



WPI

Work-related Personality Inventory

Manual

Ixly © 2012

Powered by



Authors

Drs. Heleanne Rusthoven

Drs. Diddo van Zand

Ilse Beemsterboer, MSc.

Bart Dekker, MSc.

Merel Schrijver, MSc.

Preface

This manual contains an as complete as possible overview of the relevant information on the Work-related Personality Inventory (WPI). By means of this manual we would like to provide users with information on the theory, constructions, background and statistical properties of the WPI, to get deeper insights into this questionnaire. In addition, this manual serves as an aid for the interpretation of the WPI.

The WPI was not constructed in one day. Several years of research have preceded the current WPI and several people have contributed over the years. In different phases, the following psychologists have contributed to the creation and supporting research of the WPI:

Diddo van Zand
Petrice Broeders
Mark Westerhuis
Nilka Capricorne
Annet van de Velde
Wendy Ruiterkamp
Kees van Driel
Martijn van der Woud
Lisette van Crujningen
Jolijn Hendriks
Jelle Geertsma
Ilse Beemsterboer
Heleanne Rusthoven
Bart Dekker
Mereel Schrijver

Dr. Peter Tellegen of the Rijksuniversiteit Groningen has had an advising role in the psychometrics and in writing the first version of this manual. We thank him for this!

Drs. Martijn van der Woud has made a great contribution to the construction and statistical foundation of the questionnaire.

In addition, several hundreds of respondents have cooperated in the different studies and in the original test construction.

Finally, we would like to thank our professional users for letting us use their data, their input and for thinking constructively with us to make the WPI a practical and meaningful instrument for the HRM-professional practice.

Reading guide

For your convenience, we have made a reading guide. This reading guide provides a short description of every chapter and their most important conclusions. This will provide you with a quick overview of the information which is relevant for the use the WPI.

1. Introduction

The Work-related Personality Inventory, the WPI, is a personality questionnaire that is developed for the Human Resource Management (HRM) by Ixly B.V. The WPI consists of five factors, subdivided into 25 scales.

In this chapter, a number of theories on personality are discussed. The Five Factor Model (the 'Big Five'), underlying the WPI, is discussed. From this follows the description on how personality can be measured. Finally, we elaborate on the definition of personality, as it is operationalized in the WPI.

2. Development

This chapter describes the history of the WPI. In the first stage, the purpose of the questionnaire was defined: constructing a broad, work-related personality questionnaire, that covers the informational need by psychologists in selection, career transition and career advice as fully as possible. In the second stage, the first analyses and item selection for the ProSiD-PI 35 took place. This is a *Six Dimension Personal Inventory* with 35 scales, a first (draft) version of the WPI. After a couple of years of data collection with the ProSiD-PI 35 and renewed statistical analyses eventually the current WPI, with its five factor structure, was established. In Appendix 4 all scales and factors are clearly represented, discussed and defined.

3. Norm research

This chapter describes the process of establishing the norms of the WPI. The WPI uses norms for the interpretation of the scores of an individual. The norm population of the WPI is a representative sample of the Dutch work force in terms of education, age, gender and job status. The WPI has norm groups for two testing situations, namely advice and selection. The composition of the norm groups and the procedure of establishing the norms of the WPI are discussed. Norm tables associated with the relevant norm groups are represented in Appendix 7.

4. Reliability

In order to discuss the reliability of the WPI, reliability coefficients and the generalizability of the factors and internal consistency of every scale are calculated. The WPI consists of homogeneous, reliable and stable scales. On the factor level, reliabilities are high to very high ($> .95$). There are virtually no differences in reliability between the factors. The Advice norm group has higher alpha coefficients compared to the Selection norm group, on both the scale and factor level. However, the differences are small. To investigate the stability of the factors and scales over time a test-retest analysis is conducted. The correlation between the first (test) and second (retest) test are close to the reliability of the questionnaire. The test-retest reliability of the WPI, with a time gap of 9 months, is found to be very high.

5. Validity

Validity is the extent to which a test fulfils its purpose. In this chapter, a distinction is made between criterium validity and construct validity of the WPI. Criterium validity concerns the question whether and to what extent a test score is a good predictor of behaviour, for example in a future job. Construct validity concerns the question whether and to what extent a test measures the attribute it is supposed to measure.

Criterium validity

To investigate the criterium validity Ixly has conducted a study with 360 degrees feedback reviews of student nurses who attended the BBL variant of nursing school. The goal was to determine the

predictive value of several personality characteristics of student nurses for the level of success in their education. Based on the results of the analyses it can be concluded that there is an association between the competencies, as measured in the 360 degrees method and the personality characteristics as measured in the WPI. This association is as large as one would expect in this type of research.

In addition, a large study was conducted at an employment agency. This study shows that several scales of the WPI have predictive value for the appreciation, turnover and satisfaction of employees.

Construct validity

An important indicator of the construct validity is the internal structure of the WPI. This chapter describes research which shows that no item and/or scale correlates significantly with a scale, respectively factor, other than to which it is theoretically assigned. This is support for the stability of the structure. A large study is carried out at an ICT company in which the WPI as well as the Five Factor Personality Inventory (FFPI) was filled out. The association between the factors of the FFPI and the scales and factors of the WPI provides clarity on the psychological constructs that are measured with the WPI. The scales of the WPI all correlate in an explainable way with the factors of the FFPI.

To check whether the factor- or scale scores have a relation with the background variables, for every variable, scale and factor it is investigated whether the mean scores of the different categories of the background variable differ significantly from each other. Despite the large number of people in the norm groups, only a few significant differences were found. In terms of effect size, the significant differences for the background variables are not large enough to justify the creation of different norm groups for the variables for which differences were found.

Extensive research has been conducted into the cultural justice of the WPI, which is represented in Appendix 11. The WPI does not show cultural bias and the observed differences do justice to real individual differences.

Finally, this chapter discusses the results of a correlational study between the WPI and the Career Values questionnaire, also developed by Ixly. All the mentioned studies have contributed to the construct validity of the WPI.

6. Application, interpretation and use

The WPI can be used in every situation in which it is important to gain insights into the personality of a person in working situations. The WPI can be implemented in advice- and selection situations. The graphical part of the report is represented in sten scores. This chapter discusses the interpretation of these sten scores. For a good interpretation of the questionnaire, it is important to know the meaning of the factors and scales. This is discussed in Appendix 4. To illustrate the interpretation of the WPI, a psychologist describes how he applies the WPI in selection assessments and career advice.

Substantive connections between scales belonging to one factor and scales belonging to other factors can be made, from which a certain profile can be distilled. A few examples are given that are appropriate for professional indications.

1. Introduction

The Work-related Personality Inventory, the WPI, is a personality questionnaire that is developed for the Human Resource Management (HRM) by Ixly¹. The WPI consists of five factors, subdivided into 25 scales.

First, this manual will focus on the concept of 'personality' and the theoretical background of the WPI. Following this, the development, norms, the reliability and validity of the questionnaire will be discussed. Finally, the application of the questionnaire, the interpretation of the results and which conclusions one can draw from the results will be elaborated on.

1.1. Personality

The term personality is based on the Latin word *Persona*, which means 'mask'. When we strictly define personality as mask, we could see at as the aspect of ourselves that we show the outside world (Hergenhahn, 1980). Several theorists have focused on the concept of personality. The theory of personality by Freud – with its trichotomy in the *ID*, the *EGO* and the *Superergo* – is probably the best known. The theories by Maslow, Jung and Rogers are quite well known as well. For this, see for example Hergenhahn (1980).

There is a lot of discussion on the exact definition of personality. Because of this, there are a large number of different definitions, ranging from simple versions to complex technical definitions including mathematical equations. A couple of examples:

- Personality can be describes as 'a collection of attributes of the way in which situations are distinguished, interpreted and evaluated by a person. This interpretation and evaluation influence the behaviour a person will show in the specific situation' (Hoekstra, Ormel & Fruyt, 1996).
- Personality is 'the content of what is said about someone with the intent to give a specific explanation for his or her behaviour. Including the one who makes the statement about the person in the definition is important because someone's personality often depends to a large extent on the judging person and the person that is judged him/herself' (Hofstee, Brokken & Land, 1981).

Nowadays, the most prominent theory in the research on personality is the Five Factor Model (Allport & Odbert, 1936; Cattell, 1943). This theory is also called the Big Five (Goldberg, 1981). The principle of the Five Factor Model is closest to the principle of the WPI.

The Five Factor Model (FFM)

The FFM has its origins in the (psycho)lexical hypotheses. These state that all individual differences between people that matter can be expressed in language. More concrete, this means that trait adjectives from the American dictionary were collected by Allport and Odbert (1936) and later by Cattell (1943), in order to come to a complete as possible reflection of human personality. This collection was found to be describable by a five factor model. This conclusion was based on principle component analysis on the trait adjectives, collected as self-evaluations or evaluations by others. After this a varimax rotation of these components was performed, to obtain positions of the axes for which the components are as easily interpretable as possible. The position of the axes is optimal when the adjectives load highly on one component and show a negligible loading on the other components. The components are then interpreted using the items with the highest loadings.

The theory of the FFM states that there are five main factors or dimensions of personality characteristics on which people differ from each other and can be compared on. The five factors of the FFM are:

1. *Extraversion*
2. *Agreeableness*
3. *Conscientiousness*
4. *Neuroticism*
5. *Openness to experience*. (Allport & Odbert, 1936; Cattell, 1943)

¹ Ixly (previously Orga B.V.) is a publisher of online instruments and is dedicated to the development, research and publishing of questionnaires and tests for the HRM profession. These are distributed via an online application.

Development and stability of the personality factors

During adolescence it is possible for the scores of a person on the five factors to change somewhat. Often the scores on *Agreeableness* and *Conscientiousness* increase a little, while scores on *Extraversion*, *Neuroticism* and *Openness to experience* decrease somewhat. It appears that from the age of thirty no substantial differences in scores occur. From that moment on, consecutive test scores correlate highly. A person thus has a stable profile from that age on (McCrae & Costa, 1994). McCrae and Costa (1994) found remarkable stability of the factors in longitudinal research. This stability was also found in cross-sectional research. Changes are not impossible, for example when dramatic events occur in a person's life, but in general we can say that the personality profile of a person is relatively stable.

Gender differences

There are differences in the scores on the FFM between men and women. These differences are found in different cultures. Women on average score higher on the factor *Agreeableness* (*Sociability*) and lower on the factor *Stability* (de Dusschoten, 2004). Women on average score higher on *Attentiveness* and *Need for contact* (Beutel & Marini, 1995). Men score higher on *Status* and *Influence* (Aries, 1976). Research investigating the association between personality characteristics and educational attainment has also shown that there are differences between men and women. The association between personality and education is higher for men than for women (van Eijck & de Graaf, 2001).

1.2. Measuring personality in practice

Personality can be measured with a psychological test. This concept can be described as follows:

'A systematic evaluation- and measuring procedure, which makes it possible to make statements about one or more personality characteristics of the examined person or of his or her future behaviour or future performance'.

These statements are based on an objective and comparative processing of the reactions and performance of the examined person on carefully selected assignments or questions, which are presented to his/her in a standardized manner (Drenth, 1981). The reactions of the person form the test information on which the statements are made. This test information can be obtained in several ways, for example by means of self-evaluations, observations, instrumental measurements or objective documentations (Drenth, 1981).

The personality questionnaire is an example of a psychological test. The most important distinction that is made in tested behaviour is the distinction between tests of performance and tests of behaviour or conduct (Drenth & Sijtsma, 2006). A performance test, with right or wrong answers, demands a maximum performance of the examined person. For tests of behaviour there is no pre-determined 'right/wrong key' available. Because tests of behaviour are not tests in the sense of 'aptitude', we often speak about 'questionnaire' in this context (Drenth & Sijtsma, 2006).

Since the twenties of the previous century, personality questionnaires are often used in practice. For a long time, no clear relation was found between criteria of good work behaviour and the outcomes of the questionnaires (Salgado & de Fruyt, 2005). There was a clear distinction between practice and research. Despite this fact personality questionnaires remained popular in practice. In addition to the use in (mental) health care, the so called psycho-diagnostics, they can be used in the field of HRM and personnel selection as well.

Essential for the use of personality questionnaires in the work field of HRM is *predictive validity*: to what extent does personality have predictive value for (future) job performance? In recent years research has shown that personality questionnaires have predictive value in personnel selection. Several studies have shown the predictive validity of the factors of the FFM. Two of the five factors, *Extraversion* and *Conscientiousness*, appear to be particularly good predictors of job performance (Schultz & Schultz, 2002). The predictive value appears to differ between different research populations. Research by Barrick & Mount (1991) on the FFM and job performance shows that *Conscientiousness* is a good predictor for all examined groups. The same study shows that

Extraversion and *Agreeableness* are predictors for managerial positions only and that *Openness to experience* is a predictor for trainers. *Extraversion* mainly appears to have predictive value for performance in occupations in which social interaction and sales are central (Barrick & Mount, 1991). The conclusion by Barrick and Mount is mainly that *Conscientiousness* would be a good predictor of job performance. Furthermore, they state that the FFM is a good tool to investigate personality in the workplace. The same conclusions are drawn in a smaller study by Tett, Rothstein and Jackson (in: Salgado & de Fruyt, 2005, p.176).

WPI

The initiative within Ixly for the development of a personality questionnaire came from the HRM field: there was a need for a personality questionnaire specifically developed for the use within the HRM work field. The personality questionnaires that were used in this field did not meet the needs of the assessment psychologists. The goal of the development of this questionnaire was to form to an as complete as possible questionnaire, covering all the aspects of personality that are relevant for work situations. Furthermore, there was the need to substitute the relatively time intensive questionnaires that were used to get a personality profile with one questionnaire from which a similar personality profile could be distilled but within a shorter time period.

Personality, as it is operationalized in the WPI, can be defined as follows:

'A reasonably stable behavioural tendency that is most likely to be shown in relevant (work) situations'.

With behaviour, we mean the instrumental actions one takes. *Orderliness* and *Agreeableness* are examples of this. Behaviour also encompasses the cognitive actions, sometimes called 'covert behaviour'. One can think about things like *Attentiveness* and *Originality*.

Assessment psychologists indicated that they were in need of 'narrow' constructs in personality questionnaires. With narrow constructs we mean psychological constructs that are not indicative of more personality attributes. Psychologists find differentiation very important: they often want as many constructs as possible. However, in academics, more limited and simpler questionnaires are preferred. The WPI is developed for the field of practice, making sure that no relevant differentiations that can be of interest in a work-related personality questionnaire are lost.

2. Development

This chapter describes the history of the WPI. The development of the questionnaire is described in terms of several developmental stages. In the first stage the purpose of the questionnaire is defined and the design of the questionnaire is described. In the second stage the first analyses and the item selection for the ProSiD-PI 35² take place while in the third stage the structure of the WPI is represented and how this structure came about.

2.1 First stage: history

In 1998 the need for a personality questionnaire tailored to the HRM work field arose among psychologist working in the field of personnel selection: a work-related personality questionnaire. At that time, multiple personality questionnaires were used in work-related assessments. These questionnaires did not meet the requirements of psychologist in the testing process. Furthermore, for candidates, the testing process was quite time consuming.

A research team was commissioned by Ixly to investigate and get an overview of the specific information requirements in the field. Twenty psychologists were approached, ten of which were working at Ixly while the other ten were working externally. The need turned out to be for a personality questionnaire that reports on scales that say something about the job performance or competencies of a person. If there could be one single personality questionnaire to serve this purpose – instead of multiple questionnaires – then this would save considering testing time as well. After full investigation of the specific needs, a purpose and target audience for the questionnaire was defined.

Purpose

The purpose was to construct a broad, work-related personality questionnaire that covered the informational needs by psychologists in selection, career transition and career advice as fully as possible.

Target audience

The target audience of the WPI is the Dutch work force. This resulted in the following requirements:

- Completeness
- Relevance: applied to working situations;
- Widely applicable: from lower to higher educational levels and for different types of functions;
- Narrow definitions of psychological constructs that can be used to make unambiguous statements. Furthermore, this enables automatic reporting.

The aforementioned psychologists were asked, independent from each other, to state as many terms and concepts that are work-related and/or apply in an HRM-related personality questionnaire, in their opinion. All mentioned concepts were scrutinized after which 80 concepts remained. After this, these concepts were operationalized which resulted in about 800 items. During the operationalization to the item level explanatory keywords with predictive value for the relevant concepts were sought. The goal was for a keyword in an item to be only related to one single concept. Respondents are found to respond to keywords when answering questions. This was evidenced by previous observations by the research team. For the Career Values Questionnaire, the OPF and Mobility Indicator (Orga, 2007), the research team was under the impression that people are intended to respond on the word level rather than on the sentence level. Due to possible future use of the questionnaire by others in a 360 degrees feedback version it was decided to formulate the items in the third person. Furthermore, there are indications that people tend to evaluate themselves more neutrally when the items are formulated in the third person (Hendriks, Hofstee & de Raad, 1998). All items were presented to a number of foreign persons who were proficient in the Dutch language but not native speakers. According to the results, all proverbial and typically Dutch items were adapted. The resulting 800 items were presented to a first population (N = 150) in a paper and pencil version. On the basis of analyses of the results, refinements have been made where we focused on high inter-item correlation, creating narrow constructs. Eventually, 430 items remained, corresponding to 35 constructs.

² ProSiD 35 stands for Professional Six Dimensional Personal Inventory with 35 scales. This is the name of the first (draft) version of the WPI.

2.2. Second stage: ProSiD-PI 35

Research continued with the 35 concepts and 430 corresponding items. For clarity, from now on we will call the concepts that are made up by the items, 'scales'. To create a structure, these scales are classified in factors: *Ambition*, *Work Attitude*, *Emotional Stability*, *Extraversion*, *Altruism* and *Culture*. The factor *Work Attitude* matches the factor *Conscientiousness* of the FFM, the factor *Emotional Stability* matches the factor *Neuroticism*, *Extraversion* has the same name, the factor *Altruism* matches the factor *Agreeableness* and the last FFM factor *Openness to experience* is divided up into *Ambition* and *Culture*. It is important to note that these factors in the model of this version of the WPI have an overarching, theoretical function. The fact that the names of the five factors of the FFM are changed is in the first place because this is a work-related questionnaire for which the new names are more applicable. Second, we have chosen to coin the names in a positive manner, resulting the factor *Neuroticism* to be changed to *Emotional Stability*. Third, the factor *Openness to experience* is ambiguous in the FFM. Separation of this factor into *Ambition* and *Culture* is easier to substantiate theoretically. In the classification of the factors we have taken the work-related characteristic of the personality questionnaire into account, meaning that the factors should be easily interpretable for assessment psychologist in the HRM work field.

In 2001, the first version of the WPI was created: the ProSiD-PI 35. To investigate the validity and reliability of the ProSiD-PI 35, it was administered to 350 people. A correlational study (Orga, 2002) was performed, in which four personality questionnaires were used as criteria. The four questionnaires were:

- Nederlandse Persoonlijkheid Vragenlijst (NPV: Luteijn, Starren & van Dijk, 1985)
- Edwards Personal Preference Schedule (EPPS Nederlandse bewerking: Tjoa, 1993)
- Guilford LTP Temperament Survey (GLTS: Akkerman & Buijk, 1994)
- GPP/GPI (Gordon, 1963)

Two norm groups were formed for the ProSiD-PI 35, namely Selection (persons who took the questionnaire in a personnel selection situation) and Advice (persons who took the questionnaire in a career advice situation). To assess the reliability of the ProSiD-PI 35 the internal consistency (Cronbach's alpha) for every scale, for the two norm groups was calculated. For the Selection norm group the internal consistency (α) ranged from 0.68 to 0.91. For the Advice norm group the internal consistency ranged from 0.74 to 0.92.

2.3. Third stage: Structure of the WPI

After a couple years of gathering data with the ProSiD-PI 35 and renewed statistical analyses the final version, the WPI, was finished in 2008. For users, the WPI was called the ProSiD-PI 25 after the ProSiD-PI 35. This version was also used to gather data to analyse. The number of items of the WPI is reduced in comparison with the ProSiD-PI 35 and the ProSiD-PI 25; this was done after the development of a new underlying model that supports the classification in factors, scales and items statistically. The WPI consists of 25 scales with 276 items.

In developing a psychological test, one constructs a test from theory. Subsequently, one checks whether the test meets the predetermined requirements. Unfortunately, large numbers of data are not directly available for the analyses that are necessary in order to improve the tests. In constructing a psychological test, therefore, one often works 'backwards'. The test can only be put together and restructured correctly after years of gathering data. Only then a solid statistical foundation is possible. This process has also taken place in the development of the WPI. There was a frequently used questionnaire, the ProSiD-PI 35, with a solid theoretical basis and from which a lot of data was collected. This questionnaire was adjusted and improved through statistical analyses, which resulted in the current version of the WPI. This process is discussed in more detail below.

