An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This paper presents comprehensive data on the psychometric properties of the Insights Discovery Evaluator (IDE). The IDE's capacity to affect both reliable and valid measures of the primary personality preferences coded by the model as four colours is assessed. It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. The full methodological and statistical account of this programme may be found in technical papers produced at the University of Westminster's Business Psychology Centre (bpc). 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 that based on self-report, must be able to meet certain demonstrable criteria in order to be considered as an objective measure. This paper sets out to explain these psychometric criteria in easily understandable terms. It is the authors' intention to make the statistics and arguments presented understandable to two different professional groups, both of which may have a need to work with the Insights Discovery Evaluator (IDE). Firstly, psychometricians require key statistics and supporting technical information, in order for them to make a professional assessment of the evidence presented. Secondly, the wider community of Human Resource professionals require simple non-technical explanations of these same key statistics, in order for them to assess the appropriateness of the IDE's to their business. Four categories of information are presented covering 'item analysis', 'norms data', 'reliability' and 'validity' as these form the bases for demonstrable objectivity. Key statistics have been computed for each of these four areas and they have been benchmarked against international standards. The paper concludes that the measurement of the four colours is both valid and reliable.

Contents Page

Introduction

In 1998 Andi Lothian collaborated with Jeff Davis at the University of Westminster to develop the first version of the Insights Discovery model. This work formed a core component of Jeff Davis's Ph.D. entitled "Jung's Typology – The Development of a Psychometric Tool". Dr. Stephen Benton, one of the authors of this paper, supervised the Ph.D. Since then, 9 postgraduate dissertations, supervised through the Business Psychology Centre at the University of Westminster, have further developed the model and its applications. Any evaluator claiming to be a psychometric measure must meet the international standards clearly defined by the relevant professional bodies e.g. as found in both the practices of the American Psychological Association (APA) and the British Psychological Society (BPS). This paper presents strong evidence in support of the Insights Discovery Evaluator's (IDE) claim to be a high quality psychometric tool, which meets relevant professional standards. To convey these standards in easily understandable terms, the statistics central to establishing the psychometric properties of the IDE are presented in a four segment pyramid shown in Figure 1.

Objectives

- To explain how the Insights Discovery model has been developed
- To present the evidence for the Insights Discovery Evaluator's (IDE) psychometric measurement of the four colours being both valid and reliable
- To benchmark this evidence against other comparable preference based psychometrics
- To present a high level summary of the case for the evaluator meeting the psychometric standards set out by both the APA and the BPS

Pyramid of Key Psychometric Statistics

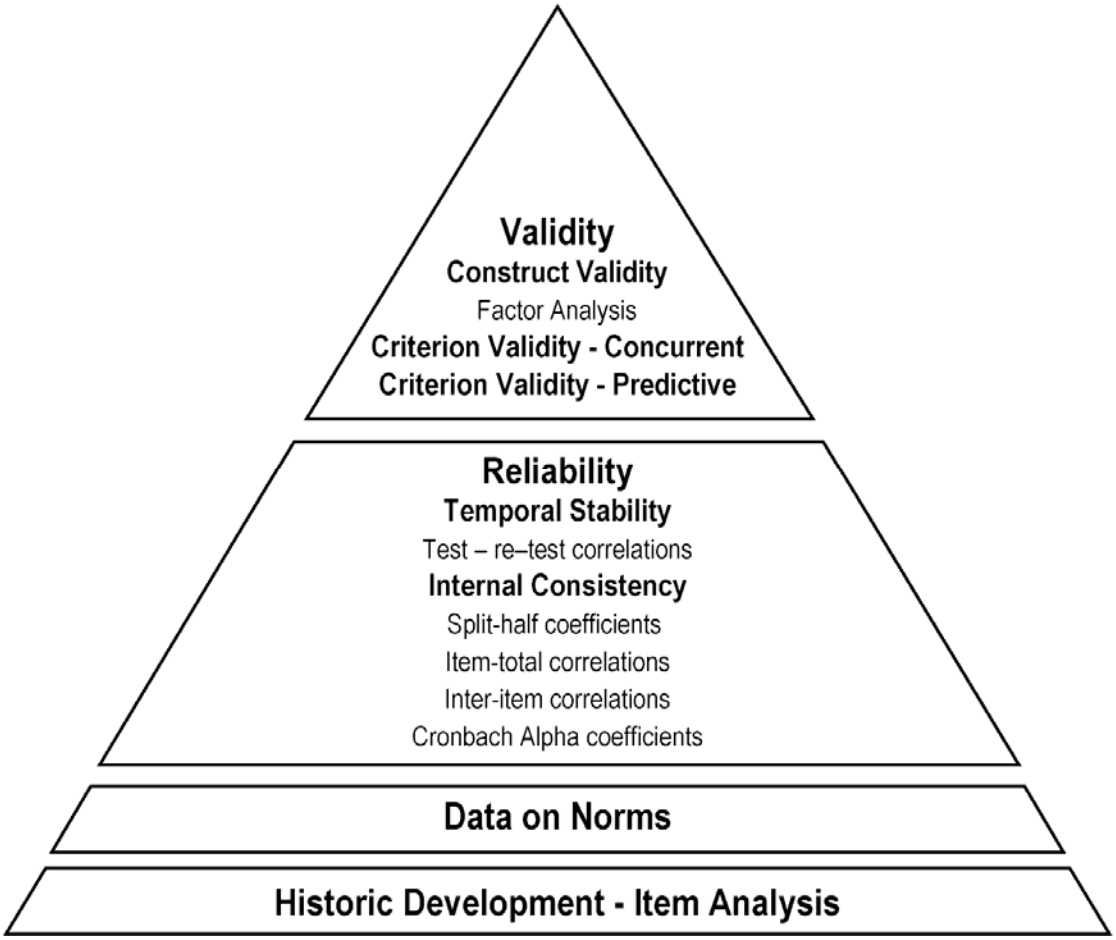


Figure 1 – Pyramid of key psychometric statistics

Methodology

The examination of the IDE in this paper has been based upon a number of samples as summarized below:

- Sample sizes of between 350 and 2,000 have been used for three separate 'four colour item analysis' conducted between 1998 and 2005, resulting in the current English version 3.0 of the evaluator. Four colour examples presented of the item analysis graphs are based on 519,467 English version 3.0 evaluators completed between 2003 and 2008.

- The norms data presented here is a small subset of an analysis of evaluators completed between 2000 and 2008. This includes evaluator data from the earlier versions 2.0 and 2.2 that have been improved through item analysis to become version 3.0. The norm data related to age groups, languages and countries are based on a set of 519,467 English version 3.0 evaluators.

- The test/re-test reliability statistics are based on 1,543 English version 3.0 evaluators completed between 2003 and 2007.

- The internal consistency reliability statistics are based on 519,467 English version 3.0 evaluators completed between 2003 and 2008.

- The construct validity data is based on 519,467 English version 3.0 evaluators and 114,670 English version 3.0 evaluators completed just in the UK. These evaluators were completed between 2003 and 2008. Covering similar time spans, smaller sample sizes ranging between 5,392 and 21,417 have been used to perform factor analysis on Dutch, German, French and Spanish translations of the evaluators.

- The face validity data is based on a bpc University of Westminster survey (Remarczyk, 2005) of 80 people who completed the IDE and were presented with their four colour scores accompanied by 50 sentences selected to describe the intensity of their personal four colour scores.

This data has been collected from people completing evaluators in connection with their experiencing an Insights Discovery workshop, coaching session or self-paced learning session i.e. the context for completing the evaluator was developmental.

This paper draws on the APA's (American Psychology Association) book entitled 'Standards for Educational and Psychological Testing' (1999) as an authoritative source detailing the objective standards that all psychometrics must meet. In addition, two of Paul Kline's seminal texts entitled the 'Handbook of Psychological Testing' (2000) and the 'Psychometric Primer' (1997) have been used to define key psychometric concepts and as a source of benchmark statistics for other comparable psychometrics.

The Insights Discovery Evaluator (IDE)

Appendix A contains version 3.0 of the English Insights Discovery Evaluator (IDE). It is an ipsative (forced choice) and normative (a range) evaluator consisting of 25 frames in which the user chooses from a choice of four word pairs a 'most', a 'least' and then scores the remaining two options in between least and most on a scale of 1 to 5. Each of the 4 items in a frame measure preferences called 'fiery red', 'sunshine yellow', 'earth green' and 'cool blue'.

A completed evaluator will have 25 colour preference scores, each giving a score between 6 (for most) and 0 (for least), for each of the four colours. A simple arithmetic mean across all 25 frames is calculated for each of the four colours. Figure 2 shows an example of the first 5 frames and an example of the colour bar chart produced from all 25 frames. In this example explaining how the evaluator works, the item's colour has been highlighted.

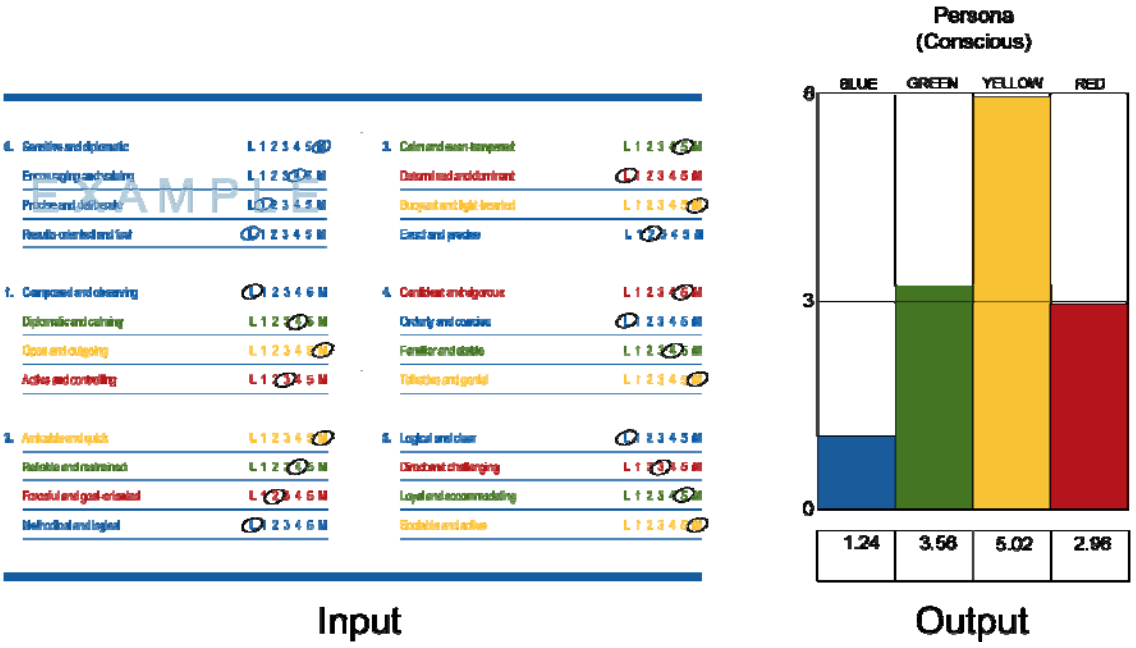


Figure 2 – Sample of evaluator frames and example of profile output

The highest colour in this bar chart is described as an individual's 'dominant' colour i.e. it is their most preferred colour.

A Definition of Personality

'Personality refers to those behaviour patterns which are characteristic of an individual and which tend to be consistent across situations and over time' (Glassman, 1995). These behaviour patterns are influenced by many things including attitudes, values, beliefs, experiences, preferences etc.

The IDE sets out to measure personality *preferences*, which are just one aspect of personality. No claim is made by the IDE to measure the personality in total. Collectively, the four colour preference bars derived from the IDE are described as an individual's *conscious persona*.

Overview of the Insights Discovery Wheel

Each of the colours is associated with a set of personal qualities that describe an individual's personality preferences. The higher an individual's colour score is, the more they will be inclined to use or have a preference for using these qualities. The converse is also true i.e. the lower a colour score is, the less they will be inclined to use these qualities.

The colours and the associated personal qualities have been displayed on the wheel in Figure 3.

The Insights Discovery model also asserts that the 'fiery red' qualities are the polar opposite of the 'earth green' qualities. It also asserts that the 'sunshine yellow' qualities are the polar opposite of the 'cool blue' qualities. This is symbolically conveyed through these colours being in polar opposite positions on the wheel i.e. the 'earth green' colour is on the opposite side of the wheel to the 'fiery red' colour.

For example, it is likely that an individual could relatively easily integrate a preference for the 'fiery red' quality of being 'competitive' with the adjacent 'cool blue' preference for being 'cautious'. However, it would be much unusual (although still quite possible) for an individual to integrate a preference for the 'fiery red' quality of being 'competitive' with the polar opposite 'earth green' preference for being 'caring'.

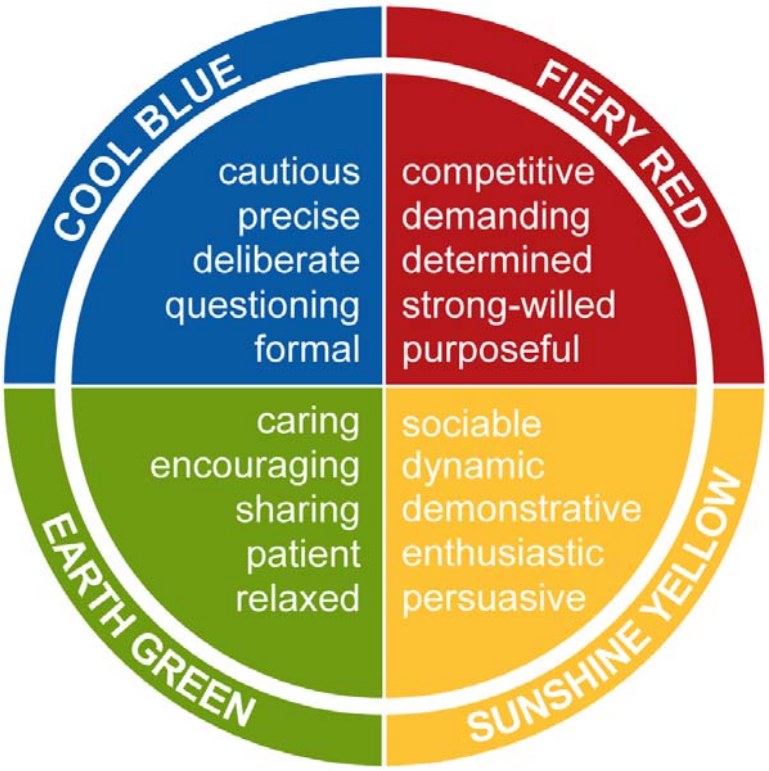


Figure 3 – The colours symbolically displayed on the Insights Discovery wheel

Further Reading about the Insights Discovery Wheel

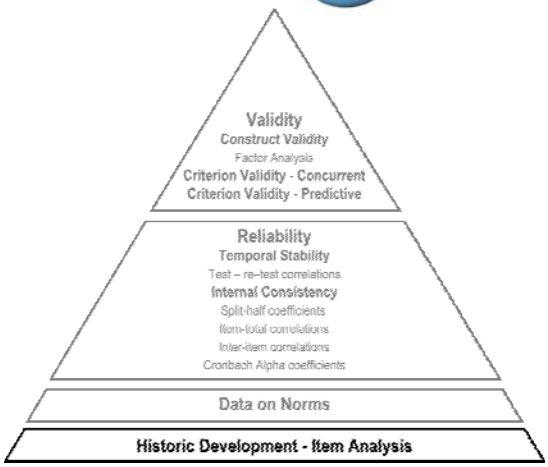
For the reader that would like more background on the Insights Discovery Wheel and how it is applied in practice, we refer you to the following authors and their public domain papers:

- Stephen Benton (2005) provides a further description of the Insights colours and their use within the discipline of Business Psychology in 'Every Individual Is the Exception to the Rule' published by the Association for Project Managers in the UK.

- Mark Beauchamp, Alan Maclachlan and Andrew M. Lothian (2005) have written an interesting paper on the application of the IDE in the field of sport. It has been published in volume 19 of 'The Sport Psychologist'.

- Mark Mullaly and Janice Thomas (2004) have written a stimulating paper exploring the inter-relationship between the Insights colours and project manager competency.

Historic Development – Item Analysis



One of the outstanding contributors to the field of psychometrics, the late Professor Paul Kline, described this technique in his 1997 book 'Psychometrics Primer' by saying:

"Item Analysis is a simple and effective method of test construction and many well-known tests have been developed using this approach"

Page 39 of the APA's (American Psychology Association) 'Standards for Educational and Psychological Testing' (APA, 1999) explains "The test developer usually assembles an item pool that consists of a larger set of items than what is required by the test specifications. This allows for the test developer to select a set of items for the test that meet the test specifications. The quality of the items is usually ascertained through item review procedures and pilot testing". Only an example of the Insights Discovery Evaluator's (IDE) item analysis procedure is provided in this paper and full documentation can be found in other technical papers produced at the University of Westminster's Business Psychology Centre (bpc) (Van Erkom Schurink, 2004).
One of the core methods underpinning the development of the IDE has been the iterative application of 'item analysis'. There are 100 colour 'items' (i.e. questions) spread over the 25 frames in the IDE. Item analysis involves empirically testing the quality of these 100 items and replacing weaker items with better ones. An example of a 'fiery red' colour item from the evaluator is 'determined and resolute'.

Item analysis has been used to produce from a wide pool of items, four colour based sub-sets of 25 items that are homogeneous, internally consistent and univariate within a colour i.e. each word pair statement measures just *one* colour.

One example of an item analysis on the 25 'sunshine yellow' items is show in figure 4. On the horizontal axis are the 25 frames. On the vertical axis is the number of respondents that highlighted 'sunshine yellow', 'fiery red', 'earth green' or 'cool blue' as the 'most' in the evaluator. However, this sample of respondents is comprised of those who, on average across all 25 frames, have scored 5 or more (out of 6) for 'sunshine yellow'. Consequently, we would expect that the 'sunshine yellow' line should always be significantly above the other three coloured lines. It can be seen that the 6th frame of the questionnaire is weak. We would expect the people in this sample to consistently select the 'sunshine yellow' item ahead of the other three colours. However, the graph at the top of figure 4 shows that in the 6th frame more people selected the 'earth green' item (*accommodating and helping*) as 'most', ahead of the 'sunshine yellow' item (*upbeat and hopeful*).

The items within the 6th frame were therefore subjected to systematic variation and re-evaluation as new word pairs were empirically tested. The best results were found when the 'earth green' item was changed from *'accommodating and helping'* to *'relating and amenable'*, combined with a change in the 'sunshine yellow' item from *'upbeat and hopeful'* to *'expressive and hopeful'*.

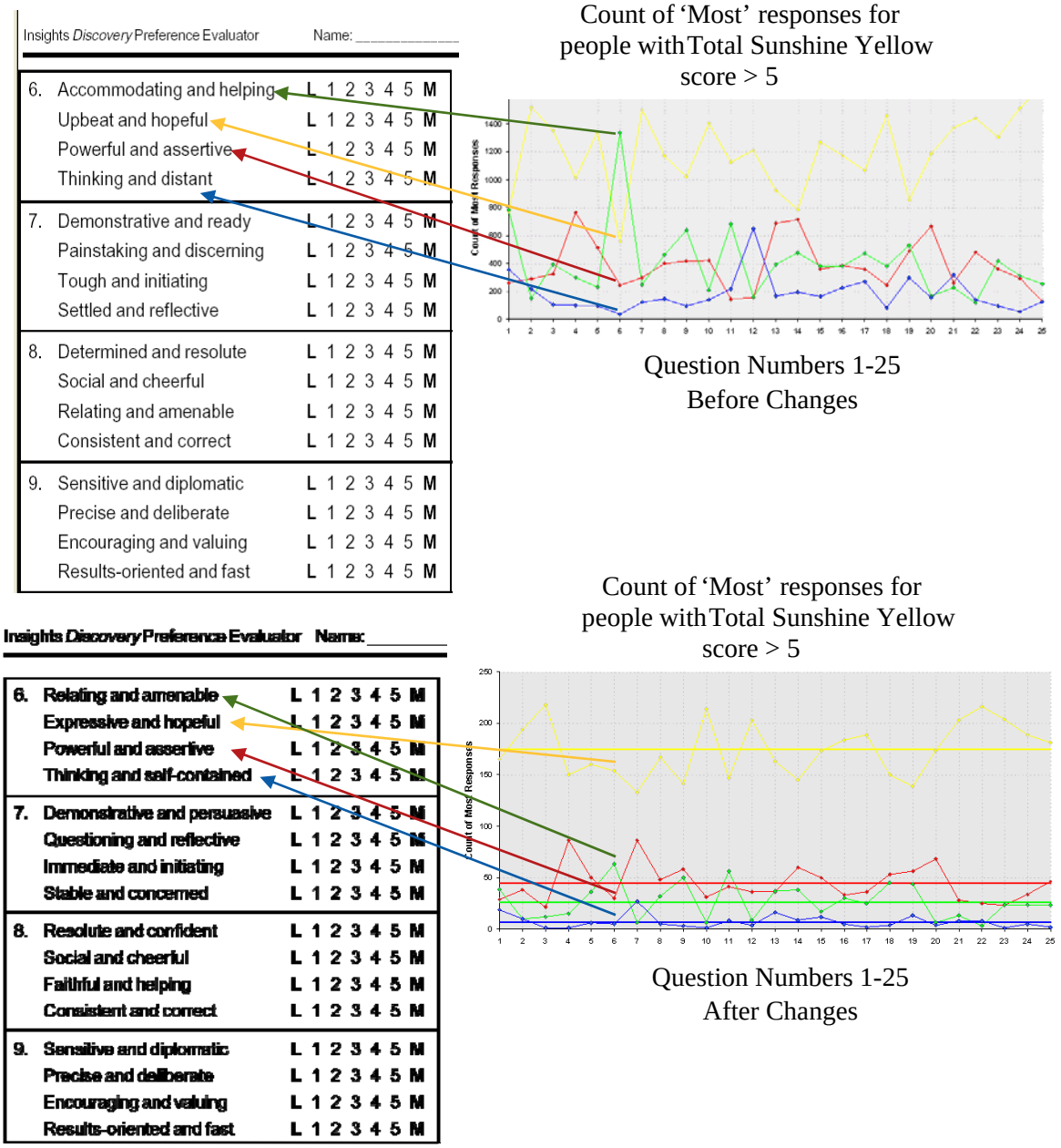


Figure 4 – Item analysis graphs before and after item changes

This procedure was repeated for all 25 frames in order to 'fine-tune' the discriminative capacity of the item pairs. Item analysis provides a learning curve built into the pool of selected items, thereby allowing the replacement of weak pairs with stronger ones.

The graph at the bottom of figure 4 visually demonstrates the benefit of making a series of word pair improvements across the evaluator. Consequently the quality of the evaluator has been systematically improved over time through the accumulative quality control associated with item analysis. The IDE is available in over 31 languages and after initial translation, each of these versions has also been developed using item analysis.

Item Analysis for the English Version 3.0 of the IDE

This data is based on 519,467 evaluators completed between 2003 and 2008.

