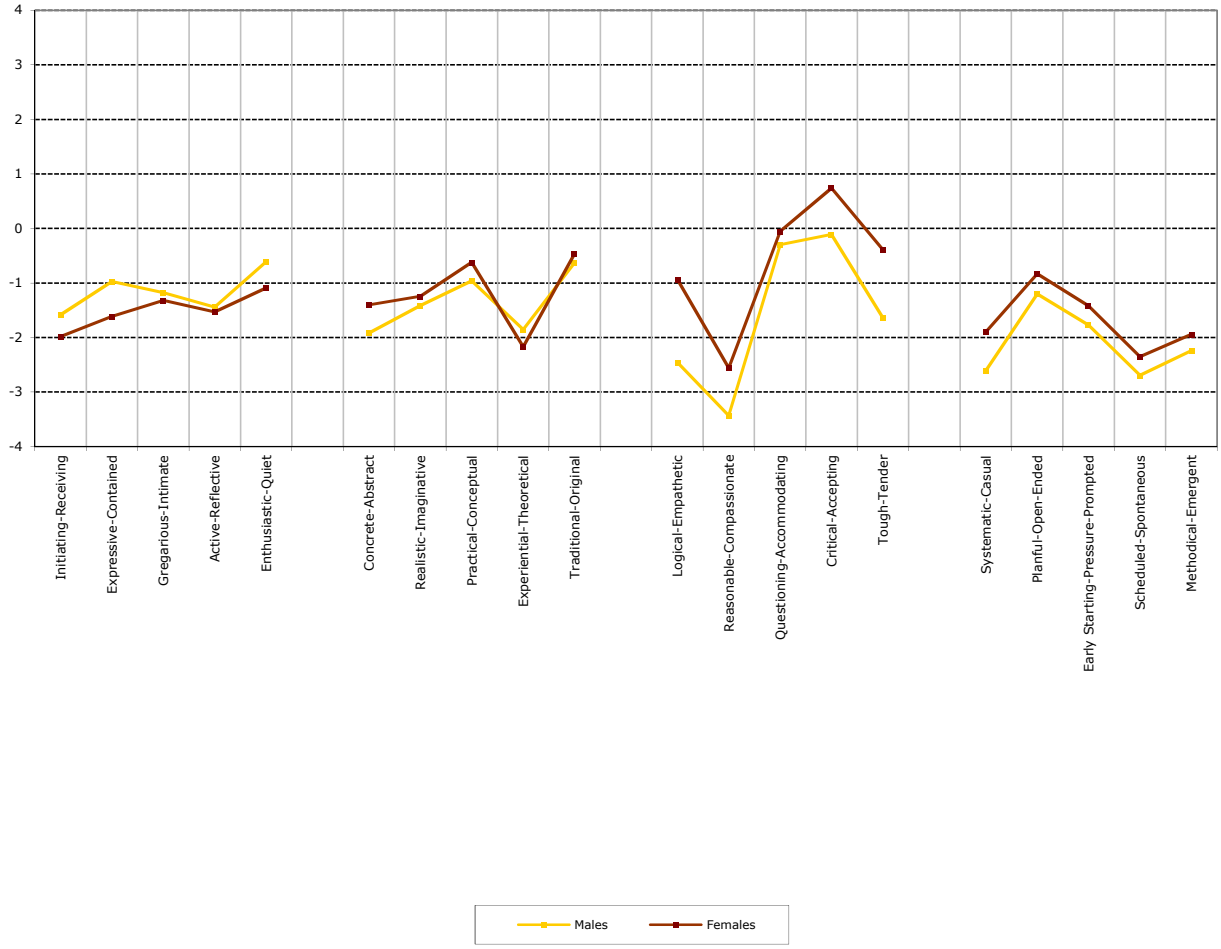
Table 11.7: Gender differences in facet scale scores

	Males (n=1,503)		Females (n=983)		Difference (M–F)[8]
Step II facet scale	**Mean**	**SD**	**Mean**	**SD**	
E–I facet scales					
Initiating–Receiving	-1.58	3.19	-1.98	3.04	0.40**
Expressive–Contained	-0.97	3.02	-1.61	2.85	0.64**
Gregarious–Intimate	-1.18	2.78	-1.32	2.45	0.14
Active–Reflective	-1.44	2.86	-1.53	2.72	0.09
Enthusiastic–Quiet	-0.61	2.87	-1.09	2.75	0.49**
S–N facet scales					
Concrete–Abstract	-1.92	2.51	-1.40	2.66	-0.52**
Realistic–Imaginative	-1.41	3.12	-1.24	3.23	-0.17
Practical–Conceptual	-0.96	2.64	-0.62	2.68	-0.35**
Experiential–Theoretical	-1.85	2.54	-2.17	2.40	0.32**
Traditional–Original	-0.63	2.87	-0.47	2.95	-0.15
T–F facet scales					
Logical–Empathetic	-2.46	2.73	-0.94	3.13	-1.52**
Reasonable–Compassionate	-3.43	2.20	-2.56	2.59	-0.87**
Questioning–Accommodating	-0.30	2.45	-0.05	2.47	-0.25*
Critical–Accepting	-0.11	2.12	0.74	2.01	-0.85**
Tough–Tender	-1.64	2.55	-0.39	2.77	-1.25**
J–P facet scales					
Systematic–Casual	-2.61	2.52	-1.90	2.58	-0.71**
Planful–Open-Ended	-1.20	3.44	-0.83	3.60	-0.37*
Early Starting–Pressure-Prompted	-1.77	3.16	-1.42	3.22	-0.35**
Scheduled–Spontaneous	-2.69	2.32	-2.35	2.48	-0.33**
Methodical–Emergent	-2.24	2.68	-1.94	2.74	-0.31**

Difference significant at:*p<0.05, **p<0.01 (based on an independent samples t-test).

[8] A positive value indicates that male scores tend more towards I, N, F or P, and a negative value indicates that female scores tend more towards I, N, F or P.

Figure 11.2: Gender differences in facet scale scores



Age

There were found to be no meaningful links between age and facet scale scores. Although correlational analysis showed several facet scales to be significantly correlated with age, the significance levels were more the result of the relatively large sample size rather than being indicative of a meaningful relationship. The highest correlation was 0.16 (Enthusiastic–Quiet).

For interpretation purposes, it is reasonable to conclude that there are no clear relationships between age and facet scale scores.

Ethnic origin

Ethnic origin information was not captured for people who completed the Spanish language version of the Step II questionnaire, so no analyses were conducted.

Occupational level

Table 11.8 shows the facet scale means and standard deviation for different occupational levels. The findings can be summarised as follows:

- A pattern was found regarding the facets relating to the S–N dimension, with significant differences found on three of the facet scales. Mean facet scores amongst people in the 'Top level' group tended less towards the S pole than those for the lower occupational level groups on these three facet scales. Mean facet scores amongst people in the 'Middle management' group also tended further towards the S pole than all the other groups on two of these facet scales.

- A pattern was also found regarding the facets relating to the T–F dimension, with significant differences found on three of the facet scales. Mean facet scores amongst people in the 'Middle management' group tended further towards the T pole than the other occupational level groups on these three facet scales.

- There were no significant differences by occupational level for any of the E–I or J–P facet scales.