Step 1: The theoretical classification of the factors, scale and items of the ProSiD-PI 35 were investigated thoroughly. The scores on the scales of a temporary version, Version X (which has never effectively been in use), of 17950 respondents were studied. Version X existed of 171 items.

Step 2: The scale scores of Version X were factor analysed, using Principal Component Analysis (PCA) with varimax rotation. This analysis indicated that a five factor structure was most suitable.

Step 3: Subsequently, for every item correlations with the factor scores resulting from step 2 were calculated. On the basis of these correlations, every item was assigned to one of the five factors.

Step 4: For each factor the assigned items were factor analysed to form scales. Again, we used PCA with varimax rotation. The number of scales within a factor was established using *scree plots* (graphical representations of the amount of explained variance by consecutive factors) and by the interpretability of the found solutions. Items were assigned to scales based on the component loadings as well as on theoretical considerations. The reason to divide items up into scales by factor rather than by all items at once, was to reduce the number of elements on which the PCA was performed. This reduction makes the found solution more stable (more replicable) and easier to interpret.

Step 5: After assignment of the items to scales, adjustments were made using 'trial and error', to optimize the scales in the following ways:

- Not too many items in one scale
- Reliability as high as possible
- Each scale correlates clearly with one single factor
- Minimize ceiling effects

This optimization procedure resulted in the deletion of items that did not have a substantial influence on the reliability of a scale score or that caused a second factor. In this process, we strived for minimization of ceiling effects by keeping the items with low mean scores and to make the number of items in a scale not too small. We did not find any floor effects, so these were not considered. In some cases, scales were combined when doing so yielded a one dimensional scale.

Step 6: After optimization of the scales, another factor analysis (PCA) was performed on the created scale scores in order to reassign the scales to factors. This was done because the optimized scales differed from the scales scores of 'version X' (step 1) in some aspects. Now, we found a six factor solution to be optimal.

Step 7: A last factor analysis of the factors was done using the Multiple Group Method (MGM) instead of the more common known PCA. This was done because of the fact that the MGM is more applicable for testing a specific hypothesis (Nunnally, 1978). Given that we wanted to know whether the previous assignment of scales into the six factor structure held for the current data, this method was preferred: the hypothesis being of course that the structure was the same as in the analyses on previous datasets. For more information on this procedure see Stuive, Kiers, Timmerman & ten Berge (2008). One thing we have to note is that in the MGM analysis we have performed, we used the formula of Steiger (1980) in order to determine the significance of the difference between dependable correlations.

The analysis was initially done for the Advice and Selection group together. This analysis showed that there were a few items that should have been assigned to another scale, indicated by a significantly higher correlation with one of the other scales than their assigned scale. Theoretically this was explicable, so the respective items were assigned to the other scales or deleted. This was the case for twenty items. After changes, deletion or some additions to prevent ceiling effects, the MGM analysis was run again. This analysis was performed for the Advice group and the Selection group separately. The results are presented in Appendix 1. For each of the groups, two items still appeared not to correlate significantly the highest with their 'own' scale. We decided to keep the items in their own scales, because the concerned scales were close to each other in terms of interpretation. An overview of the differences and overlap between the ProSiD-PI 35 and the WPV is provided in Appendix 2.

Step 8: An MGM analysis was also performed on the division of scales into factors, for the Selection group as well as for the Advice group. In this analysis, we tested whether the six factor structure was still optimal. We found earlier, using PCA and *scree plots*, that in this composition of scales a five factor and six factor structure would be defensible. In the MGM performed with five factors we could substantiate dividing the factor *Ambition* and place their scales under the factors *Influence* and *Exuberance*. An overview of the changes that have taken place in the formation of scales is provided in Appendix 3.

Due to changes in the structure and item format and following results from validation research, that will be discussed later in this manual, some changes in the names of scales and factors have occurred.

In Appendix 4 all factors and scales are clearly represented, discussed and defined. For each scale, an example item is included. By means of the description of the history of the questionnaire's structure and this appendix including the sample example items, one can get an idea of the content of the items and questionnaire. On account of protection of the questionnaire not all items of the WPI will be discussed in this manual. In Table 2.1, the final factor and scale structure are shown.

Table 2.1. Factor- and scale structure of the WPI

Factor	Scale	Number of items
Influence	Status	10
	Dominance	12
	Competition	9
	Self-presentation	8
Sociability	Need for contact	10
	Leisure contact	12
	Self-disclosure	10
	Trust	10
	Friendliness	12
	Attentiveness	19
Exuberance	Energy	12
	Personal growth	10
	Perseverance	13
	Adaptability	7
	Originality	10
	Independence	11
Structure	Orderliness	10
	Precision	10
	Regularity	10
	Conformity	12
	Decisiveness	11
Stability	Self-confidence	12
	Optimism	14
	Frustration-tolerance	11
	Resilience	11

2.4 Scoring of the WPI

Using a five point scale, candidates are asked to indicate whether he or she agrees or disagrees with the statement. For each scale, there are some items that are worded negatively. In summing the candidate's score these items are keyed in the other direction. A scale score is an unweighed sum of the items belonging to that scale. A factor score is an unweighed sum of the standardized scale scores.

3. Norm development

The WPI uses norm groups for the interpretation of the scores. This means that the scores of a candidate are being compared with the scores of a reference population. The reference population for the WPI is a representation of the work force of the Netherlands. By means of weights we have made sure that the sample population matches the Dutch work force in terms of education, age, gender and working situation as close as possible. For information on the work force we have used data from 2006, from the Central Bureau of Statistics (CBS, 2007). The questionnaire has two norm groups, namely Selection and Advice. In this chapter, we will discuss how the norm groups are formed and we describe the standardization procedure.

3.1 Norm groups

The WPI is used in two different testing contexts, namely in selection and advice situations. Advice situations include all forms of career counselling, consultancy and coaching. Candidates who take the WPI in these kinds of situations will come to different, most likely less socially desirable scores than candidates that take the test in selection situations. With selection situations we mean, for example, assessments in job application procedures. In these situations, people will be more prone to give socially desirable answers because they will think this will get them through the selection. For both situations, independently from each other, but in the same manner, norms were calculated.

The data over which the standardization procedure was run, was collected between 01-03-2004 and 29-12-2006 for the Advice group and between 22-01-2004 and 03-01-2007 for the Selection group. Data was collected for 5629 people in the Advice group, of which 47.9% was male and 52.1% was female. The average age of this group is 36 with a minimum of 15 and a maximum of 62. For the Selection group, information was available on 1514 people, 38.8% male and 61.2% female. The average age of this group is 32 with a minimum of 17 and a maximum of 59. Education levels for both groups vary from lower education to academic education.

The two research groups were weighed for four variables, in such a way that they resembled the Dutch work force in terms of these four variables. In Table 3.1 these variables are described.

Table 3.1. Description of the background variables used in the weighing procedure

Education	The education levels in which the work force is divided are: lower education, secondary education and higher education. The education programs that fall under each of the levels are presented in Appendix 5.
Age	The age variable is divided up into three categories. The work force falls within the age range of 15 to 65 years. The three age groups are 15-24 years old, 25-44 years and 45-65.
Gender	Male/female
Job status	A distinction is made between the employed and the registered unemployed. The data on the latter group is obtained via the UWV ³ , where job seekers took the WPI.

All combinations of the four variables result in 36 cells. The weight factors are chosen in such a matter that the distribution of the 36 cells corresponds as much as possible with the relative proportions in the Dutch population (See Appendix 6 for details on this procedure).

The weighted Advice- and Selection group form the norm groups for respectively the Advice- and Selection situation. In Table 3.2 the distribution in terms of the background characteristics are shown for both research groups, for the weighted norm groups and for the Dutch work force.

³ At the time of research, the UWV was known as the Centre for Work and Income (CWI).

Table 3.2. Frequencies of background characteristics in both norm groups

	Advice group	Weighted Advice group	Selection group	Weighted Selection group	Work force
Male	2699= 47.9%	402= 56.5%	585= 38.6%	209= 56.5%	56.5%
Female	2930= 52.1%	310= 43.5%	929= 61.4%	161= 43.5%	43.5%
Education:					
lower	1412= 25.1%	172= 24.2%	260= 17.2%	89= 24.2%	24.2%
secondary	2354= 41.8%	317= 44.5%	470= 31.0%	164= 44.5%	44.5%
higher	1863= 33.1%	223= 31.3%	784= 51.8%	116= 31.3%	31.3%
15-24 years	861= 15.3%	81= 11.3%	404= 26.7%	42= 11.3%	11.3%
25-44 years	3289= 58.4%	373= 52.4%	865= 57.1%	193= 52.4%	52.4%
45-65 years	1479= 26.3%	258= 36.3%	245= 16.2%	134= 36.3%	36.3%
CWI	4701= 83.5%	24= 3.4%	339= 22.4%	13= 3.4%	3.4%
non CWI	928= 16.5%	688= 96.6%	1175= 77.4%	357= 96.6%	96.6%

From Table 3.2 it becomes clear that in both norm groups, especially for the Advice group, the background characteristic *working situation* has had a large influence in the weighing procedure. In the data on which the norms have been formed there were relatively more women than men, while in the work force this is exactly the opposite. In addition, in data of the Selection group, there was a relatively large group of higher educated people.

National representativeness

The data was collected at 118 different companies from all over the Netherlands. For the Advice norm group, information was gathered from 2775 people. For the Selection norm group, information was gathered for 2708 people. In Table 3.3 the number of persons from every Dutch province are provided for both norm groups. The real distribution of persons over the Dutch provinces is provided as well, as measured by the CBS in 2011.

Table 3.3. Distribution of regions in both norm groups

	Advice norm group		Selection norm group		CBS 2011*
	N	%	N	%	%
Zuid-Holland	627	22.6	484	17.87	21.32
Noord-Holland	443	16.0	738	27.25	16.78
Brabant	362	13.0	345	12.74	14.82
Limburg	330	11.9	218	8.05	6.51
Utrecht	255	9.2	243	8.97	7.64
Gelderland	237	8.5	243	8.97	11.86
Overijssel	143	5.2	126	4.65	6.63
Groningen	106	3.8	35	1.29	3.32
Drenthe	99	3.6	32	1.18	2.76
Friesland	81	2.9	70	2.58	3.74
Flevoland	42	1.5	67	2.47	2.47
Zeeland	25	.9	12	0.44	2.16
Unknown / will not say	25	.9	95	3.51	-
Total	2775	100%	2708	100%	100.0

* CBS Key figures Work force 4th quarter of 2011, requested in 2012

Distribution of work sectors in norm population

The background characteristics of a sample of 1786 respondents from 118 companies, operating in different work sectors, were successfully retrieved. Because the sample was drawn randomly, it can be rightfully assumed to be a representative sample of the reference population.

A sample of 1786 respondents was drawn from the same companies that provided the respondents on which the norm groups were based. Of these companies the work sector, ranging over 12 different sectors, was retrieved.

Because the sample was drawn randomly, it can be rightfully assumed to be a representative sample of the Dutch work force. Since the same companies were used, it can be reasonably assumed that the distribution of the companies in sectors is the same as in the norm group.

Table 3.4. Distribution of work sectors in both norm groups

Test situation		N	%
Advice	Health, Wellbeing and Personal Care	443	23%
	Business services	314	17%
	Public administration, Safety and Law	284	15%
	Automation and ICT	201	11%
	Engineering and Production	194	10%
	Commerce and Administration	131	7%
	"Do not know"	88	5%
	Education, Culture and Science	69	4%
	Agriculture, Culture and Environment	44	2%
	Catering and Housekeeping	41	2%
	Storage and Transport	33	2%
	"Prefer not to answer"	23	1%
	Language, Media and Communication	22	1%
	Tourism and Recreation	7	1%
Total	1894	100%	
Selection	Health, Wellbeing and Personal Care	741	34%
	Public administration, Safety and Law	511	23%
	Business services	315	14%
	Engineering and Production	175	8%
	"Do not know"	125	6%
	Commerce and Administration	79	4%
	Education, Culture and Science	70	3%
	Automation and ICT	46	2%
	"Prefer not to answer"	28	1%
	Language, Media and Communication	26	1%
	Catering and Housekeeping	24	1%
	Storage and Transport	21	1%
	Agriculture, Culture and Environment	17	1%
	Tourism and Recreation	7	1%
Total	2185	100%	

In Table 3.4, the number of respondents in every work sector are represented. In the advice situation, 111 respondents of the total 1894 responded with "I do not know" or "Prefer not to answer". In the selection situation this was the case for 153 of the 2185 respondents. All the work sectors are represented in the sample.

The sectors 'Health, Wellbeing and Personal care', 'Public administration, Safety and Law' and 'Business services' contain a lot of people not because they are overrepresented, but because a relatively large number of professions fall within these categories. The distribution of people in different sectors is representative of the work force.

Our norm groups for the selection and advice situation are a good representation of the Dutch work force in terms of education, age, gender, work situation and work sector. This, in addition to the size of the norm groups contributes to the reliable standardization of the WPI.

3.2 Establishing norm groups

For each scale and for both weighed groups the raw scale scores are transformed into stenscores. These stenscores are normally distributed scores on a scale ranging from 1 to 10. The theoretical mean is 5.5 with a standard deviation of 2, the deviation of the true scores is fixed to 2. The scale is symmetrical and normally distributed. Standard scores, to which stenscores belong, provide insight in the way a certain scores relate to the mean of all scores. How the stenscores are interpreted will be discussed in the chapter "Application, interpretation and use".

The stenscores are not directly calculated discretely, but as interval scores and reported discretely. The factor scores are being calculated by summing the standardized scale scores, after which this

sum is in turn transformed into stenscores as well. The procedure that is used to transform the raw scale scores into stenscores is described in van der Woud (2007), see Appendix 7.

Standard scores and likelihood intervals

In Appendix 8 the norm tables of the 25 scales and 5 factors are shown. In these tables, for every raw score the corresponding standard score, norm score and latent score is given. In addition, for the norm scores a confidence interval is given and for the latent score a probability interval. In order to avoid confusion, both concepts are discussed in more detail next.

Clarification of statistical terms (COTAN, 2010)	
True score $\hat{T} = X$	The true score \hat{T} is not observable and is estimated simply by using the observed score X.
Standard error of measurement $S_E = \sigma_x \sqrt{1 - r_{xx}}$	The standard error of measurement (S_E) is estimated by the standard deviation (σ_x) and the reliability of the scale (r_{xx}). It is assumed that measurement errors are normally distributed. Standard errors of measurement are used to estimate a reliability interval for the true score \hat{T} .
Reliability interval $X \pm 1.96S_E$	The reliability interval (RI) is situated symmetrically around X and is estimated by the standard error of measurement and a corresponding level of reliability. The value of 1.96 is used for the 95% interval, 1.28 for the 80% interval. The RI indicates the precision of a measurement and can be used to test hypotheses on someone's true score.
True score $\hat{T} = \mu_x + \sqrt{r_{xx}}(X - \mu)$	In this case, the true score \hat{T} is estimated by means of linear regression. From the formula, it can be derived that the mean μ_x influences the true score \hat{T} . The higher the reliability of a scale, the larger the share of an individual's score X on the estimate \hat{T} . With a low reliability, μ_x will have a larger share in \hat{T} .
Standard estimation error $S_{est} = \sigma_x \sqrt{r_{xx}} \sqrt{1 - r_{xx}}$	The standard estimation error S_{est} indicates the variance of the true score. As can be derived from the formulas, the standard estimation error is $\sqrt{r_{xx}}$ times smaller than the standard error of measurement.
Likelihood interval $\hat{T} \pm 1.96S_{est}$	The standard estimation error S_{est} produces a likelihood interval (LI) which is symmetrically distributed around \hat{T} . The LI is important when one wants to get an estimation of the level of the measured variable, while taking the reliability of the variable into account.
Eta-squared $\eta^2 = S_{predictor} / S_{total}$	Indicates the explained variance in the sample by the model. η^2 usually overestimates the explained variance, but in larger samples this overestimation disappears.

For both norm groups the average stenscore of every factor is shown in Table 3.5, before and after weighing. The average stenscores of every factor for the categories of the different background variables are given as well.

Table 3.5. Average stenscore of the factors and background variables

	Average stenscore* unweighed Advice group	Average stenscore* weighted Advice group	Average stenscore* unweighed Selection group	Average stenscore* weighted Selection group
Influence	5.0 (2.2)	5.5 (2.2)	5.6 (2.6)	5.5 (2.2)
Sociability	5.4 (2.3)	5.5 (2.2)	5.6 (2.2)	5.5 (2.2)
Exuberance	5.3 (2.2)	5.5 (2.2)	5.4 (2.2)	5.5 (2.2)
Structure	5.8 (2.6)	5.5 (2.4)	5.4 (2.4)	5.5 (2.4)
Stability	5.2 (2.2)	5.5 (2.2)	5.2 (2.2)	5.5 (2.2)
Total score	5.3 (2.3)	5.5 (2.3)	5.4 (2.2)	5.5 (2.3)
Male	5.3 (2.4)	5.5 (2.3)	5.5 (2.4)	5.6 (2.4)
Female	5.2 (2.1)	5.5 (2.2)	5.4 (2.1)	5.4 (2.1)
Education:				
lower	5.2 (2.4)	5.6 (2.4)	5.3 (2.5)	5.4 (2.5)
secondary	5.3 (2.3)	5.4 (2.2)	5.3 (2.3)	5.4 (2.2)
Higher	5.4 (2.1)	5.6 (2.2)	5.5 (2.0)	5.7 (2.1)
15-24	4.8 (2.2)	5.0 (2.1)	5.5 (2.2)	5.9 (2.2)
25-44	5.3 (2.3)	5.6 (2.2)	5.4 (2.2)	5.6 (2.2)
45-65	5.4 (2.3)	5.5 (2.4)	5.2 (2.5)	5.3 (2.4)
CWI	5.2 (2.3)	5.2 (2.4)	4.8 (2.5)	4.8 (2.7)
non CWI	5.5 (2.2)	5.5 (2.3)	5.6 (2.1)	5.5 (2.2)

*Stenscore with the corresponding standard deviation in brackets.

Conclusion

The standardization of the WPI has taken place on the basis of a weighted norm group. The norm groups are a good representation of the work force in terms of education, age, work situation and work sector. In Appendix 8, for every scale and for both norm groups a detailed overview is given of the raw scores and their corresponding stenscores. Thanks to the large norm groups and the weighed distribution, the standardization of the WPI is very solid.

4. Reliability

The reliability of a questionnaire provides an indication of the precision of an instrument. The concept concerns the replicability of the measured results: to what extent do we find the same results when we use an instrument for a second (or third) time, or to what extent are the results of a comparable set of items the same?

Internal consistency

The reliability of a questionnaire can be determined in multiple ways. The same questionnaire can be administered to the same respondents twice, after which the two measurements can be compared with each other (test-retest reliability). Furthermore, scores on one half of the questionnaire can be compared with the other half of the questionnaire (split-half reliability). The most used method and suitable for a questionnaire like WPI is the use of Cronbach's alpha (α -coefficient). This is a measure for internal consistency (Nunnally, 1978). With α greater than 0.85, one can speak about a reasonably homogeneous group of items (Green, Salkind, & Akey, 2002).

In order to assess the reliability of the WPI the reliability and the generalizability of the factors were calculated, as well as the internal consistency (Cronbach's alpha) of every scale. This was done separately for the Advice and Selection group. The reliability of the factor scores was calculated with the formula for the stratified alpha (see Nunnally, p. 248). The generalizability was calculated using the formula for alpha in which however the items were not the parameters, but the scale scores (see Snijders, Tellegen & Laros, 1988). The generalizability indicates the expected correlation with a factor score based on a different, similarly large, sample of scale scores from the same domain of that specific factor.

Table 4.1. Reliability and generalizability of factors for both norm groups

Factor	Advice group (N= 712)		Selection group (N= 369)	
	Reliability*/ Generalizability**	SEM***/ SEE****	Reliability*/ Generalizability**	SEM***/ SEE****
Influence	.96 / .80	1.00/ .89	.95 / .80	1.00/ .89
Sociability	.97 / .83	.91/ .82	.96 / .81	.97/ .87
Exuberance	.97 / .82	.94/ .85	.96 / .81	.97/ 0.87
Structure	.96 / .68	1.37/ 1.13	.95 / .70	1.13/ 1.10
Stability	.96 / .82	.94/ .85	.96 / .81	.97/ .87
Mean	.96/ .79	1.03/ .91	.96/ .79	1.01/ .92

* Reliability is the stratified alpha

** Generalizability is Cronbach's alpha

*** Standard error of measurement

**** Standard error of estimate

The reliability and generalizability of the factors in both groups are shown in Table 4.1. In addition, the standard error of measurement and standard error of estimate are shown.