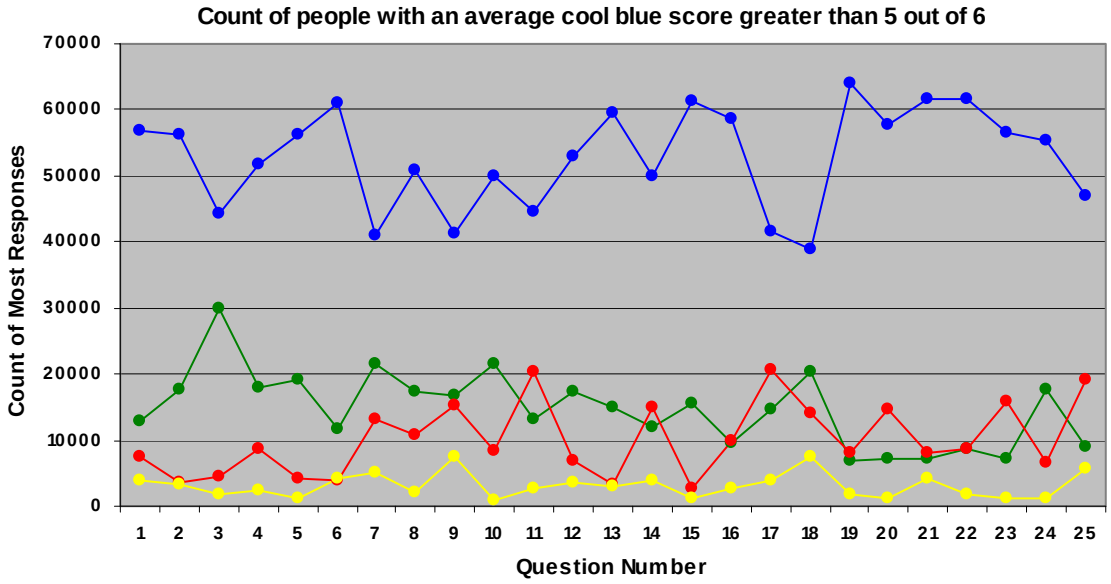


Figure 5 – Cool blue item analysis

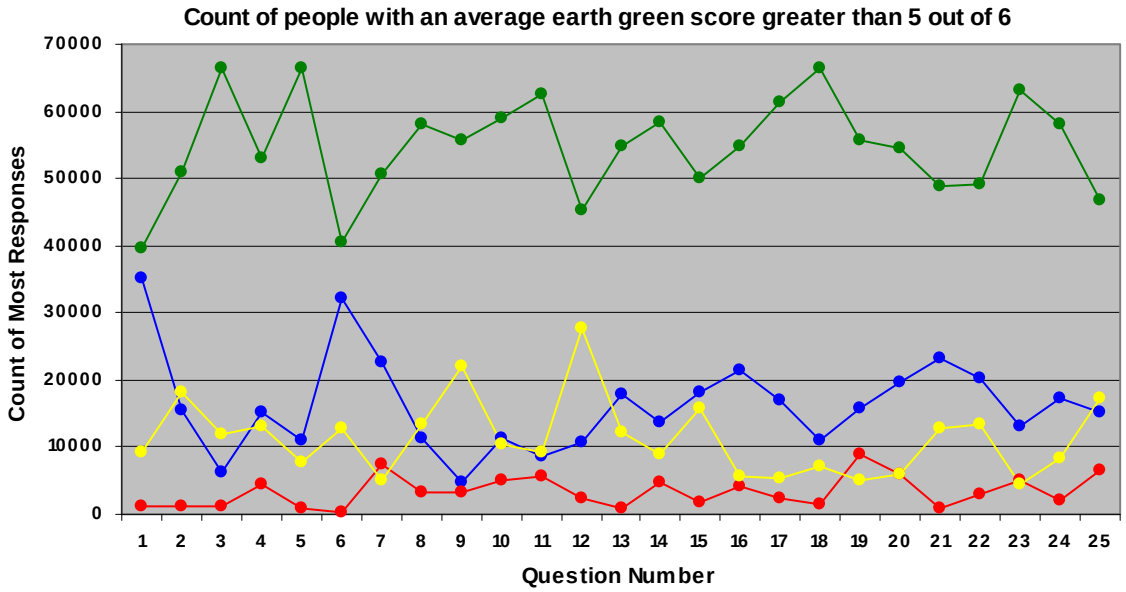


Figure 6 – Earth green item analysis

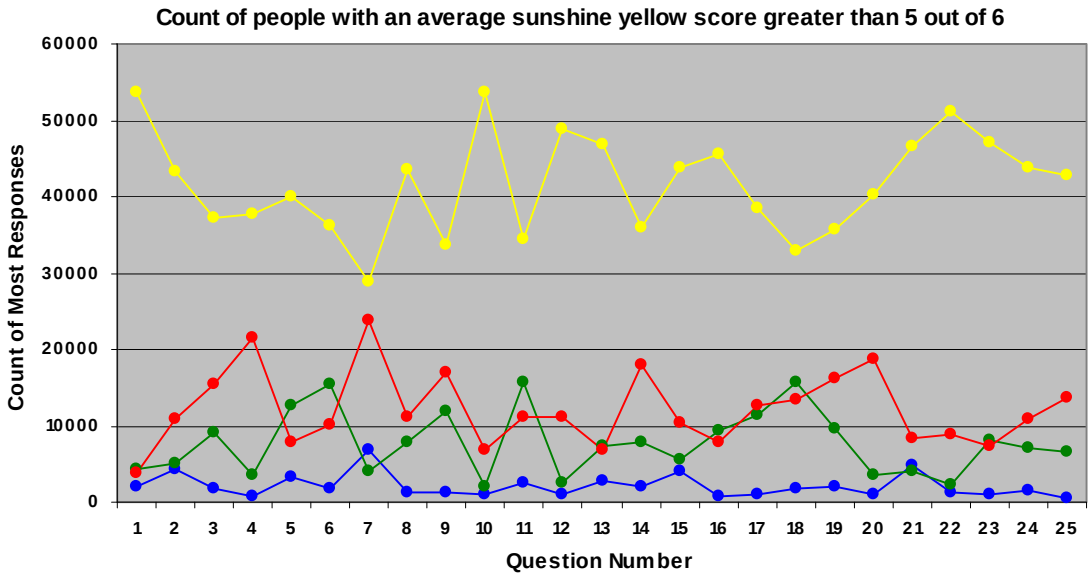


Figure 7 – Sunshine yellow item analysis

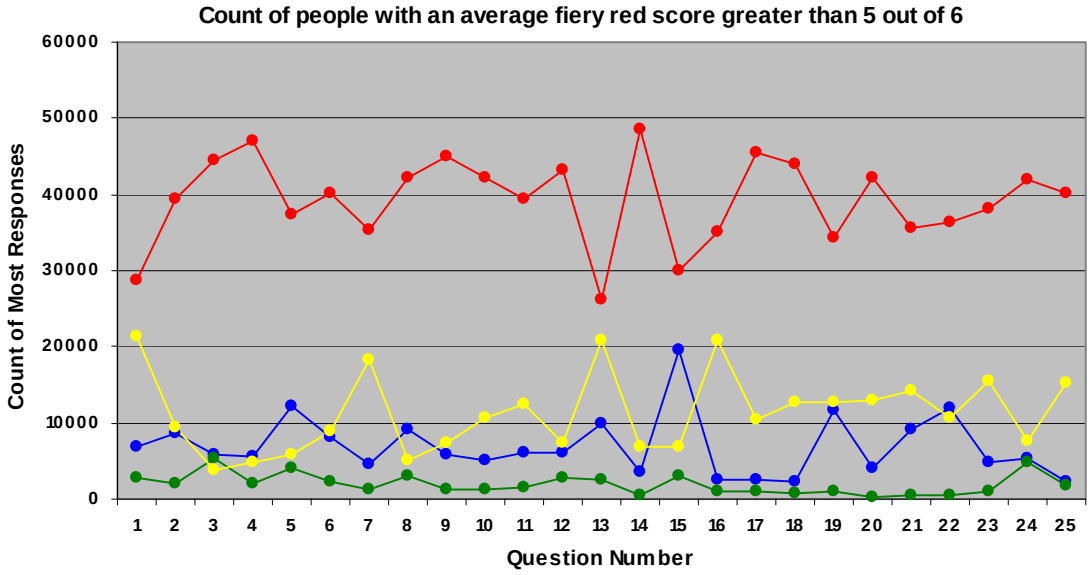


Figure 8 – Fiery red item analysis

Development of the evaluator using item analysis has significantly improved its quality. Figures 5 through 8 graphically show the benefit of evolving the item pairs over time. Although the graphs presented here are both visually impactful and intuitively appealing, there is a need for a more statistical approach to quantify the discriminative quality of the item analysis.

Consequently, t-tests were conducted on the item analysis data in order to gauge the statistical strength of the relationships plotted in the graph. A t-test allows you to determine if the distance between the two colour scores within one frame, are statistically significant or not. Below is an example for the 'cool blue' items from 4 samples comparing data across 4 continuously improving versions of the evaluators. This example is based on respondents who, on average across all 25 frames, have scored 5 or more (out of 6) for 'cool blue'.

For example, the shading of the first row of the table below shows that in the 3rd frame of the evaluator, the 'cool blue' statement *(exact and precise)* is successfully selected more often than the 'earth green' statement *(calm and even-tempered)*. However, the t-test also highlights that although the difference between them is statistically significant (which is good), they are nevertheless close to each other (which is not ideal). Any statement highlighted as a problem or potential problem is a candidate for improvement through further empirically driven item analysis.

Table 1 - T-test on item analysis of respondents who scored 5 or more (out of 6) for cool blue

		UK - English R25		UK - English S2		UK - English S2.2		UK - English S3.0	
Pair tested	Blue score>4 VS other colors	t value on paired t-test	Sig. of t (2 tailed) CI=95%	t value on paired t-test	Sig. of t (2-tailed) CI=95%	t value on paired t-test	Sig. of t (2-tailed) CI=95%	t value on paired t-test	Sig. of t (2-tailed) CI=95%
Pair 7	B3 & G3	3.11	0.002	2.73	0.006	11.85	0.000	9.53	0.000
Pair 19	B7 & G7	7.44	0.000	6.00	0.000	23.99	0.000	40.73	0.000
Pair 22	B8 & G8	14.40	0.000	10.53	0.000	25.80	0.000	26.51	0.000
Pair 25	B9 & G9	6.29	0.000	3.59	0.000	9.23	0.000	11.56	0.000
Pair 26	B9 & Y9	10.10	0.000	7.12	0.000	18.40	0.000	11.78	0.000
Pair 28	B10 & G10	11.85	0.000	9.33	0.000	32.73	0.000	36.82	0.000
Pair 37	B13 & G13	-8.71	0.000	22.98	0.000	26.45	0.000	27.99	0.000
Pair 49	B17 & G17	13.20	0.000	12.60	0.000	33.10	0.000	28.01	0.000
Pair 51	B17 & R17	10.27	0.000	15.67	0.000	38.39	0.000	39.75	0.000
Pair 52	B18 & G18	-1.09	0.276	-1.20	0.232	-5.89	0.000	10.47	0.000
Pair 70	B24 & G24	8.84	0.000	13.91	0.000	33.26	0.000	28.90	0.000

POTENTIAL PROBLEM: The two items (t-test) are very close (low positive t-value), but the distance between the two items is still significant

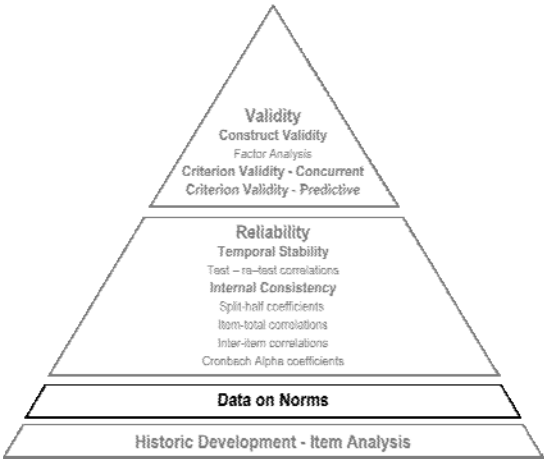
PROBLEM: The two items (t-test) are very close (low positive/negative t-value), and the distance between the two items is NOT significant

PROBLEM: The two items (t-test) are in inverted order (negative t-value), and the distance between the two items is/is not significant

However, even a high quality item analysis does not necessarily ensure an evaluator is valid. Consequently, quoting further from Paul Kline in his Psychometrics Primer (1997), "After the items have been selected by item analysis and the results replicated with a new sample, it is necessary, as has been argued, to show the test is valid and reliable".

Data on Norms

The BPS (British Psychological Society) and APA (American Psychological Association) standards state that all psychometrics must supply norms for comparative purposes, the form of which varies according to the constructs and attributes being measured. Norms must be up to date and appropriate for the intended usage and population.



The norms data for the IDE is of good quality, being segmented by; the language of the evaluator completed (Table 2); the country a respondent is based in (Table 3); age, in ten-year bands (Table 4) and occupation, with over 100 different occupations analysed (Table 16).

However, care must be taken when making use of this data, so as not to make invalid statistical interpretations. For many psychometrics, norms are used as a reference against which an individual's psychological test results can be interpreted relative to the distribution of a larger population. For example, a test of IQ might produce a score of 150. However, without a large and relevant sample population to provide a spread or distribution of scores against which to interpret the individual's score, the number 150 is without value. Yet, if we know a score of 150 places the individual in the top 5% of the distribution of IQ scores for people of age X, from a certain socio-economic group, then the information becomes more useful. Perhaps the situation would be further confused if there were more than one test claiming to measure IQ. Individuals may find themselves scored as a 150 (high) on one test and as a 50 (low) on another, then the issue of test norming becomes central to test selection.

Similarly, in clinical and related tests measuring a condition, norms provide a key reference for interpreting individual's scores and responses. Here it is important to establish if an individual's score, relative to the population, is indicative of 'more' or 'less' of the construct in question. For example; more depressed, slower reaction time or faster and/or more accurate short-term memory recall. Norms, for these forms of tests, provide a means of assessing a person's relative standing in comparison to others.

However, being able to measure whether or not someone has more or less personality isn't the aim of the IDE. The norms data presented in this paper are not intended to be used to make comparisons between an individual's colour scores and the continuous population distribution. Instead, for the purposes of personality profiling derived from a preference based evaluator, norms data is typically used to compare an individual's dominant preference with the percentage of the norm sample that have the same dominant preference. This is typically examined across ages, professions and cultures.

Figure 9 shows a sample of the data for people speaking French in both Canada (on the left) and France (on the right). This data is not a *random* sample of the population using these evaluators in different languages, but a *convenient* sample drawn from those participants that have (for whatever reason), experienced an Insights Discovery workshop or coaching session.

A French Canadian whose individual results show a dominant preference for 'fiery red', may be interested to know that just over 22% of the norm sample also has the same dominant

preference, while 78% have a preference for a different colour (see the bottom left hand graphs in Figure 9).

This example is just a small sample of a very large set of norms data available in more detailed papers at the University of Westminster. Some of this norms data provides good evidence of the 'predictive validity' of the model (see the later section in this paper on criterion validity for a fuller exploration of this). Predictive validity indicates a model can 'predict' something. Using the Insights Discovery model, we can predict that certain professions are likely to have a higher percentage of a dominant colour. Accountants, for example, are more likely to have a preference for 'cool blue'. This does *not* mean that to be a successful accountant, you must have a preference for 'cool blue' as the IDE does not measure *capability*. It is very important to remember that the IDE is only measuring *preference.*

Although this norms data provides good evidence of the 'predictive validity' of the model, it does ***not*** imply the evaluator is valid for use in recruitment. The norms data only presents aggregate data on *preferences* and says nothing about the sample's *capabilities*. If a practitioner of the Discovery system were to use this norms data as the basis for recruitment to predict whom the capable candidates may be, this would be an unethical and discriminatory practice. However, in the authors' experience this is a common error made by some inept practitioners working with other psychometric tools on the market. This error should not be made with the IDE.

What follows is a:

- graphical and tabular display of the norms data based on the language of the IDE completed.
- tabular summary of the norms data based on the country the respondent was located in.
- tabular summary and a set of bar charts showing the age band norms data.

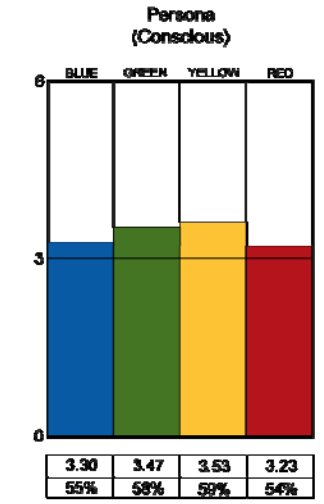
Only rows where the sample size was above 300 have been displayed.

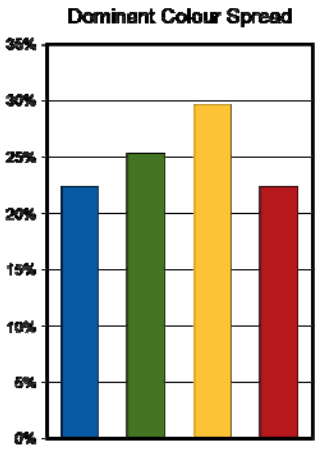
In Figure 9, the top graphs are based on the average of the four colour scores across all the people in the norm sample for French speakers in both Canada (the left hand graph) and France (the right hand graph). Table 2 shows this data for other norm samples and it is typically in the range 2.5 to 3.5

In Figure 9, the bottom graphs are based on the percentage of the norm sample with a dominant colour. Again this is shown for French speakers in both Canada and France. Table 2 also shows this data for other norm samples and it is typically in the range 20% to 35%.

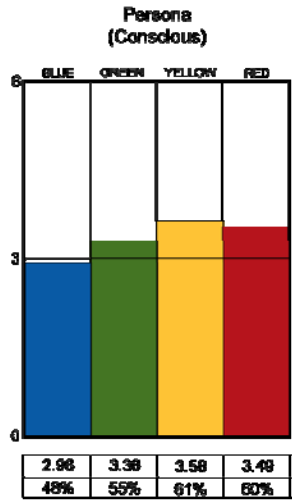
In Figure 10, please note that this is not longitudinal data and hence the variation could also be due to factors other than age e.g. it may be that the underlying driver of what appears to be an age variation may in fact be a variation in colours due to the country of the respondents. Further analysis of the data is required before any strong conclusions can be drawn from Figure 10.

6,583 French Speakers in Canada





4,883 French Speakers in France



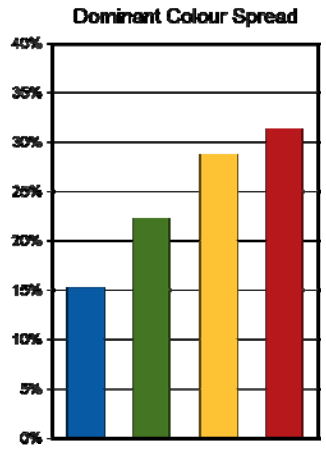


Figure 9 – Example of norms data for French speakers in both Canada and France

Table 2 – Tabular summary of the norms data based on the language of the IDE completed

		Average Colour Scores Scale 0 to 6				Percentage of Norm Sample with Dominant Colour			
Population Segment	Sample Size	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Arabic	411	3.63	3.53	3.54	3.42	26%	25%	25%	24%
Chinese	1 248	3.18	3.81	3.61	3.14	23%	28%	26%	23%
Czech	1 769	3.38	3.30	3.26	3.50	25%	25%	24%	26%
Danish	8 207	3.60	2.80	3.53	3.61	27%	21%	26%	27%
Dutch	36 851	3.79	2.93	3.41	3.42	28%	22%	25%	25%
English	519 467	3.17	3.47	3.52	3.40	23%	26%	26%	25%
English (South African)	5 652	3.18	3.75	3.57	3.37	23%	27%	26%	24%
English (Young Person)	531	3.03	2.75	3.49	3.86	23%	21%	27%	29%
Finnish	3 880	3.66	3.57	3.06	3.13	27%	27%	23%	23%
French	23 828	3.36	3.05	3.48	3.58	25%	23%	26%	27%
French (Canadian)	17 476	3.05	3.41	3.60	3.52	22%	25%	27%	26%
German	37 875	3.32	3.38	3.43	3.40	25%	25%	25%	25%
German (Swiss)	1 281	3.47	3.18	3.21	3.60	26%	24%	24%	27%
Greek	402	3.93	3.61	2.96	3.36	28%	26%	21%	24%
Hebrew	856	3.52	3.09	3.85	3.59	25%	22%	27%	26%
Hungarian	2 316	2.98	3.69	3.52	3.52	22%	27%	26%	26%
Italian	7 041	3.72	2.96	3.18	3.74	27%	22%	23%	27%
Japanese	853	2.91	3.33	3.89	3.21	22%	25%	29%	24%
Norwegian (Bokmal)	2 855	3.28	3.17	3.63	3.60	24%	23%	27%	26%
Polish	5 241	3.20	3.30	3.74	3.13	24%	25%	28%	23%
Portuguese	887	3.71	2.95	3.56	3.66	27%	21%	26%	26%
Portuguese (Brazilian)	7 140	3.99	3.11	3.40	3.43	29%	22%	24%	25%
Romanian	823	3.94	3.66	2.85	3.35	29%	27%	21%	24%
Russian	1 137	3.37	3.25	3.47	3.46	25%	24%	26%	26%
Spanish	9 110	3.48	3.30	3.65	3.41	25%	24%	26%	25%
Spanish (Mexico)	14 239	3.59	3.40	3.64	3.12	26%	25%	26%	23%
Swedish	3 358	3.16	3.10	4.01	3.37	23%	23%	29%	25%
Turkish	2 619	3.50	3.79	3.35	3.53	25%	27%	24%	25%

Table 3 – Tabular summary of the norms data based on country the respondent was located in

		Average Colour Scores Scale 0 to 6				Percentage of Norm Sample with Dominant Colour			
Population Segment	Sample Size	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Angola	305	3.25	3.58	3.40	3.45	24%	26%	25%	25%
Argentina	466	3.19	3.61	3.49	3.66	23%	26%	25%	26%
Australia	4 932	3.40	3.54	3.57	3.18	25%	26%	26%	23%
Austria	510	3.41	3.31	3.43	3.51	25%	24%	25%	26%
Bahamas	391	4.40	3.52	3.16	2.84	32%	25%	23%	20%
Belgium	1 389	3.22	3.24	3.46	3.70	24%	24%	25%	27%
Brasil	4 223	3.10	3.38	3.44	4.03	22%	24%	25%	29%
Canada	74 095	3.47	3.71	3.43	2.99	26%	27%	25%	22%
Chile	818	3.49	3.55	3.31	3.73	25%	25%	24%	26%
China	553	3.62	3.51	3.25	3.41	26%	25%	24%	25%
Czech Republic	368	3.22	3.16	3.51	3.59	24%	23%	26%	27%
Denmark	4 861	2.78	3.53	3.66	3.60	20%	26%	27%	27%
Egypt	366	3.47	3.47	3.34	3.64	25%	25%	24%	26%
Finland	488	3.41	3.09	3.30	3.75	25%	23%	24%	28%
France	11 134	3.03	3.42	3.56	3.44	23%	25%	26%	26%
Germany	8 499	3.39	3.33	3.34	3.44	25%	25%	25%	25%
Hong Kong	408	3.55	3.42	3.43	3.34	26%	25%	25%	24%
Hungary	489	3.61	3.43	3.54	3.20	26%	25%	26%	23%
India	850	3.62	3.37	3.23	3.75	26%	24%	23%	27%
Ireland	5 535	3.42	3.54	3.54	3.11	25%	26%	26%	23%
Italia	1 570	2.98	3.20	3.68	3.72	22%	24%	27%	27%
Japan	619	3.29	3.63	3.26	3.15	25%	27%	24%	24%
Malaysia	429	3.52	3.53	3.18	3.46	26%	26%	23%	25%
Mexico	9 918	3.33	3.62	3.15	3.57	24%	26%	23%	26%
Nederland	17 567	2.89	3.46	3.48	3.76	21%	25%	26%	28%
Norway	969	3.07	3.44	3.67	3.55	22%	25%	27%	26%
Poland	1 610	3.18	3.75	3.25	3.21	24%	28%	24%	24%
Portugal	628	3.01	3.53	3.59	3.79	22%	25%	26%	27%
Romania	563	3.57	2.85	3.36	4.02	26%	21%	24%	29%
Russia	448	3.10	3.22	3.58	3.69	23%	24%	26%	27%
Singapore	1 260	3.67	3.54	3.14	3.31	27%	26%	23%	24%
South Africa	15 448	3.61	3.45	3.26	3.44	26%	25%	24%	25%
Spain	4 234	3.20	3.61	3.53	3.45	23%	26%	26%	25%
Suomi	1 519	3.56	3.05	3.19	3.67	26%	23%	24%	27%
Sweden	996	2.99	3.77	3.50	3.44	22%	28%	26%	25%
Switzerland	4 817	3.11	3.41	3.65	3.46	23%	25%	27%	25%
Turkey	664	3.73	3.25	3.51	3.62	26%	23%	25%	26%
United Arab Emirates	393	3.42	3.27	3.39	3.90	24%	23%	24%	28%
United Kingdom	116 156	3.33	3.42	3.60	3.13	25%	25%	27%	23%
United States	142 105	3.47	3.54	3.39	3.27	25%	26%	25%	24%

Table 4 – Tabular summary of the age band norms data

		Average Colour Scores Scale 0 to 6				Percentage of Norm Sample with Dominant Colour			
Population Segment	Sample Size	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
<20	1 525	2.95	3.71	3.76	2.95	22%	28%	28%	22%
20-25	24 134	3.31	3.56	3.68	3.05	24%	26%	27%	22%
26-30	47 168	3.36	3.48	3.62	3.16	25%	26%	27%	23%
31-35	58 035	3.34	3.39	3.54	3.33	25%	25%	26%	24%
36-40	57 886	3.33	3.39	3.47	3.39	25%	25%	26%	25%
41-45	50 789	3.32	3.48	3.41	3.35	24%	26%	25%	25%
46-50	39 550	3.37	3.61	3.32	3.24	25%	27%	25%	24%
51-55	26 220	3.43	3.72	3.26	3.13	25%	27%	24%	23%
56-60	13 311	3.45	3.79	3.21	3.09	25%	28%	24%	23%
61-65	3 349	3.44	3.88	3.22	3.01	25%	29%	24%	22%
65+	591	3.43	3.96	3.17	3.01	25%	29%	23%	22%

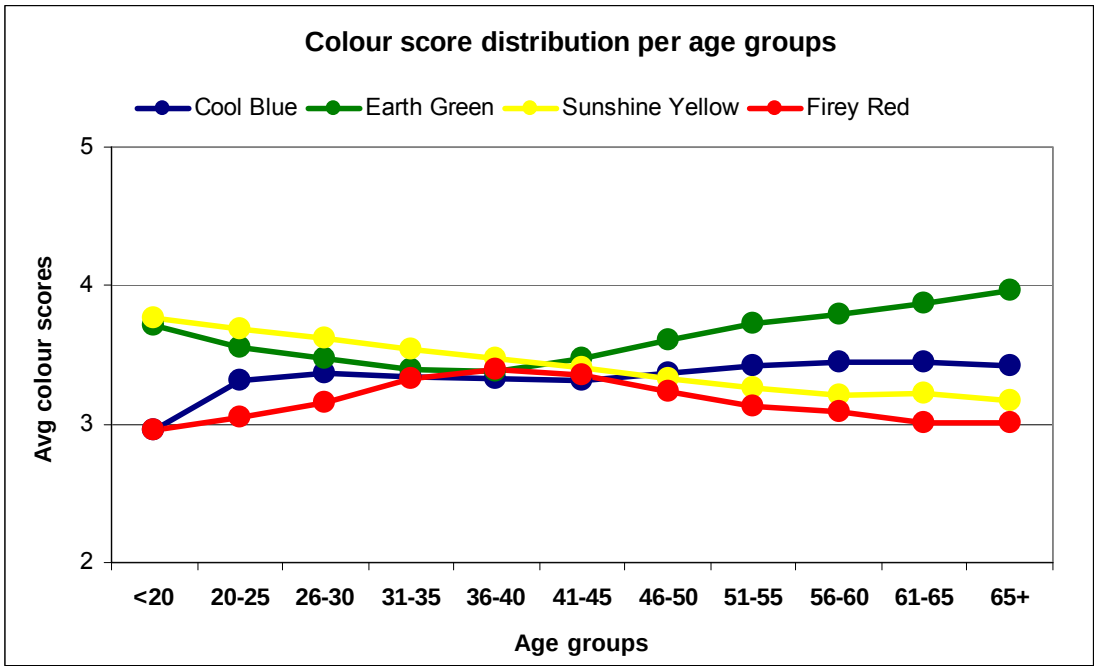
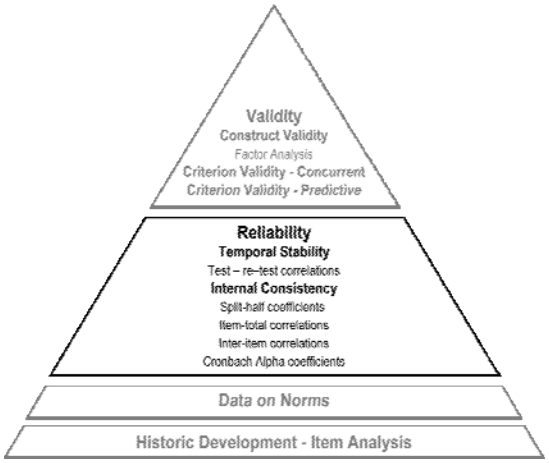


Figure 10 – Bar chart showing how the colour scores vary by age

Psychometric Reliability



Reliability has two meanings[1]:

"Does each item in the evaluator perform consistently?"