Table 11.8: Mean facet scale scores by occupational level

Step II facet scale	Top level (n=78)		Senior Executive (n=266)		Upper Middle Management (n=269)		Middle Management (n=271)		First Level Mgt/Supervisory (n=160)		Employee (n=369)		Sig.
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales													
Initiating–Receiving	-1.88	3.13	-2.00	3.00	-1.70	3.04	-1.56	3.16	-1.38	3.11	-1.81	3.25	
Expressive–Contained	-1.33	2.90	-1.24	3.07	-1.38	2.91	-1.18	3.06	-1.08	2.77	-1.37	2.95	
Gregarious–Intimate	-0.91	2.83	-1.45	2.65	-1.35	2.57	-0.97	2.69	-1.06	2.58	-1.34	2.69	
Active–Reflective	-1.24	2.62	-1.74	2.80	-1.68	2.73	-1.45	3.06	-1.02	2.82	-1.35	2.65	
Enthusiastic–Quiet	-0.69	2.70	-1.06	2.81	-0.86	2.90	-0.66	2.99	-0.57	2.67	-0.70	2.90	
S–N facet scales													
Concrete–Abstract	-1.37	2.67	-1.48	2.49	-1.82	2.64	-1.99	2.57	-1.99	2.62	-1.78	2.50	
Realistic–Imaginative	-0.59	3.20	-1.02	3.24	-1.30	3.09	-1.82	3.13	-1.66	3.07	-1.38	3.11	**
Practical–Conceptual	-0.23	2.40	-0.36	2.54	-0.72	2.75	-1.21	2.59	-0.93	2.71	-1.11	2.70	**
Experiential–Theoretical	-1.36	2.58	-1.54	2.74	-2.17	2.52	-2.17	2.44	-2.24	2.42	-2.04	2.40	**
Traditional–Original	-0.33	2.77	-0.54	2.85	-0.52	2.94	-0.79	2.96	-0.55	2.89	-0.84	2.92	
T–F facet scales													
Logical–Empathetic	-1.40	3.09	-2.41	2.85	-1.96	2.91	-2.44	2.68	-2.00	2.90	-1.57	3.12	**
Reasonable–Compassionate	-2.33	2.80	-3.47	2.27	-3.24	2.23	-3.54	2.15	-3.31	2.12	-3.16	2.35	**
Questioning–Accommodating	0.19	2.38	-0.37	2.48	-0.22	2.46	-0.29	2.41	0.03	2.35	-0.07	2.54	
Critical–Accepting	0.44	1.98	0.12	2.10	0.24	2.08	0.00	2.16	0.18	2.11	0.29	2.02	
Tough–Tender	-0.85	3.00	-1.52	2.60	-1.23	2.73	-1.61	2.47	-1.26	2.50	-0.86	2.82	**
J–P facet scales													
Systematic–Casual	-2.28	2.57	-2.31	2.63	-2.30	2.52	-2.85	2.28	-2.41	2.39	-2.30	2.58	
Planful–Open-Ended	-1.29	3.62	-1.30	3.49	-1.10	3.43	-1.28	3.39	-0.84	3.53	-1.33	3.62	
Early Starting–Pressure-Prompted	-1.27	3.40	-1.22	3.28	-1.49	3.18	-1.73	3.19	-1.69	3.39	-1.89	3.11	
Scheduled–Spontaneous	-2.29	2.61	-2.65	2.29	-2.56	2.20	-2.83	2.08	-2.57	2.44	-2.65	2.42	
Methodical–Emergent	-1.95	2.92	-1.94	2.76	-1.99	2.77	-2.45	2.59	-2.30	2.67	-2.38	2.55	

Education

Specific educational qualifications were not collected for the OPPassessment sample; however, the age at which individuals left full-time education was. No significant and meaningful correlations were found between the age at which people left full-time education and their facet scale scores, with all correlations being less than 0.1.

Work area

Information regarding the area of work people engage in was collected for the group. Many different categories were used, but for the purposes of analysis the focus was on the five most commonly occurring. These were as follows:

- Sales, customer service
- Science, engineering
- HR, training, guidance
- Finance
- Business services

Table 11.9 shows the facet scale means and standard deviations for these five work areas. The findings can be summarised as follows:

- The most noticeable pattern concerned the 'HR, training, guidance' group; mean scores for this group tended less towards the S pole for three of the five facet scales, less towards the T pole on four of the facet scales and less towards the J pole on all five of the facet scales, than all the other four groups.

Table 11.9: Mean facet scale scores by work area

Step II facet scale	Sales, customer service (n=314)		Science, engineering (n=223)		HR, training, guidance (n=212)		Finance (n=165)		Business services (n=124)		Sig.
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales											
Initiating–Receiving	-2.66	2.73	-1.13	3.33	-2.11	3.04	-1.40	3.18	-1.52	3.17	**
Expressive–Contained	-1.71	2.77	-1.01	3.15	-1.75	2.95	-1.16	2.86	-1.67	2.89	*
Gregarious–Intimate	-1.90	2.49	-1.02	2.83	-1.30	2.52	-1.18	2.66	-1.26	2.76	**
Active–Reflective	-2.08	2.63	-1.08	3.04	-1.59	2.65	-1.42	2.74	-1.84	2.58	**
Enthusiastic–Quiet	-1.44	2.64	-0.51	3.01	-1.18	2.83	-0.45	2.82	-1.02	2.82	**
S–N facet scales											
Concrete–Abstract	-2.06	2.40	-1.79	2.57	-0.94	2.83	-2.16	2.45	-1.80	2.56	**
Realistic–Imaginative	-1.70	3.01	-1.12	3.24	-0.67	3.44	-1.85	3.20	-1.43	3.27	**
Practical–Conceptual	-1.19	2.46	-0.79	2.66	-0.33	2.85	-1.01	2.68	-0.93	2.70	**
Experiential–Theoretical	-2.52	2.10	-1.52	2.71	-1.64	2.75	-2.42	2.52	-1.67	2.57	**
Traditional–Original	-0.84	2.71	-0.68	3.00	-0.26	3.21	-0.82	2.84	-0.58	3.02	
T–F facet scales											
Logical–Empathetic	-2.38	2.86	-2.42	2.74	-1.03	3.33	-1.93	2.98	-2.07	2.88	**
Reasonable–Compassionate	-3.75	1.98	-3.80	1.81	-2.40	2.75	-3.32	2.22	-3.12	2.33	**
Questioning–Accommodating	-0.39	2.42	-0.13	2.40	-0.14	2.65	-0.45	2.57	-0.29	2.38	
Critical–Accepting	0.13	1.93	-0.15	2.20	0.81	2.06	0.24	2.17	-0.05	2.25	**
Tough–Tender	-1.80	2.43	-1.35	2.44	-0.37	2.76	-1.28	2.86	-1.20	3.01	**
J–P facet scales											
Systematic–Casual	-2.82	2.39	-2.78	2.27	-1.50	2.76	-2.41	2.46	-2.57	2.32	**
Planful–Open-Ended	-1.88	3.31	-1.20	3.44	-0.29	3.79	-1.00	3.43	-1.27	3.45	**
Early Starting–Pressure-Prompted	-2.19	3.02	-2.05	3.07	-0.34	3.65	-1.39	3.27	-1.85	3.22	**
Scheduled–Spontaneous	-2.98	2.04	-3.04	2.04	-1.78	2.73	-2.48	2.44	-2.81	2.20	**
Methodical–Emergent	-2.75	2.55	-2.47	2.49	-1.16	2.94	-1.82	2.82	-2.33	2.68	**

Nationality

Information on nationality was gathered for 90% of the group. Of these, 79% were Spanish. There were insufficient people of other individual nationalities for it to be possible to conduct any analyses across groups.