The reliability of the scales for both groups is shown in Table 4.2. The questionnaire consists of predominantly homogeneous, reliable and stable scales. On factor level, we can say that the reliability is very high (>0.95). There are virtually no differences between the factors in terms of their reliability. The Advice group has higher alphas in comparison with the Selection group, on the scale level as well as on the factor level. However, the differences are negligible.

Table 4.2. Reliability of the scales for both norm groups

Scale	Number of items	Advice group (N= 712)		Selection group (N= 369)	
		Alpha*	SEM**/ SEE***	Alpha*	SEM**/ SEE***
Status	10	.90	.67/ .63	.88	.74/ .69
Dominance	12	.92	.59/ .57	.88	.74/ .69
Competition	9	.88	.74/ .69	.89	.70/ .66
Self-presentation	8	.90	.67/ .63	.87	.77/ .72
Need for contact	10	.87	.77/ .72	.86	.81/ .75
Leisure contact	12	.91	.63/ .60	.89	.70/ .66
Self-disclosure	10	.90	.67/ .63	.88	.74/ .69
Trust	10	.88	.74/ .69	.85	.84/ .77
Friendliness	12	.91	.63/ .60	.89	.70/ .66
Attentiveness	19	.94	.51/ .49	.93	.55/ .53
Energy	12	.89	.70/ .66	.86	.81/ .75
Personal growth	10	.82	.94/ .85	.78	1.06/ .94
Perseverance	13	.88	.74/ .69	.86	.81/ .75
Adaptability	7	.86	.81/ .75	.86	.81/ .75
Originality	10	.93	.55/ .53	.92	.59/ .57
Independence	11	.87	.77/ .72	.83	.91/ .82
Orderliness	10	.90	.67/ .63	.88	.74/ .69
Precision	10	.91	.63/ .60	.89	.70/ .66
Regularity	10	.92	.59/ .57	.89	.70/ .66
Conformity	12	.91	.63/ .60	.89	.70/ .66
Decisiveness	11	.90	.67/ .63	.89	.70/ .66
Self-confidence	12	.92	.59/ .57	.89	.70/ .66
Optimism	14	.90	.67/ .63	.87	.77/ .72
Frustration-tolerance	11	.90	.67/ .63	.89	.70/ .66
Resilience	11	.91	.63/ .60	.90	.67/ .63
Mean		.90	.68/ .64	.88	.75/ .70

* Internal consistency

** Standard error of measurement

*** Standard error of estimate

4.2 Stability of the scales and factors over time

A retest was performed in order to assess the stability of the scales and factors over time. In total 53 persons took the retest. This group of respondents consisted of 13 women (24.5%) and 40 men (75.5%). The average age was 36.25 years old, ranging from 23 years to 59 years old (for one of the respondents, the age was not known). The questionnaire was first taken in February of 2007 and 9 months later. This was done in an advice context.

The mean reliability (Cronbach's alpha) of the scales in the test-retest study was 0.91 in the first administration and 0.91 in the second administration as well. The reliability of the factors (stratified alpha) was 0.96 in both the first and second administration of the questionnaire. In calculating the reliability of the factors it was assumed that the total error variance was the same as for the Advice norm group. The variance on the factor level in both groups is roughly the same as in the norm group.

Correlations between the scale scores of the first administration and the second administration were calculated to assess the stability of the scales. These are represented in Table 4.3.

Table 4.3. Correlations with the second administration (N=53)

Factor	Scale	Mean stenscore first administration*	Mean stenscore second administration*	Correlation*	t-value**	p-value**
Influence	Status	5.0 (2.1)	5.1 (2.2)	.85		
	Dominance	4.9 (2.0)	5.1 (2.0)	.87		
	Competition	5.9 (2.3)	6.2 (2.1)	.87		
	Self presentation	4.8 (1.8)	4.9 (1.8)	.78		
Sociability	Need for contact	4.5 (2.1)	4.5 (2.0)	.82		
	Leisure contact	4.4 (2.1)	4.4 (2.0)	.88		
	Self-disclosure	5.2 (2.4)	5.0 (2.1)	.88		
	Trust	6.2 (2.2)	5.9 (2.4)	.78		
	Friendliness	5.0 (2.2)	4.7 (2.0)	.89		
	Attentiveness	4.9 (1.9)	4.4 (1.7)	.78	-2.96	.005
Exuberance	Energy	4.8 (1.8)	4.7 (1.9)	.77		
	Personal growth	5.1 (2.3)	5.0 (2.2)	.82		
	Perseverance	5.4 (2.0)	5.2 (1.5)	.65		
	Adaptability	5.3 (1.9)	5.1 (1.8)	.79		
	Originality	5.1 (2.2)	4.7 (2.4)	.89	-2.17	.034
	Independence	6.2 (1.6)	6.4 (1.6)	.64		
Structure	Orderliness	4.9 (2.2)	4.9 (2.0)	.88		
	Precision	5.5 (2.1)	5.0 (1.9)	.87	-3.63	.001
	Regularity	5.5 (1.9)	5.5 (1.9)	.88		
	Conformity	5.5 (1.7)	5.1 (2.0)	.85		
	Decisiveness	4.8 (1.9)	4.7 (2.1)	.82		
Stability	Self-confidence	5.5 (1.9)	5.2 (1.9)	.86		
	Optimism	5.6 (1.9)	5.1 (1.9)	.77	-3.31	.002
	Frustration-tolerance	6.0 (2.1)	5.8 (2.2)	.87		
	Resilience	5.7 (2.0)	5.6 (2.0)	.78		

* All correlation were tested at the .01 level

** Only shown for significant differences

The correlations between the factors and the average standardized factor scores in the first and second administration is represented in Table 4.4.

Table 4.4. Mean standardized factor scores and corresponding t- and p-values (N=53)

Factor	Mean stenscore first administration*	Mean stenscore second administration*	Correlation**	t-value	p-value
Influence	5.0 (2.0)	5.2 (2.0)	.90	1.81	.08
Sociability	4.8 (2.3)	4.5 (2.1)	.87	-1.95	.06
Exuberance	5.3 (2.0)	5.1 (2.0)	.79	-0.99	.33
Structure	5.0 (2.3)	4.7 (2.3)	.93	-2.73	.01
Stability	5.8 (2.1)	5.4 (2.3)	.87	-2.78	.01
Mean	5.2 (2.1)	5.0 (2.1)	.87		

* The corresponding standard deviations are presented in brackets

** All correlations were tested at the .01 level

T-tests were used in order to check whether the means were significantly different from each other. The t-tests showed significant differences for four of the scales (*Attentiveness*, *Precision*, *Optimism* and *Originality*) and two of the factors (*Structure* and *Stability*). The t-values and corresponding p value for the scales for which the first and second administration differed significantly are represented in Table 4.3. The t-values and corresponding p value for all the factors are represented in Table 4.4.

The correlations between the first and second administration are about as high as the reliabilities of the respective scales and factors. The test-retest correlations are in general lower than the reliabilities because there are real changes (in personality) in the research group. The stability of the scores on the WPI with an intervening period of nine months appears to be very high.

5. Validity

The validity of a questionnaire provides an indication of the extent to which a test actually measures the construct that it is intended to measure. In other words: does a personality questionnaire really measure personality? In the case of the WPI: does the questionnaire actually measure the work-related personality of a person? To be more specific, the validity indicates whether each factor or scale measures what it purports to measure.

Three types of validity can be distinguished. The external validity indicates to what extent the results of a test can be generalized. This mainly depends on the quality of the norm groups, which we have discussed in detail in Chapter 3. The internal validity of a questionnaire indicates to what extent a causal relation can be assumed, and whether alternative hypotheses can be dismissed. In personality questionnaires, it is very hard to actually prove this. By describing the different aspects of validity of a questionnaire this can nevertheless be made plausible.

The last type of validity is test validity; this type of validity will predominantly be discussed in this chapter. Test validity entails criterion validity and construct validity. Criterion validity indicates to what extent the results from a questionnaire are congruent with a criterium, for example the results from a different questionnaire. When the criterium is measured at the same time as the questionnaire, we talk about concurrent validity, when the criterium is measured at a future time point we talk about predictive validity.

When thinking about construct validity, researchers think about what constructs contribute to what is measured. Concerning the WPI, we ask ourselves what contributes to someone's work-related personality, and what certainly does not contribute. Constructs that do not contribute to work-related personality are therefore constructs that we do not want to measure. This is called discriminant validity.

The distinction between criterion- and construct validity is not always clear. Criterion validity focuses on what is predicted. In terms of IQ-test, an employer is not interested in whether a candidate can reproduce a certain figure, but in his or her alleged intelligence of which this is an indication. In terms of construct validity one actually is interested in the test itself, and a possible correlation with a different instrument is merely meant as support that the test indeed measures what it is supposed to measure.

5.1 Criterion validity

To support the criterion validity (split up into concurrent and predictive validity) three studies have taken place that will be discussed below. For all tests a significance level of 5% was used.

5.1.1. Concurrent validity: Study with the FFPI

As mentioned before, the concurrent validity can be indicated by concurrently administering the WPI and an instrument that is intended to measure the same construct(s), in this case the FFPI.

A large study was conducted at an ICT company in which approximately 2700 employees were asked to take tests and questionnaires. The goal of this study was to gather data for a norm group for capacity tests that the company was going to use. The WPI and the FFPI were also offered as a combined questionnaire. To exclude possible effects of the order in which the questionnaires were taken, this order was determined at random. To be sure that the employees filled out the questionnaire in a serious way, they could win prizes if they participated in the study. In total, 354 employees filled out the WPI and the FFPI.

In this study, the earlier version of the WPI, the ProSiD-PI 25, was used. In comparison with the WPI, some of the scales and factors have different names and there are four items that are not included in this version⁴. Partly on the basis of this study we have chosen to alter the names of some scales and

⁴ The problem of the four missing items was solved by repeating the standardization process for the scales and factors to which these items belonged, using the same norm population on which the official norms were based. The correlations between the scales and factors were .97 and .99, which means that the differences were negligible.

factors. In the presentation of the results we use the division into scales and factors and the naming of the WPI.

The FFPI

The Five Factor Personality Inventory (Hendriks, Hofstee & de Raad, 1999) is a personality questionnaire which can determine someone's scores on five broad dimensions. The dimensions of the FFPI are: Extraversion, Agreeableness, Conscientiousness, Emotional Stability and (intellectual) Autonomy. The FFPI consists of 100 short concretely formulated behavioural items.

The FFPI stems from the lexical approach to personality, where the goal is to come to a parsimonious but as complete as possible model for describing the most important differences in behaviour. The Abridged Big-Five dimensional Circumplex-model (AB5C-model) (Hofstee & de Raad, 1991; Hofstee, de Raad & Goldberg, 1992) was the starting point of the development of the FFPI. This AB5C-model is a combination of the previously discussed FFM and the interpersonal circumplex model, in which variables are arranged circularly according to their loadings on two orthogonal factors (Hendriks, Hofstee & de Raad, 1999).

The interpersonal circumplex model was introduced by Leary and his employees in the early 1950's. To most psychologist, this model will be known as the Interpersonal Checklist (ICL) by LaForge and Suczek (LaForge & Suczek, 1955), which is often called 'Leary's rose'. The model is mostly represented as a circular arrangement of sixteen categories of interpersonal behaviour, positioned relative to two orthogonal axes. This circular arrangement is more than a convenient presentation. Behind this simplicity lies a substantive and psychometrically interesting regularity. The sixteen categories, often combined per two creating eight octants, are not arbitrarily chosen clusters. They are categories of behaviour that were found, first by intuition and later in numerous empirical studies by several different researchers, to be important aspects of interpersonal behaviour. The nature, the number of categories and the position relative to the main axes were confirmed. All this led to the conclusion that the circumplex structure was not just a coincidental configuration. From the start on the circumplex model has served as a theoretical framework for a great number of studies, including the FFPI.

Expectations

Given that the FFPI is a well-constructed and validated personality questionnaire and that one of the underlying theories of the FFPI is the FFM, we believe the FFPI to be a good instrument to study the criterium validity of the WPI. We investigated whether the factors of the WPI correlate highly with the corresponding factors from the FFPI. On the scale level, we expect a high correlation with those factors of the FFPI of which the description are most similar to the meaning of the scales of the WPI.

Description of the psychometric qualities and results

The sample consisted of 353 people, 55 females (15.6%) and 298 males (84.4%). The average age was 34.5 years old, ranging from 21 to 61 (the age of 8 respondents was unknown).

Both the correlations between the factors of the FFPI and the factors of the WPI as well as the correlations between the factors of the FFPI and the scales of the WPI were calculated. When comparing a questionnaire with a criterium, a correlation of 0.2 is considered to be low; a correlation of 0.3 is considered to be moderate/average; and a correlation of 0.5 is considered to be high (Cohen, 1992). An attenuation correction is shown for factor correlations above .40. Attenuation is the phenomenon that the correlation between two variables decreases when the reliability of the variables is lower. This means that an estimate is given of the correlation in the hypothetical case of no attenuation. These correlations were based on the standardized scores. It is not possible to investigate the correlations between the scales of the FFPI and the scales of the WPI because the FFPI does not work with scales in the way the WPI does. However, we do know which concepts fall under the factors of the FFPI, see Appendix 9. In this way, we can assert whether the scales of the WPI correlate with the factors of the FFPI in an interpretable way.

The reliability of the factors of FFPI in this sample range from 0.80 (Cronbach's alpha, *Autonomy*) and .87 (*Extraversion*). The reliability of the scales of the WPI in this sample ranged from .82 to .94. The

reliability (stratified alpha) of the factors of the WPI ranged between .95 and .97. In calculating the reliability of the factors we have assumed that the total error variance is the same as in the Advice group which was used for standardization.

Table 5.1 shows the means of the standardized factor scores and their corresponding standard deviations. The distributions are quite similar to the distributions in the norm group.

Table 5.1. Mean stenscore and standard deviation for each factor

Factor	Mean	Standard deviation
Influence	5.8	2.1
Sociability	4.7	2.1
Exuberance	5.4	1.8
Structure	4.8	2.4
Stability	6.2	2.0

From Table 5.1 it becomes clear that each factor of the WPI corresponds highly with one factor of the FFPI, except for the factor *Influence*; this factor correlates highly with three factors of the WPI.

Table 5.2. Correlations between the factors of the FFPI and the WPI

WPI (N=353)	FFPI (N= 353)				
	Extraversion	Agreeableness	Conscientiousness	Stability	Autonomy
Stability	.27**	.11*	.13*	.69**(.85)	.19**
Structure	-.24**	.35**	.76**(.1.00)	.10	-.17**
Influence	.57**(.69)	-.46**(-.58)	.01	.07	.40**(.51)
Sociability	.77**(.91)	.09	-.04	.07	.19**
Exuberance	.26**	.09	.15**	.20**	.56**(.72)

Correlation larger than .30 are in bold, for these correlations the attenuation correction is in brackets.

** Significant at the .01 level (2- sided)

* Significant at the .05 level (2- sided)

With correlations of these magnitudes we can conclude that the factors are more or less tapping the same constructs and thus have the same meaning. To show this, we have undertaken attenuation corrections. When calculating the correlation between two instruments one has to take the unreliability of the two separate instruments into account. Attenuation correction is a way to correct for this.

- The factor *Stability* had a correlation of .69 with *Emotional Stability* of the FFPI, after attenuation correction this correlation was .85. These factors both include related scales, they are both intended to measure the emotional stability of a person.

- The factor *Structure* had a correlation of .76 with *Conscientiousness* of the FFPI, after attenuation correction this correlation was 0.997. The factor *Structure* included the scale *Orderliness* and related scales, while under the FFPI factor *Conscientiousness* concepts fall that are related to structure.

- The factor *Influence* had a correlation of .40 with *Autonomy* of the FFPI, after attenuation correction this correlation was .51. Both factors include similar scales that focus on the 'self', for example *Self representation* and *Dominance*. *Influence* also correlates with *Extraversion* (.57, after attenuation correction .69). *Extraversion* is intended to measure the degree of talkativeness, a concept that can also be found in the scales of *Self representation* and *Dominance*. Finally, *Influence* had a negative correlation of -.46 (after attenuation correction -.58) with *Agreeableness* of the FFPI. This can be explained by the fact that *Agreeableness* is directly opposed to scales such as *Competition*, *Dominance* and *Self presentation*.

- The factor *Sociability* had a correlation of .77 with *Extraversion* of the FFPI, after attenuation correction .91. The name of the WPI factor *Extraversion* has been changed into *Sociability*. Because the definition of this factor has remained the same, it is not surprising that this factor correlates highly with *Extraversion* of the FFPI.

- The factor *Exuberance* had a correlation of .56, after attenuation correction .72, with the FFPI factor *Autonomy*. Both factors include aspects of independency. This factor also has a fairly high with three other factors of the FFPI. On the scale level, we will show that this factor correlates in an interpretable way with the FFPI.

It is striking that when we look at all the correlations of the factors, four factors of the WPI correlate with one of the factors of the FFPI. Only for the factor *Influence* of the WPI, there are more correlations, however, these are – as described above – explainable. We have to note that the reverse

is not always the case. The factors of the FFPI that correlate highly with the WPI, correlate highly with other factors of the WPI as well. This will lead to correlations between the factors of the FFPI and the scales of the WPI, mostly in an explainable way.

An overview of all the correlations between the factors of the FFPI and the scales of the WPI is presented in Appendix 10. The most important and most notable results are presented in Table 5.3, for every factor of the FFPI. All correlations larger than .50 with a significance level smaller than 0.01 are presented in the first column with the corresponding correlations, where the highest correlations are presented first. In the second and third column, all other significant correlations are shown, in the order of the magnitude of the correlation. We have included descriptions of the FFPI factors in Appendix 9.

Table 5.3. Correlations between scales of the WPI and factors of the FFPI (N=353)

FFPI factor	Significant correlations with scales of the WPI >.50.	Significant* positive correlations with scales of the WPI up to .50.	Significant* negative correlations of the scales of the WPI up to .50.
Extraversion	Leisure contact (.75) Need for contact (.72) Friendliness (.67) Self-presentation (.68) Dominance (.54) Self-disclosure (.53)	Energy (.45) Attentiveness (.39) Status (.37) Self-confidence (.34) Adaptability (.33) Optimism (.31) Trust (.30) Resilience (.20) Competition (.14) Originality (.14)	Regularity (-.30) Decisiveness (-.27) Precision (-.17)
Agreeableness	Self-presentation (-.52)	Attentiveness (.41) Conformity (.38) Frustration-tolerance (.32) Decisiveness (.30) Precision (.29) Perseverance (.23) Personal growth (.16) Optimism (.15) Trust (.14)	Dominance (-.31) Competition (-.30) Status (-.26) Leisure contact (-.19) Self-confidence (-.17)
Conscientiousness	Orderliness (.66) Perseverance (.60) Precision (.55) Decisiveness (.50)	Regularity (.49) Conformity (.41) Energy (.24) Optimism (.14)	Originality (-.19) Adaptability (-.14)
Stability	Frustration-tolerance (.64) Resilience (.52) Optimism (.51)	Self-confidence (.47) Adaptability (.22) Conformity (.21) Friendliness (.19) Decisiveness (.17) Trust (.17) Energy (.17) Leisure contact (.17) Perseverance (.15)	Attentiveness (-.16)
Autonomy	Dominance (.63)	Originality (.50) Adaptability (.42) Self-confidence (.40) Energy (.38) Independence (.35) Personal growth (.33) Leisure contact (.30) Perseverance (.27) Attentiveness (.26) Self-presentation (.22) Status (.19) Competition (.18) Optimism (.16) Resilience (.16)	Regularity (-.43) Conformity (-.36)

*Significant at the .01 level

The scales of the WPI all correlate in an explainable way with the factors of the FFPI. For all the correlations it holds that roughly the same constructs are described by the scales and factors. The positive correlation between *Conformity* with the factor *Conscientiousness* of the FFPI can be explained by the fact that this factor is defined as “orderliness”. An interesting finding is that *Attentiveness* has a negative correlation with the *Stability* factor of the FFPI. This could mean that less stable people are more inclined to take care of others. Almost all scales show a significant correlation with the factors of the FFPI. The scales *Dominance*, *Self-presentation*, *Need for contact*, *Leisure contact*, *Self-disclosure*, *Friendliness*, *Perseverance*, *Orderliness*, *Precision*, *Decisiveness*, *Optimism*, *Frustration tolerance* and *Resilience* even have correlations higher than .50 with factors of the FFPI. Hereby not only the validity of the factors but of the scales as well, is shown.