(called *Internal Consistency*)

"Do we have consistent results over a period of time?"

(called *Temporal Stability*)

A highly reliable evaluator will produce consistent colour scores that are repeatable over time. The ideal is to have high internal consistency and high temporal stability. However, all measurement procedures have the potential for error that will reduce reliability. The aim is to identify where the error is coming from (e.g. an unclear question) and to minimise it. Any observed score is made up of what statisticians call the 'true score' plus the measurement of unwanted and/or unknown factors i.e. 'measurement error'. In estimating the reliability of the IDE, we need to determine how much of the variability in the colour scores is due to measurement error and how much is due to real variability in the true scores. Essentially, errors of measurement result in a person's recorded colour scores being skewed away from the true score. This may happen because of a whole raft of reasons such as; they were sick, suffering from a headache, anxious, bored or trying to give answers they think would suit the expectations of others.

Reliability: Internal Consistency

Internal Consistency applies to the consistency of the scores amongst the 25 colour items i.e. it deals with measures of homogeneity *within* the colour items. The rationale for internal consistency is that the individual 25 colour items should all be measuring the same construct and thus be highly inter-correlated (Churchill, 1979; Nunnally, 1979). Four types of internal consistency have been examined; inter-item reliability; item-total reliability; Cronbach-Alpha reliability and split–half reliability.

The Use of an Ipsative (forced-choice) Scale in Reliability and Validity Analysis

The narrow classical view about the use of Cronbach-Alpha and other reliability analysis, as well as factor analysis (for validity tests), which are both derived from a correlation matrix, is that only 'interval' data types can be used. Ipsative (forced-choice) scales are based on 'ordinal' (i.e. ranked) data types and this 'forces' a correlation between items that artificially inflates the correlations in the correlation matrix i.e. the effect on a correlation matrix of items being scored using a forced-choice ordinal level scale is to accentuate the resulting correlations.

[1] Reliability meanings based on pages 26 to 33 in 'A Psychometrics Primer' by Paul Kline (1997)

However, the IDE uses a scale which is a hybrid between a forced choice scale and a Likert scale i.e. each item is given a score between 0 and 6 (a 7 point scale) and the forced choice is over 4 items. Although this Likert scale is still 'ordinal', Jaccard & Wan (1996) comment on this type of structure as follows: 'their use in statistical procedures which assume interval type data is commonplace'. Other authors also state that, 'the use of ordinal variables such as Likert scale with interval techniques is the norm in contemporary social science' (Labovitz, 1967, 1970; Kim, 1975; Binder, 1984).

The impact on the correlation matrix of a forced choice across 4 items, using a 7-point scale is likely to be significantly less than using a 'dichotomous' scale where a choice must be made one way or another. Here is an example from the 1942 version of the Gray-Wheelwright Jungian Type Survey (Wheelwright, 1964) that uses a 'dichotomous' scale:

At a party I

(a) like to talk

(b) like to listen

Choice (a) is a forced extraverted choice and (b) is a forced introverted choice. Numerous academic literature sources (Harvey et al.,1995; Tischler, 1994; Tzeng et al., 1989; Myers et al., 1998; Sipps & Alexander, 1987; Sipps et al., 1985) refer to the use of Factor Analysis as applied to 'dichotomous' measures e.g. as found in the MBTI (Myers-Briggs Type Indicator). These powerful statistical techniques, such as Factor Analysis and Cronbach-Alpha, should be considered equally valid for the Insights forced-choice 7-point rating scale.

In summary, although at odds with the narrow classical view, there is sufficient evidence to support the valid use of these techniques on the IDE data.

'Inter-Item' and 'Item-Total' Reliability

'Inter-item' and 'item-total' correlation coefficients have been calculated using the Pearson Product-Moment Correlation. This involved creating four colour based '25 by 25' matrices showing the correlation between the 25 colour items. In addition, we have computed 'item-total' correlation coefficients by correlating the individual colour item score to the sum of all 25 scores for the same colour. A condensed summary of these correlations is shown in Table 5, with the full matrices shown in Appendix B. In 1991, Robinson et al. concluded that the mean 'inter-item' correlation coefficient should equal or exceed 0.30 and the 'item-total' correlation coefficients 0.50 for this to be good evidence of reliability. The analysis of the 519'467 evaluators shows that, for each of the four colours in the evaluator, the average 'inter-item' correlation coefficient is significantly above 0.3 and the 'item-total' correlation coefficient is significantly above 0.5, providing strong evidence of the case for reliability.

For example, in Table 17 in Appendix B, the 'cool blue' item *'methodical and logical'* (Frame 2, Question 4) has a correlation coefficient of 0.36 with the item *'orderly and concise'* (Frame 4, Question 2). As this is above 0.3, it is considered a good result.

Below is a summary of the 'inter-item' correlations for the four colours. The top row of statistics in the table below show the average 'inter-item' correlations are significantly above 0.3. In addition, a high percentage of the colour items (between 18 and 22 out of 25) are statistically considered 'strong'.

Table 5 – Inter-item correlations

N = 519'467 Inter-Item Correlations	**Colour preference**			
	Cool Blue	**Earth Green**	**Sunshine Yellow**	**Fiery Red**
Mean	**0.33**	**0.31**	**0.31**	**0.35**
Minimum	0.10	0.11	0.04	0.14
Maximum	0.59	0.60	0.61	0.55
Range Maximum minus Minimum	0.48	0.49	0.57	0.41
Maximum divided by Minimum	5.9	5.45	15.25	3.93
Variance	5.70	5.58	15.15	4.02
N of items in the scale	25	25	25	25
N of weak items N of strong items	3 22	4 21	8 17	2 23

Cronbach-Alpha Reliability

In addition to the 'inter-item' and 'item-total' correlations, another 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 inferior limit (see DeVellis, 1991; Robinson & Shaver, 1973; Robinson & al, 1991; Swailes & McIntyre-Bhatty, 2002).

Analysing the same 519'467 completed evaluators shows the four colours to have very high Cronbach-Alpha coefficients, providing further evidence of excellent reliability.

Table 6 – Cronbach-Alpha coefficients

N = 519'467	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	0.92	0.92	0.93	0.92

Split-Half Reliability

The final measure of Internal Consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly 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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The full results of the 'split-half' analysis are show in Table 21 in Appendix B.

The analysis shows high coefficients for the IDE, with a key summary being:

- Cronbach-Alpha Coefficients above 0.8 for each half (ranging from 0.83 to 0.88)
- Pearson Correlation Coefficients above 0.7 i.e. the 2 halves correlate highly (ranging from 0.79 to 0.85)

Reliability: Temporal Stability: 'Test - re-test'

'Temporal stability' or 'test - re-test' reliability is determined through the administration of the same evaluator across time and it helps us gauge how robust the items are. If the results are statistically sound, then practitioners may have confidence in both the durability of results and their applicability across situations. This type of reliability is particularly useful for measures of stable personality traits, but not for measures of aptitude, where practice effects can significantly influence scores on future administrations.

For the IDE there are 2 key reasons why an individual's re-test scores may differ from their original test. Firstly, there may be variability in their responses due to measurement error and this would signify a lack of reliability in the instrument. Secondly, the individual may have experienced personal change in this period and now genuinely have altered their colour scores. The Insights Discovery research team is continuously working to eliminate the first possible reason (instrument error). The research team is also working to understand and quantify the second reason in the belief that human beings are dynamic and evolving. This approach acknowledges the possibility that people may shift their preferences several times during their lifetime.

An analysis of 'test – re-test' data based on a *convenient sample*

A convenient sample of 1,543 people, who needed to complete the evaluator twice, had their original and re-tested colour scores assessed through a Pearson correlation analysis. This data was captured between 2003 and 2008. As this is a convenient sample, the time period between the test/re-test varies between 3 months and 5 years for different respondents in the sample. Reliability is expressed as correlation coefficients, ranging from 1 to 0. Temporal stability tests are generally expected to yield reliability coefficients ranging between 0.70 and 0.90

The results of the 'test - re-test' analysis performed on the four colour scores show a good reliability, translating into Pearson correlation coefficients ranging from 0.79 to 0.82

Table 7 – 'Test re-test' Pearson correlation coefficients based on a convenient sample of 1,543

N = 1'543	RETEST Cool Blue	RETEST Earth Green	RETEST Sunshine Yellow	RETEST Fiery Red
TEST Cool Blue	**0.82**	0.08	-0.67	-0.20
TEST Earth Green	0.10	**0.77**	-0.13	-0.63
TEST Sunshine Yellow	-0.69	-0.10	**0.82**	0.06
TEST Fiery Red	-0.23	-0.65	0.08	**0.79**

All correlations in this table are significant at the 0.01 level (2-tailed).

An analysis of 'test – re-test' data based on *1st year medical students*

In a paper entitled "Medical Students' Personality – a Typological Approach" (Halpin and Green, 2005) it is reported that over a period of six years, the IDE was administered to successive cohorts of University of Dundee 1st year medical students. Analysis of the data indicated that the distribution of colour personality preferences was significantly uniform across all six of the years. Moreover, two 'test - re-test' analyses, using two of the cohorts, suggested that these differences had significant longitudinal stability. The data, in which gender differences were modest, enabled the profile of the typical medical student to be identified.

Table 8 – 'Test re-test' Pearson correlation coefficients based on 1st year medical students

	112 students completed the IDE in October 2001 *and again in* August 2004	86 students completed the IDE in October 2000 *and again in* June 2005
'Test – re-test' for Cool Blue	0.72	0.62
'Test – re-test' for Earth Green	0.74	0.59
'Test – re-test' for Sunshine Yellow	0.67	0.59
'Test – re-test' for Fiery Red	0.73	0.64

All correlations in this table are significant at the 0.01 level (2-tailed).

‘Test re-test’ comparison with the MBTI and 16PF

As a matter of comparison, studies published on other instruments report the following range of coefficients:

- 0.69 to 0.83 - Pearson’s correlation coefficients for the MBTI instrument (Carlson, 1985; Carlyn, 1977)
- 0.52 to 0.96 - Pearson’s correlation coefficients for the 16PF instrument (Harrell and Lombardo, 1984)

Type Mobility and ‘test - re-test’

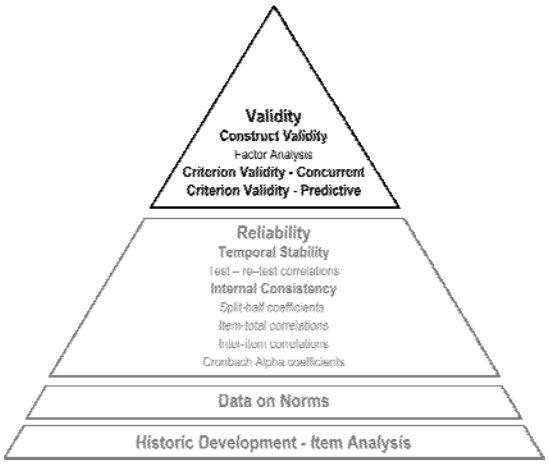
Tosey and Gregory (2002) write that ‘the age-old debate of nature versus nature refers to different hypotheses about personality development. Some believe personality traits such as temperament are genetically determined, others that all people are born with the same propensities and that experience determines the difference’. They also highlight that ‘many people believe that personality traits cannot be changed; however, this is not necessarily the case’. In the authors experience, when the outputs from the IDE are used in experiential workshops and coaching sessions, an optimistic and humanistic stance towards development is generally adopted i.e. it is assumed that *an individuals preferences can evolve over time* through increased self-awareness and a focus on personal and professional development. This evolution of an individual’s preferences has been termed ‘type-mobility’.

If these humanistic assumptions are accepted, then we would not expect the IDE ‘test - re-test’ correlations for a colour score to be too close to 1.0 (a correlation of 1.0 would suggest the people taking the IDE are unable to evolve their preferences at all). However, we would still expect the psychometric measurement of *preferences* to have a similar stability to *trait* measurements such as the 16PF instrument. We would also expect both preference and trait measurements to be much more stable than *state* measurements e.g. when measuring emotions we may expect different readings every day. For more information on this please read Paul Kline’s book ‘Handbook of Psychological Testing’ (2000).

The ‘test - re-test’ results presented in this paper are consistent with the above assumptions on ‘type mobility’ whilst simultaneously offering good evidence of the reliability of the IDE.

The degree to which a test is reliable defines the accuracy with which it elicits and assesses someone’s responses. However, just because a test is capable of delivering high reliability scores does not mean that it is valid i.e. a test, which is reliable, does not necessarily measure what it is supposed to. That is, just because a test is reliable and consistent over time, does not mean that it is valid. An effective psychometric tool requires a combination of evidence supporting both its reliability *and* its validity.

Psychometric Validity



Validity means:

"Do we measure what we say we measure?"

Although this question may sound banal, providing psychometrically sound answers that meet international standards involves a substantial amount of work. The American Psychological Association 'Standards for Educational and Psychological Testing' (1999) publication says:

'a sound validity argument integrates various strands of evidence into a coherent account of the degree to which existing evidence and theory support the intended interpretation of test scores for specific uses'.

Intended Interpretation of IDE Outputs

Validity addresses *what* the evaluator actually measures and *how well* it measures it. For the IDE, validity is concerned with what can be interpreted from the colour scores. Psychometric measurements are always validated in regards to a particular use i.e. one cannot say that the evaluator has 'high' or 'low' validity per se. However, evidence can be gathered for interpretation of the colour scores being 'valid' in a particular way in which they are used i.e. the context is always very important to any validity evidence presented and sets the conditions within which validity operates.

The following are the necessary assumptions and presuppositions that the reader needs to be aware of when considering any claim to validity for the IDE:

- The intended use of the IDE is to support an individual's personal growth through a workshop or coaching experience.
- No one colour is better than another.
- We know ourselves better than any set of questions can identify.
- The colour scores are a guide to our personality preferences only.
- Degrees of skill, or interests, are ***not*** identifiable from a person's colour scores.
- Based on their colour *preferences*, people should not be persuaded that they would be good or bad at a specific job i.e. no judgment should be made on a person's *capabilities*.
- Knowledge of the colour preferences may be used to understand the *why* of personal failings, but never to *excuse* them.

In contrast, it is common for some well-established tests to be erroneously referred to as having 'high validity' based on unspoken assumptions about how the test is used. All hidden

assumptions and presuppositions must be made visible for a claim to validity to be authenticated.

Different Types of Validity

Different types of validity discussed in this section include:

- Face Validity - do the items (inputs) and/or the measures (outputs) from a test appear plausible to the user?
- Content Validity - can an expert objective source validate the quality of the items?
- Construct Validity - the degree to which the test measures the underlying theoretical construct.
- Criterion Validity – predictive validity. This refers to the degree to which a test can predict a person's behaviours or performance on future, specified activities.
- Criterion Validity – concurrent validity. Here the validity of any test is best determined by comparing it to another test or some observable fact i.e. criterion validity is always based on external relationships.

Overview of Face Validity

There are two different applications of the term face validity. The first concerns the degree to which the items in a test *appear* to measure what the test claims to measure. The second concerns the extent to which the users of the test *believe* the outputs from a test are accurate, as defined by how the outputs match their *self-perception*.

Although it is usually considered desirable for a test's items to appear valid, this may not always be the case. For example, on measures geared toward the assessment of malingering and deception, low face validity may aid in more effective detection.

However, for personality preference questionnaires such as the IDE, having the items appear valid is desirable in that it helps ensure users are willing to fill in the evaluator. In addition, without reasonable face validity, the users' confidence in the IDE outputs may be undermined if they do not think the items are plausible when they are completing it. However, although high face validity is usually a prerequisite for achieving a practical outcome, it is of minimal value in assessing psychometric validity.

Face Validity Does Not Support Any Claim to Psychometric Validity

Although user confidence is important, it is not presentable evidence in making a psychometric claim to validity. One reason for this is the so called "Barnum Effect" (Paul, 2004) that psychologist Bertram Forer highlighted in 1949 when he gave a personality questionnaire to a group having told them the results were individually personalised. Unbeknown to the group, they actually were all given the same description based on an astrology book. The group then scored the accuracy of the results and the average score was 4.2 on a scale of 0 to 5, with over 40% scoring it 5/5. The obvious conclusion from Forer's experiment is that many factors can cause a user to erroneously perceive a questionnaire as valid when in fact it is not.

However, despite face validity carrying a relatively low level of psychometric credibility, some instruments will cite face validity as the main evidence of their validity. Clearly this is not the case for the IDE. However, although having high face validity does not demonstrate overall validity, a *lack* of face validity would be a serious obstacle to the *practical application* of the evaluator. If user confidence is inhibited by low face validity, people may choose not to use it or not to believe its output.

To provide you with confidence in the practical use of the IDE, here are some face validity results (with the caution that these statistics are not enough to establish psychometric validity). In a bpc University of Westminster survey (Remarczyk, 2005), a group of 80 people completed the IDE and were presented with their four colour scores accompanied by 50 sentences selected to describe the intensity of their personal four colour scores.

They were asked to mark out of 5, the overall accuracy of the information contained in the colour scores and the supporting descriptive statements. The mean rated accuracy score was 4.3/5.0 (86%) with a standard deviation of 0.65.

Furthermore, the group also assessed the quality of each of the 50 sentences describing their personal colour scores. Below is a summary of the results. Based on this subjective feedback, over 85% of statements were scored either 4/5 or 5/5.

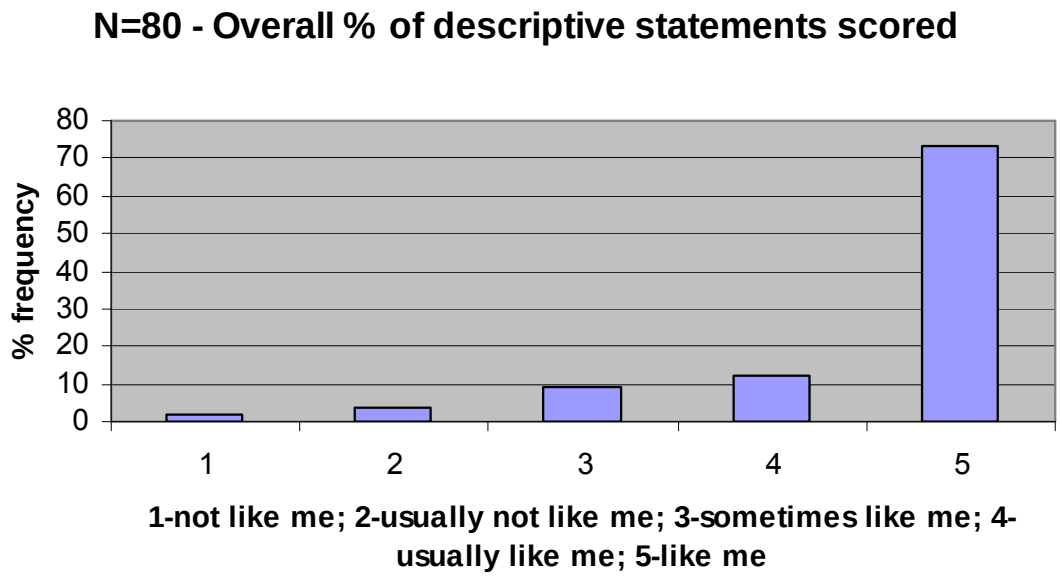


Figure 11 – Histogram of assessed quality of each of the 50 sentences describing personal scores for the colour energies

The Impact of Face Validity on the Practical Application of the Insights Discovery Model

Creating the appropriate context with the user is an essential pre-requisite to the *practical application* of the knowledge gained with the IDE. The colour based information derived from the evaluator is intended to be used as a catalyst for personal and/or professional development, through the experience of coaching, self-paced learning or a workshop.

However, just because a questionnaire has strong evidence of its psychometric validity, does not guarantee that its use will result in practical application by the user. Having psychometric validity is a *necessary but not sufficient* condition for ensuring practical application.

For any personality preference questionnaire to be a catalyst for practical application requires the user has a personally *meaningful experience*, which is based on a *psychometrically valid* model.

Objective Psychometric Measurement

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 that based on self-report, must be able to meet the objective standards highlighted earlier in Figure 1 - the 'Pyramid of Key Psychometric Statistics'. A positivist stance is appropriate in this detached and objective approach.

Subjective Meaningful Experiences

However, the context in which the IDE is used also requires that the user subjectively finds the experience meaningful. This is the domain of phenomenology which has been defined as 'a philosophical view that is interested in the way that each person experiences the world … it assumes that subjectivity of experience is important, and emphasizes the nature of experience before linguistic, conceptual labels are applied' (Tosey and Gregory, 2002). For a fuller account of phenomenology the reader is referred to 'The Interpreted World: An Introduction to Phenomenological Psychology' by Spinelli (1989).

The reader may also wish to read 'Checkland' (1999) for an interesting exploration of the evolution of the objective positivist stance and the subjective phenomenological stance over the last century.

The Interaction of Psychometric Validity with Face Validity

The psychometric validity of the model being used and a combination of face validity/user experience combine to create the matrix below. Inside each cell of the matrix is a description of the quality of the facilitator or coach.

Psychometrically Valid Model	The Impractical Bore	The Facilitator of Personally Meaningful and Practical Action
Psychometrically Invalid Model	The 'Snake-Oil' Sales-Person	The Well Intended 'Deluder'
	Low Face Validity and/or a learning experience of little personal significance	High Face Validity and a personally meaningful learning experience

Figure 12 – The interaction of psychometric validity with face validity

The Facilitator of Personally Meaningful and Practical Action

It is the responsibility of the group facilitator or coach to ensure they provide a personally meaningful learning experience, using a psychometrically valid model. In the matrix in Figure 12, this implies that facilitators/coaches using preference based psychometrics, should be operating in the top right hand box i.e. for a facilitator/coach to empower their client(s) to take personally meaningful and practical action requires they operate in this top right hand box. Clearly it is desirable for any facilitator or coach to be operating out of the top right hand box.

The Impractical Bore

The Impractical bore provides their clients with a psychometrically valid model, but fails to ignite the client's interest or provide a meaningful experience. An uninspiring facilitator can significantly impact the face validity experienced by users. This has been described by John Heron (2002) in 'The Complete Facilitators Handbook' as 'the bane of education at all levels' when students experience 'an authoritarian education system, using oppressive forms of teacher authority'. To find out how to avoid this risk, we suggest you read Heron's seminal text.

The 'Snake-Oil' Sales-Person

The 'Snake-Oil' sales-person peddles invalid models and manipulates the client to accept them as valid. The client wakes up to find they are left with belly-ache. An example of this is highlighted in Sir John Foster's 'Enquiry into Scientology' (1971) conducted by a team of psychologists from the governing council of the British Psychological Society. Commenting on Scientology's use of a questionnaire as a recruitment tool the report stated that "no reputable psychologist would accept the procedure of pulling people off the street with a leaflet, giving them a 'personality test' and reporting back in terms that show the people to be 'inadequate', 'unacceptable' or in need of 'urgent' attention". They state that "…the profiles derived from its completion are constructed in such a manner as to give the appearance of being adequate psychometric devices, whereas, in fact, they totally fail to meet the normally accepted criteria". They conclude by saying "the prime aim of the procedure seems to be to convince these people of their need for the corrective courses run by the Scientology organizations". This seems to be a classic example of the snake-oil sales-person applying a psychometrically invalid model in combination with a learning experience of little personal significance to the user.

The Deluder

The Deluder is often a well intended facilitator that creates a meaningful experience for the client, which they may even find helpful. However, in the final analysis, the invalid model used makes the facilitator a peddler in delusion. One striking example of this is the field of astrology where many astrologers genuinely believe in the predictive power of their practice despite there being "simply no reliable data establishing any of astrology's claims" (Schick and Vaughan, 2002). Here is just a small example of the failed attempts to link astrological charts to personality:

- Dr. Jonus Noblitt found no correlation between 155 volunteers taking the 16PF personality questionnaire and the angular relations among planets (Huyghe, 1989).

