

WORK PERSONALITY INDEX

The Work Personality Index®

2nd Edition Development and Norming

Donald Macnab
Shawn Bakker
Psychometrics Canada Ltd.

The WPI was designed to provide an efficient and useful measure of personality for workplace applications. To meet this goal, the development of the items and the creation of test norms conformed to a specific set of procedures.

The starting point in the development process was to set design criteria that would guide the progress of the WPI's creation. In order to meet the overall goal to develop a comprehensive but efficient work personality measure, 5 key criteria were established.

- 1.** The WPI should only measure personality traits that are directly related to the work environment and job performance.
- 2.** The WPI should be useful for applications such as personnel selection, team building, leadership development and personal development.
- 3.** The application and interpretation of the WPI should not require specialist training in psychology or personality.
- 4.** The questions should be easy to read and comprehend.
- 5.** Respondents should be able to complete the assessment in a timely manner.

The design criteria guided the development of the WPI in several ways. The personality traits measured by the WPI have been found to relate to work preferences and job performance in a direct way. This ensures that users are provided information that directly helps in the accurate selection or informed development of personnel. In traditional personality assessments, job relevant traits are measured alongside non-job relevant traits, making the interpretation of the results more difficult and thereby increasing the chances for making errors.

Alongside the selection of work related personality traits, the application and interpretation of the WPI was designed for the typical business professional. Many tests require the services of a psychologist or specially trained human resource professionals. However, the WPI avoids complex personality factors, making the results easily interpretable.

Development Process

As mentioned earlier in this chapter, the WPI is built upon the 17 personality traits identified in O*NET, the occupational classification system designed by the U.S. Department of Labor. After conducting an extensive literature review and examination of current personality measures, researchers identified 17 traits that were commonly found to relate to and predict work performance. These 17 traits provided the foundation of the WPI. After 10 years of use and research with the WPI it was established that additional scales were required over and above the original 17 traits. The traits added included Persuasion, Social Confidence, Multi-Tasking and Planfulness. Having identified the 21 traits, the following steps were taken to construct the WPI.

Scale development and item writing

To construct a measure of the 21 traits, a thorough review was conducted of research studies and current personality measures in order to develop objective, operational definitions for each trait. At this stage, the definitions developed for the O*NET model were modified to better reflect research findings. After having created definitions for each personality trait, items that provided behavioural evidence for the traits were developed. The items were written according to the following rules:

1. Items should relate directly to work behaviour. The situation presented in the item needed to examine preferences and motivations for different types of work and work environments.
2. Items were to be written in the first person (e.g. "I am... ", "At work I... ", "For me... "). Since the WPI is a self-report measure, it was necessary to write items in the first person.
3. Items should target a single construct. It was essential for the situation outlined in each item to be related to only one of the twenty-one personality traits. The items were designed around behaviours that provided evidence of one specific trait.
4. Items should be short, direct, and easy to understand.
5. A five-point Likert scale was developed to allow individuals to register the extent of their agreement or disagreement with each statement. The respondents could indicate their preferences by choosing one of the five options for each item.

For example:

People describe me as understanding.

Strongly Disagree Disagree Neutral Agree Strongly Agree

In addition to the items from the first version of the WPI more than 300 items were written and field-tested over a period of 18 months. Over 15000 individuals took various forms of the assessment during this time period. A combination of classical test analysis and Rasch analysis of the items for each scale were examined.

The Rasch model (Rasch, 1980; Wright, 1979, 1982) is a mathematical formula that indicates the relationship between persons and scores that define a trait. The model is usually referred to as a one – parameter model, but in fact looks at two parameters: people and items. These are usually defined as person logits and item logits. The analysis provides a number of fit statistics that are indices of how well the data fit the model. The fit statistics include average fit and individual item fit. Fit statistics include “infit”, a chi-squared based statistic that is sensitive to unexpected patterns of observations by persons on items; and “outfit” statistics a chi-squared based statistic that is sensitive to unexpected observations by persons on items. High outfit or outfit mean squares greater than 1.5 indicate that an item is not fitting the model reasonably well. In addition, an examination of the mean infit and outfit statistics presents overall information about whether the data acceptably fit the model. It is expected that the mean infit and outfit for both persons and items should be about 1.0. At the level of an individual item the accepted general rule regarding infit and outfit mean square fit statistics is they should be between 0.5 and 1.5. (Linacre, 1995, 2002, 2004, 2010).