Employment status

Employment status information was gathered for 72% of the group. Of these, 91% worked full-time. There were insufficient people who did not work full-time for it to be possible to conduct any analyses across groups.

Appendix 1: Sample description

Sample 1: Data from OPPassessment (representative Spanish-speaking professional and managerial sample)

This sample consists of 2486 individuals who completed the MBTI Step II instrument in Spanish via the OPPassessment system between August 2006 and July 2016. 60.5% per cent of the respondents were male and 39.5% were female. Age ranged from 17 to 65 years, with a mean of 38.5 and a median of 39.

Nationality was disclosed by 90% of respondents. Of these, 79% were Spanish. Several Central and South American nationalities were represented, but not in large numbers:

Nationality	Percentage
Spanish	78.5%
Other – Central and South America	32.5%
Other	4.7%

Seventy-two per cent of respondents stated their employment status. Of these, the majority of the group were in full-time employment:

Employment status	Percentage
Full-time	90.5%
Self-employed	5.0%
Part-time	1.5%
Unemployed	2.9%
Homemaker	0.1%
Retired	0.1%

Sixty-six per cent of respondents stated their occupational level. The majority of the group were of managerial level or above, though the largest single group was employee (23%):

Occupational level	Percentage
Top level	4.8%
Senior executive	16.3%
Upper middle management	16.4%
Middle management	16.6%
First level management/supervisor	9.8%
Employee	22.6%
Other	13.6%

Sixty-nine per cent of respondents stated their work area. A range of work areas were represented:

Work area (job type)	Percentage
Sales, customer service	18.4%
Science, engineering	13.1%
HR, training, guidance	12.4%
Finance	9.7%
Business services	7.3%
Health, social services etc	7.3%
Research and development	6.2%
IT	5.0%
Admin or secretarial	2.6%
Education	1.8%
Skilled operative	1.4%
Land, sea or air transport	0.4%
Leisure, personal service	0.4%
Unskilled operative	0.2%
Military, police, prison, fire	0.1%
Other private sector	6.2%
Other public sector	2.0%
Other	10.0%

MBTI® Step II™ instrument

European Data Supplement

Swedish

January 2018

OPP Ltd
Elsfield Hall
15–17 Elsfield Way
Oxford OX2 8EP
UK
t: +44 (0)1865 404 500

Contents

Introduction

Data collected for the European MBTI® Step II™ instrument were analysed to produce the findings in this supplement. This is the second data supplement produced for this instrument. A brief description of the sample is given below. Further details of the sample are provided in Appendix 1.

- The sample consisted of 12,343 individuals who completed the MBTI Step II instrument in Swedish via the OPPassessment™ system between 2006 and mid-2016.[1] This sample is considered to be representative of the groups of people with whom the Swedish MBTI Step II instrument has been and will be used for applications such as management development, coaching, counselling and teambuilding. As such, it is likely to represent a cross-section of the Swedish-speaking professional and managerial population.

The results of the analyses are outlined below.

[1] OPPassessment allows personality questionnaires such as the MBTI instrument to be administered via email and/or completed online.

Type distribution

Type tables are a way of illustrating the proportion of each type within a particular group. Below is a type table taken from the sample described on the previous page.

For each of the 16 different types, the number of cases, the percentage of the total that this represents and the self-selection ratio (SSR) are shown. The SSR (Myers et al., 1998) is a way of demonstrating whether a given type appears more or less often in a particular group than would be expected given its frequency in a reference group. Ideally, the type distribution from a large representative sample of the Swedish population would be used to calculate SSRs in this data supplement. However, such a sample does not currently exist. In its place, SSRs have been calculated using type data from the UK general population (Kendall, 1998), which can be justified by the fact that type distributions for comparable Swedish and British groups, such as managers and professionals, are similar. Evidence (eg Hackston and Kendall, 2004; Quenk et al., 2004) does suggest that although type-related behaviours vary a good deal from country to country and from culture to culture, the frequencies of underlying MBTI types do not.

An SSR of greater than 1 indicates that a type is over-represented, and an SSR of less than 1 denotes that it is under-represented. Asterisks are used to denote whether the over- or under-representations are statistically significant, based on the results of chi-square.[2]

[2] Chi-square analysis (often abbreviated to χ^2) is a technique used to explore whether observed frequency distributions differ significantly from other, predefined, distributions. In this case, the UK general population group is used as the reference group, and the chi-square analysis indicates whether the proportion of people of each type within a particular sample differs significantly from the proportion of people reporting the same type within the reference group.

OPPassessment data (representative Swedish-speaking professional and managerial sample)

Table 12.1: Type table for OPPassessment data (reported type, n=12,343)

ISTJ	ISFJ	INFJ	INTJ
n=1092 8.8% SSR=0.64**	n=399 3.2% SSR=0.25**	n=141 1.1% SSR=0.65	n=331 2.7% SSR=1.93**
ISTP	**ISFP**	**INFP**	**INTP**
n=400 3.2% SSR=0.50**	n=169 1.4% SSR=0.23**	n=247 2.0% SSR=0.63**	n=372 3.0% SSR=1.25
ESTP	**ESFP**	**ENFP**	**ENTP**
n=786 6.4% SSR=1.10	n=485 3.9% SSR=0.45**	n=971 7.9% SSR=1.25*	n=1089 8.8% SSR=3.14**
ESTJ	**ESFJ**	**ENFJ**	**ENTJ**
n=2694 21.8% SSR=2.10**	n=1256 10.2% SSR=0.81**	n=647 5.2% SSR=1.86**	n=1264 10.2% SSR=3.52**

Type	n	%
E	9,192	74.5%**
I	3,151	25.5%**
S	7,281	59.0%**
N	5,062	41.0%**
T	8,028	65.0%**
F	4,315	35.0%**
J	7,824	63.4%**
P	4,519	36.6%**

*Difference significant at p<0.05, based on chi-square results.
**Difference significant at p<0.01, based on chi-square results.

The most common single type preference is ESTJ (22% of the total); this is a common finding with managerial groups in other countries. The SSR results suggest that, in comparison with the reference population, those with preferences for NT are over-represented, and those with preferences for SF are under-represented. Again, this is a common finding with managerial groups.