Conclusions

The Five Factor Model is the underlying theory of the FFPI as well as the WPI. From the described results on the relation between the factors of the FFPI with the factors and scales of the WPI we have gained insights in the psychological constructs that are measured by the questionnaires. The results closely match our expectations. Each factor of the WPI correlates highly with a theoretically similar factor of the FFPI, except for *Influence*, that correlated highly with three of the FFPI factors. The scales of the WPI correlate highly and in an interpretable way with the factors of the FFPI as well.

Given that the FFPI is a well-constructed and validated personality questionnaire, this study has been a good contribution to the validity of both the FFPI and the factors and the majority of the scales of the WPI.

5.1.2. Predictive validity

To assert the predictive validity of the WPI we have conducted two studies in which certain employment criteria were measured some time after completion of the WPI. Subsequently, we have tested the predictive value of the WPI. Again, a significance level of 5% was used.

5.1.2.1. Predictive validity: nurse study with 360 degrees feedback reviews

Ixly has conducted a study in the period between September 2005 and December 2008 on student-nurses following the BBL version of nursing school. The goal of the research was: to assess the predictive value of several personal characteristics of student-nurses for the degree of success in their education. We chose to only include the students that were enrolled in the education program with an apprenticeship agreement. This was done because the curriculum of these forms of education resembles the practical job situations more than the curriculum of the more theoretical forms of education. Moreover, in practical situations – more so than in a more theoretical context – a wider range of personal characteristics is required in order to function effectively.

The personality characteristics were measured in the beginning of the educational program by means of the WPI. The degree of success in the education was measured by means of 360 degrees reviews. The content of the 360 degrees feedback questionnaire was created by an experienced psychologist and adapted to the practical situation in consultation with the participating hospitals. The students took the 360 degrees questionnaire themselves and at least two up to ten persons that had worked with the student were invited to complete the questionnaire about the student. The people that reviewed the student were classified into five categories: managers, work supervisors, trainers/teachers, certified nurses and third or fourth year student nurses.

The student was rated on fourteen competencies, each of them measured by five behavioural indicators. During the study, it became clear that administrating the questionnaire took quite a lot of time, for the trainers as well as the students. Because of this, not all the participating students completed a 360 degrees questionnaire during the education program. In total, 80 students completed 360 degrees questionnaires that could be analysed.

Sample

The total research sample included 204 student-nurses, enrolled in the nurse education program, either on secondary level (MBO) or a higher level (HBO) (35 were male, 160 female, for 8 students no information on gender was provided). When we started the analyses, 80 students had completed the 360 degrees questionnaire (12 male and 67 female, for 1 student information on gender was not provided). The average age of this sample was 24.6 years old (ranging from 18 to 44 years). Of these

students, 58 were enrolled in the program on secondary level and 14 on the higher level (for 8 of the students, the level was unknown).

Expectations

Since the study was exploratory in nature, no specific hypotheses were formulated. According to Bloemers and van der Molen, the predictive value that can be found for instruments, lies mostly between the 0.20 and 0.50 (Bloemers and van der Molen, 2004). Research conducted by Schmidt and Hunter shows that for personality instruments, relationships in the order of 0.31 are found (Schmidt and Hunter, 1998). The general expectation is therefore that a number of personality traits will show a correlation of about 0.30 with almost all competencies.

Results

The reliabilities (Cronbach's alpha) of the scales of the WPI in this sample were between 0.79 (*Independence*) and 0.94 (*Attentiveness*). The reliability of the factors were between 0.91 (*Influence*) and 0.96 (*Exuberance*).

First, the factor scores of this sample were compared with the factor scores of the reference population. With a t-test, we checked whether there were significant differences in the factor scores between the sample and reference population. Analyses showed that the factor scores of *Influence* and *Exuberance* differed significantly. The student-nurses score significantly lower on these factors than the reference population (see Table 5.4). In this light, we can conclude that student-nurses are less ambitious and less driven than the reference population.

Table 5.4. Mean stenscore in this sample in comparison with the sten scores of the norm group (N=80)

Factor	Mean stenscore*	t-value	p-value
Influence	5.4 (1.5)	2.59	.01
Sociability	6.2 (2.0)	-0.89	.37
Exuberance	5.3 (1.9)	2.85	.00
Structure	5.7 (2.6)	-0.11	.92
Stability	5.8 (1.8)	1.62	.21

* The corresponding standard deviations are in brackets

We chose to calculate, for every student, the average rating for every behavioural indicator of a competency. In this calculation, only the ratings of the respondents were used, not the self-rating of the student. The reliability of the competencies ranged from 0.89 (*Relational qualities*) to 0.95 (*Adaptability*). The magnitude of these reliabilities is partly due to the average rating that was used per behavioural indicator.

A factor analysis (PCA, varimax rotation) was conducted on the average ratings per behavioural indicator. The results showed that the first factor explained 65% of the variance. Therefore, we have initially created a total score, based on all competencies. This total score was used to calculate the correlations with the standardized scale and factor scores of the WPI. The scales *Self-disclosure* ($r = 0.38, p < 0.01$) and *Optimism* ($r = 0.37, p < 0.01$) were significantly correlated with the total score of the 360 degrees feedback questionnaire. In addition, the factor *Sociability* showed a positive correlation with the 360 degrees feedback questionnaire ($r = 0.28, p < 0.01$).

In addition to this analysis, correlations were calculated between the ratings on the fourteen separate competencies and the factor scores of the WPI (Table 5.5). Six significant relations were found. Five of these relations included the factor *Sociability* of the WPI. The competencies *Adaptability*, *Ability to influence*, *Emotional stability*, *Drive* and *Relational qualities* all correlated significantly with the factor *Sociability*. *Quality of work* showed a relation with the scales of the factor *Stability*.

Table 5.5. Correlations competencies 360test and standardized factor scores WPI (N=53)

Competencies 360	Factors WPI				
	Influence	Sociability	Exuberance	Structure	Stability
Adaptability	.04	.30*	.06	.03	.04
Analytical qualities	.07	.27	.11	-.003	.19
Ability to influence	.04	.33*	.14	.05	.22
Basic communicative skills	.09	.26	.07	.02	.02
Creativity/innovation	.04	.26	.06	-.07	.08
Discipline	.03	.17	.10	.03	.14
Emotional stability	.11	.31*	.14	.06	.26
Exuberance	.06	.28*	.26	.19	.26
Knowledge	-.02	.17	.08	.02	.11
Quality of work	-.004	.16	.13	.05	.18
Quantity of work	.02	.24	.21	.24	.32*
Organizational qualities	.05	.24	.11	.04	.24
Relational qualities	.02	.30*	.08	.14	.15
Nursing skills	.08	.20	.07	.02	.12

* Significant at the .05 level (2- sided)

Conclusions

Based on the results of the analyses we can conclude that there is an association between the competencies as measured by 360 degrees method and the personality traits as measured by the WPI. The sizes of the associations are of the magnitude that may be expected on the basis of other studies (see Schmidt and Hunter, 1998). This result adds to the criterion validity of the WPI.

Five of the fourteen competencies necessary for the work of a nurse (i.e.: Adaptability, Ability to influence, Emotional stability, Exuberance and Relational qualities) all show a significant correlation with the factor *Sociability*. This result is explainable, given that the factor *Sociability* included the scales *Need for contact*, *Leisure contact*, *Friendliness* and *Attentiveness*. These are all personality traits that may be helpful as a nurse. From this we can conclude that when student-nurses score higher on the factor *Sociability*, that they can be expected to be successful in their job in the hospital.

In addition to the score on *Sociability*, it appears that the score on *Stability* is important in the selection of student nurses as well. The competency Quantity of work shows a significant correlation with the factor *Stability*. Noteworthy is the fact that the competency Emotional stability does not correlate significantly with the factor *Stability*. A possible explanation for this is that the lack of stability for student nurses is most notable in the quantity of work and less notable for the environment by means of the behavioural indicators that measure Emotional stability. Finally, Ability to influence, Emotional stability, Exuberance and Organizational qualities show relatively high correlations with the factor *Stability*.

Suggestions for future research

For the generalizability of the results, future research should formulate hypotheses in advance that can be tested by the analyses. The hypotheses may be derived from the results as found in this study. The hypotheses would then be:

- *Sociability* is an important factor for the successful functioning of (student)nurses
- *Stability* is an important factor for the quantity of work and is reported by the individual rather than that it is perceivable by the individual's environment (co-workers).

5.1.2.2 Predictive validity: performance study within a competitive employment agency

In 2010, for a period of 6 months, a study has been performed at a competitive employment agency (hereafter called Agency X) on the relationship between personality, job satisfaction and job performance.

Method

The full employee base of Agency X (284 employees) was invited to take the WPI; an employee satisfaction survey and the Career values questionnaire via e-mail. Performance measures, i.e. individual turnover and supervisor rating, on all persons were available. First, the mean scale and factor scores for different job positions were investigated. Subsequently, using linear regression, we have tested several hypotheses in order to assess the predictive value of the WPI on three job criteria.

Sample

189 employees have gone through the entire research process. Of these, 105 were female, 84 were male. 78 employees had a contract for a definite period, 106 employees for an indefinite period and for 5 employees this was unknown. At Agency X, two general job positions could be distinguished: 54 recruiters and 73 account managers; the other 62 employees were either staff or support staff. Table 5.6 describes the sample.

Table 5.6. Description of the sample of Agency X

189 employees		
Mean age 24,5 year		
84 Male		105 Female
54 Recruiters	73 Account managers	62 Other functions
78 Definite contract		106 Indefinite contract

Expectations

Hypotheses on the relationships between personality, performance and job satisfaction were formulated prior to the analyses. Hypotheses were formulated on the basis of the expectations of the HR consultants at Agency X.

1. The WPI scales *Leisure contact* and *Need for contact* have a positive relation performance.
2. The WPI scales *Decisiveness* and *Regularity* have a negative relationship with performance.

Results

The mean factor and scale scores for both function groups are presented in Table 5.7. Remarkably high scores for both function groups are the scores on the scale *Competition* and the factor *Influence*. Remarkably low scores for both groups are the scores on the scale *Decisiveness*. Differences in mean scale scores are most apparent for the scales *Independence*, *Self-confidence*, *Structure* and *Originality*.

Table 5.7. Mean stenscores of the factors and scales for both function groups of Bureau X (norm group used: Advice)

	Recruiter (N = 54)		Account manager (N = 73)	
	M	SD	M	SD
Influence	7.06	1.58	7.96	1.60
Sociability	5.54	1.62	5.63	1.57
Exuberance	5.26	1.55	5.91	1.45
Structure	3.81	1.71	3.53	1.81
Stability	5.60	1.62	6.00	1.56
	M	SD	M	SD
Adaptability	5.08	1.43	5.80	1.35
Independence	5.60	1.46	6.53	1.35
Perseverance	5.15	1.50	5.05	1.42
Competition	8.16	1.57	8.80	1.57
Energy	5.61	1.44	5.88	1.45
Frustration-tolerance	5.20	1.80	5.51	1.78
Resilience	6.04	1.50	6.24	1.63
Attentiveness	4.66	1.57	4.39	1.54
Dominance	5.88	1.36	6.44	1.31
Status	5.99	1.44	6.67	1.57
Precision	4.37	1.57	4.26	1.21
Conformity	5.30	1.42	4.88	1.71
Self-disclosure	5.50	1.67	5.31	1.71
Orderliness	4.24	1.70	4.33	1.83
Originality	4.97	1.41	5.75	1.55
Decisiveness	4.07	1.70	4.00	1.60
Optimism	5.56	1.60	5.38	1.33
Personal growth	5.45	1.72	5.57	1.91
Regularity	4.80	1.85	4.47	1.46
Need for contact	6.27	1.76	6.34	1.33
Leisure contact	6.19	1.49	6.76	1.65
Self-presentation	6.66	1.84	7.43	1.50
Trust	4.80	1.92	4.85	1.92
Friendliness	5.81	1.31	5.88	1.38
Self-confidence	5.59	1.34	6.42	1.47

Linear regression

Based on the average scores on the scales and factors in Table 5.7, a profile of the average scores of the persons within a function group can be made. However, this does not tell us anything about the predictability of the performance and satisfaction on the basis of the WPI. For this, linear regression is used.

Based on the assumptions that underlie regression analyses, we have performed several explorative analyses on the dependent variables (y^1) *Turnover*, (y^2) *Supervisor rating* and (y^3) *Job satisfaction*. The assumptions underlying regression analyses are:

1. Linear relationship between X and Y
2. All pairs of observations (X, Y) are independent of each other
3. The residuals are normally distributed

4. The variances of the residuals are equal, independent of X (Homoscedasticity)

In the following tables, B denotes the regression coefficient, or in other words the extent to which the scale contributes to the dependent variable.

Table 5.8. Linear regression (x) WPI scales and (y¹) turnover per function for Agency X (Norm group Advice)*

Recruiter**					
	B	SE	Beta	t	Sig.
(Constant)	24.472	7.786		3.054	.004
Decisiveness	3.611	1.319	0.336	2.620	.012
Orderliness	-3.278	1.304	-0.382	-2.983	.005
Account manager					
	B	SE	Beta	t	Sig.
<i>No variables included according to stepwise inclusion of variables.</i>					

* Stepwise inclusion of (x)

** R² = .247

Table 5.9. Linear regression (x) WPI scales and (y²) supervisor rating per function for Agency X (Norm group Advice)*

Recruiter**					
	B	SE	Beta	t	Sig.
(Constant)	0.481	0.208		2.317	.031
Adaptability	0.151	0.042	0.862	3.561	.002
Competition	-0.091	0.035	-0.623	-2.575	.018
Account manager***					
	B	SE	Beta	t	Sig.
(Constant)	0.177	0.116		1.53	.135
Decisiveness	0.075	0.028	0.405	2.659	.012

* Stepwise inclusion of (x)

** R² = .388

*** R² = .164

Table 5.10. Linear regression (x) WPI scales and (y³) satisfaction per function for Bureau X (Norm group Advice)*

Recruiter**					
	B	SE	Beta	t	Sig.
(Constant)	58.362	4.650		12.552	.000
Adaptability	2.104	0.716	0.388	2.937	.005
Independence	-1.674	0.616	-0.315	-2.718	.009
Resilience	1.510	0.639	0.292	2.364	.022
Need for contact	1.149	0.508	0.26	2.259	.028
Account manager***					
	B	SE	Beta	t	Sig.
(Constant)	49.643	4.705		10.552	.000
Optimism	4.428	0.863	0.534	5.132	.000
Conformity	1.679	0.637	0.261	2.636	.010
Orderliness	-1.344	5.98	-0.223	-2.249	.028

* Stepwise inclusion of (x)

** R² = .440

*** R² = .398

Conclusions

Results from the study at Agency X show that a clear personality profile can be distilled of the two main functions within the organization (Table 5.8). The high average competitiveness of the employees is a striking but expected trend within this organization. Following the same reasoning, the average low score on *Decisiveness* is a trend that shows that employees within this competitive organization are able (or should be able to) switch quickly and that this way of working fits with the current employee base.

In terms of the predictive value of the results of the WPI on the measures of performance a couple of scales stood out. Quite remarkable was that the score of a recruiter on the scale *Competition* had a negative effect on the supervisor's rating of performance. From conversations with the HR consultants of Agency X it became clear that while a competitive attitude during the selection process seemed to be an advantage, this was not always seen as a positive attitude during working in teams.

Remarkable is the positive effect of *Decisiveness* as well, while the average employee of Agency X scored below average on this scale. A different important aspect of the functioning is the satisfaction of the employee him/herself. From our previous analyses, it became clear that for both functions different scales contributed to this satisfaction. For recruiters, this was mainly the case for *Adaptability* while for account managers this was mainly the case for *Optimism*.

The goal of this study was to assess the predictive value of the WPI by means of proposed hypotheses. It appears that several scales of the WPI are predictive of job criteria. This contributes greatly to the predictive validity of the WPI.

5.2 Construct validity

5.2.1. Construct validity: Factor and scale structure

The most important indication of the construct validity of the WPI are the results from the MGM (Multiple Group Method) which we have discussed earlier in the paragraph 'Structure of the WPI'. For both the Advice and Selection group an MGM analysis was performed. The background characteristics of these groups are described in the chapter Norms (paragraph 3.1). The analysis was performed on the data from the weighed norm groups.

When an MGM analysis is applied to the both weighed groups, the structure does not appear to differ significantly from the structure that was found in the construction phase of the questionnaire (on the unweighed groups). This holds for the scales as well as for the factors. The results of the MGM analyses on the division of items in scales are included in Appendix 1. The results of the MGM analyses on the division of scales into factors for both norm groups are presented in the following two tables. Notable results and differences between the groups will be discussed.

The presentation of the results of the MGM analyses (Table 5.1 and Table 5.2) is made easier by the use of colours. Three colours are used, each with their own meaning. Grey means that the item/scale correlates the highest with its own scale/factor (item-rest correlation) and that the difference in height with the other correlations is significant (as calculated with the t-test for differences between dependent correlations (Steiger (1980))). Yellow means that there are more correlations of the item/scale that correlate with other scales/factors (respectively), but not in a significant way. The colour red means that the item/scale has a higher correlation with a different scale/factor than with its own scale/factor and that the difference in the height of the correlation is also significant. In the presented tables there are no red cells.

In the MGM on the division of scales into factors, there are 20 scales that correlate significantly the highest with their own scale (grey) for the Advice group, and 16 scales for the Selection group (see Tables 5.11 and 5.12). It appears that no other scale correlates higher with a different factor than that it was theoretically assigned to (red). In the Advice group, there are five scales that correlate in a meaningful way with more than one factor but without a significance difference, for the Selection group this was the case for nine scales. Three of these scales overlap between the Advice and Selection group.

Information on the results of the MGM analysis based on the division of items in scales is reported in Appendix 1. For both norm groups, there is no item that has a significantly higher correlation with a scale other than the scale it was assigned to. In the Advice group, 18 of the 276 items correlated with scales other than its own scale but without a significant difference in correlation heights (yellow). In the Selection group, 44 of the 276 items were in yellow cells. All other items correlated significantly higher with their own scale than with other scales.

When both groups are compared, it appears that 17 items correlate with more than one scale – but not significantly (yellow) – in both groups. One item appears to show a yellow cell in the Selection group, but not in the Advice group. There are 27 items in the Selection group that correlate with more than one scale – but not significantly (yellow) – while these items were in grey cells in the Advice group. The fact that there are fewer items in the Selection group that differ significantly from each other in terms of their correlations is due to smaller number of candidates in this group. Because the research population is smaller, significant results are less likely to occur.

Because the MGM results show relatively few differences, and because no item or scale correlates significantly with a different scale or factor, respectively, than their own scale or factor, these analyses can be viewed as solid support for the stability of the structure of the WPI. The division of items into scales and scales into factors is supported empirically.

MGM in an independent group

Since the Advice and Selection group were not only used to test the structure of the WPI, but were also used as input for the development of the questionnaire, an additional MGM analysis was run on an independent group. Background characteristics of this group can be found in the paragraph 'Study with the FFPI'. The results of this group are compared with the advice norm group, because the WPI was administered in an advice setting. It appeared that the results MGM of the independent group only differed slightly from the results of the Advice group. The MGM results for the analyses on both the scales as on the factors are presented in Appendix 1. Notable results and differences between the independent group and the Advice group (the MGM from the previous paragraph) will be discussed.

Table 5.11. MGM results for the Advice group (N=712)

Factor	Scale	Exuberance	Stability	Influence	Sociability	Structure
Exuberance	Adaptability	.69	.59	.51	.48	-.03
	Perseverance	.54 ^a	.41	.27	.32	.51 ^a
	Energy	.71	.53	.47	.49	.27
	Independence	.37 ^a	.12	.34 ^a	.08	-.10
	Originality	.57	.42	.48	.37	-.03
	Personal growth	.63	.35	.48	.34	.16
Stability	Frustration-tolerance	.29	.59	.10	.33	.24
	Resilience	.40	.68	.29	.25	.06
	Optimism	.51	.63	.33	.52	.12
	Self-confidence	.60 ^a	.64 ^a	.53	.40	.08
Influence	Competition	.39	.16	.61	.09	-.11
	Dominance	.63 ^a	.51	.63 ^a	.48	-.10
	Status	.47	.31	.57	.21	.05
	Self-presentation	.38	.26	.67	.47	-.24
Sociability	Need for contact	.35	.37	.37	.65	.10
	Friendliness	.52	.50	.37	.71	.14
	Leisure contact	.52	.56	.52	.62	.01
	Trust	.14	.21	.10	.48	-.08
	Self-disclosure	.26	.20	.33	.59	-.06
	Attentiveness	.33	.20	.03	.53	.30
Structure	Conformity	.03	.19	-.14	.20	.36
	Precision	.35	.17	-.01	.09	.55
	Orderliness	.33	.24	.06	.19	.50
	Regularity	-.31 ^a	-.32 ^a	-.30 ^a	-.17	.33 ^a
	Decisiveness	.20	.23	-.03	-.02	.45

^a The numbers concern the yellow cells, the other cells are coloured grey.

