insights®

Insights Learning & Development

Reliability and validity of the Finnish version of the Insights Discovery Preference Evaluator

Analysed by Qiong Yang

Prepared by Insights Research Team

Jan 2020

Contents

Abstract

This summary presents data on the psychometric properties of the Insights Discovery Preference Evaluator (IDPE) in Finnish version 3.3. The data used includes the most recent extraction up to Mar 2019. Psychometric science endeavours to impose objectivity upon the measurement of both human performance and aspects of personality. In order to do so, any questionnaire based on measurement of human behaviour, especially one that is based on self-report, must be able to meet certain demonstrable criteria in order to be considered an objective measure. This paper summarises these psychometric criteria in easily understandable terms. Key statistics have been computed for these areas and have been benchmarked against international standards. The paper concludes that the measurement of the four colours of the evaluator in Finnish is both valid and reliable.

This report presents information covering reliability and validity, covering:

Norm data by gender

- Table 1.1 – Summary of the norm data on gender
- Figure 1.2 – Colour energy average on gender

Reliability

- Cronbach Alpha coefficients
  - Table 2.1 – Summary of Cronbach Alpha coefficients
- Split-half correlations
  - Table 2.2 – Summary of split-half correlations
  - Table 2.3 – Summary of split-half coefficients

Validity: confirmatory factor analysis

- Table 3.1 – Summary of item factor analysis
- Figure 3.2 – Item loadings (25 x 4 colours) on the Insights Discovery Wheel

Data on norms

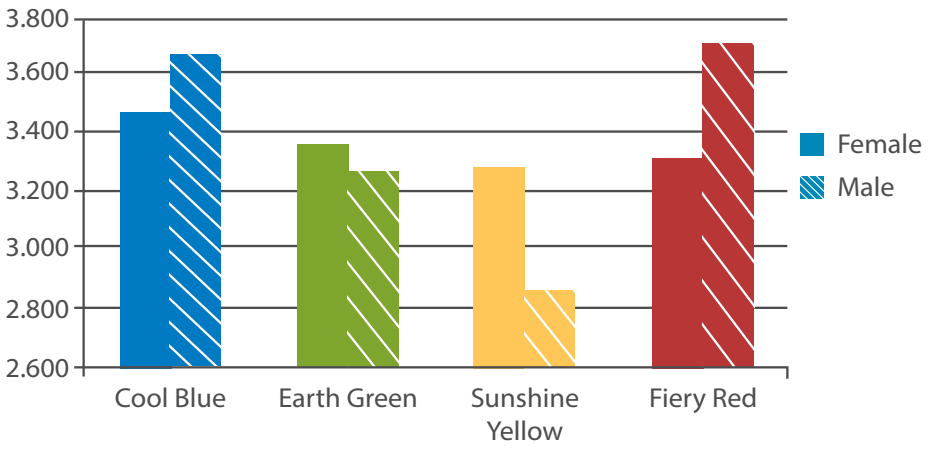
In Table 1.1, the averages of four colour scores and percentages of the norm sample with a dominant colour for female and male respondents are listed.

In Finnish, the data shows that female participants of IDPE tended to have a higher average on Sunshine Yellow, and slightly higher on Earth Green, than those of their male counterparts, whilst it's more likely for males to obtain a higher average in Cool Blue and Fiery Red than females. Overall, the average scores on Cool Blue and Fiery Red in Finnish sample are higher than other geographical regions. Furthermore, the graphical gender comparison within each colour energy is presented in Figure 1.2; this agrees with the data as described in Table 1.1.

Population segment	Sample size	Average colour scores Scale 0 to 6				Percentage of norm sample with dominant colour			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Female	1481	3.49	3.36	3.25	3.27	25%	27%	24%	24%
Male	1252	3.66	3.23	2.85	3.68	33%	20%	16%	31%
Whole	2733	3.57	3.30	3.07	3.46	29%	23%	21%	27%

Table 1.1: Summary of the norm data on gender

Figure 1.2:
Colour energy average on gender



Reliability

Reliability: Cronbach Alpha coefficients

In addition to the 'inter-item' and 'item-total' correlations, an important measure of reliability is the Cronbach-Alpha coefficient. The coefficient measures the error variance on the average inter-item correlation. When the error variance is low, which is desirable, the alpha coefficient approaches 1.0. A value of 0.70 is the commonly accepted lower limit (DeVellis, 1991; Robinson & Shaver, 1973; Robinson & al, 1991; Swailes & McIntyre-Bhatty, 2002).