- Dr. Shawn Carlson gave thirty prominent astrologers full astrological information on 116 people. However, there was no correlation found between their personality predicted by the astrologers and that measured by a self-assessment using the California Personality Inventory (Moody, 1989)

- A meta-study by Geoffrey Dean and Arthur Mather considered over three hundred scientific works on astrology and concluded "astrology presents a dazzling and technically sound superstructure supported by unproven beliefs; it starts with fantasy and then proceeds entirely logically" (Blackmore, 1993)

- Dr. Gauquelin spent years researching astrology and did find some correlations between birth sign and career. Unfortunately these were not those that astrology predicted and he was forced to conclude "every attempt, whether of astrologers or scientists, to produce evidence of astrological laws, has been in vain" (Ingber, 1981).

At the time of writing, the authors are not aware of a single scientific study that has provided good evidence of the validity of astrology's claim to be able to predict aspects of personality.

In summary, having a valid psychometric model and having a meaningful learning experience based on high face validity are two quite different dimensions. While highlighting the need for both, this paper is focused on assessing the IDE's psychometric validity.

Overview of Content Validity

This refers to the systematic determination of whether the content of a questionnaire measures the traits or preferences that it is designed to measure. The developer attempts to build this type of validity into the questionnaire when it is constructed, through the selection of appropriate items. However, establishing psychometric proof of content validity is only possible when that which is being measured is a specific skill, which independent experts agree an item or method can 100% verify e.g. an ear test involving the identification of musical notes could be proven to have content validity. Unfortunately for personality tests such as the IDE, there is no such agreement amongst experts on what constitutes good content validity for items describing psychological preferences. Consequently, content validity cannot be used as an approach to demonstrating the psychometric validity of the IDE. For the IDE, face validity and content validity is of marginal importance in establishing true validity. Consider the opinion of H.L. Mencken below in demonstrating this point:

"The most common of all follies is to believe passionately in the palpably not true. It is the chief occupation of mankind."

Overview of Construct Validity

As noted by Paul Kline in his 1997 book, A Psychometric Prime (pages 36-37):

"Face validity is not a guide to true validity, concurrent validity is applicable only where there are benchmark measures for the variables, and predictive validity, although powerful, is only effective where clear criteria can be established. … Content validity…is suited only to fields with specified skills and knowledge. To obviate these problems as far as possible, Cronbach and Meehl (1955) developed an approach to test validation, a test known as construct validity.

In establishing the construct validity of a test, the first step involves the definition and delineation of the meaning of the test variable. A construct, in the sense of construct validity, is essentially a concept. Hence, delineating the meaning of the test variable means clarifying the nature of the concept to be measured."

Construct validity is a generic name given to a class of multivariate statistical methods whose primary purpose is to define the underlying structure of a data set. The underlying structure of the data is defined by a set of dimensions known as factors. Factor analysis is a widely used powerful statistical technique that can help establish what factors actually exist in questionnaire data or test the assumption that certain hypothesized factors actually do exist. In this context, the procedure enables robust inferences to be developed which relate the data profile to those hypothesized and underlying structures taken to be measured e.g. personality preferences.

Factor analysis can be used either in an *exploratory* or a *confirmatory* mode. Both approaches have been used in this paper. An initial exploratory factor analysis has been followed by a confirmatory factor analysis to test the theoretical hypothesis which underlies the distinction between the Insights colour preferences. However, first we will explore what factor analysis is, through an artificial example.

What is Factor Analysis?

This will be demonstrated with a classic example of factor analysis on school children's exam results. Let us assume we want to determine the factors that drive exam performance based on an analysis of school children's actual exam results. We invite you now to perform your own intuitive factor analysis based on a fictional sample of 9 children's results. To assist you with this visual assessment of the data, we have clustered the 9 children's exam results into 3 batches of 3.

Review table 9 and ask yourself: *'are their any groupings of subjects that have similar exam scores for different clusters of people?'*

Now let us assume that there are different types of intelligence that drive exam performance.

Now ask yourself: *'what would I call the underlying intelligences that drive performance in these subject groupings?'*

For example, you may notice that the exam scores for Latin, French and Spanish are all highly correlated i.e. they are either always all high (see row 3 for Vivien 93%, 97%, 94%) or always

all low (see row 1 for Lynne 45%, 40%, 50%). We could hypothesize (or make the inference) that this is based upon the factor of 'linguistic intelligence'.

Can you use your intuition to name two further types of intelligence represented in the data?

Table 9 – Illustration based on just 9 children's made up exam results

Name	Algebra	Chemistry	Physics	Latin	French	Spanish	Geometry	Engineering	Technical Drawing
Lynne	95%	94%	98%	45%	40%	50%	46%	45%	51%
Doug	90%	97%	93%	43%	35%	20%	43%	35%	24%
Andy	99%	94%	94%	51%	39%	43%	41%	42%	55%
Vivien	60%	24%	43%	93%	97%	94%	30%	51%	42%
Stewart	43%	40%	50%	95%	94%	98%	60%	40%	43%
Colin	47%	31%	20%	100%	98%	90%	55%	35%	43%
Cathy	51%	32%	53%	53%	37%	43%	89%	98%	93%
Mike	40%	41%	42%	30%	61%	52%	95%	97%	98%
Russell	51%	43%	33%	41%	45%	33%	99%	100%	90%

If you were able to intuitively discern the 'intelligences' in this data, it is likely this was only possible due to the small size of the made up sample, the exaggerated polarization of the results in the example and the helpful clustering into 3 blocks of 3 by the authors. If there were 100,000 children's real exam results to analyze, an intuitive visual assessment would be impossible. It is highly likely that the data set would contain such a variation of scores and inter-score trends, that the contamination of true perceptions by erroneous ones would be inseparable through visual inspection alone.

However, powerful computers can easily perform a statistical 'factor analysis' on 100,000 exam results and summarize the variance and covariances into a simple readable table. The output from a 'factor analysis' on 100,000 exam results could look like the data in table 10.

Table 10 – Illustration based on artificial factor analysis data

Variables	**Factor 1**	**Factor 2**	**Factor 3**
Algebra	0.17	0.35	0.61
Chemistry	0.20	0.36	0.50
Physics	0.23	0.33	0.42
Latin	0.74	0.16	0.41
French	0.68	0.20	0.31
Spanish	0.77	0.23	0.27
Geometry	0.13	0.71	0.47
Engineering	0.12	0.65	0.51
Technical Drawing	0.20	0.64	0.44

The dimension for the pupil's name has been collapsed and we now have a manageable matrix with 3 columns and 9 rows. If any number is above 0.3, then the factor is considered statistically significant. Any number below 0.3 can be ignored (Hair et al., 1998).

Now let us continue to assume that there are different types of intelligence that drive exam performance.

Ask yourself: *'based on the factor analysis table, what would I call the three factors identified?'*

What is a 'Factor' and a 'Factor Loading'?

A factor is a construct that is a summary of the relationships within and between a set of variables. In the example of the school children, it is mathematically a linear combination of the exam results.

One academic definition states that a 'factor' is *'a construct operationally defined by its factor loadings'* (Royce, 1963) and that a 'factor loading' is *'the correlations of a variable with a factor'* (Kline, 1997). Factor analysis hunts for emergent patterns of relationships among the many dependent variables e.g. items eliciting scores. The intention is to discover something about the nature of the underlying and independent variables (those thought to be generating the responses to the items) that affect them, noting that the independent variables (e.g. personality) cannot be measured directly.

Table 10 simply shows the factors (1, 2 and 3) and their correlations (termed factor loadings) with the exam results (variables) in different subjects i.e. the statistics package has defined the factors and computed all the factor loadings. However, this still leaves the interpretation of the meaning of the factors down to the analyst.

We could hypothesize that factor one is 'linguistic intelligence' as it loads highly onto Latin, French and Spanish, but lowly onto Geometry, Engineering and Technical Drawing.

We could further hypothesize that factor two is 'spatial awareness' as it loads highly onto Geometry, Engineering and Technical Drawing, but much more lowly onto other subjects.

Finally, we could hypothesize that factor three is 'general reasoning and logical intelligence' as it loads onto most subjects, albeit significantly higher on Algebra, Physics and Chemistry.

Of course, more evidence would be needed before these hypotheses could themselves be claimed as valid knowledge. However, this fictitious example shows how in principle, a complex data set comprised of 100,000 exam results can be reduced to three factors and enable various hypotheses to be formed. It helps relate that which is observable to that which can be hidden or disguised by variability.

Why is Factor Analysis Important for the IDE?

Any personality questionnaire that purports to quantify personality preferences must be able to demonstrate that the preferences exist in the data as factors. The IDE purports to measure four colour preferences.

Let us now assume the same pupils from the previous exam results analysis have now completed the IDE. Their names are listed vertically with one row for each pupil. Listed horizontally are their scores from the 100 items in the IDE (25 frames multiplied by 4 colours). The table below just shows the eight items from the first two frames.

Review the data below and ask yourself *'what patterns do I notice in the data?'*

Table 11 – Illustration based on just 9 people's made up IDE item scores

	IDE Frame One				IDE Frame Two			
Name	Composed and observing	Diplomatic and calming	Open and outgoing	Active and controlling	Amiable and quick	Reliable and restrained	Forceful and goal-oriented	Methodical and logical
Lynne	M	3	L	5	3	L	5	M
Doug	M	2	L	5	L	3	M	5
Andy	L	3	M	1	M	4	1	L
Vivien	L	M	5	1	3	M	1	L
Stewart	1	2	M	L	M	3	L	1
Colin	5	3	L	M	L	1	5	M
Cathy	M	L	3	5	L	3	M	1
Mike	5	3	L	M	4	1	M	5
Russell	L	2	M	1	M	4	1	L

In particular, ask yourself *'what do I notice about the 'fiery red' and 'cool blue' scores?'*

If this pattern was repeated across 100,000 users and across all 25 frames (i.e. whenever they scored red high, they also scored blue high) a 'factor analysis' could be performed and the output could look like the data in table 12. The 'factor analysis' would correctly identify that the 'fiery red' and 'cool blue' items are measuring the same underlying construct (both have 0.6 for factor 1).

If this happened, we would be forced to conclude that there is no difference between 'fiery red' and 'cool blue' items in our data i.e. both colour items would be considered to be measuring the same factor and the hypothesis that 'fiery red' and 'cool blue' represent different aspects of personality would be invalidated by this damning evidence.

Table 12 – Illustration based on artificial factor analysis data

Illustration based on artificial factor analysis data of 100,000 students	Average Factor Loadings	
	Factor 1	Factor 2
Earth Green	-0.1	0.2
Sunshine Yellow	-0.2	**0.7**
Cool Blue	**0.6**	-0.2
Fiery Red	**0.6**	0.1

IDE Factor Analysis Results

The previous example was of course constructed to help explain the conceptual and procedural elements of factor analysis. What follows now are some key results based on a real sample of 519,467 English version S3.0 Insights Discovery evaluators and further samples using evaluators translated into French, German, Spanish and Dutch. Additional factor analysis, based on Insights Discovery Evaluators from other countries, are available in more detailed papers at the bpc at the University of Westminster (Van Erkom Schurink, 2004).

Exploratory Factor Analysis

Exploratory factor analysis was conducted to establish the number of factors present in the IDE data. The method used was the latent root criterion that involves the analysis of the Eigenvalues and the sum of squared loadings. This method calculates the amount of variance accounted for by each factor. Only the factors having an Eigenvalue equal or greater than 1 are considered significant.

The minimum number of factors to be extracted can be determined by the 'cumulative sum of squared loadings' that indicates the percentage of variance explained by the incremental factoring procedure. When the later factors do not significantly increase the total variance explained, it is debatable whether their inclusion adds value to the analysis.

Table 13 shows the Eigenvalues and the sum of squared loadings obtained for the English version S3.0 of the IDE. It is based on 519,467 IDEs completed between 2003 and 2008.

Table 13 – The Eigenvalues and the sum of squared loadings based on an exploratory factor analysis

English S 3.0 IDE		
N Factor (component)	Initial Eigenvalue (≥1)	Cumulative Squared Loadings
1	19	19%
2	14	34%
3	4	38%
4	3	40%
5	2	42%
6	2	44%
7	2	46%
8	2	48%
9	1	49%
10	1	50%
11	1	52%
12	1	53%
13	1	54%
14	1	55%
15	1	56%
16	1	57%
17	1	58%
18	1	59%

The data in this table suggests that there are two significant factors accounting for 34% of the variation in the data. Factors 3 and 4 make only a marginal contribution to the variance explained (4% and 3% respectively). Factors 5 through 18 could be considered of peripheral importance, adding between 1% and 2% with each additional factor. The Eigenvalues for factors 5 through 18 are also very low (between 1 and 2 for each factor). Although more detailed papers have explored the top four factors further, the confirmatory factor analysis that follows in this paper is primarily concerned with the top two most significant factors.

Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

• The polar opposite nature of the 'fiery red' vs. 'earth green' items is apparent

• The polar opposite nature of the 'sunshine yellow' vs. 'cool blue' items is 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 hypothesized structure and offer evidence for the construct validity of the IDE. The method used was the Varimax Principal Component. In addition, Appendix C contains an analysis on the same data using Varimax Maximum Likelihood, Oblimin Principal Component and Oblimin Maximum Likelihood. Encouragingly, all these different methods produce very similar results.

These results show that the conceptual model's hypothesized polar dynamics are represented by the colour opposition in the colour loadings. The model suggests that the polar opposite of 'cool blue' is 'sunshine yellow' and this is supported by the factor analysis i.e. it can be seen that the 'cool blue' items load *negatively* onto factor two and the 'sunshine yellow' 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 contribute to a small increase in the variance explained, is an added value but not a pre-requisite to the validation of the IDE 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 tables that follow.

Table 14 – English S3.0 IDE - factor loadings summary table

IDE English version S3.0 N=519'467 **Completed from around the globe**	**Average Factor Loadings**	
	Factor 1	**Factor 2**
Earth Green	**0.54**	-0.03
Sunshine Yellow	-0.04	**0.54**
Cool Blue	0.08	**-0.55**
Fiery Red	**-0.58**	0.05

IDE English version S3.0 N=114'670 **Completed just in the UK**	**Average Factor Loadings**	
	Factor 1	**Factor 2**
Earth Green	**0.56**	-0.03
Sunshine Yellow	-0.06	**0.53**
Cool Blue	0.11	**-0.58**
Fiery Red	**-0.61**	0.10

Both of the above tables are based on the same S3.0 English language evaluator. However, the table on the left is based on IDEs completed by people in many countries from around the globe who chose to use the English language evaluator, rather than one of the other 31 languages available i.e. they were based in the UK, USA, Canada, South Africa etc..

The table on the right is based on IDEs completed by people based in the UK, who also chose to use the English language evaluator. Analyzing the right hand table we find that the 'fiery red' items load strongly onto Factor 1 at *minus* 0.61. The 'earth green' items also load strongly onto Factor 1 at *plus* 0.56. The opposite signs of these loadings supports the theoretical construct of the model that hypothesises that the 'fiery red' and 'earth green' constructs are polar opposites.

At *minus* 0.58, the 'cool blue' items load strongly onto factor 2. The 'sunshine yellow' items load onto factor 2 at *plus* 0.52

Again, the opposite signs of these loadings support the theory that 'cool blue' and 'sunshine yellow' are polar opposites constructs.

The above table is an *average* of the factor loadings. However, it is also possible to analyze the factor loadings for *each of the 100 items* in the IDE. This detailed data is contained in appendix C. For example, in Appendix C, frame two's 'earth green' item is '*reliable and restrained*' and this item loads onto factor 1 at 0.52 and factor 2 at -0.30 (for VARIMAX rotation with Principal Component extraction method)

Figure 13 is a scatter plot of the 100 items. It shows all 100 items loading onto the top two factors. Both diagonal's scales range from 0.8 to minus 0.8, with the axis crossing at zero.

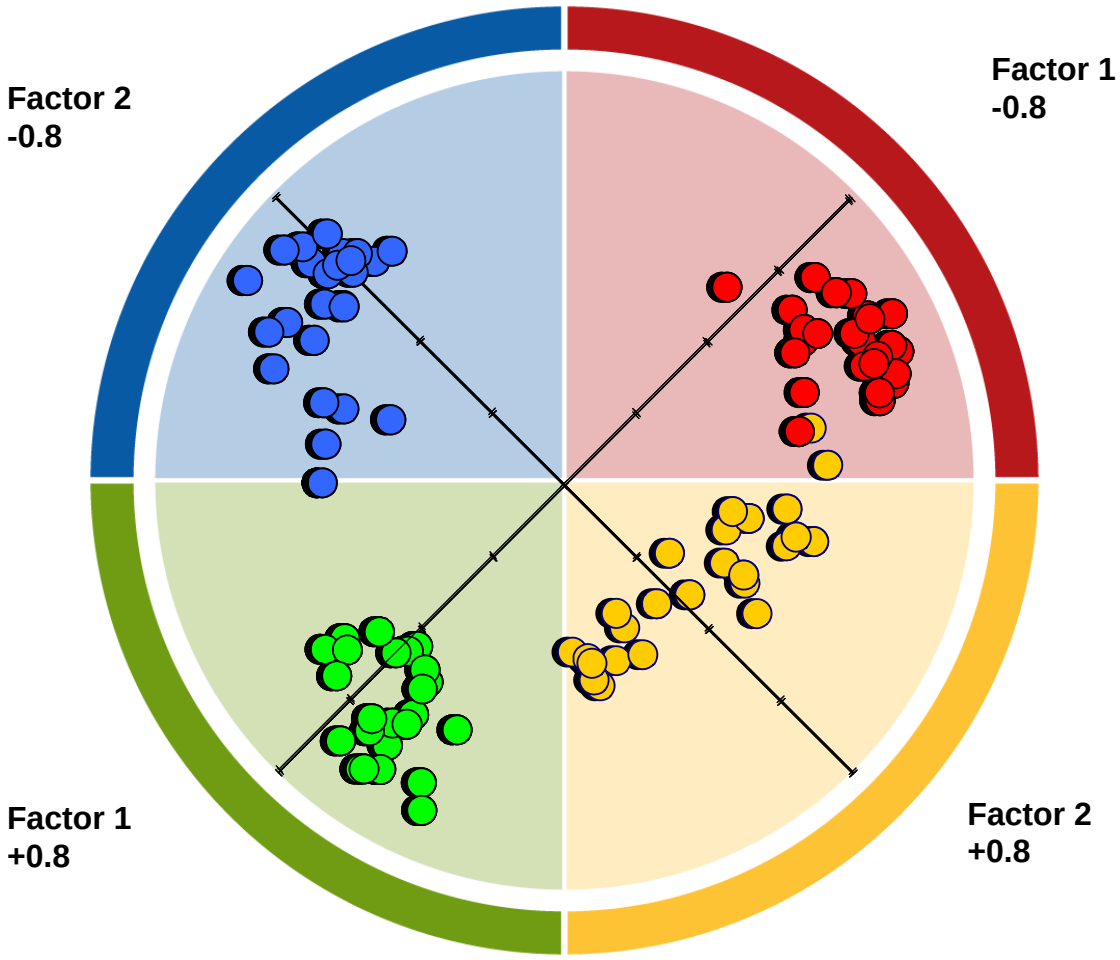


Figure 13 – English S3.0 IDE – Graph of the 100 items (25 x 4 colours) plotted against the factors

The graph in Figure 13 has been superimposed onto the Insights Discovery wheel. It shows the relationship between each of the 100 items (four colours multiplied by 25 frames) and the top two factors. It can be seen that 97 out of 100 items appear in the 'correct' quadrant. Two yellow items appear just inside the red quadrant and one 'cool blue' item appears on the border of the 'cool blue'/'earth green' quadrants.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model. Further results for other country's evaluators show a similar pattern and the results for the French, German, Spanish and Dutch evaluators are shown next. The one borderline result is the second factor based on the German data. Although the second factor shows a statistically significant loading for the 'sunshine yellow' items (0.51), the factor loading for the 'cool blue' items is -0.27 (just below the -0.3 'rule of thumb'). A (somewhat generous) argument to round the -0.27 to -0.3 could be made. However, a further analysis found that for the German S3.1 evaluator, if the third factor is examined, the 'cool blue' items load onto it at 0.48 and the 'sunshine yellow' items at –0.24. Given this data, one could also argue that factors 2 and 3 together provide evidence of the polar nature of the 'sunshine yellow' and 'cool blue' items.

Table 15 – German and Dutch factor loadings

German version S3.1 N=21'417	Average Factor Loadings	
	Factor 1	Factor 2
Earth Green	**-0.52**	-0.03
Sunshine Yellow	0.07	**0.51**
Cool Blue	-0.04	**-0.52**
Fiery Red	**0.52**	0.09

Dutch version S1.2 (beta) N=11'677	Average Factor Loadings	
	Factor 1	Factor 2
Earth Green	**-0.47**	-0.04
Sunshine Yellow	0.01	**0.55**
Cool Blue	-0.09	**-0.54**
Fiery Red	**0.60**	0.09

Table 16 – French and Spanish factor loadings

French version S2.0 N=14'435	Average Factor Loadings	
	Factor 1	Factor 2
Earth Green	**-0.44**	-0.04
Sunshine Yellow	-0.04	**0.52**
Cool Blue	-0.07	**-0.46**
Fiery Red	**0.56**	0.02

Spanish version S1.3 N=5'392	Average Factor Loadings	
	Factor 1	Factor 2
Earth Green	-0.02	**0.46**
Sunshine Yellow	**0.52**	0.01
Cool Blue	**-0.49**	0.05
Fiery Red	0.02	**-0.52**

Technical Explanation of Construct Validity

As a matter of comparison, other studies testing the validity of Jungian psychometric instruments, report the following results:

- For the Jungian 'Singer-Loomis' instrument, individual factors included in a four-factor solution account for between 4% and 8% of the variance, and all factors explain 34% of the variance (Loomis & Singer, 1980)
- Four distinct factors are correlated to the four MBTI constructs, accounting for 56% of the variance (Tzeng & al., 1989)
- A two-factor solution was found in the MBTI, corresponding to the EI and the JP dimensions (Sipps & Alexander, 1987)
- Six distinct factors were found in the MBTI, of which four resembled the four Jungian scales (Sipps et al. (1985)
- A four-factor model vs. two competing five-factor models was found in the MBTI (Harvey et al., 1995)

Criterion Validity

Criterion validity includes both predictive validity and concurrent validity. Concurrent validity studies are underway with the University of Westminster, but the results are not available yet. Predictive validity is evidenced by the data showing how aggregate data for different professions score differently across the colours. To illustrate the IDE's capacity in this aspect, data across many professions has been analysed and four data sets are presented graphically here.

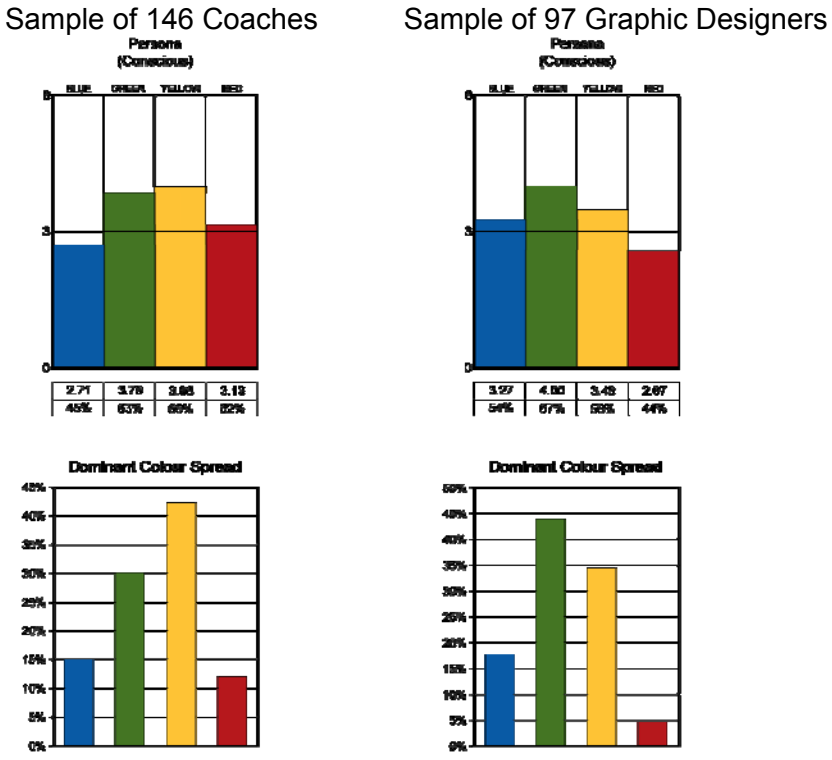


Figure 14 - A graphical view of Coaches vs. Graphic Designers, in support of the argument for predictive validity

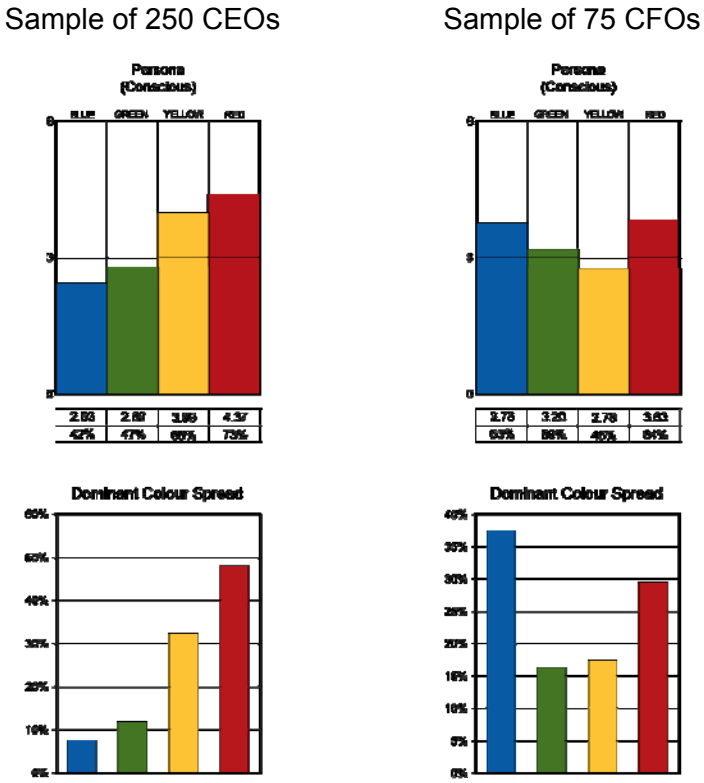


Figure 15 - A graphical view of CEOs vs. CFOs, in support of the argument for predictive validity

Table 17 shows one row for each different job description.