Table 5.12. MGM results for the Selection group (N=369)

Factor	Scale	Exuberance	Stability	Influence	Sociability	Structure
Exuberance	Adaptability	.72	.60	.38	.54	.17
	Perseverance	.65	.53	.28	.43	.51
	Energy	.71	.64	.39	.59	.41
	Independence	.23 ^a	.03	.20 ^a	.05	-.07
	Originality	.58	.49	.42	.45	.15
	Personal growth	.59	.39	.50	.43	.30
Stability	Frustration-tolerance	.39	.62	.06	.40	.38
	Resilience	.44	.70	.21	.40	.28
	Optimism	.59 ^a	.58 ^a	.31	.62 ^a	.23
	Self-confidence	.57	.64	.43	.45	.32
Influence	Competition	.30	.08	.64	.12	-.10
	Dominance	.58 ^a	.46	.57 ^a	.53 ^a	.01
	Status	.40	.23	.58	.22	.12
	Self-presentation	.32	.21	.66	.42	-.21
Sociability	Need for contact	.44	.45	.39	.66	.17
	Friendliness	.59	.60	.34	.68	.34
	Leisure contact	.52	.62 ^a	.48	.60 ^a	.12
	Trust	.27	.31	.19	.46	.03
	Self-disclosure	.19	.16	.29	.49	-.06
	Attentiveness	.48 ^a	.38	.05	.55 ^a	.40
Structure	Conformity	.29	.35 ^a	-.01	.34 ^a	.42 ^a
	Precision	.36	.32	-.01	.16	.60
	Orderliness	.44 ^a	.43 ^a	.09	.27	.51 ^a
	Regularity	-.19	-.21 ^a	-.23 ^a	-.14	.31 ^a
	Decisiveness	.25	.38 ^a	-.03	.14	.45 ^a

^a The numbers concern the yellow cells, the other cells are coloured grey.

From the results of the MGM analyses on this independent research group on the item level (Appendix 11) it became clear that no single item correlated higher with a different scale other than the scale it was assigned to (red). 241 of the 272 items correlated higher with their own scales than with any other scale. The remaining 31 items correlated with scales other than their own scale, but not in a significant way. After comparison with the Advice group it appeared that 9 of the yellow labelled items received a yellow label in the Advice group as well. We concluded that the chosen makeup of the questionnaire on the item level was upheld in the independent group.

From the results of the MGM analyses on this independent research group on the factor level it appeared that 19 of the 25 scales correlated higher with their own factor than with any other factor. Five scales correlated with other factors, but not in a significant way. There was one scale that correlated significantly higher with a different factor than with its own designated factor. However, this scale did correlate significantly higher with its own factor than with any other factor in the Advice group. In any independent group, especially when it concerns a specific organization as is the case here, there will be differences with the norm group. Since the difference only concerns one scale, it seems that the items and scales of the WPI show a very stable structure.

Correlations between factors

To check whether the factors have a shared component (or components), a correlation analysis was done on the factor level as well as a factor analysis in both research groups (Advice and Selection). The correlations are reported in Table 5.13.

Table 5.13. Correlations between factors

Factor	Correlations Advice group (N=712) / Selection group (N=369)*				
	Influence	Sociability	Exuberance	Structure	Stability
Influence	-				
Sociability	.39 / .44	-			
Exuberance	.59 / .55	.48 / .56	-		
Structure	-.12 / -.06	.10 / .16	.18 / .26	-	
Stability	.39 / .35	.47 / .51	.56 / .60	.16 / .33	-

* Every correlation is presented for the Advice Group First, for the Selection Group behind the / sign

The results of the rotated component matrices of the performed factor analyses are as follows (Table 5.14 and 5.15).

Table 5.14. Rotated component matrix for the Advice group (N=712)

	Component	
	1	2
Influence	.8	-.4
Sociability	.7	.1
Exuberance	.9	.1
Structure	.1	1.0
Stability	.8	.2

Table 5.15. Rotated component matrix for the Selection group (N=369)

	Component	
	1	2
Influence	.8	-.4
Sociability	.8	.2
Exuberance	.8	.3
Structure	.1	.9
Stability	.7	.5

Conclusion

For both the advice and selection group, as well as the independent group, MGM analyses showed that the WPI has a solid internal structure. The majority of the items and scales correlate significantly higher with its "own" designated scales and factors, respectively, than with other scales or factors. A few items and scales show correlations with other scales or factors, but not in a significant way. Furthermore, all factors showed a shared component, with the exception of the factor *Structure*. Future research should shed more light on what this component is. However, it is clear that all factors have a meaning, independent from one another.

5.2.2. Construct validity: relationship with background variables

To assert whether the factor and scale scores have a relationship with the background variables, we have analysed for all the variables, whether the mean scores on each scale and each factor for the different categories of these variables differed significantly from each other. For both the Advice and

Selection group this was done by means of ANOVA (Table 5.16 through 5.19). The data was weighed during these analyses. The weighed N of the Advice group is 712 (unweighed N = 5629), for the Selection group the N is 369 (unweighed N = 1514).

In addition, the eta (see page 18 for an overview of the statistical terms) was calculated. Eta can be used as a measure of association for variables with more than two categories. When the difference is significant in a two sided test on the 5%-level, this is indicated by a * next to the eta of the corresponding factor/scale.

Finally, results of a study are reported in which job sector differences in the mean sten scores of some factors and scales were identified (Table 5.10). The significance of the differences is determined by T tests.

Gender

The mean scale and factor sten scores for males and females are represented in Table 5.16.

Table 5.16. Mean stenscores by gender including eta-value

Factor	Advice group (N=712)			Selection group (N=369)		
	eta*	Male	Female	eta*	Male	Female
Influence	.17*	5.8	5.1	.16*	5.8	5.1
Sociability	.14*	5.2	5.9	.11*	5.3	5.8
Exuberance	.04	5.6	5.4	.03	5.5	5.6
Structure	.09*	5.3	5.7	.03	5.4	5.6
Stability	.08*	5.7	5.3	.10	5.7	5.3
Scale						
Status	.10*	5.7	5.3	.06	5.6	5.4
Dominance	.13*	5.7	5.2	.13*	5.7	5.2
Competition	.23*	5.9	4.9	.22*	5.9	5.0
Self-presentation	.08*	5.6	5.3	.10	5.7	5.3
Need for contact	.08*	5.4	5.7	.07	5.4	5.7
Leisure contact	.01	5.5	5.5	.01	5.5	5.5
Self-disclosure	.08*	5.4	5.7	.02	5.5	5.6
Trust	.10*	5.3	5.7	.02	5.5	5.5
Friendliness	.12*	5.3	5.8	.15*	5.2	5.9
Attentiveness	.23*	5.1	6.0	.22*	5.1	6.0
Energy	.00	5.5	5.5	.05	5.4	5.6
Personal growth	.01	5.5	5.5	.06	5.4	5.7
Perseverance	.06	5.4	5.6	.04	5.4	5.6
Adaptability	.03	5.6	5.4	.03	5.5	5.6
Originality	.13*	5.7	5.2	.12*	5.7	5.2
Independence	.07	5.6	5.3	.04	5.4	5.6
Orderliness	.10*	5.3	5.7	.07	5.4	5.7
Precision	.09*	5.3	5.7	.01	5.5	5.5
Regularity	.08*	5.4	5.7	.08	5.4	5.7
Conformity	.05	5.4	5.6	.06	5.4	5.6
Decisiveness	.02	5.5	5.5	.10	5.7	5.3
Self-confidence	.15*	5.8	5.1	.20*	5.9	5.0
Optimism	.07*	5.4	5.7	.05	5.4	5.6
Frustration-tolerance	.02	5.5	5.5	.07	5.6	5.3
Resilience	.17*	5.8	5.1	.13*	5.7	5.2

* Significant results of the ANOVA are indicated by a * for the eta values.

In the Advice group, there are 4 factors and 16 scales that differ significantly from each other on the gender variable. In the Selection group, this is the case for 2 factors and 7 scales.

For the background variable gender, the found differences were maximally 1.0 sten; this corresponds to a difference of 0.5 times the standard deviation. The mean difference over all significant scales and factors was 0.56 sten; this corresponds to 0.28 times the standard deviation. The differences reported for the gender variable are differences that are also reported in the literature. Examples are women's higher scores on *Sociability*, *Attentiveness* and *Leisure contact* and men's higher scores on *Status*, *Trust* and *Influence* (Beutel & Marini, 1995).

Table 5.17. Mean stenscore by age including eta-value

Factor	Advice group (N=712)			Selection group (N=369)				
	eta*	15-24 years	25-44 years	45-65 years	eta*	15-24 years	25-44 years	45-65 years
Influence	.08	5.5	5.7	5.3	.14*	5.9	5.7	5.1
Sociability	.01	5.5	5.5	5.5	.06	5.8	5.5	5.4
Exuberance	.15*	4.6	5.6	5.6	.08	5.7	5.6	5.3
Structure	.08	5.0	5.5	5.6	.08	6.0	5.5	5.4
Stability	.05	5.2	5.5	5.6	.03	5.7	5.5	5.5
Scale								
Status	.18*	6.1	5.7	5.0	.21*	6.2	5.7	4.9
Dominance	.12*	4.9	5.5	5.7	.07	5.3	5.4	5.7
Competition	.11*	5.4	5.7	5.2	.17*	5.8	5.8	5.0
Self-presentation	.07	5.4	5.6	5.3	.13	5.7	5.7	5.1
Need for contact	.16*	6.4	5.5	5.2	.15*	6.0	5.7	5.1
Leisure contact	.08	5.0	5.5	5.6	.02	5.4	5.5	5.5
Self-disclosure	.07	5.1	5.6	5.5	.06	5.5	5.4	5.7
Trust	.05	5.5	5.4	5.6	.12	5.4	5.3	5.8
Friendliness	.03	5.4	5.6	5.4	.19*	6.0	5.8	5.0
Attentiveness	.05	5.8	5.4	5.5	.07	5.8	5.4	5.6
Energy	.10*	5.0	5.6	5.5	.12	6.0	5.6	5.2
Personal growth	.08	5.4	5.7	5.3	.20*	6.2	5.8	4.9
Perseverance	.16*	4.6	5.5	5.7	.05	5.6	5.6	5.4
Adaptability	.07	5.1	5.5	5.6	.05	5.7	5.6	5.4
Originality	.09*	5.0	5.5	5.7	.04	5.5	5.4	5.6
Independence	.21*	4.3	5.6	5.7	.10	4.9	5.5	5.6
Orderliness	.15*	4.6	5.6	5.7	.03	5.7	5.5	5.5
Precision	.12*	4.8	5.5	5.6	.07	5.8	5.5	5.4
Regularity	.03	5.4	5.5	5.6	.05	5.7	5.4	5.6
Conformity	.15*	6.3	5.5	5.2	.24*	6.4	5.8	4.9
Decisiveness	.13*	4.8	5.5	5.7	.12	5.4	5.3	5.8
Self-confidence	.10*	5.0	5.5	5.7	.02	5.6	5.5	5.5
Optimism	.07	5.1	5.5	5.6	.01	5.6	5.5	5.5
Frustration-tolerance	.03	5.4	5.5	5.6	.04	5.7	5.5	5.5
Resilience	.04	5.6	5.6	5.4	.02	5.6	5.5	5.5

* Significant results of the ANOVA are indicated by a * for the eta values.

Age

The mean scale and sten scores for the different age categories are represented in Table 5.17. In the standardization process, the age variable was divided into three categories (15-24, 25-44 and 45-65 years). For this variable, ANOVA analyses were run as well to check whether differences were significant. Significant differences are marked with a * next to the eta.

In the Advice group, there were 1 factor and 13 scales that differed significantly from each other in terms of age, for the Selection group this was the case for 1 factor and 6 scales. The largest difference in means is 1.5 sten; this corresponds to 0.75 times the standard deviation. The average difference over all the significant scales and factors is 0.98 sten. This corresponds to a difference of 0.49 times the standard deviation. From Table 5.17 it appears that the difference between the two oldest groups (25-44 years and 45-65 years) is generally small. The found significant differences are mainly caused by the mean sten score of the group up to 25 years.

Education

The mean scale and sten scores for the education variable are represented in Table 5.18.

Table 5.18. Mean stenscore by education including eta-values

Factor	Advice group (N=712)				Selection group (N=369)			
	eta*	Education			eta*	Education		
		Lower	Second	Higher		Lower	Second	Higher
Influence	.12*	5.3	5.4	5.9	.21*	5.2	5.2	6.2
Sociability	.07	5.5	5.4	5.7	.12	5.3	5.3	5.9
Exuberance	.10*	5.3	5.4	5.8	.09	5.4	5.3	5.8
Structure	.20*	6.1	5.6	4.9	.19*	5.8	5.8	4.8
Stability	.07	5.5	5.3	5.7	.04	5.5	5.4	5.6
Scale								
Status	.03	5.5	5.4	5.6	.10	5.4	5.3	5.8
Dominance	.13*	5.2	5.4	5.9	.22*	5.1	5.2	6.2
Competition	.11*	5.2	5.4	5.8	.19*	5.2	5.3	6.1
Self-presentation	.12*	5.3	5.3	5.9	.16*	5.4	5.2	6.0
Need for contact	.03	5.6	5.4	5.5	.08	5.4	5.4	5.8
Leisure contact	.06	5.4	5.4	5.7	.15*	5.3	5.3	6.0
Self-disclosure	.14*	5.1	5.4	5.9	.11	5.3	5.4	5.8
Trust	.18*	5.3	5.2	6.1	.17*	5.2	5.3	6.0
Friendliness	.05	5.7	5.5	5.4	.03	5.5	5.4	5.6
Attentiveness	.10*	5.8	5.4	5.3	.04	5.4	5.5	5.6
Energy	.02	5.4	5.5	5.5	.04	5.5	5.4	5.6
Personal growth	.13*	5.2	5.4	5.9	.08	5.4	5.3	5.8
Perseverance	.05	5.7	5.5	5.4	.05	5.7	5.4	5.5
Adaptability	.10*	5.3	5.4	5.8	.07	5.6	5.3	5.6
Originality	.06	5.5	5.4	5.7	.11	5.2	5.4	5.8
Independence	.17*	5.1	5.4	6.0	.09	5.3	5.4	5.8
Orderliness	.11*	5.8	5.6	5.2	.11	5.7	5.6	5.2
Precision	.15*	5.8	5.6	5.1	.20*	5.7	5.8	4.9
Regularity	.22*	6.1	5.6	4.9	.20*	5.8	5.8	4.9
Conformity	.22*	6.1	5.6	4.9	.14*	5.7	5.7	5.1
Decisiveness	.04	5.5	5.4	5.6	.01	5.5	5.5	5.5
Self-confidence	.03	5.5	5.4	5.6	.04	5.6	5.5	5.5
Optimism	.10*	5.3	5.4	5.8	.14*	5.2	5.4	6.0
Frustration-tolerance	.09	5.6	5.3	5.7	.01	5.5	5.5	5.5
Resilience	.03	5.6	5.5	5.5	.11	5.9	5.3	5.4

* Significant results of the ANOVA are indicated by a * for the eta values.

In the Advice group, there were 3 factors and 14 scales that differed significantly from each other in terms of education, for the Selection group this was the case for 2 factors and 9 scales. The largest difference in means is 1.2 sten, which corresponds to a difference of 0.6 times the standard deviation. The average distance over all significant scales and factors is 0.77 sten: this corresponds to a difference of 0.39 times the standard deviation.

'Stereotypical' differences in scores are found for the different educational levels. Examples are the fact that people with lower educational levels have a higher need for *Structure* and that people with higher educational levels find *Personal growth* more important.

Table 5.19. Mean sten score by job status including eta-values

Factor	Advice group (N=712)			Selection group (N=369)		
	eta*	non-UWV	UWV	eta*	non-UWV	UWV
Influence	.06	5.5	4.7	.09	5.5	4.4
Sociability	.02	5.5	5.3	.04	5.5	5.0
Exuberance	.03	5.5	5.2	.08	5.5	4.6
Structure	.05	5.5	6.1	.07	5.5	6.4
Stability	.02	5.5	5.2	.08	5.5	4.6
Scale						
Status	.05	5.5	4.9	.04	5.5	5.1
Dominance	.06	5.5	4.9	.12*	5.5	4.1
Competition	.05	5.5	4.9	.07	5.5	4.8
Self-presentation	.04	5.5	5.0	.07	5.5	4.7
Need for contact	.02	5.5	5.3	.06	5.5	4.8
Leisure contact	.02	5.5	5.2	.07	5.5	4.7
Self-disclosure	.02	5.5	5.2	.02	5.5	5.3
Trust	.03	5.5	5.2	.05	5.5	5.0
Friendliness	.01	5.5	5.6	.03	5.5	5.2
Attentiveness	.03	5.5	5.8	.03	5.5	5.9
Energy	.02	5.5	5.2	.06	5.5	4.9
Personal growth	.04	5.5	5.0	.06	5.5	4.8
Perseverance	.01	5.5	5.7	.02	5.5	5.3
Adaptability	.02	5.5	5.2	.08	5.5	4.6
Originality	.02	5.5	5.3	.08	5.5	4.7
Independence	.02	5.5	5.3	.04	5.5	5.0
Orderliness	.03	5.5	5.8	.00	5.5	5.5
Precision	.02	5.5	5.8	.05	5.5	6.0
Regularity	.06	5.5	6.2	.13	5.4	6.9
Conformity	.05	5.5	6.1	.08	5.5	6.4
Decisiveness	.00	5.5	5.5	.01	5.5	5.4
Self-confidence	.03	5.5	5.2	.07	5.5	4.8
Optimism	.04	5.5	5.0	.12*	5.5	4.1
Frustration-tolerance	.01	5.5	5.6	.02	5.5	5.3
Resilience	.01	5.5	5.4	.03	5.5	5.2

* Significant results of the ANOVA are indicated by a * for the eta values.

Two significant differences were found for the job status variable (Table 5.19). People not registered at the UWV scored significantly higher on *Dominance* and *Optimism* than people registered at the UWV in the Selection group.

Job sector

Using t-tests, we have checked which scales and factors show significant differences in means between the different job sectors. The distribution and sizes of the different sectors in the sample are represented in Table 3.4. The significant differences in mean scores are represented in Table 5.20, 5.21, 5.22 and 5.23. This study was performed on a sample of 712 people in the Advice group and 369 people in the Selection group.

Table 5.20. Differences in mean sten scores on the factor level between sectors in the Advice group (N=719)

Factor	Sector I	Sten I	Sector J	Sten J	Difference (I-J)	p*	SD _p **	Effect size***
Structure	Business and administration	6.12	Commercial services	4.99	1.13	.00	1.64	0.69
Structure	Business and administration	6.12	Public administration, Safety and Law	5.07	1.06	.00	1.59	0.66

* Due to the large number of significance tests performed in this study, a correction of the significance level is needed. Only significant differences after the Bonferonni correction are presented.

** Pooled standard deviation.

*** The difference between mean scores divided by the pooled standard deviation.

On the factor level, only the factor *Structure* shows significant differences between the different job sectors. The sector 'Business and Administration' differs significantly from 'Commercial services' and 'Public administration, Safety and Law' on this factor.

Table 5.21. Differences in mean sten scores on the factor level between sectors in the Selection group (N=1067)

Factor	Sector I	Sten I	Sector J	Sten J	Difference (I-J)	p*	SD _p **	Effect size***
Influence	Commercial services	6.68	Health, Wellbeing and Personal Care	5.92	.75	.00	1.30	0.58
Influence	Commercial services	6.68	Public administration, Safety and Law	5.90	.78	.00	1.29	0.61
Influence	Engineering and Production	6.94	Public administration, Safety and Law	5.90	1.04	.00	1.35	0.74
Influence	Engineering and Production	6.94	Health, Wellbeing and Personal Care	5.92	1.01	.00	1.33	0.74
Exuberance	Commercial services	6.44	Public administration, Safety and Law	5.78	.66	.00		

* Due to the large number of significance tests performed in this study, a correction of the significance level is needed. Only significant differences after the Bonferonni correction are presented.