In addition, the Rasch model allows the examination of how the items perform for different groups of people. This is known as Differential Item Functioning (DIF) (Zwick and Thorpe, 1996). DIF is carried out by comparing two groups of interest (for example, males and females). One set of parameters used to judge this are call DIF Contrast Logits. In general logits under 0.43 are considered to indicate negligible differences between groups; those between 0.44 and 0.64 to be slight, and those over 0.65 to be moderate or large (Linacre, 2010).

Through this process, the research questionnaire consisting of 221 items (10 items per scale) was used to collect normative data. Examination of the large data pool from the normative study was once again analyzed using Rasch technology. This analysis suggested that the scales could function as well if not better by being reduced by 1 item per scale. Internal consistency and item reliability analysis resulted in the dropping of 21 items, bringing the total to 189 or 9 items per scale. The items that were retained contributed positively to coefficient alpha reliability and ensured that the number of items on the

questionnaire remained manageable for users. Internal consistency reliabilities range from .70 to .89. The Rasch analysis for the final 189 item version of the Work Personality Index is summarized in Table 1.

Table 1 Rasch Analysis Summary for the Work Personality Index (n=5808)

WPI Scale (n=5808)	Average Mean Square Outfit	Average Mean Square Infit	Outfit	Infit	DIF Contrast Logits above negligible
Ambition	1.03	1.03	0	1	1
Analytical Thinking	1.02	1.02	0	0	0
Attention to Detail	1.03	1.04	0	0	0
Concern for Others	1.06	1.03	0	0	0
Democratic	1.01	1.01	0	0	0
Dependability	1.02	1.02	0	0	1
Energy	1.02	1.01	0	0	0
Flexibility	1.00	1.00	0	0	0
Initiative	1.02	1.02	0	0	0
Innovation	1.02	1.02	0	0	0
Leadership	1.03	1.02	0	0	0
Multi-Tasking	0.99	0.99	0	0	0
Outgoing	1.02	1.01	0	0	0
Persistence	1.02	1.02	0	0	0
Persuasion	1.01	1.02	0	0	0
Planning	1.03	1.03	0	0	0
Rule-Following	1.00	1.00	0	0	0
Self-Control	1.02	1.01	0	0	0
Social Confidence	1.03	1.03	0	0	0
Stress Tolerance	1.02	1.02	0	0	0
Teamwork	0.99	1.02	0	0	0

The table above indicates that each of the 21 WPI scales shows very good model fit with average mean square outfit and average mean square infit statistics being extremely close to 1.00. Table 5.1 also displays the number of items that do not fit the model as measured by the item mean squared (infit or outfit). One item on the Ambition scale has a marginally high Outfit mean square. Considering that there

are 189 item items in the total test this is an extremely positive result. Two items show DIF that is above the negligible level (both would be classified as slight to moderate); one in the Dependability scale (0.44) (Females higher than Males) and one on the Ambition scale (-0.51)(males higher than females).

Comparing versions of the Work Personality Index

As part of the development process we did not wish for the new version of the WPI to fundamentally deviate fundamentally in psychometric characteristics from the previous versions. For the common scales we therefore expected that the alternate form reliability would be high, as would the factorial structure of the two assessments. The previous version of the Work Personality Index consists of 153 items representing 17 scales. The new version consists of the same 17 scales with the addition of Multi-Tasking, Planning, Social Confidence, Persuasion and an 8 item validity scale. The first step in judging the comparability of the new and previous versions of the WPI was to look at alternate forms of reliability for the scales. This is a measure of how the individual scales compare across versions. 1852 individuals completed both versions of the WPI. In general, observed correlations between the two forms are shown in Table 2. Correlations are very high indicating that the traits measured by both versions are similar and that any resulting interpretation would also be similar.