Table 2.1: Summary of Cronbach Alpha coefficients

Sample size: 2733	**Colour preference**			
	Cool Blue	**Earth Green**	**Sunshine Yellow**	**Fiery Red**
Cronbach-Alpha coefficient	0.918	0.926	0.936	0.934
95% confidence interval	[0.913, 0.922]	[0.923 0.929]	[0.932, 0.938]	[0.930, 0.937]

Analysing the same 2733 completed evaluators, the Finnish evaluator shows the four colour energies do have very high Cronbach-Alpha coefficients (shown in Table 2.1), providing strong evidence of reliability. Therefore, we can say that, overall, the items of our Finnish evaluator have strong reliability.

Reliability: Split-half correlations

Another important measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability, we divide all items that are thought to measure the same construct into two sets, e.g. we create two sets of Fiery Red items. We test the evaluator on a sample of people and compute the total score for each randomly divided half. The split-half assessment of reliability is based on how well these two total scores correlate.

The split-half measures for the IDPE were achieved by splitting the 25 frames into two groups. The colour results are computed for each of the two groups and then correlated. A high correlation suggests high reliability i.e. the higher the association (correlation coefficient) between the two data sub-sets, the higher the internal consistency of the scale.

For each of the four colour preferences, the same procedure is conducted. The 25 frames for these four colour preference are randomly divided into two equal halves, a scale mean is computed for each half, and then the two sets of scale means are correlated to estimate a split-half correlation. The split-half correlation is adjusted by the spearman-brown prophecy formula to create a split-half reliability. This procedure is repeated 1000 times and the mean of the split-half correlations is reported as the best estimate of the reliability of a single item, while the mean of the split-half reliabilities is reported as the best estimate of the reliability of the composite of all items. The standard error of the reliability estimate is also reported which demonstrate the variability/dispersion and accuracy of the estimates.

Table 2.2: Summary of split-half correlations

Sample size: 2733	**Colour preference**			
	Cool Blue	**Earth Green**	**Sunshine Yellow**	**Fiery Red**
Mean of all split-half correlations	0.852	0.867	0.885	0.879
Mean of all split-half reliabilities	0.920	0.929	0.939	0.936
Standard error of split-half reliabilities	0.061	0.071	0.065	0.058

The results of the split-half analysis are shown in Table 2.2 above. It is clear the reliability is strong within each colour preference in the Finnish IPDE evaluator. In particular, all of the Cronbach-Alpha coefficients are greater than 0.7, and the correlation coefficients are also all above 0.7, indicating split-half reliability.

In addition, we conducted the split-half correlation in the usual way, which is randomly splitting the 25 frames into two groups of 13 and 12. The colour energy results (like the averages) are computed for each of the two groups and then correlated. A high correlation suggests high reliability, i.e. the higher the association (correlation coefficient) between the two data sub-sets, the higher the internal consistency of the scale. The full results of the split-half analysis are shown in the table below (Table 2.3).

Table 2.3: Summary of split-half coefficients

Group	Question number	Split-half reliability statistics	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
1	1 3 4 5 6 7 8 11 12 16 17 20 22	Cronbach coefficient	0.815	0.888	0.895	0.878
2	2 9 10 13 14 15 18 19 21 23 24 25	Cronbach coefficient	0.877	0.841	0.862	0.869
		Correlation coefficient	0.839	0.821	0.864	0.891

As you can see, all of the Cronbach-Alpha coefficients are greater than 0.7, and the correlation coefficients are also all above 0.7, indicating the Finnish evaluator shows good split-half reliability. Furthermore the results agree with those shown in Table 2.2 using a simulation tool.

To summarise; these results show that the Finnish IDPE is considered reliable by the tests applied above. However, when developing an evaluator we need to strike a balance between diverse questions which provide a richer description of personality and those questions which simply give us the highest numerical correlations, which could easily be achieved if we asked the same question 25 times, although of very little practical use.

Validity: confirmatory factor analysis

Confirmatory factor analysis was used to test the hypothesised factor structure of the Insights Discovery model. Specifically, it is hypothesised that the four sets of 25 colour energy based items in the IDPE, should load onto the factors such that:

- The polar opposite nature of the Fiery Red vs. Earth Green items are apparent
- The polar opposite nature of the Sunshine Yellow vs. Cool Blue items are apparent
- Fiery Red items should not load significantly onto any factor that Cool Blue and/or Sunshine Yellow items load onto
- Earth Green items should not load significantly onto a factor that Cool Blue and/or Sunshine Yellow items load onto
- Sunshine Yellow items should not load significantly onto any factor that Fiery Red and/or Earth Green items load onto
- Cool Blue items should not load significantly onto a factor that Fiery Red and/or Earth Green items load onto

The results that follow successfully confirm this hypothesised structure and offer evidence for the construct validity of the IDPE. (The methodology used was the Varimax Principal Component.)

These results show that the conceptual model's hypothesised polar dynamics are represented by the colour energies opposition in the colour loadings. The model suggests that the polar opposite of Cool Blue energy is Sunshine Yellow energy, which is supported by the factor analysis; i.e. the Cool Blue items can be seen to load negatively onto factor two, and the Sunshine Yellow energy items load positively onto factor two.