While this data indicates that people in certain roles tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they are doing they're job or how capable they are in fulfilling that role.

Table 17 - A tabular summary of the job description norms data

Job Title	Sample	Average Colour Scores Scale 0 to 6				Percentage of Norm Sample with Dominant Colour			
	Size	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Account Director	253	2.7	2.9	4.0	4.1	10%	17%	34%	40%
Account Executive	758	3.1	3.3	3.9	3.6	17%	19%	36%	27%
Account Manager	1,119	2.9	3.3	3.9	3.6	15%	19%	36%	30%
Account Supervisor	112	3.3	3.7	3.8	3.1	22%	30%	31%	17%
Accountant	213	4.3	3.9	2.8	2.6	44%	32%	16%	8%
Administrative Assistant	753	3.6	4.2	3.4	2.4	23%	45%	23%	9%
Administrator	235	3.6	4.1	3.4	2.2	23%	43%	20%	13%
Analyst	157	3.9	3.6	3.1	3.0	41%	26%	13%	20%
Area Manager	252	3.0	3.2	3.5	3.8	18%	19%	20%	43%
Associate	168	3.5	3.5	3.4	3.0	32%	23%	26%	20%
Associate Director	168	3.1	3.2	3.6	3.6	19%	21%	27%	33%
Branch Manager	250	3.1	3.6	3.5	3.4	18%	30%	25%	28%
Business Analyst	385	3.9	3.6	3.2	2.7	38%	24%	24%	14%
Business Development Executive	133	2.7	3.5	4.3	3.2	16%	16%	53%	15%
Business Development Manager	347	3.0	3.1	3.8	3.7	18%	15%	30%	38%
Business Manager	294	3.1	3.2	3.7	3.6	20%	23%	26%	31%
CEO	250	2.5	2.8	4.0	4.4	8%	12%	32%	48%
CFO	75	3.8	3.2	2.8	3.8	37%	16%	17%	29%
Coach	146	2.7	3.8	4.0	3.1	15%	30%	43%	12%
Commercial Manager	141	3.7	2.9	3.2	3.6	33%	16%	19%	32%
Construction Manager	130	3.8	3.1	3.0	4.0	32%	13%	12%	43%
Consultant	1,334	3.0	3.4	3.8	3.4	19%	23%	35%	23%
Controller	180	4.2	3.3	2.6	3.5	51%	14%	9%	26%
Customer Service	105	3.6	3.7	3.4	3.0	28%	31%	25%	16%
Customer Service Manager	149	3.3	3.8	3.6	2.7	27%	31%	30%	13%
Customer Service Representative	156	3.8	4.3	3.2	2.3	30%	43%	15%	12%
Desk Based Account Manager	119	3.0	3.8	3.9	2.9	16%	35%	34%	16%
Directeur	220	2.5	3.4	3.8	3.9	11%	24%	21%	44%
Director	1,306	2.9	3.1	3.7	3.9	16%	17%	28%	39%
Director Human Resources	144	3.1	3.9	3.4	3.2	15%	36%	21%	28%
District Sales manager	116	2.8	3.4	4.0	3.6	13%	23%	37%	27%
Editor	100	3.7	3.1	3.3	3.3	34%	20%	24%	22%
Engineer	126	3.8	3.3	2.9	3.3	40%	19%	19%	23%
Executive Assistant	392	3.5	4.1	3.5	2.5	21%	42%	26%	12%
Executive Director	228	2.6	3.5	3.9	3.7	11%	26%	35%	29%

Job Title	Sample	Average Colour Scores Scale 0 to 6				Percentage of Norm Sample with Dominant Colour			
	Size	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Field Sales Consultant	282	2.9	3.3	4.0	3.4	18%	17%	41%	24%
Finance Director	114	4.1	2.8	2.7	3.8	47%	7%	11%	35%
Finance Manager	145	4.0	3.3	2.9	3.2	43%	21%	18%	17%
Financial Analyst	188	4.1	3.3	3.0	3.1	43%	19%	15%	23%
Financial Controller	101	4.1	3.3	2.6	3.4	47%	23%	7%	24%
General Manager	493	3.0	3.0	3.5	4.1	15%	14%	24%	47%
HR Administrator	113	3.7	4.2	3.3	2.3	26%	50%	20%	4%
HR Advisor	143	3.1	3.6	3.9	3.0	13%	31%	38%	19%
HR Business Partner	118	2.9	3.4	4.1	3.2	22%	20%	38%	20%
HR Consultant	176	3.2	3.5	3.8	3.0	22%	27%	36%	15%
HR Director	124	2.7	3.4	3.8	3.7	14%	22%	36%	28%
HR Manager	417	3.2	3.6	3.5	3.1	23%	27%	30%	20%
Human Resources Consultant	114	3.2	4.0	3.6	2.8	18%	33%	34%	15%
Human Resources Manager	339	3.2	3.8	3.6	3.1	20%	33%	30%	17%
IT Manager	122	3.5	3.3	3.3	3.5	27%	24%	13%	36%
Key Account Manager	145	3.2	3.0	3.5	3.9	19%	19%	28%	34%
Manager	1,090	3.2	3.5	3.4	3.3	22%	29%	26%	24%
Managing Director	607	2.5	2.8	3.9	4.3	10%	12%	29%	49%
Marketing Director	123	2.6	2.7	4.0	4.2	12%	11%	33%	45%
Marketing Manager	559	3.0	3.1	3.9	3.6	20%	17%	39%	24%
MD	106	2.6	2.3	3.9	4.7	5%	5%	31%	59%
Medical Representative	175	3.1	3.5	3.9	3.1	17%	26%	39%	18%
Office Manager	247	3.5	3.9	3.4	2.8	23%	37%	22%	17%
Operations Director	112	3.2	2.9	3.4	3.9	21%	18%	19%	42%
Operations Manager	404	3.6	3.4	3.1	3.5	24%	30%	12%	35%
Operator	109	3.6	3.9	3.2	2.8	28%	37%	19%	17%
Owner	181	2.8	3.3	3.8	3.8	11%	19%	30%	40%
Partner	394	2.9	3.1	3.5	3.8	18%	18%	25%	39%
Personal Assistant	161	3.7	4.1	3.5	2.2	24%	44%	24%	9%
Physical Therapist	146	3.4	4.4	3.5	2.0	16%	56%	24%	3%
President	936	2.6	3.0	3.9	4.1	10%	14%	33%	43%
Principal	268	2.8	3.5	3.8	3.6	15%	25%	31%	29%
Product Manager	370	3.5	3.1	3.5	3.5	29%	15%	27%	30%
Production Manager	114	3.7	3.6	3.0	3.3	32%	27%	18%	23%
Program Coordinator	134	3.1	4.1	3.6	2.8	14%	40%	32%	13%
Program Manager	289	3.4	3.6	3.5	3.2	22%	30%	26%	23%
Programme Manager	117	3.2	3.2	3.7	3.5	27%	13%	33%	27%
Project Leader	101	3.4	3.4	3.4	3.2	26%	25%	22%	28%
Project Manager	1,294	3.5	3.3	3.3	3.4	28%	21%	23%	29%

Job Title	Sample	Average Colour Scores Scale 0 to 6				Percentage of Norm Sample with Dominant Colour			
	Size	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Receptionist	193	3.2	4.4	3.5	2.2	13%	58%	22%	7%
Regional Director	127	3.0	3.2	3.6	3.9	17%	20%	24%	39%
Regional manager	177	3.0	3.3	3.7	3.6	14%	20%	29%	37%
Regional Sales Manager	191	2.8	3.0	3.8	3.9	14%	19%	37%	30%
RN	321	3.5	4.5	3.4	2.0	17%	59%	17%	7%
Sales	328	2.9	3.5	4.0	3.5	14%	22%	39%	26%
Sales Associate	138	3.0	3.8	4.0	3.1	16%	33%	41%	11%
Sales Consultant	393	2.9	3.3	4.0	3.6	13%	23%	42%	23%
Sales Director	121	2.6	2.7	3.9	4.5	10%	9%	28%	53%
Sales Executive	185	3.0	3.4	3.9	3.3	20%	26%	30%	24%
Sales Manager	742	2.9	3.1	3.7	3.9	15%	18%	28%	40%
Sales Rep	225	2.6	3.1	4.2	3.9	9%	14%	44%	32%
Sales Representative	623	3.1	3.3	3.9	3.5	14%	22%	39%	26%
Secretary	298	3.6	4.5	3.2	1.8	17%	62%	17%	4%
Senior Account Manager	133	2.8	3.2	4.0	3.7	11%	19%	39%	31%
Senior Auditor	144	3.7	3.3	3.3	3.1	35%	19%	22%	24%
Senior Consultant	207	3.1	3.1	3.6	3.8	24%	14%	24%	38%
Senior Manager	202	3.2	3.2	3.3	3.8	24%	16%	23%	37%
Senior Project Manager	105	3.6	3.3	3.2	3.5	30%	22%	16%	32%
Service Leader	139	3.6	3.8	3.6	2.5	29%	30%	30%	12%
Service Manager	170	3.2	3.5	3.6	3.4	21%	29%	22%	28%
Store Manager	103	3.3	3.2	3.5	3.7	27%	16%	23%	34%
Student	1,321	3.0	3.5	3.7	3.2	15%	29%	33%	23%
Supervisor	432	3.4	3.9	3.2	3.1	18%	41%	16%	25%
Systems Analyst	131	4.1	3.5	3.0	3.1	46%	22%	17%	15%
Systems Engineer	126	4.1	3.9	3.0	2.5	45%	29%	12%	14%
Teacher	415	3.2	4.1	3.7	2.5	20%	41%	29%	11%
Team Leader	904	3.3	3.8	3.5	2.9	20%	36%	28%	17%
Team Manager	280	3.1	3.8	3.7	3.0	18%	36%	30%	17%
Trainer	357	2.8	3.7	4.1	2.9	13%	30%	45%	12%
Training Consultant	149	2.6	3.6	4.3	3.0	12%	23%	48%	17%
Training Manager	264	2.6	3.6	4.1	3.1	11%	28%	43%	17%
Underwriter	102	3.8	3.9	2.9	2.8	32%	43%	12%	13%
Unit Supervisor	130	3.5	4.3	3.0	2.5	27%	47%	19%	8%
Vice President	381	3.2	3.1	3.5	3.8	23%	18%	23%	36%

Conclusion

There is strong evidence of construct validity as demonstrated by the factor analysis. Two statistically significant factors have been identified and these explain around 34% of the variance. The factor analysis data also provides evidence to support the interpretation that 'cool blue' and 'sunshine yellow' are polar opposites (as are 'fiery red' and 'earth green'), as evidenced by the factors loading both positively and negatively respectively onto the polar opposite colours.

Evidence of predictive validity is provided through an analysis of how the colour scores vary strongly by profession, e.g. accountants tend to have higher 'cool blue' scores and CEOs tend to have higher 'fiery red' scores.

There is strong evidence of the reliability of the measure of the four colours, as demonstrated by the Cronbach-Alpha. Scores of between 0.92 and 0.93 compare favourably when benchmarked against other personality tests where research shows range between 0.7 and 0.9

Large samples of interesting norm data are available.

The development of the model through item analysis has been completed to a high standard.

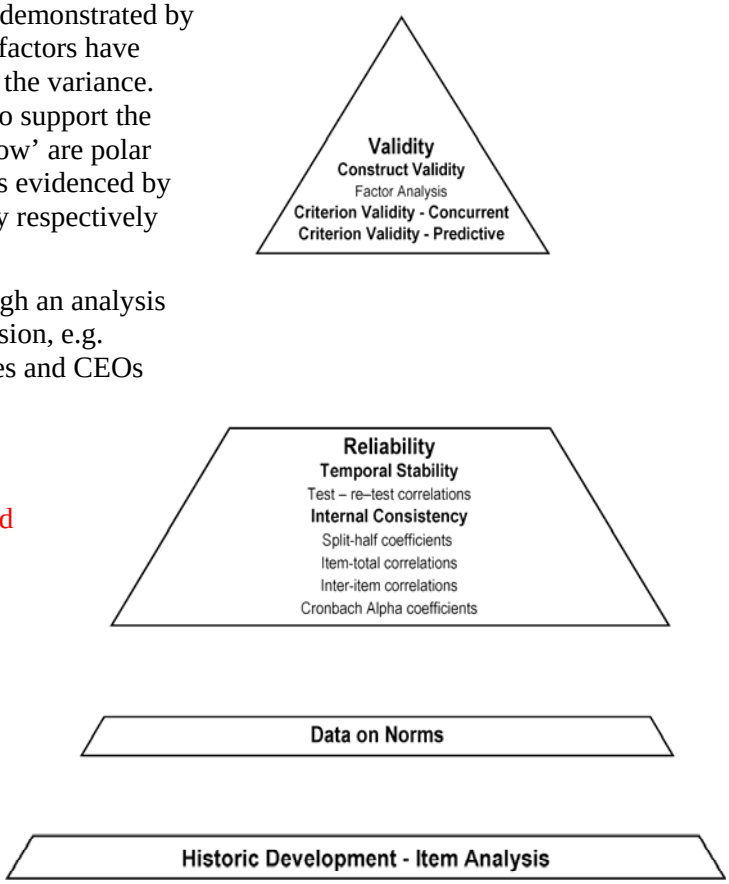


Figure 1 Repeated – Pyramid of Key Psychometric Statistics

This paper has explained how the Insights Discovery model has been developed through item analysis, supported by a large quantity of good quality data on norms. Building on this base, strong evidence of the model's reliability has been presented through the internal consistency tests and the test/re-test temporal stability data. The construct validity has been demonstrated through factor analysis and there is good predictive validity data by profession. These results all compare favourably with other Jungian based instruments that are held in high regard by psychometricians and which also meet the standards set out by both the American Psychological Association and the British Psychological Society. In summary, we have strong evidence to support the four colour measures calculated from the Insights Discovery model being both reliable and valid.

References

General References

Beauchamp, M.R., Maclachlan, A. and Lothian, A.M., (2005) Communication Within Sport Teams: Jungian Preferences and Group Dynamics: The Sport Psychologist, Vol 19, Number 2, Pages 203 to 220, June 2005, ISSN 0888-4781, ISBN 0-736011072, Human Kinetics Publishers, Inc.

Benton, S. (2005) 'Every Individual Is the Exception to the Rule', The Association for Project Managers Handbook, UK

Blackmore, S., (1993) 'Dying to Live: Near-Death Experiences', Prometheus Books UK

Glassman, W. E. (1995) 'Approaches to Psychology', Buckinghamshire Open University Press

Heron, J. (2002) 'The Complete Facilitators Handbook', Kogan Page Ltd, London, UK

Huyghe, P. (1989) 'Moody's Crystal Ball', Omni, p90

Ingber, D, (1981) 'Visions of an Afterlife', Science Digest, Jan/Feb 1981, page 142

Moody, R. (1989) 'The Light Beyond', Bantam Publishers

Mullaly, M.E. and Thomas, J. (2004) 'Linking Personality & Project Success: Exploring the inter-relationship of psychological type and project manager competency', paper presented at IRNOP VI – Turku, Finland, 25-27 August 2004

Tosey, P. and Gregory, J. (2002) 'Dictionary of Personal Development', London: Whurr

Foster, Sir J, K.B.E., Q.C., M.P., (1971) 'Enquiry into the Practice and Effects of Scientology', published by Her Majesty's Stationery Office, London, December 1971

References Concerning the use of Ipsative Scales

Binder, A (1984) 'Restrictions on Statistics Imposed by Method of Measurement: Some Reality, Some Myth', Journal of Criminal Justice, Vol. 12: 467-481

Harvey, R. J., Murry, W. D., & Stamoulis, D. T. (1995) 'Unresolved Issues in the Dimensionality of the Myers-Briggs Type Indicator', Educational and psychological Measurement, 55, 535-544

Jaccard, J, Wan, C. K. (1996) 'LISREL approaches to interaction effects in multiple regression', Thousand Oaks, CA, Sage Publications

Kim, J. O. (1975) 'Multivariate Analysis of Ordinal Variables' American Journal of Sociology, Vol. 81: 261-298

Kim, J. O, Mueller, C. W. (1978) 'Factor Analysis: Statistical methods and practical issues', Thousand Oaks, CA: Sage Publications, Quantitative Applications in the Social Sciences

Labovitz, S. (1967) 'Some Observations on Measurement and Statistics', Social Forces, Vol. 46: 151-160Series, No. 14

Labovitz, S. (1970) 'The Assignment of Numbers to Rank Order Categories', American Sociological Review, Vol. 35: 515-524

Myers, I. B., McCalley, M. H., Quenk, N. I., & Hammer, A. L. (1998)' MBTI Manual: A guide to the development and use of the Myers-Briggs Type Indicator' 3rd Ed. Palo Alto., CA: Consulting Psychologists Press

Sipps, G. J., Alexander, R. A., & L. Friedt. (1985) 'Item Analysis of the Myers-Briggs Type Indicator', Educational and Psychological Measurement. 45(4), pp. 789-796

Sipps, G. J., & Alexander, R. A. (1987) 'The multifactorial nature of extraversion-introversion In the Myers-Briggs Type Indicator and Eysenck personality inventory', Educational and Psychological Measurement, 47, 543-522

Tischler, L. (1994) 'The MBTI factor structure', Journal of Psychological Type, 31, 24-31

Tzeng, O., Landis, D., & Chen, J. M. (1989) 'Measurement and utility of continuous unipolar rating for the Myers-Briggs Type Indicator', Journal of Personality Assessment, 53, 727-738

References for both Reliability and Validity

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

Kline, P (1997) 'A Psychometrics Primer', Free Association Books, NYC

Kline, P (2000) 'Handbook of Psychological Testing', Routledge, 11 New Fetter Lane, London, EC4P 4EE

Van Erkom Schurink, C. (2004) 'The Issue of Reliability and Validity in Psychometric Instruments' Working Paper, The Analytical Research Bureau, p 40

References for Reliability – Internal Consistency

APA (1999), 'Standards for educational and psychological testing', prepared by a joint committee comprised of the American Psychological Association, the American Educational Research Association and the National Council on Measurement in Education

Buley, J. (2000) 'Reliability, Validity and Correlation'

Churchill, G. A. (1979) 'A Paradigm for Developing Better Measures of Marketing Constructs', Journal of Marketing Research, 16 February, 64-73

Cronbach, L. J., & Meehl, P. E. (1955) 'Construct Validity in Psychological Tests' Psychological Bulletin, 52, 281 – 302

Dawis, R. V. (1987) 'Scale construction', Journal of Counselling Psychology, 34, 481-489

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

Green, S.B., Lissit, R.W., Mulaik, S.A. (1977) 'Limitation of coefficient alpha as an index of test unidimentionality', Educational and Psychological Measurement, 37, 827-838

Henson, R. K., Kogan, L. R., & Vacha-Haase, T. (2001) 'A reliability generalization study of the Teacher Efficacy Scale and related instruments', Education and Psychological Measurement, 61, 404-420

Nunnaly, J. C. (1972) 'Educational Measurement and Evaluation', 2nd ed. McGraw-Hill, New York, NY

Nunnaly, J. C. (1979) 'Psychometric Theory', McGraw-Hill, New York, NY

Peter, J.P. (1979) 'Reliability: A Review of Psychometric Basics and Recent Marketing Practices', Journal of Marketing Research, 16 (2), 6-17

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

Vacha-Haase, T. (1998) 'Reliability generalization: Exploring variance in measurement error affecting score reliability across studies', Educational and Psychological Measurement, 58, 6-20

Van Erkom Schurink, C. (August 2004) 'Measure of Validity Through Factor Analysis, with a Focus on the Assessment of the Various Versions of the Insights Discovery Evaluator', working paper

Wheelwright, J.B., Wheelwright, J.H., and Buehler, J.A. (1964) 'Jungian Type Survey (The Gray Wheelwright Test)', San Francisco: Society of Jungian Analysts of Northern California

References for Reliability - Test-Re-test Reliability Scales

Capraro, M.R., & Capraro, M.M., (2000?) 'Myers Briggs Type Indicator Score Reliability Across Studies: A Meta-Analytic Reliability Generalization Study', Working Paper (rcapraro@coe.tamu.edu)

Carlson, J.G., (1985) 'Recent Assessment of the Myers-Briggs Type Indicator', Journal of Personality Assessment, 49 (4), 356-365

Carlyn, M., (1977) 'An Assessment of the Myers-Briggs Type Indicator', Journal of Personality Assessment, 41 (5), 461-466

Halpin, N., Green, M., (2005) 'Medical Students' Personality – a Typological Approach': a study of University of Dundee 1st year medical students: working paper

Harrell, T.H., & Lombardo, T.A., (1984) 'Validation of an Automated 16PF Administration Procedure', Journal of Personality Assessment, 49 (6), 638-642

References for Validity

Cattell, R.B., (1966) 'The Scree Test for the Number of Factors', Multivariate Behavioural Research, 1 April, 245-76.

Cliff, N., Hamburger, C.D., (1967) 'The Study of Sampling Error in Factor Analysis by Means of Artificial Experiments' Psychology Bulletin, 68, 430-45

Gorsuch, R.L., (1983) 'Factor Analysis', Hillsdale, N.J.; Lawrence Erlbaum Associates

Kaiser,H.F., (1970) 'A Second Generation Little Jiffy', Psychometrika, 35, 401-15

Kaiser,H.F., (1974) 'Little Jiffy, Mark IV', Educational and Psychology Measurements, 34, 111-17

Loomis, M., & Singer, J. (1980) 'Testing the Bipolarity Assumption in Jung's Typology', Journal of Analytical Psychology, 25, 4

Paul, A.M. (2004) 'The Cult of Personality', New York, NY: Free Press

Remarczyk, M. (2005) MSc dissertation 'Face Validity of Insights Discovery', available from the BPC, University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Royce, J.R. (1963) 'Factors as theoretical constructs', in D.N. Jackson and S. Messick (eds) Problems in Human Assessment. New York: McGraw-Hill

Appendix A - Version 3.0 of the English Insights Discovery Preference Evaluator

Insights® Learning & Development Ltd
Jack Martin Way, Clearbrook Business Park
Dundee DD4 9FF, Scotland
Tel: +44 (0) 1382 908050 Fax: +44 (0) 1382 908051
E-mail insights@insights.com
www.insights.com

University of Westminster

bpc

INSIGHTS DISCOVERY PREFERENCE EVALUATOR **Version S3.0 (UK)**

Personal Details (please use BLOCK CAPITALS)

Date :____/____/____(DD/MM/YY)
Title :_____ Male: ☐ Female: ☐
First Name :_____
Last Name :_____
Job Title :_____
Department :_____
Company :_____
Address :_____
:_____
:_____Postcode_____

Telephone :_____
Fax :_____
E-mail :_____
Date of Birth :____/____/____(DD/MM/YY)
Staff No. :_____

Insights Use

0.	Sensitive and diplomatic	L 1 2 3 4 5 (M)
	Encouraging and valuing	L 1 2 3 (4) 5 M
	Precise and deliberate	L (1) 2 3 4 5 M
	Results-oriented and fast	(L) 1 2 3 4 5 M
1.	Composed and observing	L 1 2 3 4 5 M
	Diplomatic and calming	L 1 2 3 4 5 M
	Open and outgoing	L 1 2 3 4 5 M
	Active and controlling	L 1 2 3 4 5 M
2.	Amicable and quick	L 1 2 3 4 5 M
	Reliable and restrained	L 1 2 3 4 5 M
	Forceful and goal-oriented	L 1 2 3 4 5 M
	Methodical and logical	L 1 2 3 4 5 M
3.	Calm and even-tempered	L 1 2 3 4 5 M
	Determined and dominant	L 1 2 3 4 5 M
	Buoyant and light-hearted	L 1 2 3 4 5 M
	Exact and precise	L 1 2 3 4 5 M
4.	Confident and vigorous	L 1 2 3 4 5 M
	Orderly and concise	L 1 2 3 4 5 M
	Familiar and stable	L 1 2 3 4 5 M
	Talkative and genial	L 1 2 3 4 5 M
5.	Logical and clear	L 1 2 3 4 5 M
	Direct and challenging	L 1 2 3 4 5 M
	Loyal and accommodating	L 1 2 3 4 5 M
	Sociable and active	L 1 2 3 4 5 M

When you have completed the Evaluator, please ensure every frame has been allocated **ONE 'M', ONE 'L'** and **Two Different Values** selected from 1, 2, 3, 4, or 5.