Table 2 Alternate forms reliability of the Work Personality Index (n=1852)

Scale	r
Ambition	0.89
Analytical Thinking	0.98
Attention to Detail	0.84
Concern for Others	0.96
Democratic	0.97
Dependability	0.94
Energy	0.93
Flexibility	0.93
Initiative	0.91
Innovation	0.96
Leadership	0.93
Outgoing	0.96
Persistence	0.96
Rule-Following	0.98
Self-Control	0.95
Stress Tolerance	0.95
Teamwork	0.97

Additionally, the factor structures of the 17 common scales across the two forms were compared. The inter-correlation matrices of the 17 WPI scales for each version of the WPI were calculated. A principal components analysis was applied to each correlation matrix with varimax rotation. A comparison was made between the factor structures using procedures and methods outlined by Barrett, P. T., Petrides, K. V., Eysenck, S. B. G., & Eysenck, H. J. (1998). Comparison of the factor solutions with varimax rotation for the two assessments was made using congruence analysis, following the procedures outlined by Barrett (1986). In the analysis the North American norming sample was used as the target matrix with comparisons made to data gathered on the new version of the Work Personality Index. Results are shown in Table 5.3. All congruence coefficients are above 0.90 which is indicative of congruence between factors (Barrett, 1986; Ten Berge, 1986). The overall coefficient of congruence (0.99) for the analyses demonstrates high similarity for all five factors. The present study sought to demonstrate the

initial validity and factor invariance of the Work Personality Index scales across the two versions of the assessment. The number and content of the factors are similar. All five factors showed near perfect equivalence across versions. The results suggest that the factor structure of the Work Personality Index is consistent across versions. At this level, participants responded to the two versions of the Work Personality Index in a highly similar fashion. Overall, this study supports the validity of the Work Personality Index factor structure. This gives administrators confidence that the Work Personality Index may be utilized across versions with similar interpretations.

Table 3 Coefficients of Congruence for Work Personality Index Version 1 and 2

	Congruence Coefficients	Congruence Coefficients between the target (WPI Version 1) and maximally congruent comparison matrix (WPI Version 2)					
		Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	
<i>Energy and Drive</i>							
Ambition	0.98						
Initiative	0.99	1	0.98	0.23	0.23	0.51	0.35
Flexibility	0.97	2	0.22	0.99	0.22	0.08	0.18
Energy	0.95	3	0.21	0.04	0.99	0.05	0.28
Leadership	0.98	4	0.50	0.08	0.17	0.98	0.24
		5	0.35	0.18	0.39	0.25	0.97
<i>Work Style</i>							
Persistence	1.00	Overall Solution Congruence = 0.98					
Attention to Detail	0.96						
Rule-Following	0.97						
Dependability	0.99						
<i>Working with Others</i>							
Teamwork	1.00						
Concern for Others	0.99						
Outgoing	0.97						
Democratic	0.97						
<i>Problem Solving</i>							
Innovation	0.99						
Analytical Thinking	0.99						
<i>Dealing with Pressure and Stress</i>							
Self-Control	0.99						
Stress Tolerance	0.99						

Norming of the WPI

Norming is a key step in test development. The norms establish the baseline by which all test results are measured against, which allows the comparison of different individual scores. Norms identify the below average, average, and above average performance on the test, and help the test user appropriately interpret a person's results and make decisions. The more people that are included in the norm sample the more we can ensure that the test norms represent the actual distribution of personality traits of the people in the population. In turn, this allows the test results to be more accurate and informative when comparing different individuals.

The WPI was standardized using large sample of 8360 people (4180 females and 4180 males). The large number of participants in the norm sample ensures that the WPI results accurately represent personality traits in the target population.

Sten Scores

A person's results on the WPI are reported in a standard score format known as Sten Scores. Standard scores are converted raw scores that help with the interpretation of the test results by allowing the comparison of an individual's results with the norm group. Standard scores also help compare a person's primary scale scores against each other. This allows us, as an example, to determine if the person scores higher on Flexibility than on Persistence. Sten scores range from 1 to 10, have a Mean of 5.5, and a Standard Deviation of 2. This means that an individual with a Sten score of 5.5 precisely reflects the average score of the norm population. As a result, 50 percent of the norm sample would score above and below the individual.

Means and Standard Deviations

Table 4 provides a detailed description of the WPI raw scale scores for the norming sample. The means and standard deviations shown provide the norms which individuals who complete the WPI are compared against. The mean raw score for each scale represents the "average" score of people in North America. The standard deviation indicates the spread of scores found among people in the normative sample. Approximately 68% of the population will obtain scores within one standard deviation above and below the mean, while 95% of the population will score within two standard deviations of the mean.