This may lead to the conclusion that the fundamental explanation of the four Insights colour preferences is contained in the first two factors that clearly account for the bulk of the variance. The presence of satisfactory loading values in further factors, which contributes to a small increase in the variance explained, is an added value but not a pre-requisite to the validation of the IDPE psychometric tool.

Please remember that in general, researchers use a 'rule of thumb' that considers factor loadings greater than 0.30 or below -0.30 as meeting the minimal level for significance (Hair, et al., 1998). Using these criteria the statistically significant factor loadings have been highlighted in a larger bold font in the table that follows.

Summary of item factor analysis

Table 3.1: Summary of item factor analysis

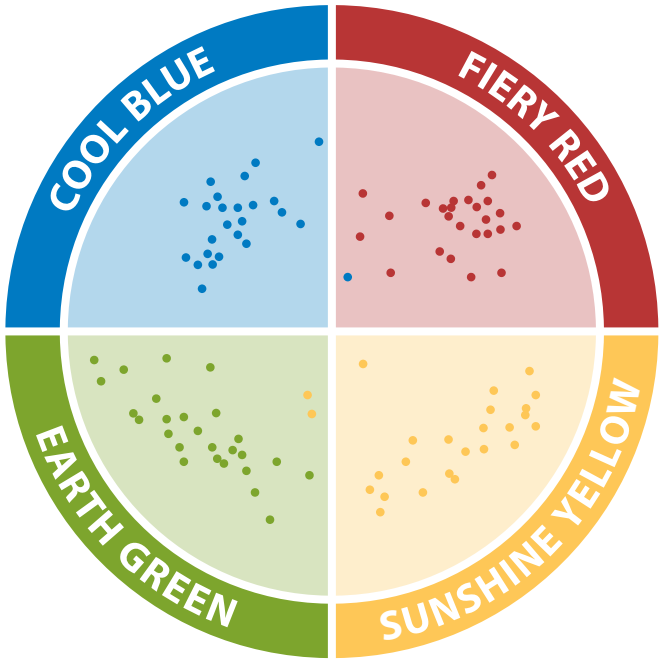
Sample size: 2733	Item average factor loadings			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Factor 1	-0.067	**-0.540**	0.056	**0.576**
Factor 2	**-0.496**	-0.088	**0.593**	-0.004

Table 3.1 is based on the Finnish version 3.3. evaluator. Analysing the table, we find that the Fiery Red items load strongly onto Factor 1 at plus 0.576. The Earth Green items also load strongly onto Factor 1 at minus 0.540 The opposite signs of these loadings support the theoretical construct of the model that hypothesises that the Fiery Red and Earth Green constructs are polar opposites. On the other hand, at minus 0.496, the Cool Blue items load strongly onto factor 2. The Sunshine Yellow items load onto factor 2 at plus 0.593. Again, the opposite signs of these loadings support the theory that Cool Blue and Sunshine Yellow are polar opposite constructs.

The above table is an average of the factor loadings. It is also possible to analyse the factor loadings for each of the 100 items in the IDPE individually. (This detailed data is contained in a separate document.) Figure 3.2 below is a scatter plot of the 100 items. It shows all 100 items loading onto the top two factors. The graph in Figure 3.2 has been superimposed onto the Insights Discovery Wheel. It shows the relationship between each of the 100 items (four colour energies multiplied by 25 frames) and the top two factors.

You will notice that all the items appear in the 'correct' quadrant except for one Cool Blue item and two Sunshine Yellow items. In general the graph provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

Figure 3.2:
Item loadings (25 x 4 colours) on the Insights Discovery Wheel



Conclusion

In conjunction with the detailed analysis of the English version of the IDPE(in a separate report), this brief paper offers good evidence of the internal reliability (using Cronbach Alpha and split-alf analysis) and construct validity (using Confirmatory Factor Analysis) of the Finnish version 3.3.

References

DeVellis, R. F. (1991) 'Scale Development: Theory and Applications', Sage Publications, Newbury Park, CA.

Hair, J.F., Anderson, R.E., Tatham, R.L., Black, W.C., (1998) 'Multivariate Data Analysis', 5th ed, Prentice-Hall, Inc.

Robinson, J.P., Shaver, P.R. (1973) 'Measure of Psychological Attitudes' MI: Survey Research Centre Institute for Social Research, University of Michigan.

Robinson, J.P., Shaver, P.R., Wrightsman, L. S (1991) 'Criteria for Scale Selection and Evaluation In Measure of Personality and Social Psychological Attitudes' Calif: Academic Press, San Diego.

Swailes, S., & McIntyre-Bhatty, T. (2002) 'The "Belbin" team role inventory: reinterpreting reliability estimates', Journal of Managerial Psychology, 17, 6, 529 – 536.

www.insights.com