** Pooled standard deviation.

*** The difference between mean scores divided by the pooled standard deviation.

In the Selection group, significant differences were only found for the factors *Influence* and *Exuberance* (see Table 5.21).

Table 5.22. Significant differences in mean sten scores on the scale level between sectors in the Advice group

Scale	Sector I	Sector J	Difference (I-J)	p
Regularity	Business and administration	Public administration, Safety and Law	1.50	.00
Dominance	Business and administration	Public administration, Safety and Law	-1.53	.00
Preciseness	Business and administration	Public administration, Safety and Law	1.54	.00
Competition	Health, Wellbeing and Personal Care	Commercial services	-1.09	.00
Self-confidence	Health, Wellbeing and Personal Care	Public administration, Safety and Law	-1.07	.00
Regularity	Business and administration	Commercial services	1.20	.00
Dominance	Business and administration	Commercial services	-1.22	.00
Preciseness	Business and administration	Commercial services	1.22	.00

Table 5.23. Significant differences in mean sten scores on the scale level between sectors in the Selection group

Scale	Sector I	Sector J	Difference (I-J)	p
Competition	Health, Wellbeing and Personal Care	Commercial services	-1.24	.00
Competition	Health, Wellbeing and Personal Care	Engineering and Production	-2.00	.00
Competition	Public administration, Safety and Law	Engineering and Production	-1.99	.00
Competition	Public administration, Safety and Law	Commercial services	-1.23	.00
Status	Public administration, Safety and Law	Commercial services	-.96	.00
Status	Public administration, Safety and Law	Commercial services	-.76	.00

The most important question that needs to be answered is whether the WPI can be used sector independently for selection and advice purposes. In other words: how large is the effect of the sector on the scale and factor scores? In answering this question, it is not important how one sector relates to the other sectors, but more how a sector relates to the total norm group. The distribution of people over the different sectors for both norm groups is represented in Appendix 12, as well as the mean factor and scale scales and the corresponding effect difference with the total group. From this it becomes clear that people in some sectors score significantly higher on some scales and factors than the people in the total norm group, but that this does not outweigh the information that the groups contain.

Conclusion

Despite the large numbers of people in the norm groups, we found only a limited amount of significant effects for the background variables gender, age, job status and education. In terms of effect size (Cohen, 1992), the significant differences for the background variables are not of such an order that separate norm groups should be created. It can be interesting to include this information in the interpretation of the results since it concerns real differences, but this is not absolutely necessary. The study on differences between sectors has shown a number of results that can be considered as real, given the nature of the sectors. This contributes to the construct validity of the WPI, and indicates why separate norm groups are not desirable.

5.2.3. Construct validity: Cultural Bias

Since people with different cultural backgrounds live and work in the Netherlands, it is necessary that the WPI is a culturally fair questionnaire. Culturally fair means that no unfair biases occur in individual outcomes, and that only real differences between individuals are evident in relation to the entire work force. After all, these differences have a significant meaning for the Dutch labour market.

An extensive study was done to investigate the cultural bias of the WPI. The results of this study are described in Appendix 13.

The most important finding is that significant differences for only two scales were found; this was the case for *Status* and *Trust*. The native Dutch respondents have a lower score on *Status* and a higher score on *Trust*, compared with their non-native counterparts. This is consistent with the expectations one can derive from the literature on Dutch cultural characteristics. From this we can conclude that the WPI does not show cultural bias and that the scores and texts in the report represent real individual differences.

5.2.4 Construct validity: The WPI and the Career Values questionnaire by Ixly

The Career Values questionnaire (CV) is a personality questionnaire that provides insight into the aspects of work that can motivate a person. The questionnaire was developed for the HRM work field and can be used in both advice and selection contexts. In advice situations, the questionnaire provides insights into what a person finds motivating in a certain job. This way, it is easier for someone

to look for a suitable job. In selection contexts, the CV provides insights into the match between the career values of the candidate and the characteristics of the job the candidate is applying for. When there is a match, this tells something about the motivation a candidate has for this specific function (Orga, 2007). The questionnaire that was used in this study (CV ipsative) consists of 190 items which are presented in a forced choice format: two items are presented opposed to each other, with a four point scale in between. The candidate has to indicate which of the two items applies more to him/her. The questionnaire consists of 20 scales. For a description of the scales, see Appendix 14.

This study also contributes to the criterium validity of the WPI. Although the CV is not what the WPI intends to measure, finding no or unsuspected relations between the two would question the criterium validity of the WPI.

The CV and the WPI are complementary during assessment procedures, because the WPI concerns the personality of the candidate while the CV concerns the personal preferences and values in terms of the candidate's professional career. From experience we know that when someone scores high on a certain scale of the WPI, this does not necessarily have to mean that someone values this specific characteristic in a job. It is interesting to investigate how the results on both questionnaires correlate since one can expect that there are similarities in personal characteristics and the characteristics people find important in their job. The expectation is that when someone scores high or low on a scale of the WPI, this corresponds with certain value on the CV. The reason for this is that people often find aspects of work interesting or fun because they are good at it or because it they with their personality.

Data was available of 1329 people who took both the WPI and the CV questionnaires. All the information came from the Ixly's dataset and was retrieved at different companies. The age of 761 people from the dataset was known, varying from 16 to 62 years old, with a mean age of 35,4. Of all the people of which the gender was known 539 were male and 669 female. The questionnaire was taken by 1139 people in an advice context, while 190 people took the questionnaire in a selection context. The questionnaires were completed during the beginning of 2004 and mid-2006.

An overview of the mean sten score on the factors of the WPV and their standard deviations are reported in Table 5.24. The mean sten scores and the standard deviations are very similar to the norm group.

Table 5.24. Mean sten scores and corresponding standard deviations (N=1329)

Factor	Mean sten score	Mean standard deviation
Influence	5	2.2
Sociability	5.2	2.2
Exuberance	5	2.2
Structure	5.3	2.5
Stability	5	2.2

Correlations were calculated between the factors and scales of the WPI and the scales of the CV. The reliabilities of the scales of the CV in this sample were between .65 and .89 (Cronbach's alpha). The reliabilities of the scales of the WPI were between .83 and .94 (α) and the reliabilities of the factors of the WPI were between .96 and .97 (stratified alpha). In calculating the reliability of the factors we have assumed that the total error variance was the same as in the norm group.

The correlations between the factors of the WPI and the CV are reported in Table 5.25. The correlations larger than .30 and smaller than -.30 are in bold. In studies like this, correlations that are larger than .30 (absolute) are considered as of average size and correlation larger than .50 of large size.

Table 5.25. Correlations between the factors of the WPI and the scales of the CV

CV-scales (N=1329)	WPI-factors (N=1329)				
	Influence	Sociability	Exuberance	Structure	Stability
Balance private life - work	-.23**	-.03	-.23**	.12**	-.22**
Financial reward	.27**	-.03	.05	.09**	-.05
Praise and recognition	.14**	.16**	.13**	.16**	-.13**
Tangible results	.21**	.15**	.37**	.31**	.11**
Useful contribution	.03	.19**	.14**	.12**	.01
Career	.68**	.08**	.40**	.04	.17**
Quality	.24**	.11**	.39**	.45**	.15**
Creative thinking	.37**	.19**	.56**	-.07*	.19**
Influencing	.67**	.26**	.43**	-.10**	.25**
Enterprising	.50**	.25**	.44**	-.13**	.24**
Helping others	-.06*	.39**	.13**	.23**	.06*
Analysing	.32**	.18**	.53**	.14**	.25**
Developing	.30**	.24**	.53**	.10**	.23**
Being physically active	-.06*	-.00	-.03	.03	-.05
Autonomy	.28**	-.03	.27**	-.28**	-.03
Security and stability	-.03	.03	-.05	.44**	-.05
Challenging tasks	.36**	.24**	.60**	-.02	.33**
Cooperation	.16**	.50**	.23**	.24**	.23**
Attention	.63**	.21**	.23**	-.13**	.05
Hectic situations	.26**	.28**	.46**	.15**	.23**

** Significant at $p < .01$ (2-sided)

* Significant at $p < .05$ (2-sided)

The factor Influence shows average (.30) to large (>.50) correlations with the following career values: *Attention, Influencing, Creative thinking, Enterprising, Developing, Career, Analysing and Challenging* tasks. People that score highly on this factor find these career values very important. All of these career values are, in a logical and explainable way, consistent with scoring high on the Influence factor of the WPI.

The factor Sociability shows average to large correlations with the following career values: *Helping others* and *Cooperation*. Social people often find it important to be able to help others in their job and social people tend to value the cooperative aspects of work as well.

The factor Exuberance shows average to large correlations with the career values: *Influence, Creative thinking, Dynamics, Entrepreneurship, Development, Quality, Career, Analysing, Tangible results* and *Security and stability*. All these career values can be linked, again, in an interpretable way with the factor Exuberance. When we look at scales this factor comprises, it is easy to see that scales such as *Self-development, Adaptability and Originality* fit well with these career values.

The factor Structure shows average to large correlations with the following career values: *Quality, Tangible results* and *Security and stability*. Structured people like to deliver quality and specific, tangible results.

The factor Stability only shows an average relation with the career value *Task challenge*. These people are stable and therefore need a challenge in their job. The fact that this factor does not correlate with the career value *Security and stability* can be explained by the reasoning that people who are stable do not need this in their job. The factor Stability of the WPI concerns the (emotional) stability of a person while the career value *Security and stability* mainly concerns the rewards and the materialistic (monetary) aspects of a job (career). Thus, to what extent are you stable, and to what extent do you need stability, respectively.

The table with all correlations between the WPI and the scale scores of the CV is included in Appendix 15. Below in Table 5.26, an overview is provided of the correlations between the scales of the WPI and the CV that are at least .30. The career values with the highest correlations with the corresponding scales of the WPI are reported first.

Table 5.26 Relationships between the scales of the WPI and CV.

WPI scale	Career Values	
	Positive relation	Negative relation
Status	Career, Attention, Influencing, Enterprising	Balance private life – work, Helping others, Security and stability, Analysing, Being physically active
Dominance	Influencing, Enterprising	Security and stability, Balance private life – work
Competition	Career	Balance private life – work, Helping others
Self-presentation	Attention, Influencing, Career, Enterprising	Security and stability, Balance private life – work.
Leisure contact	Cooperation	
Attentiveness	Helping others	
Energy	Hectic situations, Career	Balance private life – work, Security and stability
Personal growth	Developing, Career, Challenging tasks	Being physically active
Perseverance	Quality	
Adaptability	Challenging tasks, Developing, Creative thinking	Security and stability, Balance private life – work
Originality	Creative thinking	Security and stability
Independence	Autonomy	Cooperation, Helping others
Precision	Quality	Enterprising, Attention, Autonomy, Influencing
Regularity	Security and stability, Balance private life – work, Quality	Enterprising, Influencing, Challenging tasks, Creative thinking
Conformity		Autonomy
Decisiveness	Quality	
Self-confidence	Influencing	Security and stability
Optimism		Security and stability

The scales *Self-disclosure*, *Trust*, *Friendliness*, *Orderliness*, *Optimism*, *Frustration-tolerance* and *Resilience* are not shown in Table 5.26 because they did not show an average (>.30) or high (>.50) relation with one of the career values. The associations between the CV-scales and WPI-scales are not difficult to explain. Interestingly, *Security and stability* and *Balance private life – work* both show mostly negative correlations with the WPI-scales, while the WPI-scales mostly show positive correlations with *Enterprising*. The WPI is a work-related personality questionnaire and people who find enterprising important in their career, appear to have less of a need for security and stability and a well-balanced private and work life.

The research described here strongly contributes to the construct validity of the WPI. All the relationships found were either expected or easy to explain.

5.3. Concluding remarks on the validity of the WPI

In the research on the validity of the WPI, a distinction was made between predictive validity and construct validity.

Predictive validity

Research focusing on concurrent validity shows that the factors and the majority of the scales of the WPI show high and theoretically justifiable correlations with the factors of the FFPI. The average correlation between the respective factors is .650

Research on student nurses shows that particularly the Sociability and Stability factors (and their underlying scales) are predictive of the academic performance in nursing school. Research at an

employment agency has shown that the WPI has predictive power in this context as well. Different criteria of job performance were predicted by scores on the WPI. These two studies show the predictive validity of the WPI in two very different populations.

Construct validity

Research on the internal structure showed that the Multiple Group Method indicated that the questionnaire consists of relatively independent, homogeneous, reliable and stable scales. At the factor level, the reliability is high to very high ($>.95$). The stability of the scales is .87 on average.

Research on different background characteristics and their relations with WPI scores shows that there are only small differences in scores between different subpopulations. Furthermore, we believe that the small differences found reflect real and relevant differences which are a good reflection of the reality. If, for example, different norm groups would be formed for men and women separately, then a male with a low score for Attentiveness would not receive a very large sten score. While in reality, we are still dealing with an inattentive person. Because of this we have chosen to use one single norm group for the different subpopulations.

Finally, a study was conducted investigating the relations between the WPI and the Career Values questionnaire. This study showed that all factors of the WPI had an average to strong relation with one or more scales of the CV questionnaire. Similarly, at the scale level, a number of theoretically justifiable correlations were found between the WPI and the CV. These results also support the construct validity of the WPI.

6. Application, interpretation and use

The application, interpretation and use of the WPI will be discussed in this chapter. First, we will elaborate on the possible applications of the questionnaire, the test material, the instructions, the knowledge required for the use of the WPI and the limitations of the questionnaire. The interpretation of the test scores will be clarified by means of an example casus. Finally, we will discuss the competencies that are derived from the results of the WPI, how they are established and the value one can attach to these competencies.

6.1. Applications

The WPI can be used in any setting in which it is important to have information on someone's personality. The WPI is a personality questionnaire related to personnel selection, personal development and career advice. It can be assumed that the questionnaire can be used for the entire Dutch work force.

6.2. Test material and test instructions

Material

All the questionnaires by Ixly are administered in the Ixly Test-Toolkit. This is an online environment that offers qualitatively high instruments to professionals and consultants in the field of Human Resource Management. The toolkit can be reached via <http://l.test-toolkit.nl>. Consultants receive a username and password when they subscribe to the Test-Toolkit. We ask them to regularly change this information, due to security reasons. The online questionnaires provided by Ixly are thus not available to people that do not have a subscription to the Test-Toolkit. Consultants can assign one or multiple questionnaires to a candidate. The candidates are informed through an e-mail sent from the Test-Toolkit. This e-mail included the login details of the candidate. The questionnaires are therefore not available without login information and thus not available to everyone. After completion of the questionnaire, the results are only retrievable by the consultant. Information about the operation and possibilities of the Test-Toolkit are described in Appendix 16.

Instruction

The candidate is able to login to <http://l.test-toolkit.nl> with the login information obtained in the e-mail. On the front page, there is a short description on who has administered the questionnaire and how the Test-Toolkit works. In addition, the candidate gets a short (pre-)questionnaire in which the candidates fills out demographical information. It is indicated that the data is processed anonymously and that the information will be used for research purposes only. After the pre-questionnaire, the candidate arrives at an overview of all the tests and questionnaires made available to him/her, including a short description of these tests and questionnaires. The candidate can take the tests in any order they prefer.

Since the WPI is a personality questionnaire, it is administered without any time pressure. This means that the candidate can take as much time as he or she needs for the questionnaire. It takes the average candidate about 45 minutes to complete the questionnaire. However, in the instructions it is indicated that it is important that the candidate does not think too long about the questions: what comes to mind first should be filled out. In addition, the candidates are instructed to give answers based on typical, general work-related situations, not based on atypical, exceptional situations. Both instructions are important in getting a reliable as possible image of the candidate. Because the WPI is administered online, candidates always receive standard instructions.

The WPI consists of 276 statements for which the candidates is asked to indicate to what extent he or she agrees with the statement. Answers are given on a 5 point Likert scale ranging from totally do not agree to totally agree. How the candidate should interpret the 5 answering categories is explained by an example. The instructions can be requested on every page.

The candidate has the possibility to close the questionnaire at any moment. The answers that are given up to that point are saved locally on the computer of the candidate. In this way, it is not possible to leave the software without saving the given answers. It is, however, necessary for the candidate to

complete the WPI on the same computer and in the same browser as it was started on. The candidate is notified of this before starting the questionnaire. After completion of the whole questionnaire, the results are automatically sent to the database of the Test-Toolkit.

We have chosen not to give the candidates the option to go back to previous screens when completing the questionnaire. Although this is possible in paper-and-pencil questionnaires, we believe that the automation of the questionnaire is an improvement to this. When a person is able to go back to previous screens, it is possible to make answers more consistent with each other while this is not a reflection of the true response tendency of the person. The candidates are therefore advised not to think too long about their answers and to answer with what first comes to mind. This is not facilitated by giving the candidates the possibility to alter previously given answers.

The WPI does not need to be administered in a controlled setting. This is possible due to the fact that it concerns a self-report test which means that there are no correct or incorrect answers. The candidate can take the WPI at home, which saves time in the advice and selection procedure.

Software and support

The WPI can be taken on any computer that has a connection to the internet and a functioning browser. In addition, only a keyboard and mouse are necessary. The computer must meet the following system requirements:

Windows XP Professional SP2, Windows Vista, Windows 7, Apple OS X 10.4 or higher

Minimum of 512 MB RAM-memory and 10 MB free disk space

A minimum screen resolution of 1024x768 pixels

Internet Explorer^a (IE 6 or higher), Firefox (3.x or higher), Safari (3.x or higher)

The browser should be able to execute Javascript^b and Flash^c (8 or higher) needs to be installed

Internet connection with a minimum average download speed of 2 Mbps / 256 KBps^d

An "open connection" via port 80 (no Proxy)

ad a. Some minor adjustments need to be made in the cookie-settings of Internet Explorer. Ixly's helpdesk can lend support for this, if needed.

ad b. The way in which this can be set depends on the type of browser used. The website <http://www.browserchecker.nl/javascript-aanzetten.php> can be used to check whether javascript is enabled or not. The previously stated website show how javascript can be enabled.

ad c. On www.adobe.com/products/flash/about/ one can download Flash or check which version is installed.

ad d. This is an approximation of the average bandwidth required, for each user on the network. For the average user, a basis ADSL connection suffices. The available bandwidth can be checked on www.speedtest.nl/TestSuite/.

The tests and questionnaire are locally saved on the computer of the candidate when the candidate clicks <start> in the online environment. The internet connection is only needed when starting the questionnaire and when sending the responses to the server (when the candidate clicks on the <end> button). The candidate does not run into problems should the internet connection be lost during completion of the questionnaire. When there is no internet connection at the time of sending the responses to the server, the candidate will be noted about this. As soon as the internet connection is restored, the responses can still be sent to the server. Until that time, they will be stored locally on the computer.

For questions on the system requirements and for technical support, candidates can contact the Ixly Toolkit's helpdesk. The helpdesk can be reached during business days from 8 a.m. to 5.30 p.m., via helpdesk@ixly.nl or 088-4959000.

Frequently asked questions

A list of frequently asked questions is enclosed in Appendix 17.

6.3 Knowledge required for the use of the WPI

Test user

When the WPI is used by a professional in order to advise others, it should be ensured that:

- This person is competent, qualified, licensed or authorized to use psychological tests for different purposes, such as assessment, coaching, training and Human Resource Management, in which he/she works. This should be in line with the laws and regulations of the respective country.
- This person will use the product in accordance with the national or international professional standards and ethics.
- This person will use the product in accordance with national or international laws and regulations, instructions and guidelines and all other applicable governmental or quasi-governmental rules.
- This person will only use the product for the organization he or she works for or for his or her own company, in his/her own name and on his/her own account. It is not allowed to sell, lease, copy, give or transfer the product in any way to whomever or whatever company, except for the use of products and services as an integral part of the service to clients or the use within the organization that is the direct employer of the professional.

For certain services or products, Ixly will test the reliability and knowledge as a professional before one can get access to these services or products. Ixly has the right to deny this access, without stating a reason.

Candidate

The Ixly Test-Toolkit is very user friendly. The candidate is only required to be able to start up a computer, to visit a web page and to operate a mouse. When the candidate has these skills, he or she is able to complete the WPI in the Test-Toolkit. When the instructions and questions are read aloud to the candidate, for example in the case of blind or candidates with impaired vision, it is important that the reader has the aforementioned skills.