Table 4 Raw Score Means and Standard Deviations for the norming sample (n=8360)

	Minimum	Maximum	Mean	Std. Deviation
Ambition	12	45	31.81	5.50
Analytical Thinking	9	45	33.90	5.39
Attention to Detail	11	45	31.60	5.85
Concern for Others	9	45	34.88	5.81
Democratic	9	45	25.43	4.32
Dependability	10	45	32.51	5.53
Energy	9	45	30.24	5.37
Flexibility	10	45	28.35	4.99
Initiative	12	45	34.16	4.84
Innovation	9	45	33.33	5.93
Leadership	9	45	30.81	6.24
Multi-Tasking	9	45	27.48	6.45
Outgoing	9	45	30.07	5.98
Persistence	9	45	32.71	5.66
Persuasion	9	45	28.79	6.89
Planning	11	45	31.07	5.51
Rule Following	9	45	27.41	6.49
Self-Control	9	45	30.40	6.11
Social Confidence	9	45	31.48	7.14
Stress Tolerance	9	45	28.96	6.45
Teamwork	9	45	30.16	5.39

Standard Error of Measurement

Every time someone completes the WPI, their scores are comprised of two components. The first component is their true score – the amount of their score represents the personality trait being measured. The second component is random score – the proportion of their score that is attributable to external conditions which have nothing to do with the trait being measured, known as chance or external influence. For example, if an individual is distracted or tired, it may impact their responses, and thus the score they obtain on the WPI. Since this random score has a negative influence on the consistency of an individual's results, it is usually called random error. The Standard Error of Measurement (SEM) gives an indication of how much an individual's obtained score might vary from their true score. The SEM can be interpreted as meaning that an individual's true score will fall within +1

or- 1 SEM 68% of the time, and within +or- 2SEM's 95% of the time. To illustrate, consider the following. Paul receives a Sten score of 6 on the Ambition scale. We know that 68% of the time, his true score will lie in between 5.09 and 6.91 (+-1 SEM). We also know that 95% of the time, his true score will lie between 4.17 and 7.83(+2 SEM). Table 5 shows the Standard Error of Measurement (SEM) for the 21 WPI scales.

Table 5 Standard Error of Measurement for WPI Scales and Confidence Intervals (n=8360)

	SEM (as sten score)	68% Confidence Interval	95% Confidence Interval
Ambition	.87	1	1.5
Analytical Thinking	.87	1	1.5
Attention to Detail	.80	1	1.5
Concern for Others	.72	1	1.5
Democratic	1.13	1	2
Dependability	.85	1	1.5
Energy	.89	1	2
Flexibility	.94	1	2
Initiative	.87	1	1.5
Innovation	.69	.5	1.5
Leadership	.72	.5	1.5
Multi-Tasking	.69	.5	1.5
Outgoing	.80	1	1.5
Persistence	.80	1	1.5
Persuasion	.66	.5	1.5
Planning	.89	1	1.5
Rule-Following	.69	.5	1.5
Self-Control	.82	1	1.5
Social Confidence	.60	.5	1.5
Stress Tolerance	.77	1	1.5
Teamwork	.85	1	1.5

Demographics of the Norming Sample

The final sample consists of 8360 individuals (4180 females and 4180 males). A detailed breakdown of the Age, Education Level, Ethnicity, Employment Status, Position Level, Work Experience and Occupational Area of the norming sample is provided in Tables 6 to Table 12.

Table 6 – Age Distribution of Norming Sample (n=8360)

Age Group	Total	Percent	% Female	% Male
15-17	450	5.4	4.5	6.3
18-20	430	5.1	4.7	5.6
21-24	920	11.0	10.5	11.5
25-28	917	11.0	10.8	11.1
29-34	1299	15.5	14.7	16.3
35-44	1779	21.3	21.8	20.8
45-54	1505	18.0	20.5	15.5
55-65	911	10.9	11.4	10.4
65+	149	1.8	1.0	2.5

Table 7 – Highest Education Level Achieved by Subjects in Norming Sample (n=8360)