Insights Learning & Development, Tel: +44 (0)1382 908050 Fax: +44 (0)1382 908051

INSIGHTS DISCOVERY PREFERENCE EVALUATOR **Version S3.0 (UK)**

Name: _____ Date: ____/____/____ (DD/MM/YY)

6.	Relating and amenable	L 1 2 3 4 5 M
	Expressive and hopeful	L 1 2 3 4 5 M
	Powerful and assertive	L 1 2 3 4 5 M
	Thinking and self-contained	L 1 2 3 4 5 M
7.	Demonstrative and persuasive	L 1 2 3 4 5 M
	Questioning and reflective	L 1 2 3 4 5 M
	Initiating and self-confident	L 1 2 3 4 5 M
	Stable and concerned	L 1 2 3 4 5 M
8.	Resolute and confident	L 1 2 3 4 5 M
	Social and cheerful	L 1 2 3 4 5 M
	Faithful and helping	L 1 2 3 4 5 M
	Considered and correct	L 1 2 3 4 5 M
9.	Sensitive and diplomatic	L 1 2 3 4 5 M
	Precise and deliberate	L 1 2 3 4 5 M
	Encouraging and valuing	L 1 2 3 4 5 M
	Results-oriented and fast	L 1 2 3 4 5 M
10.	In-charge and firm	L 1 2 3 4 5 M
	Reserved and cooperative	L 1 2 3 4 5 M
	Outgoing and gregarious	L 1 2 3 4 5 M
	Meticulous and detailed	L 1 2 3 4 5 M

11.	Team-focused and Impulsive	L 1 2 3 4 5 M
	Accurate and rational	L 1 2 3 4 5 M
	Even-tempered and amiable	L 1 2 3 4 5 M
	Task-oriented and direct	L 1 2 3 4 5 M
12.	Analysing and painstaking	L 1 2 3 4 5 M
	Friendly and entertaining	L 1 2 3 4 5 M
	Competitive and robust	L 1 2 3 4 5 M
	Unassuming and responsive	L 1 2 3 4 5 M
13.	Constant and attentive	L 1 2 3 4 5 M
	Influencing and expressive	L 1 2 3 4 5 M
	Analytical and evaluating	L 1 2 3 4 5 M
	Bold and objective	L 1 2 3 4 5 M
14.	Strong-willed and purposeful	L 1 2 3 4 5 M
	Reasoned and particular	L 1 2 3 4 5 M
	Eager and engaging	L 1 2 3 4 5 M
	Concerned and sensitive	L 1 2 3 4 5 M
15.	Systematic and principled	L 1 2 3 4 5 M
	Fun-loving and popular	L 1 2 3 4 5 M
	Steadying and moderating	L 1 2 3 4 5 M
	Fast and reinforcing	L 1 2 3 4 5 M

Once you have allocated 'M' and 'L', and are weighting the two remaining word pairings on the scale 1, 2, 3, 4, or 5, please do not choose the same weighting twice.

University of Westminster

bpc

INSIGHTS DISCOVERY PREFERENCE EVALUATOR **Version 83.0 (UK)**

Name: _____ Date: _____/_____/_____(DD/MM/YY)

16. Persuasive and animated	L	1	2	3	4	5	M
Decisive and immediate	L	1	2	3	4	5	M
Discreet and analytical	L	1	2	3	4	5	M
Tolerant and laid-back	L	1	2	3	4	5	M
17. Empathetic and patient	L	1	2	3	4	5	M
Contained and controlled	L	1	2	3	4	5	M
Task-focused and competitive	L	1	2	3	4	5	M
Discussing and spontaneous	L	1	2	3	4	5	M
18. Influential and informal	L	1	2	3	4	5	M
Considerate and empathetic	L	1	2	3	4	5	M
Impartial and evaluating	L	1	2	3	4	5	M
Challenging and determined	L	1	2	3	4	5	M
19. Prepared and systematic	L	1	2	3	4	5	M
Courageous and independent	L	1	2	3	4	5	M
Responsive and extraverted	L	1	2	3	4	5	M
Counselling and caring	L	1	2	3	4	5	M
20. Articulate and strong	L	1	2	3	4	5	M
Spontaneous and spirited	L	1	2	3	4	5	M
Studious and reasoned	L	1	2	3	4	5	M
Peaceful and harmonious	L	1	2	3	4	5	M

21. Organised and thoughtful	L	1	2	3	4	5	M
Patient and supportive	L	1	2	3	4	5	M
Strong and self-assured	L	1	2	3	4	5	M
Interacting and open	L	1	2	3	4	5	M
22. Objective and daring	L	1	2	3	4	5	M
Relaxed and peaceful	L	1	2	3	4	5	M
Factual and conventional	L	1	2	3	4	5	M
Lively and congenial	L	1	2	3	4	5	M
23. Animated and enthusiastic	L	1	2	3	4	5	M
Driving and realistic	L	1	2	3	4	5	M
Compassionate and considerate	L	1	2	3	4	5	M
Detailed and attentive	L	1	2	3	4	5	M
24. Supporting and steady	L	1	2	3	4	5	M
Independent and bold	L	1	2	3	4	5	M
Reflective and thorough	L	1	2	3	4	5	M
Good-mixer and lively	L	1	2	3	4	5	M
25. Cautious and accurate	L	1	2	3	4	5	M
Forthright and straightforward	L	1	2	3	4	5	M
Optimistic and upbeat	L	1	2	3	4	5	M
Accepting and loyal	L	1	2	3	4	5	M

When you have completed the Evaluator, please ensure every frame has been allocated ONE 'M', ONE 'L' and Two Different Values selected from 1, 2, 3, 4, or 5.

INSIGHTS DISCOVERY PREFERENCE EVALUATOR

Introduction

This Evaluator forms the basis of your Insights Discovery profile. It is not a pass or fail test. It simply records your perception of your work preferences.

Instructions - Please read carefully

Find a time and place where you will not be interrupted.

1. Fill in the personal details section. Enter your name and the date on all three pages of the evaluator.
2. In each frame, read each word pair carefully. Select the word pair that MOST describes you in your work environment and circle M next to this.
3. From the remaining three word pairs, select the pair that LEAST describes you in your work environment and circle L next to this.
4. For each of the remaining two word pairs circle a weighting from the values 1, 2, 3, 4 and 5, where 1 represents 'not likely to describe me', and 5 represents 'very likely to describe me'. Please do NOT choose the same weighting twice. Select those weightings which you believe best represent the relative intensity of the description in your working personality.
5. Continue until all 25 frames have been completed. Please ensure every frame has been scored, and each of the four word pairs has been allocated an M, an L, or a value selected from 1, 2, 3, 4 or 5.

Guidance Notes

- Remember, this is NOT a test! There are no right or wrong answers.
- Respond to the Evaluator based on your perception of yourself. Do not discuss your choices with others.
- Choose your responses quite quickly, as your first impression is often best. As a guide, this Evaluator typically takes between 10-20 minutes to complete.
- If returning this evaluator by fax, only the pages containing the word pairs are required.
- Word pairs in different language versions of this evaluator may not be compared directly. International versions are developed independently to ensure cultural differences are considered.

Appendix B – Item-item reliability data & Split-half reliability data

Table 18 - Correlation coefficients for cool blue items in English vs. S3.0 IDE N=519'467

Cool Blue colour preference																														
Frame & statement code	**Item-Total Correlation Coefficients**	Squared Item-Item Correlation Coefficients	1_1	2_4	3_4	4_2	5_1	6_4	7_2	8_4	9_2	10_4	11_2	12_1	13_3	14_2	15_1	16_3	17_2	18_3	19_1	20_3	21_1	22_3	23_4	24_3	25_1
1_1	0.45	0.30	1.00	0.23	0.20	0.27	0.28	0.43	0.25	0.22	0.17	0.23	0.22	0.21	0.26	0.27	0.21	0.36	0.37	0.29	0.26	0.31	0.25	0.23	0.30	0.35	0.30
2_4	0.56	0.37	0.23	1.00	0.38	0.35	0.47	0.29	0.19	0.30	0.31	0.38	0.37	0.35	0.43	0.34	0.40	0.37	0.24	0.26	0.39	0.36	0.27	0.34	0.36	0.35	0.29
3_4	0.59	0.45	0.20	0.38	1.00	0.47	0.33	0.26	0.12	0.42	0.44	0.55	0.41	0.38	0.34	0.30	0.36	0.35	0.22	0.21	0.41	0.33	0.34	0.37	0.50	0.37	0.40
4_2	0.60	0.42	0.27	0.35	0.47	1.00	0.36	0.30	0.14	0.39	0.34	0.48	0.33	0.33	0.31	0.30	0.37	0.35	0.30	0.22	0.48	0.36	0.46	0.36	0.46	0.38	0.38
5_1	0.55	0.37	0.28	0.47	0.33	0.36	1.00	0.30	0.14	0.35	0.26	0.32	0.40	0.31	0.40	0.37	0.38	0.34	0.27	0.27	0.40	0.35	0.30	0.33	0.33	0.36	0.28
6_4	0.52	0.35	0.43	0.29	0.26	0.30	0.30	1.00	0.24	0.30	0.24	0.26	0.26	0.29	0.35	0.33	0.26	0.41	0.39	0.31	0.31	0.36	0.26	0.30	0.33	0.37	0.34
7_2	0.29	0.16	0.25	0.19	0.12	0.14	0.14	0.24	1.00	0.12	0.10	0.15	0.14	0.20	0.26	0.19	0.14	0.25	0.14	0.20	0.14	0.22	0.13	0.12	0.15	0.31	0.19
8_4	0.55	0.34	0.22	0.30	0.42	0.39	0.35	0.30	0.12	1.00	0.36	0.40	0.38	0.30	0.30	0.32	0.33	0.31	0.27	0.23	0.40	0.34	0.30	0.37	0.41	0.35	0.36
9_2	0.49	0.30	0.17	0.31	0.44	0.34	0.26	0.24	0.10	0.36	1.00	0.39	0.36	0.32	0.28	0.31	0.30	0.27	0.24	0.20	0.35	0.29	0.23	0.31	0.38	0.28	0.32
10_4	0.62	0.49	0.23	0.38	0.55	0.48	0.32	0.26	0.15	0.40	0.39	1.00	0.38	0.41	0.36	0.30	0.38	0.37	0.24	0.21	0.43	0.35	0.40	0.36	0.59	0.42	0.42
11_2	0.54	0.33	0.22	0.37	0.41	0.33	0.40	0.26	0.14	0.38	0.36	0.38	1.00	0.32	0.35	0.35	0.33	0.33	0.26	0.24	0.35	0.33	0.27	0.33	0.37	0.34	0.33
12_1	0.56	0.38	0.21	0.35	0.38	0.33	0.31	0.29	0.20	0.30	0.32	0.41	0.32	1.00	0.51	0.31	0.38	0.43	0.21	0.25	0.36	0.36	0.25	0.33	0.40	0.37	0.38
13_3	0.59	0.44	0.26	0.43	0.34	0.31	0.40	0.35	0.26	0.30	0.28	0.36	0.35	0.51	1.00	0.36	0.40	0.50	0.24	0.32	0.38	0.41	0.25	0.37	0.37	0.40	0.33
14_2	0.54	0.32	0.27	0.34	0.30	0.30	0.37	0.33	0.19	0.32	0.31	0.30	0.35	0.31	0.36	1.00	0.32	0.36	0.32	0.33	0.38	0.39	0.25	0.35	0.35	0.35	0.30
15_1	0.56	0.36	0.21	0.40	0.36	0.37	0.38	0.26	0.14	0.33	0.30	0.38	0.33	0.38	0.40	0.32	1.00	0.37	0.23	0.23	0.47	0.37	0.32	0.38	0.37	0.37	0.30
16_3	0.62	0.43	0.36	0.37	0.35	0.35	0.34	0.41	0.25	0.31	0.27	0.37	0.33	0.43	0.50	0.36	0.37	1.00	0.35	0.35	0.38	0.43	0.32	0.38	0.42	0.44	0.41
17_2	0.47	0.30	0.37	0.24	0.22	0.30	0.27	0.39	0.14	0.27	0.24	0.24	0.26	0.21	0.24	0.32	0.23	0.35	1.00	0.30	0.32	0.34	0.27	0.31	0.32	0.31	0.32
18_3	0.44	0.23	0.29	0.26	0.21	0.22	0.27	0.31	0.20	0.23	0.20	0.21	0.24	0.25	0.32	0.33	0.23	0.35	0.30	1.00	0.26	0.31	0.20	0.26	0.28	0.30	0.26
19_1	0.64	0.47	0.26	0.39	0.41	0.48	0.40	0.31	0.14	0.40	0.35	0.43	0.35	0.36	0.38	0.38	0.47	0.38	0.32	0.26	1.00	0.44	0.47	0.44	0.48	0.42	0.39
20_3	0.61	0.39	0.31	0.36	0.33	0.36	0.35	0.36	0.22	0.34	0.29	0.35	0.33	0.36	0.41	0.39	0.37	0.43	0.34	0.31	0.44	1.00	0.32	0.45	0.40	0.43	0.37
21_1	0.51	0.34	0.25	0.27	0.34	0.46	0.30	0.26	0.13	0.30	0.23	0.40	0.27	0.25	0.25	0.25	0.32	0.32	0.27	0.20	0.47	0.32	1.00	0.29	0.43	0.37	0.33
22_3	0.57	0.36	0.23	0.34	0.37	0.36	0.33	0.30	0.12	0.37	0.31	0.36	0.33	0.33	0.37	0.35	0.38	0.38	0.31	0.26	0.44	0.45	0.29	1.00	0.42	0.37	0.37
23_4	0.66	0.50	0.30	0.36	0.50	0.46	0.33	0.33	0.15	0.41	0.38	0.59	0.37	0.40	0.37	0.35	0.37	0.42	0.32	0.28	0.48	0.40	0.43	0.42	1.00	0.46	0.45
24_3	0.63	0.41	0.35	0.35	0.37	0.38	0.36	0.37	0.31	0.35	0.28	0.42	0.34	0.37	0.40	0.35	0.37	0.44	0.31	0.30	0.42	0.43	0.37	0.37	0.46	1.00	0.41
25_1	0.58	0.36	0.30	0.29	0.40	0.38	0.28	0.34	0.19	0.36	0.32	0.42	0.33	0.38	0.33	0.30	0.30	0.41	0.32	0.26	0.39	0.37	0.33	0.37	0.45	0.41	1.00

Correlation of items with themselves (perfect correlation)

Weak items ('item to total' correlation coefficient ≤0.50 as well as 'Item to Item ≤0.30)

Acceptable coefficients for 'item to total' correlation (≥0.50)

Acceptable coefficients for 'item to item' correlation (≥0.30)

Table 19 - Correlation coefficients for earth green items in English vs. S3.0 IDE N=519'467

Earth Green colour preference																												
Frame & statement code	**Item-Total Correlation Coefficients**	Squared Item-Item Correlation Coefficients	1_2	2_2	3_1	4_3	5_3	6_1	7_4	8_3	9_1	10_2	11_3	12_4	13_1	14_4	15_3	16_4	17_1	18_2	19_4	20_4	21_2	22_2	23_3	24_1	25_4
1_2	0.44	0.30	1.00	0.22	0.43	0.22	0.20	0.24	0.17	0.16	0.41	0.25	0.36	0.20	0.19	0.24	0.31	0.23	0.32	0.26	0.25	0.30	0.31	0.30	0.28	0.24	0.11
2_2	0.50	0.30	0.22	1.00	0.30	0.32	0.37	0.15	0.36	0.29	0.23	0.37	0.24	0.26	0.37	0.27	0.30	0.26	0.25	0.28	0.23	0.33	0.34	0.32	0.28	0.30	0.25
3_1	0.50	0.43	0.43	0.30	1.00	0.30	0.20	0.22	0.23	0.16	0.32	0.35	0.53	0.26	0.25	0.21	0.36	0.34	0.33	0.25	0.21	0.37	0.36	0.42	0.28	0.28	0.12
4_3	0.48	0.29	0.22	0.32	0.30	1.00	0.32	0.19	0.39	0.28	0.23	0.38	0.28	0.26	0.28	0.26	0.32	0.27	0.24	0.24	0.24	0.32	0.32	0.31	0.27	0.32	0.22
5_3	0.55	0.39	0.20	0.37	0.20	0.32	1.00	0.25	0.35	0.45	0.28	0.35	0.25	0.24	0.37	0.36	0.25	0.28	0.30	0.38	0.32	0.35	0.39	0.30	0.36	0.33	0.46
6_1	0.39	0.20	0.24	0.15	0.22	0.19	0.25	1.00	0.17	0.20	0.30	0.17	0.31	0.18	0.24	0.27	0.19	0.22	0.31	0.33	0.26	0.24	0.27	0.19	0.29	0.21	0.18
7_4	0.50	0.33	0.17	0.36	0.23	0.39	0.35	0.17	1.00	0.31	0.21	0.36	0.24	0.25	0.38	0.36	0.28	0.27	0.25	0.31	0.26	0.37	0.35	0.31	0.30	0.33	0.27
8_3	0.51	0.34	0.16	0.29	0.16	0.28	0.45	0.20	0.31	1.00	0.26	0.33	0.19	0.23	0.31	0.39	0.22	0.22	0.32	0.38	0.37	0.35	0.39	0.29	0.39	0.32	0.37
9_1	0.55	0.39	0.41	0.23	0.32	0.23	0.28	0.30	0.21	0.26	1.00	0.28	0.40	0.21	0.26	0.47	0.28	0.30	0.44	0.44	0.40	0.36	0.39	0.32	0.44	0.25	0.23
10_2	0.55	0.36	0.25	0.37	0.35	0.38	0.35	0.17	0.36	0.33	0.28	1.00	0.31	0.35	0.32	0.32	0.35	0.35	0.28	0.29	0.26	0.40	0.38	0.40	0.32	0.32	0.25
11_3	0.54	0.41	0.36	0.24	0.53	0.28	0.25	0.31	0.24	0.19	0.40	0.31	1.00	0.26	0.27	0.30	0.32	0.39	0.38	0.34	0.29	0.39	0.37	0.40	0.35	0.25	0.18
12_4	0.42	0.21	0.20	0.26	0.26	0.26	0.24	0.18	0.25	0.23	0.21	0.35	0.26	1.00	0.29	0.24	0.27	0.26	0.23	0.23	0.19	0.29	0.27	0.28	0.24	0.24	0.18
13_1	0.52	0.32	0.19	0.37	0.25	0.28	0.37	0.24	0.38	0.31	0.26	0.32	0.27	0.29	1.00	0.33	0.26	0.26	0.31	0.38	0.27	0.36	0.38	0.33	0.32	0.34	0.29
14_4	0.62	0.48	0.24	0.27	0.21	0.26	0.36	0.27	0.36	0.39	0.47	0.32	0.30	0.24	0.33	1.00	0.25	0.30	0.47	0.54	0.53	0.44	0.47	0.36	0.54	0.32	0.33
15_3	0.47	0.28	0.31	0.30	0.36	0.32	0.25	0.19	0.28	0.22	0.28	0.35	0.32	0.27	0.26	0.25	1.00	0.24	0.27	0.24	0.25	0.30	0.31	0.30	0.26	0.34	0.14
16_4	0.51	0.33	0.23	0.26	0.34	0.27	0.28	0.22	0.27	0.22	0.30	0.35	0.39	0.26	0.26	0.30	0.24	1.00	0.34	0.32	0.27	0.42	0.37	0.48	0.34	0.19	0.27
17_1	0.62	0.48	0.32	0.25	0.33	0.24	0.30	0.31	0.25	0.32	0.44	0.28	0.38	0.23	0.31	0.47	0.27	0.34	1.00	0.58	0.48	0.43	0.54	0.40	0.52	0.31	0.26
18_2	0.64	0.52	0.26	0.28	0.25	0.24	0.38	0.33	0.31	0.38	0.44	0.29	0.34	0.23	0.38	0.54	0.24	0.32	0.58	1.00	0.51	0.44	0.51	0.38	0.59	0.33	0.33
19_4	0.58	0.45	0.25	0.23	0.21	0.24	0.32	0.26	0.26	0.37	0.40	0.26	0.29	0.19	0.27	0.53	0.25	0.27	0.48	0.51	1.00	0.42	0.50	0.34	0.57	0.34	0.32
20_4	0.65	0.49	0.30	0.33	0.37	0.32	0.35	0.24	0.37	0.35	0.36	0.40	0.39	0.29	0.36	0.44	0.30	0.42	0.43	0.44	0.42	1.00	0.50	0.60	0.46	0.32	0.32
21_2	0.68	0.50	0.31	0.34	0.36	0.32	0.39	0.27	0.35	0.39	0.39	0.38	0.37	0.27	0.38	0.47	0.31	0.37	0.54	0.51	0.50	0.50	1.00	0.47	0.54	0.41	0.34
22_2	0.61	0.47	0.30	0.32	0.42	0.31	0.30	0.19	0.31	0.29	0.32	0.40	0.40	0.28	0.33	0.36	0.30	0.48	0.40	0.38	0.34	0.60	0.47	1.00	0.43	0.29	0.27
23_3	0.65	0.52	0.28	0.28	0.28	0.27	0.36	0.29	0.30	0.39	0.44	0.32	0.35	0.24	0.32	0.54	0.26	0.34	0.52	0.59	0.57	0.46	0.54	0.43	1.00	0.36	0.35
24_1	0.51	0.30	0.24	0.30	0.28	0.32	0.33	0.21	0.33	0.32	0.25	0.32	0.25	0.24	0.34	0.32	0.34	0.19	0.31	0.33	0.34	0.32	0.41	0.29	0.36	1.00	0.26
25_4	0.45	0.30	0.11	0.25	0.12	0.22	0.46	0.18	0.27	0.37	0.23	0.25	0.18	0.18	0.29	0.33	0.14	0.27	0.26	0.33	0.32	0.32	0.34	0.27	0.35	0.26	1.00

Correlation of items with themselves (perfect correlation)

Weak items ('item to total' correlation coefficient ≤0.50 as well as 'Item to Item ≤0.30)

Acceptable coefficients for 'item to total' correlation (≥0.50)

Acceptable coefficients for 'item to item' correlation (≥0.30)

Table 20 - Correlation coefficients for fiery red items English vs. S3.0 IDE N=519'467