Properties of the Step II facet scales

Facet scale score distribution

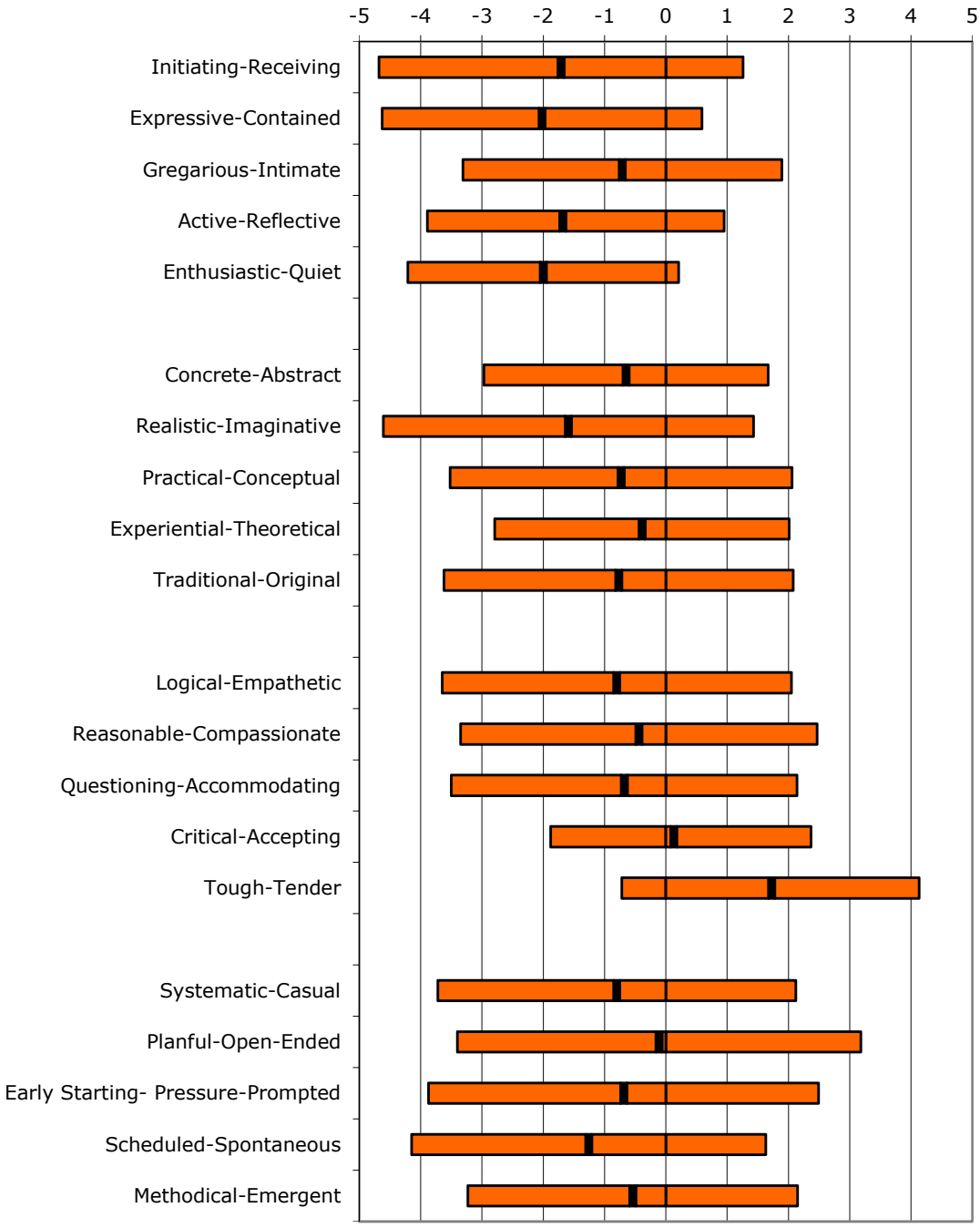
The means and standard deviations of the Step II facets are shown in Table 12.2 and illustrated graphically in Figure 12.1. The mean score for a scale is calculated by adding together the scores of each individual in the sample and then dividing the sum by the number of individuals. Note that negative values indicate that mean scores are nearer the left-hand pole of the facet and positive values indicate that the mean scores are nearer the right-hand pole. The standard deviation (SD) is a statistical measure describing the degree to which the scores from the sample either bunch up close to, or are scattered widely around, the mean for the sample.

Table 12.2: Means and standard deviations of the facet scales

Step II facet scale	Mean[3]	SD
E–I facet scales		
Initiating–Receiving	–1.71	2.97
Expressive–Contained	–2.02	2.61
Gregarious–Intimate	–0.71	2.60
Active–Reflective	–1.68	2.63
Enthusiastic–Quiet	–2.00	2.21
S–N facet scales		
Concrete–Abstract	–0.65	2.32
Realistic–Imaginative	–1.59	3.02
Practical–Conceptual	–0.73	2.80
Experiential–Theoretical	-0.39	2.40
Traditional–Original	–0.77	2.85
T–F facet scales		
Logical–Empathetic	–0.80	2.85
Reasonable–Compassionate	–0.44	2.91
Questioning–Accommodating	–0.68	2.82
Critical–Accepting	0.13	2.01
Tough–Tender	1.73	2.40
J–P facet scales		
Systematic–Casual	–0.80	2.92
Planful–Open-Ended	-0.11	3.29
Early Starting–Pressure-Prompted	–0.69	3.17
Scheduled–Spontaneous	–1.26	2.89
Methodical–Emergent	–0.54	2.69

[3] Scale means: in Step II reports, scores are given from 5 on one pole, through 0, to 5 on the opposite pole. In compiling statistical information, however, one pole needs to be clearly distinguished from the other. Throughout this data supplement, therefore, a convention has been adopted of scores 5, 4, 3, 2 or 1 on the left-hand pole being assigned values of –5, –4, –3, –2, or –1 and scores on the right-hand pole being assigned positive scores. This does not, of course, imply any suggestion that positive scores are 'better' than negative scores.

Figure 12.1: Facet scale mean scores[4]



[4] For each facet scale, the central line shows the mean and the coloured bars on either side indicate the standard deviation.

In general, the means centre quite close to the midpoint of each scale. Only two of the scales have a mean greater than 2 points from the midpoint point in either direction, and each has a standard deviation of between 2.2 and 2.7. However, a clear pattern does emerge. The mean scores generally tend towards the E, S, T and J direction (with the exception of Tough–Tender and Critical–Accepting), which is consistent with the most common four-letter type preference amongst this sample.

Facet scale intercorrelations

Correlations among the Step II facet scales are shown in Table 12.3. Facet scales within each dichotomy usually correlate more highly (often substantially so) with the other scales of the same dichotomy than they do with scales in the other three dichotomies. For example, every E–I facet scale correlates higher with the other four E–I scales than with any of the S–N, T–F or J–P facet scales.

There is only one scale that correlates **significantly** more highly with scales in other dichotomies. The T–F scale Questioning–Accommodating correlates with many scales, across dichotomies. For example, it correlates at a level of 0.21 with Enthusiastic–Quiet on the E–I dichotomy. It also correlates negatively at between –0.20 and –0.35 with all of the S–N scales on the S–N dichotomy except for Realistic–Imaginative, and at -0.19 with Early Starting–Pressure-Prompted on the J–P dichotomy. By way of comparison, it correlates at between 0.16 and 0.25 with the other four scales on the same dichotomy as itself (T–F). The negative correlations between Questioning–Accommodating and the S–N scales are consistent with findings with the US version of the Step II instrument, and suggest that a questioning approach to differences of opinion seems to be related to a range of Intuitive (N) facets. These patterns are similar to those found when the European Step II instrument was developed.

