



Test Partnership

Perceptions Series

Psychometric Properties

2018

1 Construct Validity

Conscientiousness

Of the constructs included in the Five-Factor Model of personality, Conscientiousness is arguably one of the most predictive of job performance (Barrick & Mount, 1991; Rothstein & Goffin, 2006; Salgado, 2003). For this reason, we opted to measure Conscientiousness along with its proposed sub-scales. The IPIP-NEO-120 (Johnson, 2014) is a validated measure of the Big-5 model of personality, specifically the OCEAN model (Costa & McCrae, 1992). The full questionnaire is comprised of 120 items, with multiple sub-scales per superordinate construct.

To validate Test Partnership's Perceptions Series, the following sub-scales were taken from the IPIP-NEO-120, Conscientiousness (Johnson, 2014): Self-Efficacy (C1), Orderliness (C2), Dutifulness (C3), Achievement-Striving (C4), Self-Discipline (C5) and Cautiousness (C6).

Conscientiousness is a composite of the sub-scales shown and provides a measure of general Conscientiousness.

Table 1.1: Correlations between Perceptions SJTs and Conscientiousness

IPIP-NEO-120 Scale	Perceptions SJT			
	Admin	Graduates	Managers	Sales
C1	0.40***	0.35***	0.25**	0.51***
C2	0.28***	0.20^	0.10^	0.22*
C3	0.43***	0.19^	0.33***	0.39***
C4	0.44***	0.44***	0.27**	0.43***
C5	0.41***	0.37**	0.25**	0.35***
C6	0.20***	0.20*	0.18*	0.32***
Conscientiousness	0.44***	0.37***	0.27***	0.44***

Note: ^ = Non-Significant, * = Significant at $p < 0.05$, ** = Significant at $p < 0.01$, *** = Significant at $p < 0.001$. All correlations have been corrected for reliability, but not restriction of range.

Each Situational Judgement Test within the Perceptions Series demonstrates significant correlations with a validated measure of Conscientiousness. All significant correlation coefficients for can be interpreted as ranging from small to moderate, excluding the correlation

between Perceptions Sales and Self-Efficacy which can be interpreted as large (Cohen, 1988).

Additional Personality Traits

The Big-5 model of personality is comprised of the following constructs: Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Emotional Stability. A short form of the IPIP-NEO was administered to gain measures of the four additional constructs within the Big-5. The mini-IPIP (Donnellan, Baird & Lucas, 2006), which is a short-form questionnaire derived from the IPIP-NEO, was chosen as a means of reducing the risk of candidate fatigue. The trait “Neuroticism” was reversed to provide a measure of Emotional Stability, and the trait “Imagination” provided a measure of Openness to Experience.

Table 1.2: Correlations between Perceptions SJTs and Additional Big-5 Personality Traits

Aptitude Test	Perceptions SJT			
	Admin	Graduates	Managers	Sales
Extraversion	0.13***	0.15*	0.21**	0.11^
Agreeableness	0.20***	0.06^	0.23***	0.23**
Emotional Stability	0.18***	0.14*	0.18**	0.23**
Openness to Experience	0.22***	0.24***	0.21**	0.25***

Note: ^ = Non-Significant, * = Significant at $p < 0.05$, ** = Significant at $p < 0.01$, *** = Significant at $p < 0.001$. Correlations have not been corrected for reliability or restriction of range.

Each Situational Judgement Test within the Perceptions Series demonstrates some significant correlations with a validated measure of the Big-5 personality traits. All significant correlation coefficients for can be interpreted as having magnitudes within the range of small to moderate, (Cohen, 1988).

Aptitude Tests

The Test Partnership aptitude suite is a selection of tests that when combined, provide a measure of general cognitive ability (Schwencke & Guy, 2015). There are four individual aptitude tests within the suite:

- *Critical Thinking*: This test is comprised of logical syllogisms which require deductions to be made, based only upon the information provided.
- *Verbal Reasoning*: This test is a measure of both understanding a passage of information, as well as what conclusions or assumptions can be drawn from it.
- *Numerical Reasoning*: This assesses the ability to interpret and comprehend numerical information, in order to complete the required calculations.
- *Inductive Reasoning*: This assesses the ability to recognise patterns and sequences in order to determine the next item in the sequence.