Fiery Red colour preference																										
Frame & statement code	**Item-Total Correlation Coefficients**	Squared Item-Item Correlation Coefficients	1_4	2_3	3_2	4_1	5_2	6_3	7_3	8_1	9_4	10_1	11_4	12_3	13_4	14_1	15_4	16_2	17_3	18_4	19_2	20_1	21_3	22_1	23_2	24_2	25_2
1_4	0.50	0.33	1.00	0.42	0.48	0.25	0.40	0.42	0.16	0.24	0.32	0.37	0.27	0.31	0.31	0.36	0.34	0.31	0.34	0.37	0.15	0.21	0.35	0.29	0.25	0.30	0.24
2_3	0.64	0.45	0.42	1.00	0.53	0.41	0.46	0.52	0.29	0.37	0.42	0.42	0.33	0.44	0.38	0.44	0.38	0.34	0.46	0.45	0.24	0.30	0.38	0.38	0.34	0.37	0.29
3_2	0.66	0.49	0.48	0.53	1.00	0.41	0.47	0.55	0.26	0.37	0.39	0.45	0.31	0.43	0.38	0.50	0.39	0.37	0.43	0.48	0.26	0.34	0.42	0.40	0.34	0.41	0.33
4_1	0.61	0.46	0.25	0.41	0.41	1.00	0.38	0.49	0.48	0.51	0.35	0.42	0.20	0.42	0.36	0.41	0.35	0.34	0.35	0.37	0.32	0.39	0.36	0.42	0.32	0.41	0.31
5_2	0.63	0.44	0.40	0.46	0.47	0.38	1.00	0.50	0.22	0.41	0.34	0.45	0.29	0.38	0.39	0.42	0.38	0.35	0.35	0.46	0.25	0.31	0.43	0.42	0.35	0.41	0.39
6_3	0.71	0.53	0.42	0.52	0.55	0.49	0.50	1.00	0.37	0.44	0.39	0.52	0.29	0.47	0.44	0.51	0.40	0.41	0.43	0.48	0.28	0.43	0.47	0.43	0.36	0.42	0.37
7_3	0.45	0.31	0.16	0.29	0.26	0.48	0.22	0.37	1.00	0.40	0.30	0.30	0.18	0.31	0.27	0.29	0.26	0.31	0.28	0.28	0.26	0.31	0.22	0.29	0.23	0.30	0.22
8_1	0.59	0.41	0.24	0.37	0.37	0.51	0.41	0.44	0.40	1.00	0.31	0.45	0.23	0.38	0.34	0.39	0.31	0.32	0.33	0.33	0.29	0.38	0.39	0.39	0.37	0.41	0.35
9_4	0.57	0.40	0.32	0.42	0.39	0.35	0.34	0.39	0.30	0.31	1.00	0.32	0.38	0.37	0.31	0.36	0.46	0.39	0.47	0.40	0.23	0.26	0.32	0.32	0.34	0.34	0.27
10_1	0.63	0.43	0.37	0.42	0.45	0.42	0.45	0.52	0.30	0.45	0.32	1.00	0.25	0.42	0.37	0.44	0.36	0.39	0.36	0.37	0.25	0.40	0.42	0.40	0.41	0.39	0.37
11_4	0.42	0.27	0.27	0.33	0.31	0.20	0.29	0.29	0.18	0.23	0.38	0.25	1.00	0.21	0.21	0.28	0.25	0.28	0.43	0.31	0.14	0.20	0.23	0.21	0.25	0.25	0.24
12_3	0.61	0.44	0.31	0.44	0.43	0.42	0.38	0.47	0.31	0.38	0.37	0.42	0.21	1.00	0.39	0.42	0.36	0.34	0.54	0.45	0.25	0.34	0.39	0.38	0.36	0.36	0.30
13_4	0.56	0.34	0.31	0.38	0.38	0.36	0.39	0.44	0.27	0.34	0.31	0.37	0.21	0.39	1.00	0.39	0.36	0.31	0.30	0.36	0.27	0.28	0.39	0.42	0.29	0.42	0.32
14_1	0.65	0.44	0.36	0.44	0.50	0.41	0.42	0.51	0.29	0.39	0.36	0.44	0.28	0.42	0.39	1.00	0.35	0.37	0.42	0.49	0.33	0.38	0.45	0.40	0.36	0.44	0.34
15_4	0.55	0.35	0.34	0.38	0.39	0.35	0.38	0.40	0.26	0.31	0.46	0.36	0.25	0.36	0.36	0.35	1.00	0.38	0.34	0.37	0.24	0.26	0.34	0.37	0.29	0.37	0.27
16_2	0.55	0.33	0.31	0.34	0.37	0.34	0.35	0.41	0.31	0.32	0.39	0.39	0.28	0.34	0.31	0.37	0.38	1.00	0.37	0.39	0.23	0.30	0.30	0.33	0.33	0.33	0.31
17_3	0.60	0.48	0.34	0.46	0.43	0.35	0.35	0.43	0.28	0.33	0.47	0.36	0.43	0.54	0.30	0.42	0.34	0.37	1.00	0.50	0.21	0.31	0.36	0.34	0.38	0.33	0.27
18_4	0.63	0.44	0.37	0.45	0.48	0.37	0.46	0.48	0.28	0.33	0.40	0.37	0.31	0.45	0.36	0.49	0.37	0.39	0.50	1.00	0.27	0.31	0.41	0.39	0.35	0.38	0.28
19_2	0.42	0.29	0.15	0.24	0.26	0.32	0.25	0.28	0.26	0.29	0.23	0.25	0.14	0.25	0.27	0.33	0.24	0.23	0.21	0.27	1.00	0.23	0.29	0.35	0.24	0.50	0.22
20_1	0.52	0.32	0.21	0.30	0.34	0.39	0.31	0.43	0.31	0.38	0.26	0.40	0.20	0.34	0.28	0.38	0.26	0.30	0.31	0.31	0.23	1.00	0.42	0.34	0.30	0.34	0.34
21_3	0.61	0.41	0.35	0.38	0.42	0.36	0.43	0.47	0.22	0.39	0.32	0.42	0.23	0.39	0.39	0.45	0.34	0.30	0.36	0.41	0.29	0.42	1.00	0.40	0.37	0.44	0.37
22_1	0.60	0.39	0.29	0.38	0.40	0.42	0.42	0.43	0.29	0.39	0.32	0.40	0.21	0.38	0.42	0.40	0.37	0.33	0.34	0.39	0.35	0.34	0.40	1.00	0.34	0.47	0.34
23_2	0.53	0.30	0.25	0.34	0.34	0.32	0.35	0.36	0.23	0.37	0.34	0.41	0.25	0.36	0.29	0.36	0.29	0.33	0.38	0.35	0.24	0.30	0.37	0.34	1.00	0.34	0.32
24_2	0.61	0.46	0.30	0.37	0.41	0.41	0.41	0.42	0.30	0.41	0.34	0.39	0.25	0.36	0.42	0.44	0.37	0.33	0.33	0.38	0.50	0.34	0.44	0.47	0.34	1.00	0.36
25_2	0.50	0.30	0.24	0.29	0.33	0.31	0.39	0.37	0.22	0.35	0.27	0.37	0.24	0.30	0.32	0.34	0.27	0.31	0.27	0.28	0.22	0.34	0.37	0.34	0.32	0.36	1.00

Correlation of items with themselves (perfect correlation)

Weak items ('item to total' correlation coefficient ≤0.50 as well as 'Item to Item ≤0.30)

Acceptable coefficients for 'item to total' correlation (≥0.50)

Acceptable coefficients for 'item to item' correlation (≥0.30)

Table 21 -Correlation coefficients for the sunshine yellow items in English vs. S3.0 IDE N=519'467

Sunshine Yellow colour preference																										
Frame & statement code	**Item-Total Correlation Coefficients**	Squared Item-Item Correlation Coefficients	1_3	2_1	3_3	4_4	5_4	6_2	7_1	8_2	9_3	10_3	11_1	12_2	13_2	14_3	15_2	16_1	17_4	18_1	19_3	20_2	21_4	22_4	23_1	24_4	25_3
1_3	0.65	0.47	1.00	0.25	0.32	0.43	0.45	0.29	0.24	0.44	0.21	0.61	0.26	0.42	0.31	0.33	0.42	0.37	0.35	0.19	0.45	0.41	0.43	0.51	0.43	0.48	0.32
2_1	0.39	0.19	0.25	1.00	0.33	0.30	0.30	0.21	0.07	0.31	0.15	0.29	0.13	0.31	0.19	0.21	0.27	0.14	0.23	0.15	0.21	0.23	0.23	0.31	0.26	0.27	0.20
3_3	0.49	0.32	0.32	0.33	1.00	0.40	0.38	0.29	0.04	0.44	0.23	0.38	0.19	0.41	0.19	0.25	0.41	0.15	0.24	0.15	0.23	0.28	0.27	0.38	0.31	0.37	0.28
4_4	0.60	0.41	0.43	0.30	0.40	1.00	0.48	0.28	0.17	0.50	0.20	0.50	0.22	0.44	0.27	0.27	0.43	0.30	0.33	0.17	0.35	0.33	0.36	0.49	0.40	0.47	0.25
5_4	0.60	0.50	0.45	0.30	0.38	0.48	1.00	0.29	0.09	0.61	0.20	0.52	0.21	0.50	0.23	0.30	0.55	0.21	0.24	0.14	0.32	0.30	0.34	0.49	0.34	0.54	0.30
6_2	0.42	0.20	0.29	0.21	0.29	0.28	0.29	1.00	0.10	0.31	0.22	0.32	0.20	0.30	0.24	0.25	0.27	0.22	0.23	0.10	0.21	0.25	0.24	0.30	0.28	0.26	0.26
7_1	0.33	0.27	0.24	0.07	0.04	0.17	0.09	0.10	1.00	0.10	0.06	0.24	0.20	0.11	0.35	0.18	0.12	0.46	0.23	0.26	0.28	0.26	0.18	0.21	0.30	0.19	0.17
8_2	0.63	0.53	0.44	0.31	0.44	0.50	0.61	0.31	0.10	1.00	0.23	0.53	0.22	0.55	0.25	0.30	0.57	0.23	0.26	0.14	0.32	0.32	0.34	0.51	0.38	0.54	0.34
9_3	0.31	0.15	0.21	0.15	0.23	0.20	0.20	0.22	0.06	0.23	1.00	0.22	0.20	0.25	0.23	0.18	0.21	0.16	0.15	0.11	0.10	0.15	0.20	0.22	0.18	0.19	0.23
10_3	0.73	0.58	0.61	0.29	0.38	0.50	0.52	0.32	0.24	0.53	0.22	1.00	0.30	0.50	0.34	0.36	0.52	0.39	0.36	0.22	0.50	0.45	0.42	0.57	0.49	0.56	0.36
11_1	0.41	0.21	0.26	0.13	0.19	0.22	0.21	0.20	0.20	0.22	0.20	0.30	1.00	0.23	0.24	0.22	0.26	0.27	0.31	0.15	0.25	0.38	0.23	0.26	0.30	0.27	0.18
12_2	0.60	0.48	0.42	0.31	0.41	0.44	0.50	0.30	0.11	0.55	0.25	0.50	0.23	1.00	0.28	0.28	0.60	0.24	0.25	0.16	0.28	0.31	0.30	0.50	0.34	0.50	0.28
13_2	0.50	0.34	0.31	0.19	0.19	0.27	0.23	0.24	0.35	0.25	0.23	0.34	0.24	0.28	1.00	0.25	0.26	0.45	0.28	0.37	0.29	0.30	0.29	0.33	0.35	0.29	0.29
14_3	0.47	0.24	0.33	0.21	0.25	0.27	0.30	0.25	0.18	0.30	0.18	0.36	0.22	0.28	0.25	1.00	0.29	0.27	0.25	0.18	0.31	0.29	0.32	0.35	0.35	0.33	0.31
15_2	0.62	0.52	0.42	0.27	0.41	0.43	0.55	0.27	0.12	0.57	0.21	0.52	0.26	0.60	0.26	0.29	1.00	0.24	0.24	0.17	0.32	0.35	0.30	0.51	0.36	0.57	0.29
16_1	0.52	0.44	0.37	0.14	0.15	0.30	0.21	0.22	0.46	0.23	0.16	0.39	0.27	0.24	0.45	0.27	0.24	1.00	0.34	0.31	0.38	0.38	0.30	0.38	0.50	0.31	0.26
17_4	0.50	0.34	0.35	0.23	0.24	0.33	0.24	0.23	0.23	0.26	0.15	0.36	0.31	0.25	0.28	0.25	0.24	0.34	1.00	0.24	0.35	0.48	0.36	0.35	0.40	0.30	0.22
18_1	0.33	0.20	0.19	0.15	0.15	0.17	0.14	0.10	0.26	0.14	0.11	0.22	0.15	0.16	0.37	0.18	0.17	0.31	0.24	1.00	0.23	0.20	0.22	0.20	0.26	0.19	0.20
19_3	0.56	0.37	0.45	0.21	0.23	0.35	0.32	0.21	0.28	0.32	0.10	0.50	0.25	0.28	0.29	0.31	0.32	0.38	0.35	0.23	1.00	0.40	0.37	0.42	0.43	0.42	0.27
20_2	0.58	0.41	0.41	0.23	0.28	0.33	0.30	0.25	0.26	0.32	0.15	0.45	0.38	0.31	0.30	0.29	0.35	0.38	0.48	0.20	0.40	1.00	0.35	0.44	0.48	0.38	0.28
21_4	0.54	0.32	0.43	0.23	0.27	0.36	0.34	0.24	0.18	0.34	0.20	0.42	0.23	0.30	0.29	0.32	0.30	0.30	0.36	0.22	0.37	0.35	1.00	0.41	0.37	0.39	0.31
22_4	0.69	0.52	0.51	0.31	0.38	0.49	0.49	0.30	0.21	0.51	0.22	0.57	0.26	0.50	0.33	0.35	0.51	0.38	0.35	0.20	0.42	0.44	0.41	1.00	0.48	0.57	0.34
23_1	0.64	0.46	0.43	0.26	0.31	0.40	0.34	0.28	0.30	0.38	0.18	0.49	0.30	0.34	0.35	0.35	0.36	0.50	0.40	0.26	0.43	0.48	0.37	0.48	1.00	0.43	0.36
24_4	0.67	0.52	0.48	0.27	0.37	0.47	0.54	0.26	0.19	0.54	0.19	0.56	0.27	0.50	0.29	0.33	0.57	0.31	0.30	0.19	0.42	0.38	0.39	0.57	0.43	1.00	0.32
25_3	0.47	0.25	0.32	0.20	0.28	0.25	0.30	0.26	0.17	0.34	0.23	0.36	0.18	0.28	0.29	0.31	0.29	0.26	0.22	0.20	0.27	0.28	0.31	0.34	0.36	0.32	1.00

Correlation of items with themselves (perfect correlation)

Weak items ('item to total' correlation coefficient ≤0.50 as well as 'Item to Item ≤0.30)

Acceptable coefficients for 'item to total' correlation (≥0.50)

Acceptable coefficients for 'item to item' correlation (≥0.30)

Table 22 –Split-half coefficients in English vs S3.0 IDE N=519'46

Cool Blue colour preference	Earth Green colour preference	Fiery Red colour preference	Sunshine Yellow colour preference
Split-Half reliability statistics	**Split-Half reliability statistics**	**Split-Half reliability statistics**	**Split-Half reliability statistics**
Cronbach-Alpha Coefficients	Cronbach-Alpha Coefficients	Cronbach-Alpha Coefficients	Cronbach-Alpha Coefficients
Part 1 Value 0.85	Part 1 Value 0.83	Part 1 Value 0.84	Part 1 Value 0.88
N of Items 13	N of Items 13	N of Items 13	N of Items 13
Part 2 Value 0.87	Part 2 Value 0.88	Part 2 Value 0.86	Part 2 Value 0.87
N of Items 12	N of Items 12	N of Items 12	N of Items 12
Tot. N of Items 25	Tot. N of Items 25	Tot. N of Items 25	Tot. N of Items 25
Part 1	Part 1	Part 1	Part 1
The items are: 1_1, 2_4, 3_4, 4_2, 5_1, 6_4, 7_2, 8_4, 9_2, 10_4, 11_2, 12_1, 13_3.	The items are: 1_2, 2_2, 3_1, 4_3, 5_3, 6_1, 7_4, 8_3, 9_1, 10_2, 11_3, 12_4, 13_1.	The items are: 1_4, 2_3, 3_2, 4_1, 5_2, 6_3, 7_3, 8_1, 9_4, 10_1, 11_4, 12_3, 13_4.	The items are: 1_3, 2_1, 3_3, 4_4, 5_4, 6_2, 7_1, 8_2, 9_3, 10_3, 11_1, 12_2, 13_2.
Part 2	Part 2	Part 2	Pat 2
The items are: 14_2, 15_1, 16_3, 17_2, 18_3, 19_1, 20_3, 21_1, 22_3, 23_4, 24_3, 25_1.	The items are: 14_4, 15_3, 16_4, 17_1, 18_2, 19_4, 20_4, 21_2, 22_2, 23_3, 24_1, 25_4.	The items are: 14_1, 15_4, 16_2, 17_3, 18_4, 19_2, 20_1, 21_3, 22_1, 23_2, 24_2, 25_2.	The items are: 14_3, 15_2, 16_1, 17_4, 18_1, 19_3, 20_2, 21_4, 22_4, 23_1, 24_4, 25_3.

Correlation coefficients between sub-sets (parts) of the test		Correlation coefficients between sub-sets (parts) of the test		Correlation coefficients between sub-sets (parts) of the test		Correlation coefficients between sub-sets (parts) of the test	
Pearson Correlation Coefficient	0.84	Pearson Correlation Coefficient	0.79	Pearson Correlation Coefficient	0.82	Pearson Correlation Coefficient	0.85
Spearman-Brown Coefficient	0.91	Spearman-Brown Coefficient	0.88	Spearman-Brown Coefficient	0.90	Spearman-Brown Coefficient	0.92
Guttman Split-Half Coefficient	0.91	Guttman Split-Half Coefficient	0.88	Guttman Split-Half Coefficient	0.90	Guttman Split-Half Coefficient	0.92

University of Westminster

bpc

Appendix C – Item Level Validity Factor Analysis Data

Table 23 – Varimax and Oblimin factor analysis of UK residents using version S3.0 of the English IDE.
Sample Size = 114'670

		VARIMAX Principal Component		VARIMAX Maximum Likelihood		OBLIMIN Principal Component		OBLIMIN Maximum Likelihood	
Colour preference		F1	F2	F1	F2	F1	F2	F1	F2
Frame 1 Item 1	B	0.35	-0.49	0.34	-0.48	-0.39	-0.51	-0.38	-0.50
Frame 2 Item 1	G	0.44	-0.11	0.42	-0.11	-0.45	-0.14	-0.43	-0.14
Frame 3 Item 1	Y	-0.15	0.62	-0.14	0.62	0.20	0.63	0.20	0.63
Frame 4 Item 1	R	-0.58	0.01	-0.57	0.01	0.58	0.06	0.57	0.06
Frame 1 Item 2	Y	0.17	0.45	0.17	0.44	-0.13	0.44	-0.13	0.42
Frame 2 Item 2	G	0.52	-0.20	0.51	-0.20	-0.54	-0.24	-0.52	-0.24
Frame 3 Item 2	R	-0.67	0.16	-0.67	0.16	0.69	0.21	0.68	0.21
Frame 4 Item 2	B	0.07	-0.58	0.08	-0.56	-0.12	-0.58	-0.13	-0.56
Frame 1 Item 3	G	0.46	-0.16	0.45	-0.16	-0.47	-0.19	-0.46	-0.19
Frame 2 Item 3	R	-0.68	0.14	-0.67	0.14	0.69	0.20	0.68	0.20
Frame 3 Item 3	Y	0.23	0.55	0.23	0.54	-0.19	0.53	-0.18	0.52
Frame 4 Item 3	B	0.05	-0.60	0.06	-0.58	-0.10	-0.60	-0.11	-0.58
Frame 1 Item 4	R	-0.63	0.26	-0.62	0.25	0.65	0.30	0.64	0.30
Frame 2 Item 4	B	0.05	-0.64	0.06	-0.63	-0.11	-0.65	-0.11	-0.63
Frame 3 Item 4	G	0.47	-0.12	0.45	-0.12	-0.48	-0.15	-0.46	-0.16
Frame 4 Item 4	Y	0.11	0.58	0.12	0.58	-0.06	0.57	-0.06	0.57
Frame 1 Item 5	B	-0.01	-0.59	0.00	-0.58	-0.04	-0.59	-0.05	-0.58
Frame 2 Item 5	R	-0.68	0.03	-0.68	0.02	0.68	0.08	0.67	0.08
Frame 3 Item 5	G	0.54	-0.01	0.53	-0.02	-0.54	-0.05	-0.52	-0.06
Frame 4 Item 5	Y	0.22	0.59	0.22	0.58	-0.17	0.57	-0.17	0.56
Frame 1 Item 6	G	0.46	0.13	0.45	0.12	-0.45	0.10	-0.44	0.09
Frame 2 Item 6	Y	0.09	0.42	0.09	0.41	-0.06	0.42	-0.06	0.40
Frame 3 Item 6	R	-0.72	0.16	-0.72	0.16	0.73	0.22	0.73	0.22
Frame 4 Item 6	B	0.23	-0.57	0.23	-0.56	-0.28	-0.59	-0.28	-0.57
Frame 1 Item 7	Y	-0.43	0.30	-0.43	0.30	0.46	0.34	0.45	0.33
Frame 2 Item 7	B	0.18	-0.34	0.18	-0.33	-0.21	-0.35	-0.21	-0.34
Frame 3 Item 7	R	-0.47	0.23	-0.46	0.22	0.49	0.26	0.48	0.26
Frame 4 Item 7	G	0.55	-0.18	0.54	-0.18	-0.57	-0.22	-0.55	-0.22
Frame 1 Item 8	R	-0.63	0.07	-0.62	0.07	0.64	0.12	0.63	0.12
Frame 2 Item 8	Y	0.23	0.61	0.23	0.61	-0.18	0.59	-0.18	0.59
Frame 3 Item 8	G	0.52	-0.02	0.50	-0.03	-0.52	-0.06	-0.50	-0.07
Frame 4 Item 8	B	0.00	-0.61	0.01	-0.60	-0.05	-0.61	-0.06	-0.60
Frame 1 Item 9	G	0.61	0.08	0.60	0.07	-0.60	0.03	-0.59	0.03
Frame 2 Item 9	B	0.00	-0.53	0.01	-0.51	-0.04	-0.53	-0.05	-0.51
Frame 3 Item 9	Y	0.13	0.36	0.12	0.34	-0.10	0.35	-0.09	0.33

Frame 4 Item 9	R	-0.64	0.12	-0.63	0.12	0.65	0.17	0.64	0.17
Frame 1 Item 10	R	-0.61	0.04	-0.61	0.04	0.61	0.09	0.61	0.08
Frame 2 Item 10	G	0.55	-0.28	0.53	-0.28	-0.57	-0.32	-0.56	-0.32
Frame 3 Item 10	Y	-0.11	0.71	-0.11	0.71	0.17	0.72	0.17	0.72
Frame 4 Item 10	B	0.08	-0.61	0.09	-0.59	-0.13	-0.61	-0.14	-0.60
Frame 1 Item 11	Y	-0.11	0.50	-0.11	0.48	0.15	0.51	0.15	0.49
Frame 2 Item 11	B	0.04	-0.57	0.04	-0.55	-0.09	-0.57	-0.09	-0.55
Frame 3 Item 11	G	0.58	0.05	0.57	0.04	-0.58	0.00	-0.56	0.00
Frame 4 Item 11	R	-0.52	-0.12	-0.51	-0.12	0.51	-0.08	0.50	-0.07
Frame 1 Item 12	B	0.00	-0.58	0.01	-0.56	-0.05	-0.57	-0.06	-0.56
Frame 2 Item 12	Y	0.23	0.61	0.23	0.61	-0.18	0.60	-0.18	0.59
Frame 3 Item 12	R	-0.64	0.18	-0.64	0.18	0.65	0.23	0.65	0.23
Frame 4 Item 12	G	0.46	-0.13	0.44	-0.14	-0.47	-0.17	-0.45	-0.17
Frame 1 Item 13	G	0.58	-0.07	0.57	-0.07	-0.58	-0.11	-0.57	-0.11
Frame 2 Item 13	Y	-0.23	0.46	-0.23	0.45	0.27	0.48	0.27	0.47
Frame 3 Item 13	B	0.02	-0.59	0.02	-0.57	-0.06	-0.59	-0.07	-0.57
Frame 4 Item 13	R	-0.58	0.25	-0.57	0.25	0.60	0.29	0.59	0.29
Frame 1 Item 14	R	-0.68	0.10	-0.67	0.10	0.68	0.15	0.68	0.15
Frame 2 Item 14	B	0.09	-0.56	0.10	-0.55	-0.14	-0.57	-0.14	-0.55
Frame 3 Item 14	Y	-0.06	0.44	-0.06	0.43	0.09	0.45	0.09	0.43
Frame 4 Item 14	G	0.65	0.07	0.64	0.07	-0.64	0.02	-0.63	0.02
Frame 1 Item 15	B	-0.06	-0.59	-0.05	-0.58	0.01	-0.58	0.00	-0.57
Frame 2 Item 15	Y	0.21	0.61	0.21	0.61	-0.15	0.60	-0.15	0.59
Frame 3 Item 15	G	0.47	-0.24	0.45	-0.24	-0.49	-0.27	-0.47	-0.28
Frame 4 Item 15	R	-0.60	0.22	-0.59	0.21	0.61	0.26	0.60	0.26
Frame 1 Item 16	Y	-0.41	0.46	-0.40	0.45	0.44	0.49	0.44	0.48
Frame 2 Item 16	R	-0.56	0.05	-0.55	0.05	0.56	0.10	0.55	0.10
Frame 3 Item 16	B	0.27	-0.61	0.27	-0.60	-0.32	-0.63	-0.32	-0.62
Frame 4 Item 16	G	0.56	0.10	0.55	0.09	-0.55	0.05	-0.54	0.05
Frame 1 Item 17	G	0.63	0.04	0.62	0.03	-0.62	-0.01	-0.61	-0.02
Frame 2 Item 17	B	0.28	-0.52	0.28	-0.51	-0.33	-0.54	-0.32	-0.53
Frame 3 Item 17	R	-0.66	-0.01	-0.65	-0.01	0.65	0.04	0.65	0.04
Frame 4 Item 17	Y	-0.22	0.50	-0.21	0.49	0.26	0.51	0.26	0.50
Frame 1 Item 18	Y	-0.21	0.34	-0.20	0.32	0.23	0.35	0.23	0.34
Frame 2 Item 18	G	0.68	0.11	0.67	0.11	-0.67	0.06	-0.66	0.05
Frame 3 Item 18	B	0.24	-0.47	0.24	-0.46	-0.28	-0.49	-0.28	-0.47
Frame 4 Item 18	R	-0.69	0.06	-0.68	0.06	0.69	0.11	0.68	0.12
Frame 1 Item 19	B	0.01	-0.67	0.01	-0.66	-0.06	-0.67	-0.07	-0.66
Frame 2 Item 19	R	-0.41	0.12	-0.40	0.12	0.42	0.15	0.41	0.15
Frame 3 Item 19	Y	-0.29	0.55	-0.29	0.54	0.34	0.57	0.33	0.57
Frame 4 Item 19	G	0.61	0.14	0.60	0.13	-0.60	0.09	-0.58	0.08
Frame 1 Item 20	R	-0.56	0.08	-0.55	0.08	0.56	0.12	0.55	0.12
Frame 2 Item 20	Y	-0.29	0.59	-0.29	0.59	0.34	0.61	0.34	0.61
Frame 3 Item 20	B	0.12	-0.63	0.12	-0.62	-0.17	-0.64	-0.17	-0.63
Frame 4 Item 20	G	0.67	-0.06	0.66	-0.06	-0.68	-0.11	-0.66	-0.11
Frame 1 Item 21	B	0.17	-0.52	0.18	-0.50	-0.22	-0.53	-0.22	-0.52
Frame 2 Item 21	G	0.67	0.00	0.65	0.00	-0.66	-0.05	-0.65	-0.06
Frame 3 Item 21	R	-0.66	0.02	-0.65	0.02	0.66	0.07	0.65	0.07

Frame 4 Item 21	Y	-0.10	0.54	-0.10	0.52	0.15	0.54	0.15	0.53
Frame 1 Item 22	R	-0.62	0.15	-0.62	0.15	0.64	0.20	0.63	0.20
Frame 2 Item 22	G	0.61	-0.07	0.60	-0.08	-0.62	-0.12	-0.60	-0.12
Frame 3 Item 22	B	0.05	-0.62	0.05	-0.61	-0.10	-0.63	-0.10	-0.61
Frame 4 Item 22	Y	-0.04	0.67	-0.03	0.67	0.09	0.67	0.09	0.67
Frame 1 Item 23	Y	-0.26	0.61	-0.26	0.60	0.31	0.63	0.31	0.62
Frame 2 Item 23	R	-0.57	-0.05	-0.56	-0.04	0.56	0.00	0.56	0.00
Frame 3 Item 23	G	0.65	0.16	0.63	0.15	-0.63	0.11	-0.62	0.10
Frame 4 Item 23	B	0.17	-0.67	0.17	-0.66	-0.23	-0.68	-0.23	-0.67
Frame 1 Item 24	G	0.48	-0.10	0.46	-0.11	-0.48	-0.14	-0.47	-0.14
Frame 2 Item 24	R	-0.61	0.12	-0.61	0.12	0.62	0.17	0.61	0.17
Frame 3 Item 24	B	0.19	-0.65	0.19	-0.64	-0.24	-0.66	-0.25	-0.65
Frame 4 Item 24	Y	0.04	0.65	0.04	0.65	0.02	0.65	0.02	0.64
Frame 1 Item 25	B	0.24	-0.60	0.24	-0.59	-0.29	-0.62	-0.29	-0.60
Frame 2 Item 25	R	-0.55	0.07	-0.53	0.08	0.55	0.11	0.54	0.12
Frame 3 Item 25	Y	-0.12	0.47	-0.12	0.46	0.15	0.48	0.15	0.46
Frame 4 Item 25	G	0.48	0.14	0.47	0.13	-0.47	0.10	-0.45	0.09

Average Factor Loadings

		VARIMAX Principal Component		VARIMAX Maximum Likelihood		OBLIMIN Principal Component		OBLIMIN Maximum Likelihood	
		F1	F2	F1	F2	F1	F2	F1	F2
Earth Green	G	0.56	-0.03	0.54	-0.03	-0.56	-0.07	-0.54	-0.08
Sunshine Yellow	Y	-0.06	0.53	-0.05	0.52	0.10	0.53	0.10	0.52
Cool Blue	B	0.11	-0.58	0.12	-0.56	-0.16	-0.58	-0.17	-0.57
Fiery Red	R	-0.61	0.10	-0.60	0.10	0.61	0.14	0.61	0.15

Appendix D – About the Authors

Dr Stephen Benton, Ph.D.