Table 12.3: Intercorrelations of Step II facet scales

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Initiating–Receiving	**1.00**																			
2. Expressive–Contained	**0.55**	**1.00**																		
3. Gregarious–Intimate	**0.48**	**0.43**	**1.00**																	
4. Active–Reflective	**0.64**	**0.52**	**0.48**	**1.00**																
5. Enthusiastic–Quiet	**0.54**	**0.55**	**0.50**	**0.58**	**1.00**															
6. Concrete–Abstract	-0.09	-0.07	-0.03	-0.06	-0.17	**1.00**														
7. Realistic–Imaginative	-0.08	-0.10	-0.08	-0.08	-0.20	**0.61**	**1.00**													
8. Practical–Conceptual	-0.08	-0.02	-0.02	-0.04	-0.13	**0.57**	**0.55**	**1.00**												
9. Experiential–Theoretical	-0.04	-0.02	0.00	-0.02	-0.08	**0.53**	**0.43**	**0.46**	**1.00**											
10. Traditional–Original	-0.11	-0.06	-0.03	-0.08	-0.18	**0.56**	**0.55**	**0.57**	**0.42**	**1.00**										
11. Logical–Empathetic	-0.07	-0.21	-0.03	-0.07	-0.09	0.20	0.26	-0.01	0.02	0.05	**1.00**									
12. Reasonable–Compassionate	-0.02	-0.12	0.01	0.00	-0.03	-0.18	0.22	0.00	0.03	0.03	**0.67**	**1.00**								
13. Questioning–Accommodating	0.14	0.09	0.07	0.16	0.21	-0.20	-0.15	-0.30	-0.21	-0.35	**0.18**	**0.22**	**1.00**							
14. Critical–Accepting	-0.16	-0.22	-0.14	-0.13	-0.14	0.12	0.16	-0.03	-0.02	0.00	**0.48**	**0.49**	**0.25**	**1.00**						
15. Tough–Tender	-0.02	-0.13	-0.01	0.00	-0.07	0.16	0.16	0.05	0.02	0.02	**0.42**	**0.51**	**0.16**	**0.39**	**1.00**					
16. Systematic–Casual	-0.06	-0.12	-0.07	-0.08	-0.17	0.42	0.47	0.28	0.22	0.38	0.39	0.34	-0.02	0.26	0.27	**1.00**				
17. Planful–Open-Ended	0.05	0.05	0.05	0.02	-0.01	0.27	0.27	021	0.19	0.33	0.13	0.09	-0.09	0.04	0.07	**0.52**	**1.00**			
18. Early Starting–Pressure-Prompted	-0.07	-0.05	-0.09	-0.10	-0.17	0.25	0.26	0.22	0.23	0.33	0.01	-0.03	-0.19	-0.03	-0.05	**0.37**	**0.40**	**1.00**		
19. Scheduled–Spontaneous	0.00	-0.02	-0.01	-0.03	-0.09	0.34	0.35	0.27	0.25	0.40	0.16	0.10	-0.12	0.05	0.04	**0.58**	**0.65**	**0.51**	**1.00**	
20. Methodical–Emergent	0.06	0.02	0.00	0.02	-0.02	0.20	0.21	0.13	0.15	0.19	0.12	0.10	-0.03	0.05	0.05	**0.45**	**0.44**	**0.41**	**0.52**	**1.00**

Correlations of Step II facet scales with Step I scales

Correlations between Step II facet scales and the continuous scores from the MBTI Step I™ instrument are shown in Table 12.4.[5] Positive correlations between facet scales and the corresponding Step I dimension (e.g. between Initiating–Receiving and the E–I continuous score) indicate relationships in the expected direction, with higher facet scores tending to be associated with higher continuous scores, and vice versa. Negative correlations would indicate an inverse relationship between the facet scale scores and the corresponding Step I dimension.

Table 12.4: Correlations of Step II facet scales with Step I continuous scores

	Step I continuous score			
Step II facet scales	**E–I**	**S–N**	**T–F**	**J–P**
E–I facet scales				
Initiating–Receiving	**0.82**	–0.10	–0.10	0.03
Expressive–Contained	**0.74**	–0.07	–0.23	0.00
Gregarious–Intimate	**0.60**	-0.03	–0.05	0.00
Active–Reflective	**0.77**	–0.07	–0.08	–0.01
Enthusiastic–Quiet	**0.77**	–0.18	–0.12	–0.09
S–N facet scales				
Concrete–Abstract	–0.12	**0.82**	0.22	0.38
Realistic–Imaginative	–0.15	**0.75**	0.25	0.39
Practical–Conceptual	–0.09	**0.68**	–0.01	0.30
Experiential–Theoretical	–0.06	**0.64**	0.01	0.27
Traditional–Original	–0.15	**0.74**	0.03	0.42
T–F facet scales				
Logical–Empathetic	–0.13	0.15	**0.86**	0.17
Reasonable–Compassionate	–0.05	0.12	**0.77**	0.13
Questioning–Accommodating	0.19	–0.29	**0.28**	–0.12
Critical–Accepting	–0.19	0.07	**0.62**	0.07
Tough–Tender	–0.04	0.10	**0.56**	0.08
J–P facet scales				
Systematic–Casual	–0.12	0.45	0.42	**0.70**
Planful–Open-Ended	0.05	0.32	0.12	**0.81**
Early Starting–Pressure-Prompted	–0.12	0.31	-0.01	**0.67**
Scheduled–Spontaneous	–0.04	0.39	0.14	**0.80**
Methodical–Emergent	0.04	0.22	0.11	**0.65**

The E–I facet scales correlate at 0.60 to 0.82 with the E–I continuous scores from Step I; the S–N facet scales correlate at 0.74 to 0.82 with the S–N continuous scores; the T–F facet scales correlate at 0.28 to

[5] Continuous scores place an individual's score on each dimension onto a continuous scale with a midpoint of 100. To calculate continuous scores, Preference Clarity Index (PCI) scores for each dimension are either subtracted or added to 100, depending on which direction the overall preference is. PCI scores in the direction of E, S, T or J are subtracted from 100. PCI scores in the direction of I, N, F or P are added to 100.

0.86 with the T–F continuous scores; and the J–P facet scales correlate at 0.65 to 0.81 with the J–P continuous scores. The scale that showed the lowest correlation with its associated Step I continuous score was Questioning–Accommodating at 0.28. This scale is also lowest on internal consistency (see page 275), and has been previously found to have the lowest test-retest reliability (Quenk, Hammer and Majors, 2004), which would result in the true correlations being underestimated.

These correlations are very similar to those found during the development of the Step II questionnaire. This consistency, alongside the fact that the correlations between Step II facet scales and Step I continuous scores associated with their own dimension are substantially higher than correlations with the other three dimensions, provides compelling evidence for the theoretical hierarchical structure of the Step II facet scales in relation to the Step I scales.

Out-of-preference scores

It is known that the five facet scales relating to each type dimension do not describe the dimension in its entirety; there will not be a precise and exact overlap between, for example, an individual's score on E–I and their total score across the five facet scales that relate to this dimension. For example, it is not uncommon to see an Enthusiastic Introvert or an Intimate Extravert. Such apparent inconsistencies are usually known as 'out-of-preference scores' (OOPS).

Although it is usual to have a number of OOPS in any one profile, it should be unusual to find that OOPS make up more than half of the facet scores relating to any one type dichotomy. Table 12.5 demonstrates that the proportion of individuals for whom this occurs ranges from 3.0% for the T–F block to 1.0% for the E–I block. It is therefore a very infrequent occurrence.

Table 12.5: Proportion of OOPS by type dichotomy[6]

	Proportion of 'reported type' OOPS					
	None	One	Two	Three	Four	Five
E–I	75.1%	19.2%	4.7%	0.9%	0.1%	0.0%
S–N	67.5%	22.6%	8.1%	1.6%	0.2%	0.0%
T–F	51.0%	34.4%	11.6%	2.7%	0.3%	0.0%
J–P	57.8%	31.1%	9.7%	1.3%	0.0%	0.0%

[6] In this table, an OOPS is defined as a score of 2, 3, 4 or 5 on the 'wrong' pole of the facet when compared with the reported type.