6.4. Limitations of the questionnaire

To what extent the WPI can be used in distinct (poli)clinical settings (i.e. for patients) is yet unclear. No research in this setting has been conducted up to now. In addition, the WPI has not been tested yet in school populations with an age of 15 years or younger; therefore, we do not know the merits that the WPI can have for these groups. It would be a nice addition to have more information on this in the future, although these groups do not necessarily fall under the work force, which is the target audience of the WPI. Nevertheless, it is possible that the WPI can be used in these settings as well.

6.5 Interpretation of scores

The WPI is reported in sten scores. This scale ranges from 1 to 10. Sten scores are a form of standard scores with a mean of 5.5 and a standard deviation of 2. Sten scores present an image of how a certain score relates to the mean of all scores. Stenscore 4, 5, 6 and 7 all fall within 1 standard deviation of the mean. Stenscore 2, 3 and 8, 9 fall between 1 and 2 standard deviation of the mean. Stenscore 1 and 10 is located more than two standard deviations from the mean. The mean score of the norm group is exactly located on the border between the fifth and the sixth sten (5.5). The percentages corresponding to the separate sten scores are shown in Table 6.1.

Table 6.1. Sten scores with corresponding percentages

Sten	Percentage	Cumulative percentage
1	2.3%	2.3%
2	4.4%	6.7%
3	9.2%	15.9%
4	15%	30.9%
5	19.1%	50.0%
6	19.1%	69.1%
7	15%	84.1%
8	9.2%	93.3%
9	4.4%	97.7%
10	2.3%	100%

When someone scores an 8 on a certain scale, this means that this person scores between 1 and 2 standard deviations higher than the mean which corresponds to 6.7% (4.4 + 2.3) of the people from the norm group have scored higher. We have chosen to report on sten scores because they are easy to interpret and understand, and because they are detailed enough for the application of the WPI. A possible disadvantage is that they are sometimes confused with school grades. A sten score of 5, for example, does not mean a “bad” score, but an average score that is very common in the norm group.

The factor scores are also reported as sten scores. This sten score is not a simple average of the scales belonging to a factor, rather, the factor scores are standardized separately. The interpretation of the scales and factors is the same.

6.5.1 Definitions of the factors and scales

For a good interpretation of the WPI it is important to know the meaning (content) of the factors and scales. An overview is included in Appendix 4.

6.5.2 Interpretation of the scores in a selection and an advice situation

To illustrate the interpretation of the WPI, a psychologist has described how he uses the WPI in selection assessments and in career advice.

6.5.2.1 Selection assessment

In the selection process of candidates, personality is of course not the only aspect that is assessed. In selection assessment, a suitable combination of several selection techniques is important: a combined assessment of intelligence, personality, motivation, ambition and career values. For example, in congruence with research findings that intelligence has higher predictive value for more complex job functions, cognitive tests get a larger weight in the selection procedure of expert positions than in selection procedures of positions with a smaller degree of complexity (Salgado, 2003). For some functions personality questionnaires will have a larger share in the process than for other functions, for example in sales and management functions, for which personality is an equally strong predictor as intelligence.

The goal of the WPI is to provide a questionnaire that is relevant in the professional field of HRM. This translates to a large emphasis on the ‘Conscientiousness’ factor. In the WPI, this factor is split up into two factors: *Exuberance* and *Structure*. *Exuberance* concerns motivation, the power with which people want to make a contribution. *Structure* in the WPI corresponds more to the classical interpretation of ‘Conscientiousness’, meaning conscientious and precise.

The social factors *Influence* and *Sociability* will be mostly relevant for the functions in which people work with other people and in which leadership is asked. A direct relationship of these factors with performance is not always clear. In service providing job functions, the factor *Sociability* also seems to play a role (see paragraph 5.2.4. *Validity study with 360 degrees feedback*).

Stability appears to be relevant for a lot of job functions, especially when pressure is put on people. It appears to mainly influence the output, i.e. the quantity of work. However, a linear relation with job

performance is not always the case. High scores can also lead to a degree of indifference and in consequence inertia, as we have seen quite some times in the field.

It is important for the selection psychologist to check and test any remarkable results of the selection procedure in the interview. We recommend the criterium, competence or behavioural interview method, which is also called the STAR-method. Direct use of personality questionnaires to determine the suitability of a candidate for a specific function is not advisable. Questionnaires are not tests in the strictest sense of the word. As we emphasize in each report, we are still dealing with self-reports.

In the selection processes, one first needs to make clear which competencies are important for the job concerned, and which scales of the personality questionnaire correspond to these competencies. With the results of the questionnaire, one can look whether these correspond with the competencies one deems necessary for the function. In the interview and in possible simulations, this needs to be tested.

Selection assessment case

A medium-sized automation company is looking for a business consultant. The job function is described as follows:

Business consultant (M/F)

Job description

You will carry out a wide range of tasks at different types of clients. You find it important to help your clients with innovative IT-solutions to their business problems. You have highly developed analytical skills that enable you to find useful solutions and implement them at your clients. You work independently or with colleagues in projects on location or in-house at your client. You are able to establish and maintain a professional network. As an all-round professional, you will be fulfilling multiple roles in a single or multiple projects. You have a large drive for growth, alone, but together with your colleagues in a team as well.

What we are looking for

A minimum of 4 to 10 years experience as an (external) advisor in the fields of business and IT in a consultancy firm or in an internal IT department;
Excellent counseling skills;
Pragmatic insights and customer and results oriented in counseling.

Competencies:

- *Analytical skills*
- *Innovative capacity and creativity*
- *Learning capacity*
- *Influencing ability*
- *Networking*
- *Representing the company externally*
- *Cooperation*
- *Planning and organizing*

Analysis of the function profile

Employers often have their own competency language, often through self-developed varieties of available systems, which have to be interpreted by understanding what is exactly meant. In the following, we will analyse all the competencies.

Analytical skills. Since these skills cannot be measured in a personality questionnaire, these skills are not included in the WPI. Tests for cognitive capacities are more suited for this. Since it is an intellectually complex function, it is justified that great importance is attached to this. Therefore, determination of the intellectual capacities should be included in the test program.

Innovative capacity and creativity cannot be derived directly from the WPI either. Again, this is because intelligence plays an important role. In combination with intelligence one can look at the scales *Adaptability*, *Independence* and *Originality*. These indicate the tendency to be innovative,

creative and original. Whether these translate into the competence is partly dependent on intelligence. It is quite possible to be open to all sorts of solutions, but to still always choose the wrong one.

Learning capacity. Again, intelligence, in combination with interests, plays a role. It is interesting to look at the scores on *Self-development* and the factor *Exuberance*. It is advised to minimally include one cognitive ability test and preferably the Career Values questionnaire as well.

Influencing ability, networking, cooperation and representing the company externally are all social competencies. In terms of factors and scales, *Influence, Sociability* and *Exuberance* are important. For cooperation the factor *Sociability* is mostly important, and within this factor, specifically the scales *Trust, Friendliness* and *Attentiveness*; to be able to trust others and to be willing to conform to the wishes and the roles of others. For the competence cooperation it is also important that one does not score too high on *Competition*, has a fair degree of *Stability* and does not score too low on *Conformity*. Scores that are too low on this scale can lead to too individualistic behaviour. Some degree of *Exuberance* and *Influence* are desirable, while scores that are too high can lead to less cooperative behaviour. In short, a lot of scales and factors play a role in the competence of cooperation. It is important to come to understanding from the interview how a person behaves in teams and what role the individual plays in the dynamics of a team. Furthermore, it is advisable to use simulations or role plays in order to assess these communicative competencies. Through this, direct observation of behaviour is possible.

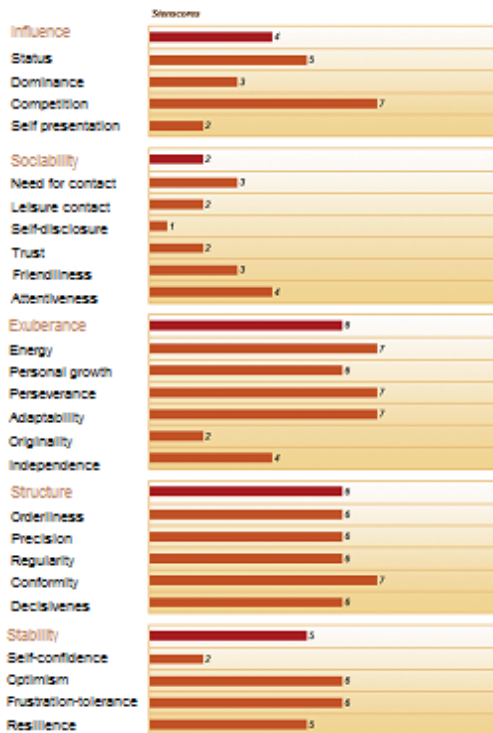
Planning and organizing. The factors *Structure* and *Exuberance* are very important for this competence. In addition, we want to emphasize that planning and organizing requires intelligence as well, especially when the planning is complex.

Whenever you are involved in a selection procedure, it is important to do the analysis as described here, in order to make clear, beforehand and unprejudiced by the results, what the relevant scales and factors are for the selection process.

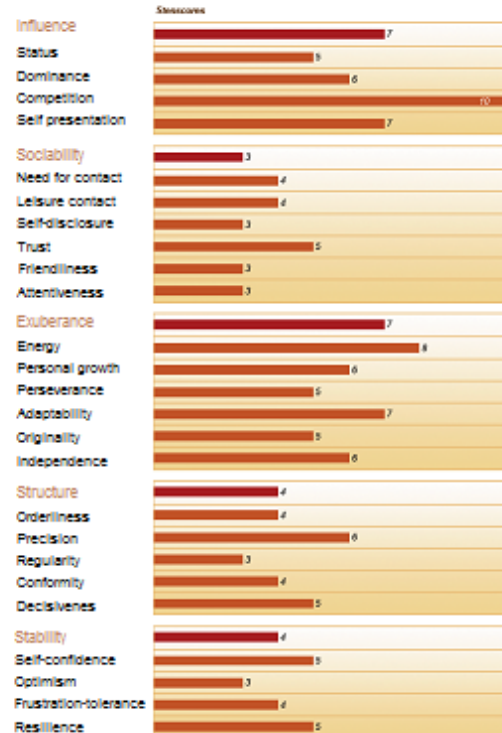
Candidates

The automation company directs two potential candidates to an assessment company, where the WPI is administered, as part of a program that additionally consists of a cognitive ability test, the Career Values questionnaire, a role play and a criterium related interview. The psychologist prepares the interview with the results of the tests, which are completed by the candidate on the internet the day prior to the day of assessment. Below, the results of the two candidates are graphically presented.

Results candidate 1:



Results candidate 2:



Candidate 1

When looking at the results, it is important to first look at the factors. It becomes apparent that *Influence* is just below average with a low score on *Self presentation*, but with a large need for competition. *Sociability* is downright low, this person is very introverted and has a low need for contact. *Exuberance*, a factor that is part of 'Conscientiousness' in other personality questionnaires, is above average, with a high score on *Perseverance* and *Energy*. *Originality* is less well pronounced. We can call the score on *Structure*, which is also part of 'Conscientiousness', average, just as the score on *Stability*, while the low score on *Self-confidence* is striking.

The following scores for the function relevant competencies are notable:

Innovative capacity and Creativity

Scores on the scales *Independence* (4) and *Originality* (2).

Apart from intelligence, this candidate appears to be less inclined to contribute in an original and creative manner. One really needs to pay attention to this and pose questions on it in the criterium related interview.

Learning capacity

The scale *Self-development* (6) and the factor *Exuberance* (6)

Independent of intelligence, this person appears to have a just above average learning capacity, in which the wish to make the best of himself/herself and to put energy in this is absolutely present. In the interview, we have to ask about the underlying motivation and the willingness to put energy in gaining new knowledge and insights.

Influencing ability

The factor *Influence* is below average, in which the low score on *Dominance* (3) is remarkable. This leads to the expectation that this candidate will not excel in influencing others. In the interview, as well as in the role play and simulation, this needs to be paid attention to.

Networking, representing the company externally

The factors *Exuberance* (6), *Influence* (4) and *Sociability* (2).

This competency appears to be the largest risk for this candidate. This candidate is an introvert rather than an extrovert. Networking and external representation of the company will mostly be motivated because the function demands it: it is not something that this candidate will do spontaneously. Some people will be able to overcome this in their job, while being more solistic and introverted at home. This should be discussed in the interview.

Cooperation

The factors *Sociability* (2) and *Exuberance* (7) and the scales *Trust* (2) and *Conformity* (7).

This candidate is not a teamplayer by nature, but might be able to cooperate functionally as the result of a large drive. There is, however, willingness to conform to the wishes and styles of others, something that is important for this function. The low level of trust will possibly play a role in the candidates functioning in groups. At first, he will probably keep some distance and question the intentions and performance of others.

Planning and organizing.

The factors *Exuberance* (6) and *Structure* (6).

This candidate appears to be conscientious enough to be able to adequately plan and organize work, although caution is warranted for solistic behaviour and a lack of delegating tasks. In addition, intelligence should be taken into consideration when looking at this task.

Summary candidate 1.

The profile indicates a couple of risks that should be investigated in the interview and role play. These risks are the following, in order of severance:

- Cooperation
- Creativity
- Influencing ability
- Networking

Candidate 2

Again, let us first look at the general overall impression of the profile. We are dealing with a driven, very competitive candidate. The levels of *Influence* and *Dominance* are above average, combined with a lower level of *Sociability*. *Exuberance* is above average, in combination with a high level of adaptability and a large need for change. Attentiveness is somewhat lower: this candidate believes that everyone has his/her own responsibilities. In terms of *Structure*, this person has a below average level of conscientiousness, with a larger need for change than for routine. *Emotional Stability* is below average, with a relatively low level of *Optimism*.

Let's look at the relevant scales for the function:

Innovative capacity and Creativity

Scores on the scales *Independence* (6) and *Originality* (5).

These average scores show an average susceptibility for creative contributions, for which the intelligence – as noted before – will also be of importance. We cannot expect this candidate to show a lot of innovativeness.

Learning capacity

The scale *Self-development* (6) and the factor *Exuberance* (7)

The above average self-development score indicates that this candidate is willing to develop himself, partly because of his exuberance.

Influencing ability

The scales *Influence* (7) and *Exuberance* (7)

The just above average scores, in combination with the high score on *Competition*, show that this candidate will be able to develop a certain level of influencing ability. Again, this needs to be questioned in the interview.

Networking, representing the company externally

The factors *Exuberance* (6), *Influence* (4) and *Sociability* (2).

The scores on *Sociability* are quite low, below average. It can be expected that this candidate will network because of his exuberance and drive and his need for influence, i.e. because of functional reasons. There is a slightly below average score on *Need for contact* while the candidate is quite relaxed in social situations (*Leisure contact*). Again, these scores should be addressed in the interview.

Cooperation

The factors *Sociability* (4) and *Exuberance* (7) and the scales *Trust* (5), *Independence* (6) and *Conformity* (7).

As mentioned with Candidate 1, low *Sociability* influences the cooperation. The need for contact is just below average. However, the candidate appears to have sufficient levels of exuberance and an average trust in the intentions and contributions of others. The *Conformity* score is just below average, *Independence* is average. We can expect this candidate to be able to function well in a group, but to adapt an autonomous and independent role within this group.

Planning and organizing

The factor *Structure* (4) and *Exuberance* (7).

The just below average score on *Structure* and above average score on *Exuberance* lead us to expect that the candidate is just sufficiently inclined to structure his work, without being a perfectionist. He is more concerned with the bigger picture; details will not always be paid attention to. Intelligence also needs to be considered when reviewing this competency.

Summary candidate 2.

The candidate appears to have some strong points and a couple of risks that need to be addressed in the interview and role play. The qualities appear to be:

- Learning ability (but keep intelligence in mind)
- Influencing ability (needs to be confirmed in the role play and interview)
- Cooperation

On average, the following competencies appear to be risk factors:

- Networking and representing the company externally
- Planning and organizing

In short, a candidate that needs to confirm and reinforce some aspects in terms of the interview, role play and intelligence. However, the candidate appears to be a serious contender.

Epilogue

Of both candidates, candidate 2 appears – only on the basis of the WPI – to be the best candidate. However, a lot depends on the interview, intellectual ability tests and on the role plays. In this situation,

the intelligence of the candidates appeared to be quite similar, i.e. at a higher educational level. In the role play, which included a situation in which influence needed to be exerted, candidate 2 did a good job. Candidate 1, on the other hand, did not perform as well. All things considered, the whole process resulted in a negative advice for candidate 1 and a positive advice for candidate 2.

6.5.2.2 Career advice

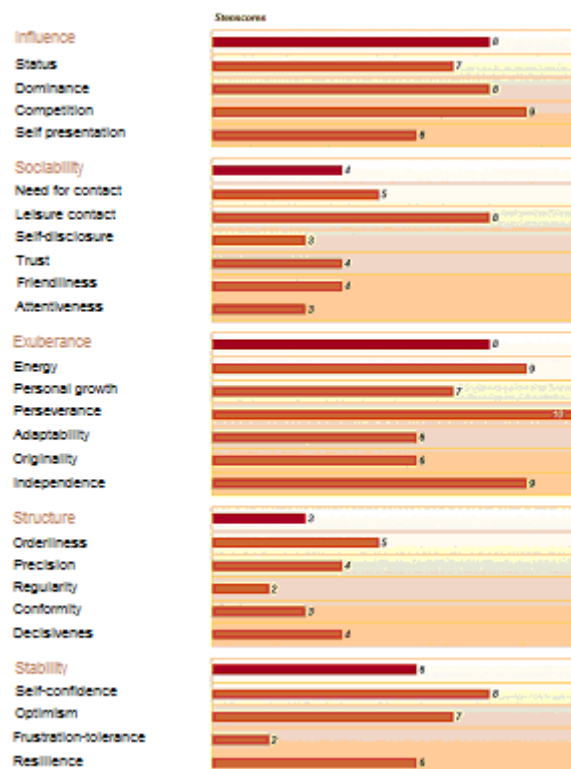
In career advice situations, the WPI is often administered in combination with the Career Values questionnaire, interests questionnaires and intellectual capacity tests: in this situation, the focus is more on the candidate. In selection processes the question is whether a candidate is suited for a certain function, while in career advice settings the question is which functions or professionals are most suitable with the personality and motivations of the candidate. Of course, other factors such as working experience, education and practical considerations such as working conditions and commuting distances are important as well. Again, we will discuss the application of the WPI, this time in the context of an outplacement program.

Career advice candidate

Situation

A 48 year old manager is signed up for an outplacement program, let's call him J. Doe. J. Doe was working for the municipality as head of cleaning services, but got into a conflict with the city councillor. The specific reason for this conflict was that he had ordered a number of garbage trucks, without following the proper internal procurement process. Mainly, no consultation took place with the responsible councillor before taking the eventual decision. However, procurement rules were taken into account. What really was going on was that J. Doe and the councillor did not really like each other. The councillor found Doe to be stubborn, while J. Doe had little respect for the councillor, because he had never given any supervision and was appointed as councillor straight from unemployment. J. Doe, on the other hand, had been working since he was 18 years old, and made a career on the job by working hard and taking several training courses at night.

We observe the following scores for the WPI:



The full report of this person is included in Appendix 18.

J. Doe has a very high score on *Influence*. He is very ambitious, dominant and competitive. He has a high need for proving himself. He has an average need to put himself in the spotlight and will have little difficulties in doing so. *Sociability* is, in general, just below average, while the relatively high score on *Leisure contact* is notable. He is not tense in groups of people and will not have social anxiety. In addition, he is more of an introvert than an extrovert; he is especially not so open about what goes on in his mind and not so much focused on others, based on his low score on *Attentiveness*. His trust in others is limited. His high score on *Exuberance* is remarkable: he is energetic, persevering and independent. He also wants to develop himself. In addition, he is averagely innovative and focused on variation and new tasks.

His need for *Structure* is, in general, below average. He has little need for *Regularity*. He is sufficiently orderly and precise for management positions. He is somewhat non-conformist; he will not be so focused on adapting to the group and its norms and values. J. Doe has an average score on *Stability*, in which both his low score on frustration-tolerance and his high score on self-confidence are remarkable.

Situational analysis

Based on J. Doe's scores, it is not hard to understand how his career has unfolded itself. In school he had problems with adapting to the school regime. In addition, he lacked discipline then. He then started working for a cleaning company, where he made a career rapidly by following several training programs and courses. His drive and influence were apparent. After a working as a manager in the cleaning sector he applied for a job as head of cleaning in a middle sized municipality and was offered the job. J. Doe gave the higher salary and status of the job as a motive to take the job: working for a municipality rather than a private cleaning company appealed more to him. In the beginning, everything went well when he was working under a councillor that gave him a lot of freedom. He reorganized the cleaning service successfully with a lot of energy and effort. Under the new councillor, things started to go bad quickly, though it still took two years before the conflict resulted in the dismissal. He competed with the councillor, a competition he was bound to lose, due to their differences in authority and power. The conclusion of this analysis is that it was important for J. Doe to work in an environment in which he was able to work sufficiently autonomous and which was less focused on rules and procedures. Especially the supervisor was important: J. Doe did not respect someone because of his function title, but only when he deserved it. A boss is only a boss when he is right, was his motto.