Table 1.3: Correlations between Perceptions SJTs and Aptitude Test Scores

Aptitude Test	Perceptions SJT			
	Admin	Graduates	Managers	Sales
Critical Thinking	0.24***	0.33**	0.10 [^]	0.24***
Verbal Reasoning	0.17***	0.13 [^]	0.03 [^]	0.17*
Numerical Reasoning	0.16***	0.19 [^]	0.01 [^]	0.20**
Inductive Reasoning	0.18***	0.14 [^]	0.17 [^]	0.11 [^]
Combined Score	0.20***	0.21*	0.07[^]	0.20**

Note: [^] = Non-Significant, * = Significant at $p < 0.05$, ** = Significant at $p < 0.01$, *** = Significant at $p < 0.001$. All correlations have been corrected for reliability, but not restriction of range.

Each Situational Judgement Test within the Perceptions Series (Excluding Managers), demonstrates significant correlations with validated measures of cognitive ability. All significant correlation coefficients for can be interpreted as ranging from small to moderate, (Cohen, 1988).

2 Group Differences

The following protected groups were investigated for evidence of average group differences and unfair discrimination:

- *Gender*: Participants that reported to be male were compared against participants reporting to be female.
- *Ethnicity*: Participants that reported being white were compared against those who reported to be BME participants.
- *Age*: Participants that reported being under the age of 40 were compared against the average scores of participants reporting to be over the age of 40.

Perceptions SJT	Group Differences in terms of Cohen's <i>d</i>		
	Gender	Age	Ethnicity
Admin	-0.05	-0.23	0.21
Graduates	-0.13	-0.21	0.20
Managers	-0.09	-0.02	0.37
Sales	0.02	-0.08	0.22

Table 2.1: Average Score Effect Sizes Across Protected Groups

Summary

Effect sizes for each Perceptions Situational Judgement Test can be interpreted as being small or non-existent (Cohen, 1988). Group differences across Age, Gender and Ethnicity generally report small to negligible group differences between focal and reference groups. None of the effect sizes were greater than 0.5, and therefore do not suggest evidence of practically significant differences between the protected groups included in analyses.

3 Reliability

Internal Consistency and Cronbach's Alpha

Measures of internal consistency represent some of the most commonly used estimates of reliability in psychometrics and questionnaire design. Internal consistency concerns the degree to which items within a questionnaire measure the same underlying psychological construct, implying a certain amount of unidimensionality.

Reliability of the Perceptions Suite

Perceptions SJT	No. Items	Sample Size	SEM	Cronbach's Alpha
Admin	60	2,783	3.07	.85
Graduates	64	2,068	3.52	.85
Managers	68	1,363	3.35	.86
Sales	60	2,475	3.14	.85

Summary

A general recommendation for Cronbach's Alpha is a minimum of .70. The observed reliabilities range from .85 to .86, exceeding the required standard. This suggests that the Situational Judgement Tests within the Perceptions Series can be considered to be sufficiently reliable in terms of their psychometric properties.



References

- Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis.
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. 2nd.
- Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (Neo-PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional manual*. Odessa, FL: Psychological Assessment Resources.
- Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: tiny-yet-effective measures of the Big Five factors of personality. *Psychological assessment*, 18(2), 192.
- International Personality Item Pool. (n.d.). International Personality Item Pool: A Scientific Collaboratory for the Development of Advanced Measures of Personality Traits and Other Individual Differences (<http://ipip.ori.org/>). Internet Web Site.
- Johnson J. A. (2014). Measuring thirty facets of the Five Factor Model with a 120-item public domain inventory: Development of the IPIP-NEO-120. *Journal of Research in Personality*, 51, 78-89.
- Rothstein, M. G., & Goffin, R. D. (2006). The use of personality measures in personnel selection: What does current research support?. *Human Resource Management Review*, 16(2), 155-180.
- Salgado, J. F. (2003). Predicting job performance using FFM and non-FFM personality measures. *Journal of Occupational and Organizational Psychology*, 76(3), 323-346.
- Schwencke, B., Guy, L. (2015). *Test Partnership's Adaptive Suite: Psychometric Properties and Validity*, London: UK.