For any individual facet scale, an OOPS tends to occur in approximately 10% of profiles, with the exception of Questioning–Accommodating (21%) and Tough–Tender (25%).

Reliability

The reliability of a test or questionnaire relates to how consistent and precise it is. Internal consistency reliability addresses the question of whether all the questions in a scale measure the same construct. For example, are the Step II facet scales consistent within themselves, and do they hold together well as scales? A common measure of internal consistency reliability is coefficient alpha (Cronbach, 1951). The alpha coefficients for the Step II facet scales are shown in Table 12.6.

Table 12.6: Internal consistency reliability

Step II facet scale	No. of items	Coefficient alpha
E–I facet scales		
Initiating–Receiving	8	0.81
Expressive–Contained	7	0.72
Gregarious–Intimate	7	0.59
Active–Reflective	8	0.66
Enthusiastic–Quiet	9	0.56
S–N facet scales		
Concrete–Abstract	9	0.67
Realistic–Imaginative	7	0.73
Practical–Conceptual	8	0.52
Experiential–Theoretical	8	0.57
Traditional–Original	8	0.70
T–F facet scales		
Logical–Empathetic	9	0.79
Reasonable–Compassionate	8	0.68
Questioning–Accommodating	7	0.47
Critical–Accepting	8	0.43
Tough–Tender	8	0.55
J–P facet scales		
Systematic–Casual	8	0.74
Planful–Open-Ended	6	0.75
Early Starting–Pressure-Prompted	6	0.70
Scheduled–Spontaneous	8	0.71
Methodical–Emergent	8	0.59
	Median	0.67

The internal consistency reliability of most scales is good, and the average (median) reliability is close to 0.7. However, two scales (Questioning–Accommodating and Critical–Accepting) do have lower alpha coefficients. This pattern is consistent with what was found during the development of the Step II instrument, where these two

facet scales were also found to show lower reliability than the others (Quenk, Hammer and Majors, 2004).

It should be noted, however, that coefficient alpha reliability statistics will give an underestimate of the reliability of the Step II facet scales, given the use of item response theory (IRT) methods in the actual scoring process.[7]

Group differences

This section looks at the way in which people from different groups responded to the Swedish version of the European Step II questionnaire.

Gender

The means and standard deviations of the Step II facet scales are shown separately for males and females in Table 12.7, along with the difference in mean scores. This is represented graphically in Figure 12.2. Statistically significant differences were found between male and female mean scores for 18 of the 20 facet scales, with some consistent patterns emerging.

- On the E–I facet scales, all five mean scores tended slightly toward the E pole for both males and females. There was no consistent gender difference across the dimension, with females tending further towards the E pole than males on two facet scales, and males tending more towards the E pole than females on three facet scales.

- On the S–N facet scales, all five mean scores tended slightly toward the S pole for both males and females. There were statistically significant gender differences on three of the five facet scales, with male mean scores tending more towards the S pole than females on two of these, while females tended further towards the S pole on the other. There was no significant gender difference on the two remaining facet scales.

- On the T–F facet scales, four of the five mean scores tended slightly toward the T pole for males, whereas two did so for females. The Tough–Tender facet stands out because both males and females tended to score towards the F pole. There were statistically significant gender differences on all five facet scales, with male mean scores tending more towards the T pole (or less towards the F pole) than females.

- On the J–P facet scales, all five mean scores tended toward the J pole for females, whereas four did so for males. There were statistically significant gender differences on all five facet scales,

[7] Item response theory (IRT) is an approach to measurement that is concerned with modelling the relationship between item responses and the underlying characteristic assessed by the scale or test the item is designed to measure. IRT can be used to select items for a test and/or to score the items.

with male mean scores tending more towards the J pole than females on one facet (Systematic–Casual) and females tending more towards the J pole on the other four facets.

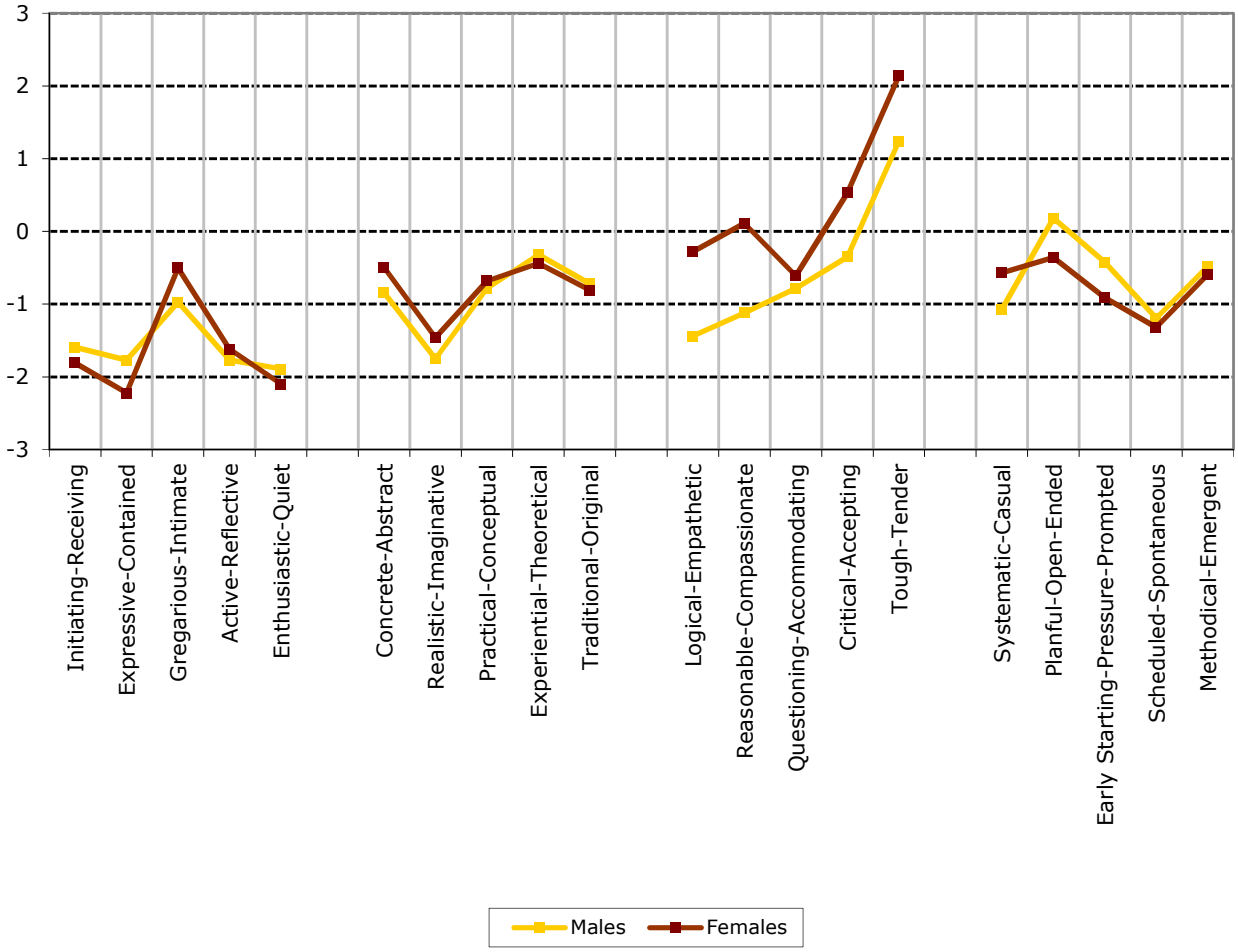
Table 12.7: Gender differences in facet scale scores