Advice

We have advised J. Doe to look for work in the profit sector again, in an organization in which managers have a large amount of autonomy. Working for the government can carry too many risks, because of his unconventional character, in combination with his competitive spirit. Managing positions are very suitable for J. Doe, due to his high scores on *Influence* and *Exuberance*. His average scores on *Sociability*, and especially on *Self-disclosure*, *Trust* and *Attentiveness*, make him a task driven supervisor. We advise J. Doe to find a job in which these aspects are required, but to pay attention to developing more people-oriented supervising skills. In the outplacement process itself, we do not expect to encounter a lot of dilemmas. His *Exuberance* is high and his *Stability* is high enough to be able to deal with the layoff. However, an outlet for his frustration needs to be sought. The factor *Influence* is positive for the outplacement process as well, since we can expect that J. Doe will take action and will be able to influence contacts to his success. Because of an average score on *Sociability* we can expect that network meetings will go sufficiently well, although those will have more instrumental meaning. He has a low *Need for contact* and finds it hard to be open about himself. On the other hand, he is self-confident and not socially anxious.

Epilogue

J. Doe found a new job within four months, again as head of cleaning, but now for a large municipality. We could have expected that he would ignore (some of) the advice. He did indicate that this municipality works with integrated management and that he would be fulfilling a more autonomous role. He was not directly involved with politicians, but was supervised by another managing director, which he knew well and respected. In addition, he indicated that he was willing to work on his people-oriented managing skills by taking courses.

6.5.3. The relations between factors and scales with regard to occupational indications

Clear positive or negative correlations between the factors are avoided by the way in which the WPI is set up. High scores on scales within a certain factor are not automatically correlated with high or low scores within a different factor. However, it is possible to distinguish substantive relationships between scales within one factor and scales of another different factor, by which a certain profile can be distilled. Some examples are given in the following:

The project management image

The following scores fit in the profile of a project manager: High scores on the factor *Exuberance*, and more specifically on the scales *Energy*, *Influence*, *Adaptability*, in combination with high scores on the factor *Stability* and more specifically *Resilience* and *Frustration-tolerance*. In addition, lower scores on the *Regularity* scale fit with the project management profile, when they are combined with moderately high scores on the *Orderliness* and *Decisiveness* scales of the *Structure* factor. However, scores that are too high on these scales can form a risk, since it can be a sign of perfectionism. High scores on *Influence* and especially *Dominance* are desirable as well. In short: this is the image of someone who wants to put his or her energy in a job with varying tasks, in which initiative is asked and the will and ability to guide and influence processes. In addition, the ability to respond flexibly to different situations and a smaller need for regularity and predictability, but a larger need for changing projects with different (kinds of) people.

The general management image

High scores particularly on the *Influence*, *Sociability*, *Exuberance* and *Stability* factors fit with the image of a general manager. Within the *Influence* factor, too high scores on *Competition* can be harmful, especially in relation with low scores on *Friendliness*, *Trust* and *Attentiveness*. High scores on these last three scales indicate a people oriented management style. When this style is apparent, one needs to pay attention to the influence of *Decisiveness*. Excessive scores on *Decisiveness* in relation with *Friendliness*, *Trust* and *Attentiveness* can lead to a manager that finds it hard to make decisions that can have negative consequences. In the personality literature, the 'Conscientiousness' factor is found to be an important predictor of job performance. In the WPI, this factor is split up into two factors, namely *Exuberance* and *Structure*. Successful managers often show high scores on *Exuberance* in combination with low scores on *Structure*. Low scores on *Structure* are not necessarily harmful: however, it depends on the role the manager is expected to fulfil. When internal control, task orientation, organizing, setting up and controlling processes are important, scores on *Structure* should not be too low. Scores on *Structure* that are too high can also have negative consequences: due to risks of perfectionism and too low levels of flexibility. Finally, *Stability* is generally a positive characteristic, but should not lead to detachment and unaffectedness when scores are extremely high. Very high scores on *Stability* can indicate that the person is untouched or unmoved by anything, which can lead to inertia.

The commercial image (sales)

In a commercial image focused on 'sales' fit, just like in the management image, particularly high scores on the factors *Influence*, *Exuberance*, *Sociability* and *Stability*. Commercial people often score somewhat lower on *Structure*, although average and above average scores can contribute to commercial success. Within the factors the following scales are most important: *Dominance*, *Competition* and *Self presentation*. These scales contribute to commercial success particularly in competitive markets. High scores on *Need for contact* and *Leisure contact* indicate that people establish contact easily. Very high scores mean that one wants to make new contacts all the time, which is mainly important for acquisition (hunting). These scores combined with low scores on *Attentiveness* and *Trust* indicate a more "hard" sales style. Average scores on *Need for contact* and *Leisure contact* and higher scores on *Attentiveness* and *Trust* indicate a preference for relation management and account management, focused on a smaller group of clients

The advising image

The image of someone who is mostly in an advising role (counsellor or consultant) is mainly made up by high scores on scales such as *Adaptation*, *Originality* and *Independence* belonging to the

Exuberance factor. These scales indicate an 'intellectual' interest. For a lot of advice functions intelligence is of course an important aspect. In addition, *Sociability* is important, although the scores do not need to be extremely high. Average scores are often sufficient. The extent to which this factor is expressed strongly depends on the degree to which contacts are needed. We often see average scores for the *Structure* factor: structured enough to work carefully, but not getting caught up in details. The scale *Decisiveness* should not be too low, because this might lead to advice being given too impulsively. Scores on *Decisiveness* that are too high can lead to a consultant that gets caught up in the analyses and is afraid of giving any advice.

The specialistic image

High scores on the *Structure* factor and somewhat lower scores on *Influence* fit with the image of someone who operates substantively as a specialist. High scores on *Structure* ensure that someone delves deeper into the content of the job, to achieve a high degree of knowledge and perfection. The other factors depict more the type of specialist. A controller, for example, should have some influence, in order to advise and control supervisors. An engineer will sometimes need to be original, to come to new solutions. Too low scores on *Sociability* indicate a solistic view of work, which can become a problem when working in project teams. *Adaptability*, the need for new tasks, can sometimes be fed in the specialism, by continuously gaining and using new knowledge. But when the need for change is too large, one would often want to become a generalist rather than a specialist.

The service image

The factor *Sociability* is very important for service occupations. The *Influence*-scales can be low to average. Someone with very low scores on *Influence* is subservient (in a positive way), but can sometimes show a lack of initiative. For teachers, for example, one would expect at least some level of *Independence*, *Self-presentation* and *Dominance*. *Stability* appears to be important as well, to be able to deal with the work load and pressure of others. In general, it is important to know which services are being provided and how the scales influence these services. The *Structure* factor is often helpful in providing services in a careful and planned way. Again, the type of work will determine what the desirable scores are.

6.5.4. Consistency, self-image and response tendency

Three measures are added to the WPI in order to limit biases caused by social desirability, response tendencies and inconsistencies.

Consistency indicates the degree to which someone has given a similar response to an item pair that correlates highly within the norm group. The Self-image measure indicates the degree to which someone shows a critical or positive image of him/her self. Low scores indicate that the candidate has predominantly chosen for the less positive options, and therefore comes across as more modest. High scores indicate that a candidate is not afraid of choosing the extreme options and probably is trying to provide a very good image of him/her self. Whether one's scores are accurate cannot be assessed by this measure. The response tendency measure indicates whether someone has chosen the extreme options or has chosen to be on the safe side and choose for the middle options. The percentages indicate whether the candidate has chosen more often or less often for a particular option on the 5-point Likert scale, compared with the norm group. For a more detailed description of these measures, we refer the interested reader to <http://www.test-toolkit.nl/consistentie-zelfbeeld-en-antwoordtendientie-van-de-wpv>.

6.5.5 Competency model of the WPI

The WPI includes a competency model. We have included this model in order to satisfy the need, coming from the HR field, to directly get an overview of the competencies of a candidate. In constructing this model, we have adopted the same strategy as for the personality questionnaire. Assessment psychologists were asked to write down competencies of which they wanted to know their developability. Of all the noted competencies, 40 were selected. The criteria for this selection were that all domains and client criteria should be covered, and that the competencies were not defined too narrowly or too widely. Subsequently, the development team of Ixly determined whether scales were

positive or negative predictors of a certain competency. The results of this process are presented in the competency matrix.

The scores on the scales (high, average or low) have different levels of predictive value for a certain competency. Certain scores can also have a negative influence on a competency. This is also included in the matrix. In the first stage, we used a linear model. A low sten score on a certain scale indicated a low level of developability of a certain competency. When the sten score increased, the developability increased in a linear way as well. This model was tested in practice by Ixly. At that time, the model was not used in career advice situations. Psychologists verified whether the developability of the competencies of the candidates as provided by the model of the WPI corresponded to their own ideas about the developability. In addition, candidates were asked whether they recognized themselves in the profile. In the cases where the results of the model and the ideas of the psychologists did not correspond, the cause of this discrepancy was investigated. On the basis of this investigation, the model was adapted. After this, the new model was tested with new candidates and again compared with the assessments of psychologists, etcetera. The model is thus constructed in an iterative way by using concrete, real-life cases.

The main finding was that the linear model did not hold and that a non-linear, curvilinear model with an optimum per sten score was more suitable. The main idea is that “more” is not always better (or worse), but that there are one or more optimal scores that are not necessarily positioned at the ends of the scale. In practice, extreme scores on personality characteristics are rarely optimal. In the development of this competency model, we have gone through the first steps. The model is constructed on the basis of theory and experiences in the field. Again, for this model, a process will be followed in which the results are compared with the findings of assessment psychologists. The final 29 competencies are listed in Appendix 19, with a description of each competency. The competency scores that follow from the model should be used as an aid, and not as a selection tool. This is why we have chosen the name “Competency Indicator”. By continuing to do research on the competencies, it would be possible to eventually use it in selection situations in the future.

Literature

- Akkerman, A. E., & Buijk, C. A. (1994). *Handleiding GLTS '94*. Lisse: Swets en Zeitlinger.
- Allport, G. W., & Odbert, H. S. (1936). Trait names: A psycho-lexical study. *Psychological Monographs*, 47 (211), 171.
- Aries, E. (1976). Interaction patterns and themes of male, female, and mixed groups. *Small Group Behavior*, 7 (1), 7-18.
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: a meta-analysis. *Personnel Psychology*, 44, 1-26.
- Beutel, A., & Marini, M. M. (1995). Gender and values. *American Sociological Review*, 60 (3), 436-448.
- Bloemers, W., & Molen, H. van der (2004). Het (genegeerde) belang van intelligentie voor arbeidsprestaties. *Opleiding & Ontwikkeling*, mei 2004, 3-9.
- Bunk, R., Slooter, M., Vergert, M. ten, & Verheul, C. (2002). *Help competentie management, handreiking bij de invoering*. Amersfoort: CPS onderwijsontwikkeling en advies.
- Catell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology*, 38 (4), 476-406.
- Centraal Bureau voor de Statistiek, StatLine. Op 20 december 2007 ontleend aan <http://statline.cbs.nl/statweb/>
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112 (1), 155-159.
- Drenth, P. J. D. (1981). *De Psychologische test*. In H. C. J. Duijker & P. A. Vroon (red.), *Codex Psychologicus*. Amsterdam/ Brussel: Elsevier.
- Drenth, P. J. D., & Sijtsma, K. (2006). *Testtheorie: inleiding in de theorie van de psychologische test en zijn toepassingen*. Houten: Bohn Stafleu van Loghum.
- Dusschooten, C. de. (2004). *Mensen maken het verschil: de relatie tussen persoonlijkheid, waarden en organisatiecultuur*. Assen: Van Gorcum.
- Eijck, K. van, Graaf, P. M. de (2001). De invloed van persoonlijkheidskenmerken op het bereikte opleidingsniveau. *Mens en Maatschappij*, 76 (4), 285-302.
- Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed.), *Review of Personality and Social Psychology*, 2, 141-165. Beverly Hills, CA: Sage.
- Gordon, L. V. (1963). *Gordon Personal Inventory/ Gordon Personal Profile (GPP/ GPI)*. New York: Harcourt, Brace & World.
- Green, S. B., Salkind, N. J., & Akey, T. M. (2002). *Using SPSS for windows: Analyzing and understanding data (2nd ed.)*. New Jersey: Prentice Hall.

- Hendriks, A. A. J., Hofstee, W. K. B., & Raad, B. de (1999). Handleiding bij de Five- Factor Personality Inventory (FFPI). Amsterdam: Harcourt Test Publishers.
- Hergenhahn, B. R. (1980). An introduction to theories of personality. New Jersey: Prentice Hall.
- Hoekstra, H. A., Ormel, J., & Fruyt, F. de (1996). Handleiding NEO Persoonlijkheids vragenlijsten NEO-PI-R en EO-FFI. Lisse: Swets & Zeitlinger.
- Hofstee, W. K. B., Brokken, F. B., & Land, H. (1981). Constructie van een Standaard Persoonlijkheids-Eigenschappen-Lijst (SPEL). Nederlands Tijdschrift voor de Psychologie en haar Grensgebieden, 36 (6), 443-452.
- Hofstee, W. K. B., & Raad, B. de (1991). Persoonlijkheidsstructuur: de AB5C-taxonomie van Nederlandse eigenschapstermen. Nederlands Tijdschrift voor de Psychologie, 46, 262-274.
- Hofstee, W. K. B., Raad, B. de, & Goldberg, L. R. (1992). Integration of the Big Five and circumplex approaches to trait structure. Journal of Personality and Social Psychology, 63, 146-163.
- LaForge, R. (1985). The early development of the Freedman-Leary-Coffey Interpersonal System. Journal of Personality Assessment, 49, 613-621.
- LaForge, R., & Suczek, R. R. (1955). The interpersonal dimension of personality: III. An interpersonal check list. Journal of Psychology, 24, 94-112.
- Luteijn, F., Starren, J., & Dijk, H. van (1985). Handleiding Nederlandse Persoonlijkheids Vragenlijst. Lisse: Swets & Zeitlinger.
- McCrae, R. R., & Costa, P. T. Jr. (1994). The stability of personality: Observations and evaluations. Current Directions in Psychological Science 3 (6), 173-175.
- Nunnally J. C. (1978). Psychometric Theory. USA: McGraw-Hill, inc.
- Orga (2002). Handleiding ProSiD-PI: Professional Six Dimensional Personality Inventory. Niet gepubliceerd. Zeist: Orga, intern verslag.
- Orga. Op 20 december 2007 ontleend aan <http://www.orga-toolkit.nl>
- Salgado, J. F., & Fruyt, F. de (2005). Personality in personnel selection. In A. Evers, N. Enderson & O. Smit-Voskuil (Eds.), The Blackwell handbook of personnel selection (pp. 175-197). USA: Blackwellpublishing.
- Salgado, J. F. (2003). Predicting job performance using FFM and non-FFM personality measures. Journal of Occupational and Organizational Psychology, 76, 323-346.
- Schmidt, F. L., & Hunter, J. H. (1998). The Validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. Psychological Bulletin, 124, 262-274.
- Schultz, D. P., & Schultz, S. E. (2002). Psychology & Work Today. New Jersey: Prentice Hall.
- Snijders, J. Th., Tellegen, P. J., & Laros, J. A. (1988). S.O.N.-R 5½-17 Verantwoording en Handleiding. Groningen: Wolters-Noordhoff BV.
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. Psychological Bulletin, 87, 245-251.
- Stuive, I., Kiers, H. A. L., Timmerman, M. E., & Berge, J. M. F. ten (2008). The empirical verification of an assignment of items to subtests: Oblique Multiple Group Method versus Confirmatory Common Factor Analysis. Unpublished manuscript.

Tjoa, A. S. H. (1993). Herziene handleiding Edwards Personal Preference Schedule. Lisse: Swets & Zeitlinger.

Woud, M. E. A. van der (2007). Normering: Normering van testcores: toepassing van een semi-parametrische fitprocedure, en standaardisatie van de verdeling van ware scores. Niet gepubliceerd. Zeist: Orga, intern verslag.

Tables

Table 2.1	Factor- and scale structure of the WPI
Table 3.1	Description of the background variables used in the weighing procedure
Table 3.2	Frequencies of background characteristics in both norm groups
Table 3.3	Distribution of regions in both norm groups
Table 3.4	Distribution of work sectors in both norm groups
Table 3.5	Average stenscore of the factors and background variables
Table 4.1	Reliability and generalizability of factors for both norm groups
Table 4.2	Reliability of the scales for both norm groups
Table 4.3	Correlations with the second administration (N=53)
Table 4.4	Mean standardized factor scores and corresponding t- and p-values (N=53)
Table 5.1	Mean stenscore and standard deviation for each factor
Table 5.2	Correlations between the factors of the FFPI and the WPI
Table 5.3	Correlations between scales of the WPI and factors of the FFPI (N=353)
Table 5.4	Mean stenscore in this sample in comparison with the sten scores of the norm group (N=80)
Table 5.5	Correlations competencies 360test and standardized factor scores WPI (N=53)
Table 5.6	Description of the sample of Agency X
Table 5.7	Mean stenscores of the factors and scales for both function groups of Bureau X (norm group used: Advice)
Table 5.8	Linear regression (x) WPI scales and (y1) turnover per function for Agency X (Norm group Advice)*
Table 5.9	Linear regression (x) WPI scales and (y2) supervisor rating per function for Agency X (Norm group Advice)*
Table 5.10	Linear regression (x) WPI scales and (y3) satisfaction per function for Bureau X (Norm group Advice)*
Table 5.11	MGM results for the Advice group (N=712)
Table 5.12	MGM results for the Selection group (N=369)
Table 5.13	Correlations between factors
Table 5.14	Rotated component matrix for the Advice group (N=712)
Table 5.15	Rotated component matrix for the Selection group (N=369)
Table 5.16	Mean stenscores by gender including eta-value
Table 5.17	Mean stenscore by age including eta-value
Table 5.18	Mean stenscore by education including eta-values
Table 5.19	Mean sten score by job status including eta-values
Table 5.20	Differences in mean sten scores on the factor level between sectors in the Advice group (N=719)
Table 5.21	Differences in mean sten scores on the factor level between sectors in the Selection group (N=1067)
Table 5.22	Significant differences in mean sten scores on the scale level between sectors in the Advice group
Table 5.23	Significant differences in mean sten scores on the scale level between sectors in the Selection group
Table 5.24	Mean sten scores and corresponding standard deviations (N=1329)
Table 5.25	Correlations between the factors of the WPI and the scales of the CV
Table 6.1	Sten scores with corresponding percentages

Appendices

- Appendix 1 MGM results: items in scales for the Advice group and the Selection group (Available on request*)
- Appendix 2 Differences between the ProSiD-PI 35 en the WPI
- Appendix 3 Differences between the ProSiD-PI 25 en the WPI
- Appendix 4 Definitions of the factors and the scales of the WPI
- Appendix 5 Dutch labour force in terms of gender, age, level of education and work status
- Appendix 6 Details for the weights of the reference groups
- Appendix 7 Woud, M. E. A. van der (2007). Normering: Normering van testcores: toepassing van een semi-parametrische fitprocedure, en standaardisatie van de verdeling van ware scores. Niet gepubliceerd. Zeist: Orga, intern verslag. (Available on request*)
- Appendix 8 Normed scale scores for both norm groups (Available on request*)
- Appendix 9 Factors of the FFPI
- Appendix 10 Correlations between the factors of the FFPI and the scales of the WPI (Available on request*)
- Appendix 11 Sogeti MGM (Available on request*)
- Appendix 12 Sten scores for each work sector, by norm group (Available on request*)
- Appendix 13 Study on cultural bias
- Appendix 14 Definitions of the Career Values
- Appendix 15 Correlations between the scales of the WPI and the scales of the Career Values (Available on request*)
- Appendix 16 Manual for consultants – operation and possibilities of the Toolkit (*in revision*)
- Appendix 17 List of frequently asked questions for the use of the Toolkit (*in revision*)
- Appendix 18 Report for career advice
- Appendix 19 Definitions of the competencies

*These appendices are not included here because of their large sizes. They can be requested digitally via helpdesk@ixly.nl