	Total	Percent	% Female	% Male
Some High School	515	6.2	5.1	7.2
High School	666	8.0	7.1	8.9
Trade	274	3.3	2.8	3.7
Some College	1871	22.4	20.3	24.5
Associate Degree	365	4.4	4.5	4.2
Community College	365	4.4	5.3	3.4
Bachelors	2166	25.9	29.7	22.1
Masters	1592	19.0	19.3	18.8
Professional	277	3.3	3.6	3.0
Doctorate	271	3.2	2.3	4.2

Table 8 – Ethnic Origin of Subjects in Norming Sample (n=8360)

	Frequency	Percent	% Female	% Male
Asian	262	3.1	3.1	3.2
African American/Canadian	682	8.0	7.6	8.8
Latino/Latina	397	4.7	4.6	4.9
Middle Eastern	109	1.3	0.8	1.8
Native American/Canadian	144	1.7	1.9	1.5
South East Asian	101	1.2	0.9	1.5
White/Caucasian	5652	67.6	68.7	66.6
Other	1013	12.1	12.5	11.8

Table 9 – Employment Status of Subjects in Norming Sample (n=8360)

	Frequency	Percent	% Female	% Male
High School Student	485	5.8	4.6	6.8
College Student	728	8.7	8.5	8.9
Employed Fulltime	4554	54.5	53.3	55.6
Employed Part-time	622	7.4	9.2	5.6
Retired	78	0.9	0.9	1.0
Homemaker	137	1.6	2.8	0.5
Self-employed	808	9.7	9.6	9.7
Seeking	948	11.3	10.8	11.9

Table 10 – Position Level of Subjects in Norming Sample (n=8360)

	Frequency	Percent	% Female	% Male
Entry level	1068	12.8	12.5	13.0
Non-supervisory employee	2257	27.0	30.2	23.8
Management	1486	17.8	16.4	19.1
Supervisor	851	10.2	8.2	12.2
Executive	478	5.7	4.5	6.9
Top Executive	234	2.8	1.8	3.8
None Given	1986	23.8	26.4	21.1

Table 11 – Years of Work Experience of Subjects in Norming Sample (n=8360)

	Frequency	Percent	% Female	% Male
Less than 1 year	742	8.9	8.4	9.4
1-2 years	831	9.9	10.0	9.9
3-5 years	1250	15.0	14.8	15.1
5-10 years	1586	19.0	19.7	18.2
More than 10 years	3024	36.2	35.2	37.2

None Given	927	11.1	11.9	10.2
------------	-----	------	------	------

Table 12 – Occupational Area of Subjects in Norming Sample (n=8360)

	Frequency	Percent	% Female	% Male
Agriculture, Fishing and Forestry	44	0.5	0.6	0.4
Arts or Design	170	2.0	2.2	1.8
Building and Maintenance	45	0.5	0.1	1.0
Business or Financial	785	9.4	9.5	9.3
Community and Social Services	394	4.7	6.5	2.9
Construction	77	0.9	0.4	1.5
Customer Service-Call Center Support	148	1.8	2.4	1.1
Customer Service-Technical support	76	0.9	0.6	1.2
Education or Training	1315	15.7	17.1	14.4
Engineering	224	2.7	1.4	3.9
Entertainment	80	1.0	0.9	1.0
Food Preparation and Serving	223	2.7	2.7	2.7
Healthcare Practitioner	230	2.8	3.3	2.2
Healthcare Support	272	3.3	4.7	1.8
Information systems and technology	381	4.6	2.2	6.9
Installation, Maintenance and Repair	69	0.8	0.2	1.5
Journalism or Media	104	1.2	1.5	1.0
Legal Occupations	141	1.7	1.7	1.7
Library Sciences	38	0.5	0.7	0.2
Life or Physical Science	125	1.5	1.9	1.1
Management	877	10.5	10.5	10.5
Manufacturing	146	1.7	1.1	2.4
Mathematics	28	0.3	0.2	0.5
Military	467	5.6	2.1	9.1
Mining	26	0.3	0.1	0.5
Office and Administrative Support	531	6.4	10.4	2.3
Personal Care and Service	110	1.3	1.9	0.7
Protective Services	98	1.2	0.5	1.9
Retail Sales	458	5.5	5.4	5.6
Sales Management	173	2.1	1.5	2.6
Social Science	268	3.2	4.3	2.2
Sports	59	0.7	0.4	1.1
Transportation	133	1.6	0.9	2.2
Wholesale Sales	45	0.5	0.2	0.9