	Males (n=5,581)		Females (n=6,762)		Difference (M–F)[8]
Step II facet scale	**Mean**	**SD**	**Mean**	**SD**	
E–I facet scales					
Initiating–Receiving	-1.59	3.03	-1.81	2.92	0.22**
Expressive–Contained	-1.77	2.71	-2.22	2.51	0.46**
Gregarious–Intimate	-0.97	2.68	-0.50	2.51	-0.47**
Active–Reflective	-1.77	2.63	-1.62	2.62	-0.15**
Enthusiastic–Quiet	-1.89	2.25	-2.10	2.17	0.20**
S–N facet scales					
Concrete–Abstract	-0.84	2.36	-0.49	2.28	-0.34**
Realistic–Imaginative	-1.75	3.04	-1.46	3.00	-0.29**
Practical–Conceptual	-0.78	2.74	-0.68	2.83	-0.10
Experiential–Theoretical	-0.32	2.40	-0.44	2.40	0.12**
Traditional–Original	-0.72	2.88	-0.81	2.83	0.09
T–F facet scales					
Logical–Empathetic	-1.44	2.67	-0.28	2.89	-1.16**
Reasonable–Compassionate	-1.12	2.90	0.11	2.79	-1.24**
Questioning–Accommodating	-0.78	2.82	-0.61	2.83	-0.17**
Critical–Accepting	-0.35	1.95	0.53	1.97	-0.89**
Tough–Tender	1.24	2.46	2.14	2.28	-0.91**
J–P facet scales					
Systematic–Casual	-1.07	2.96	-0.57	2.87	-0.50**
Planful–Open-Ended	0.18	3.25	-0.36	3.30	0.53**
Early Starting–Pressure-Prompted	-0.42	3.17	-0.91	3.15	0.49**
Scheduled–Spontaneous	-1.19	2.93	-1.32	2.86	0.13*
Methodical–Emergent	-0.48	2.77	-0.59	2.63	0.11*

Difference significant at: *p<0.05, **p<0.01 (based on an independent samples t-test).

[8] A positive value indicates that male scores tend more towards I, N, F or P, and a negative value indicates that female scores tend more towards I, N, F or P.

Figure 12.2: Gender differences in facet scale scores



Age

There were found to be no clear meaningful links between age and facet scale scores. Although correlational analysis showed several facet scales to be significantly correlated with age, the significance levels may have been more the result of the sample size rather than being indicative of a meaningful relationship. The highest correlation were 0.12 (Concrete–Abstract) and even this is too small to be considered meaningful.

For interpretation purposes, it is reasonable to conclude that there are no clear relationships between age and facet scale scores.

Ethnic origin

Ethnic origin information was not captured for the individuals who completed the Swedish language version of the questionnaire, so no analyses were conducted.

Occupational level

Research using the Swedish version of the MBTI Step I questionnaire has demonstrated that individuals in higher-level jobs in organisations are more likely to have preferences for Extraversion, Intuition and (to some extent) Thinking than those in lower-level jobs.[9]

Table 12.8 shows the facet scale means and standard deviation for different occupational levels. The findings can be summarised as follows:

- There were no clear overall patterns regarding the facets in the E–I dimension. However, mean scores for 'Employee' and 'First level management/supervisory' levels tended to be less towards the E pole than the more senior level positions.
- There were no clear overall patterns regarding the facets in the T-F dimension. However, mean scores for 'Employee' level tended to be less towards the T pole than the more senior level positions on three of the facet scales.
- There were no clear overall patterns regarding the S-N and J-P dimensions.

[9] See the MBTI Step I European Data Supplement for details.

Table 12.8: Occupational level differences in facet scale scores

Step II facet scale	Top level (n=69)		Senior Executive (n=259)		Upper Middle Management (n=269)		Middle Management (n=320)		First Level Mgt/ Supervisory (n=239)		Employee (n=476)		Sig.
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales													
Initiating–Receiving	-2.83	2.33	-2.32	2.55	-2.30	2.67	-2.00	2.89	-1.49	2.97	-1.32	3.05	**
Expressive–Contained	-2.84	2.19	-2.37	2.37	-2.40	2.31	-2.43	2.60	-2.01	2.72	-1.95	2.58	**
Gregarious–Intimate	-1.01	2.76	-1.22	2.62	-1.09	2.45	-1.13	2.52	-0.94	2.46	-0.69	2.64	*
Active–Reflective	-2.23	2.40	-2.41	2.34	-2.35	2.25	-2.17	2.53	-1.70	2.66	-1.56	2.79	**
Enthusiastic–Quiet	-2.57	1.93	-2.52	2.12	-2.50	2.06	-2.41	2.11	-2.30	2.23	-1.92	2.19	**
S–N facet scales													
Concrete–Abstract	-0.13	2.19	-0.19	2.33	-0.54	2.10	-0.58	2.44	-0.56	2.18	-0.66	2.35	
Realistic–Imaginative	-1.14	3.03	-1.27	3.12	-1.65	2.90	-1.60	3.04	-1.17	3.12	-1.55	3.10	
Practical–Conceptual	0.01	2.48	-0.10	2.63	-0.74	2.62	-0.49	2.72	-0.38	2.72	-0.68	2.89	*
Experiential–Theoretical	-0.07	2.35	0.07	2.31	-0.23	2.41	-0.49	2.43	-0.45	2.40	-0.49	2.49	*
Traditional–Original	0.07	2.81	-0.07	2.82	-0.39	2.73	-0.56	2.92	-0.40	2.88	-0.63	2.91	
T–F facet scales													
Logical–Empathetic	-0.81	2.52	-1.50	2.67	-1.30	2.58	-1.02	2.79	-1.13	2.78	-0.62	2.92	**
Reasonable–Compassionate	-0.55	2.55	-1.22	2.66	-1.12	2.70	-1.01	2.83	-0.50	2.75	-0.32	2.95	**
Questioning–Accommodating	-1.23	2.60	-1.43	2.64	-1.38	2.54	-1.11	2.71	-1.11	2.75	-0.59	2.88	**
Critical–Accepting	0.22	1.77	0.03	1.93	0.03	1.84	0.21	2.00	0.14	1.97	0.09	2.04	
Tough–Tender	1.75	2.23	1.43	2.23	1.22	2.13	1.50	2.42	1.87	2.32	1.92	2.46	**
J–P facet scales													
Systematic–Casual	-0.87	2.76	-0.82	2.91	-0.94	2.73	-0.98	2.77	-0.46	2.89	-0.47	3.01	
Planful–Open–Ended	-0.41	2.92	-0.23	3.15	-0.18	3.19	-0.26	3.33	0.16	3.08	0.14	3.40	
Early Starting–Pressure–Prompted	-0.35	3.03	0.71	3.01	0.28	2.96	-0.23	3.08	-0.25	3.01	-0.77	3.18	**
Scheduled–Spontaneous	-1.62	2.80	-1.27	2.88	-1.12	2.74	-1.26	2.76	-1.12	2.78	-1.12	2.98	
Methodical–Emergent	-0.62	2.44	-0.81	2.57	-0.70	2.50	-0.60	2.66	-0.48	2.40	-0.31	2.71	

Difference significant at: *p<0.05, **p<0.01 (based on an independent samples t-test).

Education

Specific educational qualifications were not collected for the OPPassessment sample; however, the age at which individuals left full-time education was. No significant and meaningful correlations were found between the age at which people left full-time education and their facet scale scores. All the correlations were less than 0.5 and this is too small to be considered meaningful.

Work area

Information regarding the area of work people engage in was collected for the group. Many different categories were used, but for the purposes of analysis the focus was on the four most commonly occurring. These were as follows:

- Finance
- Science, engineering
- Sales, customer service
- HR, training, guidance

Table 12.9 shows the facet scale means and standard deviations for these four work areas. The findings can be summarised as follows:

- There were no clear overall patterns regarding the facets relating to the E–I dimension. However, it is worth noting that mean scores for the 'Sales, customer service' group tended further towards the E pole than any of the other groups on three of the five facet scales.
- An interesting pattern was noticed regarding the facets in the S-N dimension. Mean scores for the 'HR, training guidance' group tended further towards the N pole (or less towards the S pole) than any of the other groups on four of the five facet scales. Conversely, mean scores for the 'Science, engineering' group tended further towards the S pole than any of the other groups on all five facets.
- With regards the T-F dimension, the mean scores for the 'HR, training guidance' group tended further towards the F pole than any of the other groups on three of the five facet scales.
- The pattern regarding the J-P dimension was similar to that of the S-N dimension in that mean scores for the 'HR, training guidance' group tended further towards the P pole (or less towards the J pole) than any of the other groups on three of the five facet scales, while mean scores for the 'Science, engineering' group tended further towards the J pole than any of the other groups on all five facets.

Table 12.9: Mean facet scale scores by work area

Step II facet scale	Finance (n=246)		Science, engineering (n=179)		HR, training, guidance (n=160)		Sales, customer service (n=135)		Sig.
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
E–I facet scales									
Initiating–Receiving	-1.91	3.01	-1.09	3.31	-2.37	2.70	-2.99	2.21	**
Expressive–Contained	-2.35	2.57	-2.02	2.82	-2.74	2.45	-2.68	2.33	*
Gregarious–Intimate	-1.20	2.56	-0.87	2.79	-0.81	2.58	-1.35	2.48	
Active–Reflective	-2.01	2.78	-1.93	2.92	-2.24	2.40	-2.92	2.14	**
Enthusiastic–Quiet	-2.37	2.19	-2.09	2.24	-2.53	2.20	-2.82	1.90	*
S–N facet scales									
Concrete–Abstract	-0.33	2.34	-1.24	2.20	0.11	2.62	-0.33	2.21	**
Realistic–Imaginative	-1.62	2.90	-2.11	2.90	-0.93	3.16	-1.22	3.21	**
Practical–Conceptual	-0.58	2.72	-1.11	2.75	-0.07	2.79	-0.56	2.63	*
Experiential–Theoretical	-0.05	2.51	-0.79	2.43	0.51	2.47	-0.48	2.38	**
Traditional–Original	-0.47	2.81	-1.27	2.85	-0.05	2.92	-0.02	2.88	**
T–F facet scales									
Logical–Empathetic	-1.25	2.69	-1.49	2.59	-0.09	3.04	-0.83	2.84	**
Reasonable–Compassionate	-0.99	2.64	-1.15	2.97	0.27	2.97	-1.03	2.91	**
Questioning–Accommodating	-1.16	2.78	-0.83	2.77	-0.72	2.85	-1.55	2.56	*
Critical–Accepting	-0.07	1.93	-0.31	1.97	0.84	2.04	0.26	1.79	**
Tough–Tender	1.39	2.30	1.58	2.42	1.93	2.33	1.29	2.12	
J–P facet scales									
Systematic–Casual	-0.78	2.88	-1.24	2.93	0.21	3.19	-0.63	2.71	**
Planful–Open-Ended	-0.20	3.20	-0.50	3.42	0.36	3.38	-0.12	3.24	
Early Starting–Pressure-Prompted	0.09	3.14	-0.55	3.14	0.35	3.27	0.06	3.06	
Scheduled–Spontaneous	-0.96	2.98	-1.58	2.95	-0.36	3.10	-1.27	2.66	**
Methodical–Emergent	-0.32	2.60	-0.90	2.71	-0.04	2.82	-0.66	2.58	*

Difference significant at: *p<0.05, **p<0.01 (based on an independent samples t-test).

Nationality

Information on nationality was gathered for 16% of the group. Of these, 96% were Swedish. There were insufficient people of other nationalities for it to be possible to conduct any analyses across groups.

Employment status

Employment status information was gathered for 14% of the group. Of these, 95% worked full-time. There were insufficient people who did not work full-time for it to be possible to conduct any analyses across groups.

Appendix 1: Sample description

Sample 1: Data from OPPassessment (representative Swedish-speaking professional and managerial sample)

This sample consists of 12,343 individuals who completed the MBTI Step II instrument in Swedish via the OPPassessment system between May 2006 and June 2016. Forty-five per cent of the respondents were male and 55% were female. Age ranged from 21 to 65 years, with a mean of 44 and a median of 44.

Nationality was disclosed by 16% of respondents. Of these, 96% were Swedish. Several other nationalities were represented, but each one formed less than 1% of the total group, except for Finnish which totalled 1.5%:

Nationality	Percentage
Swedish	95.9%
Finnish	1.5%
Other	2.6%

Fourteen per cent of respondents stated their employment status. Of these, the majority of the group were in full-time employment:

Employment status	Percentage
Full-time	95.1%
Part-time	2.5%
Self-employed	1.8%
Unemployed	0.6%

Fourteen per cent of respondents stated their occupational level. The majority of the group were of managerial level or above, although the largest single group was employee (28%):

Occupational level	Percentage
Top level	4.1%
Senior executive	15.3%
Upper middle management	15.9%
Middle management	18.9%
First level management/supervisor	14.2%
Employee	28.2%
Other	3.4%

Fourteen per cent of respondents stated their work area. Amongst these, a range of work areas were represented:

Work area (job type)	Percentage
Finance	14.5%
Science, engineering	10.5%
Sales, customer service	9.4%
HR, training, guidance	7.9%
Research and development	5.2%
IT	5.1%
Admin or secretarial	2.2%
Education	1.4%
Land, sea or air transport	1.1%
Business services	1.0%
Skilled operative	0.5%
Unskilled operative	0.2%
Leisure, personal service	0.1%
Military, police, fire,	0.1%
Other private sector	14.6%
Other public sector	4.8%
Other	14.8%

MBTI® Step II™ instrument

European Data Supplement

Bibliography

January 2018

unlocking potential

OPP Ltd
Elsfield Hall
15–17 Elsfield Way
Oxford OX2 8EP
UK
t: +44 (0)1865 404 500

Bibliography

Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika,* 16(3), 297–334, 1951.

Hackston J and Kendall E. Step II and Culture. Paper presented at the APT international conference, Toronto, July 2004.

Kendall E. *Myers-Briggs Type Indicator European English Edition Step I Manual Supplement*. Mountain View, CA: CPP, Inc., 1998.

Myers IB, McCaulley MH, Quenk NL and Hammer AL. *MBTI Manual: A Guide to the Development and Use of the Myers-Briggs Type Indicator*. Mountain View, CA: CPP Inc., 1998.

Quenk NL, Hammer AL and Majors MS. *MBTI Step II Manual: Exploring the Next Level of Type, European Edition*. Mountain View, CA: CPP, Inc., 2004.