Stephen is director of the Business Psychology Centre at the University of Westminster, which specialises in applying Business Psychology within both public and private sector organisations. He is also course leader for the innovative M.Sc. in Business Psychology and the M.A. in Psychology for Project Management. In addition, he is a member of the accreditation and validation panel of the Association of Project Managers (APM) and of the management board for the Association of Business Psychology (ABP).

After graduating in psychology from Brunel University, Stephen was awarded a SERC postgraduate scholarship to study for his Ph.D. in psychophysics at Chelsea College, University of London with Professor Geoff Leventhall. In 1985 he left a position of Research Fellow in the Department of Environmental Acoustics at Chelsea to take up a post-doctoral Research Fellowship, researching visual psychophysics, at the University College of London with Professor Michael Morgan. Stephen joined the University of Westminster in 1988 and is currently a Principal Lecturer in Business Psychology. For the last 11 years he has developed business psychology programmes that highlight the way in which applied psychology can be used to support the individual and individual differences within the workplace. In 1997, he created the M.Sc. in Business Psychology programme at the University of Westminster, the first of its kind in Europe.

Stephen is also the Director of the Human Factors Research Group. The group was started in 1992 in order to provide an environment from which a multi-disciplinary approach could be made towards the study of individual differences as they operated within and across a range of performance domains. Research conducted in this group has included investigations into speech acquisition in the hearing impaired; effects of Myalgic Encephalomyelitis (M.E.) upon social competence; psychophysics of sun glass performance and preference; effects of age upon foot-fall estimation; perception of risk by the elderly; low frequency noise and performance at work and cross-cultural aspects of negotiation effectiveness. The group currently houses three Ph.D. students with research in: Management excellence - an international perspective; the changing relationship/interaction between large and small scale industries in Indonesia; the application of Jungian typology as an effective approach to negotiation and an investigation into the cognitive components underlying cross-culturally effective advertising.

Stephen has researched and worked in South East Asia over the past fifteen years, conducting studies into; Management Competence, Cross Cultural Issues within the multinational environment and specifically the problems encountered by expatriates working/managing Indonesian management teams. He is currently visiting Professor in Business Psychology at Atma Jaya University, Jakarta Indonesia where his work includes the design and delivery of workshops on cross cultural teams, the role of conflict in team performance, typology and interpersonal stress, negotiating and managing within the multi national workplace. Stephen can be contacted on +44 (0) 774 077 6483 or by e-mail at bentons@wmin.ac.uk

Corine van Erkom Schurink, Ph.D.

Corine is the director of The Analytical Research Bureau (Pty) Ltd, a consulting firm that she founded in 1993, focusing on Qualitative & Quantitative Research and Data Mining. Her speciality is market related statistical analysis & mathematical modelling and the bureau services a large portfolio of organisations from various industries e.g. Retail, Auditing, Financial Services, Petroleum, Telecommunication, Government, Health, Insurance and IT.

Corine was born in Switzerland where she studied Biology-Sciences and obtained a M.Sc. at the University of Neuchâtel. She then pursued her academic education in South Africa accepting a bursary from the University of Cape Town to first complete a Doctorate in 'Physiology & Bio-Energetics', and secondly to follow the MBA curriculum (Master in Business Administration) at the university. Through her work with the Analytical Research Bureau she enjoys the challenge of combining her rigorous analytical side and with her highly intuitive business acumen.

Corine is also an academic at the University of Cape Town with the following responsibilities:

- Senior Researcher in the fields of 'Epidemiology', 'Ethics', 'Industrial Marketing' and 'Entrepreneurship' at the Graduate School of Business
- Visiting Lecturer and Supervisor for Post Graduate thesis in 'Statistics & Data Mining' at the Department of Applied Statistics
- External Examiner for the final year 'Marketing' papers at the Commerce Faculty

Corine also works at Peninsula University for Technology in Cape Town as:

- Lecturer of 'Research Methodology' for the Nedcor Management Development Program
- Supervisor for Master Degree thesis in 'Statistics & Applied Mathematics'

Corine has published in numerous International Journals and has received the Editor's Award for 'the best article' for work published in the American Journal of Shellfish Research (1993, Vol.9. No.1, pp 75-85). For a marketing article, she received the 'Literacy Excellence Award' from MCB University Press (England) for an article published in the Journal of Business & Industrial Marketing (1994, Vol.8, No.3, pp 28-43). Corine also received a distinction for a technology paper presented at the Global Information Technology Management World Conference held in New York in June 2002, and published in "Managing E-Business in the 21st Century" (Ch.4, Heidelberg Press Australia, 2003).

Corine lives in Cape Town, South Africa, the perfect place to satisfy her passion for horse riding, rock climbing and sailing. Missing skiing in the snowy Swiss Alps, she regularly visits her country of birth where she keeps training in Slalom & Giant Slalom. Corine can be contacted on +27 21 531 0253 or by e-mail at arbureau@iafrica.com

Stewart Desson B.Sc., M.Sc.

As a rounded Organisation Development consultant, Stewart works internationally as a facilitator, inspirational speaker, trainer, coach and change agent. His approach to business is a blend of the 'analytical/intellectual' and the 'experiential'. This is typified by his mathematical modelling Masters degree in Operational Research and his ability to use the left side of his brain. He is also educated and experienced in the practical application of psychology and the need for attention to the people aspects of business. This is illustrated by the Masters degree he is currently completing in 'Change Agent Skills and Strategies' - specialising in the application of humanistic psychology to organisational change. Stewart is also part of the faculty and leads on delivery on the prestigious Masters in Business Psychology at the University of Westminster.

Stewart's passion is change – helping to create and manage the process within both individuals and the organisations in which they live. Stewart's fifteen years with British Airways and five years with Insights Learning & Development have taught him that personal and organisational changes are inextricably linked.

As Head of Learning for Insights, Stewart has shaped the development of a carefully thought through portfolio of learning resources, web based psychometrics and e-learning products – all designed to support individual development and organisational effectiveness. One of his specialist areas is leadership and he has implemented several customised leadership models, supported by psychometrics with key clients.

Stewart also designs the curriculum for Insights' global learning programmes, as well as managing and mentoring the faculty of professionals that run Insights accreditations. Stewart supports over 2,000 Insights learning consultants and practitioners in the field.

While with British Airways, Stewart was both a consultant and people manager. He was one of BA's youngest ever Senior Managers when he ran the Heathrow Customer Services Change Programme. He led a team of fifty project managers/change agents and significantly increased their capability to manage change through structuring, influencing, and understanding the psychology of change. Earlier in his career he was part of the leadership team that ran BA's internationally acclaimed Operational Research group.

Stewart has worked in many parts of the world and in addition to travelling through Europe, regularly collaborates with colleagues in North America, Africa the Middle East and Asia. He enjoys playing football and practising judo. He lives in Wokingham in the UK with his wife and four 'organic alarm clocks' – his rather energetic young sons!

He can be contacted on +44 (0)1189 783 729 or by email at sdesson@insights.com

Appendix E

Testimonials and professional feedback on this paper

"This is a fascinating paper that provides the hard evidence required to support the use of the Insights Discovery Evaluator in both my counselling practice as well as for organisational development in the corporate world. I often find that organisations need and demand confirmation of validity before they will invest their time and money in using a psychometric. I am pleased to say that this paper provides the answers to their questions and, furthermore, gives a sense of the brilliance behind the instrument itself".

Professor Peter Smyth, M.Sc.Ed., M.S.W., Ph.D., (C) OACCPP,
Lecturer at York University Toronto and
Faculty at the Centre for Excellence in Critical Care Medicine,
Mount Sinai Hospital, University of Toronto, Canada.

"The integrity of the Insights Discovery Evaluator (IDE) rests solely upon its thorough and rigorous measures of validity/reliability and its congruence with the colour scores measuring personality preferences that embody Jung's concept of opposites. Instruments similar to the IDE have come and gone as each fails to meet one or more of these vigorous measures. Only the best survive. This present landmark effort assures the entire professional community of researchers, developers, practitioners and consumers alike, that their belief in the integrity in the IDE is well placed and positively confirmed. If one was ever skittish about pressing questions of the reliability and validity of the IDE, one can now move onto other concerns".

Professor B. Bradley West, B.S., A.M., Ph.D.
Research Specializations in MBTI, Psychometrics and Projective Techniques
Michigan State University
East Lansing, Michigan, 48864
U.S.A.

"This paper details the competent use of classic psychometric techniques, using the Insights Discovery Evaluator to illustrate these. The evaluator is carefully examined using a large, carefully collected dataset. I concur with the authors in their conclusion that there is strong evidence of the reliability and validity of the colour scores measured in the IDE."

Mike Green, B.A., M.Sc., Ph.D.
Honorary Lecturer in the Division of Mathematics,
University of Dundee, UK.
Fellow of the Royal Statistical Society.

University of Westminster

bpc

Appendix E Continued…

Testimonials and professional feedback on this paper

"Having used the Insights Discovery Evaluator extensively in workshops, research and client work at the University of Dundee, I read the recent publication by the University of Westminster on the psychometric properties of the Insights model with great interest. In presenting a thoughtful, in-depth and clear analysis, the authors have provided a persuasive argument for pursuing typological research with confidence. They vindicate my long-held opinion that the instrument has solid statistical merit, and an underlying robustness that compares very favourably with similar systems."

Nick R.Halpin, Ph.D., C. Psychol, Dip. Couns., MBACP (Accred)
Head of Counselling,
University of Dundee, UK.

"This is a robust body of work that documents the development of a novel psychometric tool and reports the theoretical principles of reliability and validity with factual accuracy"

Professor Tony Towell Ph.D.
Research Methods, University of Westminster
Senior Research Fellow
Department of Clinical Neurophysiology
Hospital for Sick Children, Great Ormond Street,
London, UK.

"This paper provides an important and significant contribution to demonstrating the validity of the Insights Discovery framework as a psychometric tool that meets the criteria of the APA and BPS. It is an essential foundation for the continued development of Insights Discovery, and positions the Insights Discovery framework as a unique and valuable basis for further research into Carl Jung's concept of 'psychological preferences' and in particular, their application in an 'at work' context. Reviewing the paper has served to reinforce for me both the power and simplicity of Discovery as well as its validity as a psychometric evaluator"

Mark E. Mullaly,
Lecturer in the School of Business, University of Alberta and
PMP, President Interthink Consulting Incorporated,
10080 Jasper Avenue, Suite 702, Edmonton, Alberta T5J 1V9, Canada.

Appendix F

Addendum Danish Validity and Reliability vs 2.01_Client Paper

Addendum Dutch S1.2 (beta) Validity and Reliability vs 2.01_Client Paper

Addendum English (Young Person) S1.0 Validity and Reliability vs 2.01_Client Paper

Addendum English S3.0 (Australia) Validity and Reliability vs 2.01_Client Paper

Addendum English S3.0 (Canada) Validity and Reliability vs 2.01_Client Paper

Addendum English S3.0 (Ireland) Validity and Reliability vs 2.01_Client Paper

Addendum English S3.0 (South Africa) Validity and Reliability vs 2.01_Client Paper

Addendum English S3.0 (USA) Validity and Reliability vs 2.01_Client Paper

Addendum Finnish Validity and Reliability vs 2.01_Client Paper

Addendum French Canadian S2.1 Validity and Reliability vs 2.01_Client Paper

Addendum French S2.0 Validity and Reliability vs 2.01_Client Paper

Addendum German S3.1 Validity and Reliability vs 2.01_Client Paper

Addendum Italian R2 Validity and Reliability vs 2.01_Client Paper

Addendum Japanese Validity and Reliability vs 2.01_Client Paper

Addendum Norwegian (Bokmal) R22 Validity and Reliability vs 2.01_Client Paper

Addendum Polish Validity and Reliability vs 2.01_Client Paper

Addendum Portuguese Brazilian POR2 Validity and Reliability vs 2.01_Client Paper

Addendum Portuguese PIR2 Validity and Reliability vs 2.01_Client Paper

Addendum Spanish 'Mexican' S1.0 Validity and Reliability vs 2.01_Client Paper

Addendum Spanish S1.3 Validity and Reliability vs 2.01_Client Paper

Addendum Swedish Validity and Reliability vs 2.01_Client Paper

Addendum Turkish R2 Validity and Reliability vs 2.01_Client Paper

Country Specific Addendum: The Development, Validity and Reliability of the Danish Version of the Insights Discovery Evaluator

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the Danish version of the Insights Discovery Evaluator (IDE). It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability
    - Cronbach Alpha coefficients
    - Split-half correlations
- Validity - Factor Analysis

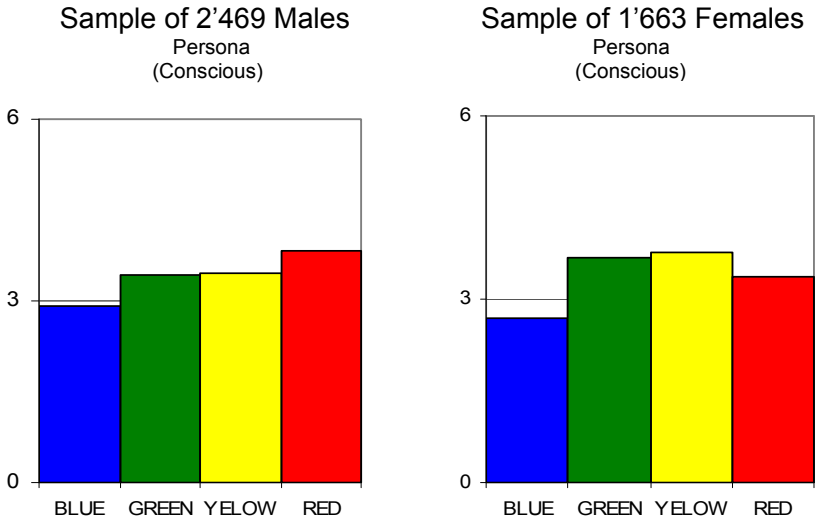
Data on Norms

Presented in Figure 1, is an example specific to the Danish IDE, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and Danish organisations. This sample is not intended to be representative of Danish people overall; it is however, a very useful overview of the Danish clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[2]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[3].



2.91	3.44	3.47	3.82
49%	57%	58%	64%

2.69	3.67	3.76	3.37
45%	61%	63%	56%

Figure 1 - A graphical view for the norms of Males vs. Females for the Danish IDE

Reliability: Cronbach-Alpha Coefficients

This 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

[2] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[3]Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

commonly accepted inferior limit[4]. Analysing 4'132 completed Danish IDE reported in Table 1 shows the four colours to have high Cronbach-Alpha coefficients between 0.85 and 0.90, providing evidence of excellent reliability.

N=4'132 Danish IDE	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	0.86	0.85	0.89	0.90

Table 1 – Cronbach-Alpha coefficients

Reliability: Split-Half Coefficients

The final measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly 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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The analysis of the Danish IDE reported in Table 2 shows reliable coefficients for the IDE, with coefficients being:

- Cronbach-Alpha Coefficients between 0.73 and 0.84 for each half
- Pearson Correlation Coefficients between 0.70 and 0.80 i.e. the 2 halves correlate highly

	N=4'132 Danish IDE	Colour preference			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Part 1	Cronbach-Alpha coefficient	0.74	0.73	0.82	0.84
	N of Items	13	13	13	13
Part 2	Cronbach-Alpha coefficient	0.78	0.77	0.82	0.77
	N of Items	12	12	12	12
	Pearson Correlation coefficient between halves	0.76	0.70	0.70	0.80
	Tot. N of Items	25	25	25	25

Table 2 – Split-Half coefficients

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

- The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent
- The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the Danish IDE.

A four factors solution accounts for 34% of the variance, while a two factors solution accounts for 27% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[5]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 3 below.

The four factor solution highlights the presence of the four constructs (translating into the four colour preferences), while the two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

	Four factor solution				Two factor solution	
Danish IDE	Average Factor Loadings				Average Factor Loadings	
N=4'132	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2
Earth Green	-0.25	-0.04	0.34	-0.04	-0.40	0.00
Sunshine Yellow	-0.01	0.49	-0.02	-0.17	0.03	0.50
Cool Blue	-0.11	-0.31	0.00	0.33	-0.10	-0.43
Fiery Red	0.45	-0.03	-0.24	-0.06	0.50	-0.01

Table 3 – Factor Analysis

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

While Table 3 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the Danish IDE most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

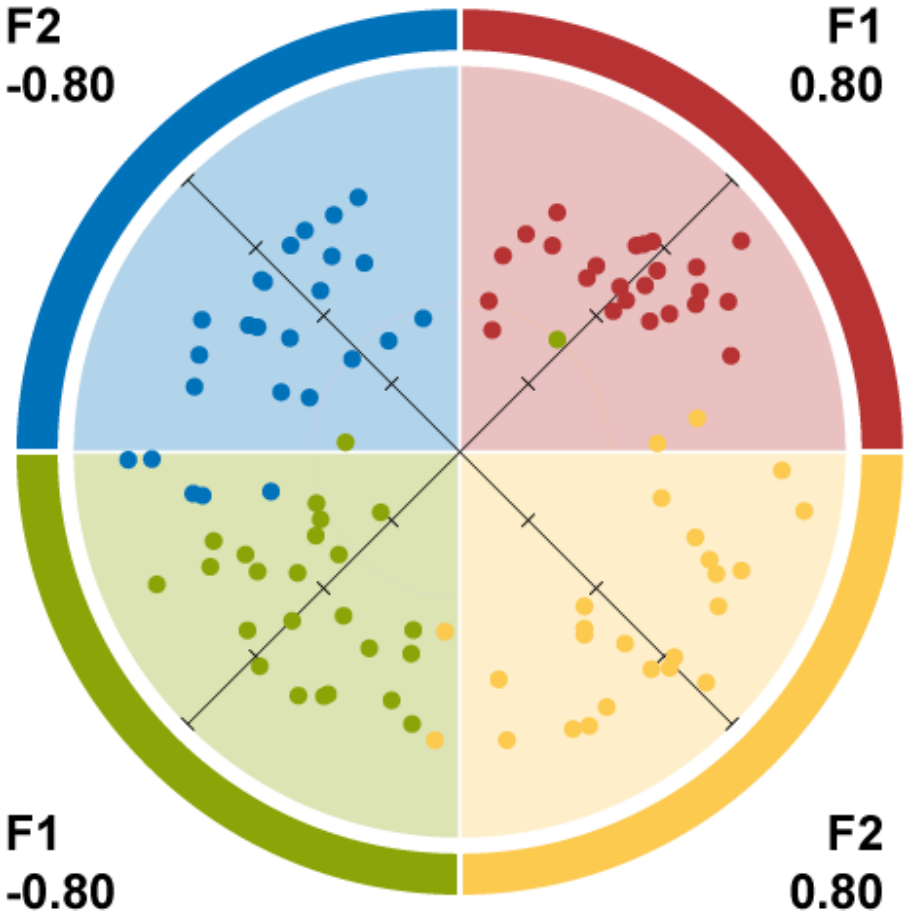


Figure 2 – Danish IDE – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Cronbach Alpha and Split-Half analysis) and construct validity (using Factor Analysis) of the Danish IDE.

Country Specific Addendum: The Development, Validity and Reliability of the Dutch Version S1.2 (beta) of the Insights Discovery Evaluator

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the Dutch version of the Insights Discovery Evaluator (IDE). It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability
    - Split-half correlations
    - Cronbach Alpha coefficients
- Validity - Factor Analysis

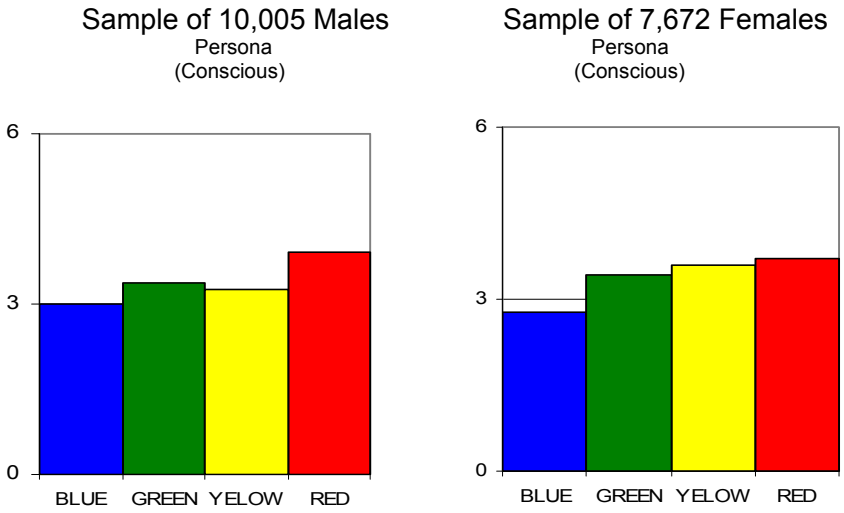
Data on Norms

Presented in Figure 1, is an example specific to the Dutch S1.2 (beta) IDE, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and Dutch organisations. This sample is not intended to be representative of Dutch people overall; it is however, a very useful overview of the Dutch clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[6]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[7].



3.01	3.37	3.27	3.92
50%	56%	54%	65%

2.78	3.41	3.58	3.70
46%	57%	60%	62%

Figure 1 - A graphical view for the norms of Males vs. Females for the Dutch S1.2 (beta) IDE

[6] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[7] Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

Reliability: Cronbach-Alpha Coefficients

This 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 inferior limit[8]. Analysing 17,677 completed Dutch S1.2 (beta) IDE reported in Table 1 shows the four colours to have high Cronbach-Alpha coefficients between 0.90 and 0.94, providing evidence of excellent reliability.

N=17,677 Dutch S1.2 (beta) IDE	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	0.92	0.90	0.92	0.94

Table 1 – Cronbach-Alpha coefficients

Reliability: Split-Half Coefficients

The final measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly 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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The analysis of the Dutch IDE reported in Table 2 shows high coefficients for the IDE, with coefficients being:

- Cronbach-Alpha Coefficients between 0.80 and 0.89 for each half
- Pearson Correlation Coefficients between 0.75 and 0.87 i.e. the 2 halves correlate highly

	N=11'677 Dutch S1.2 (beta)	Colour preference			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Part 1	Cronbach-Alpha coefficient	0.85	0.80	0.84	0.89
	N of Items	13	13	13	13
Part 2	Cronbach-Alpha coefficient	0.86	0.84	0.86	0.88
	N of Items	12	12	12	12
	Pearson Correlation coefficient between halves	0.85	0.75	0.83	0.87
	Tot. N of Items	25	25	25	25

Table 2 – Split-Half coefficients

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

- The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent
- The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the Dutch S1.2 (beta) IDE.

A four factors solution accounts for 41% of the variance, while a two factors solution accounts for 34% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[9]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 3 below.

The four factor solution highlights the presence of the four constructs (translating into the four colour preferences), while the two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

Dutch S1.2 (beta) N=11'677	Four factor solution Average Factor Loadings				Two factor solution Average Factor Loadings	
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2
Earth Green	-0.07	-0.28	0.44	-0.07	-0.47	-0.04
Sunshine Yellow	0.55	0.01	-0.03	-0.11	0.01	0.55
Cool Blue	-0.45	-0.15	-0.03	0.32	-0.09	-0.54
Fiery Red	0.04	0.57	-0.23	-0.08	0.60	0.09

Table 3– Factor Analysis

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

While Table 3 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the Dutch IDE most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

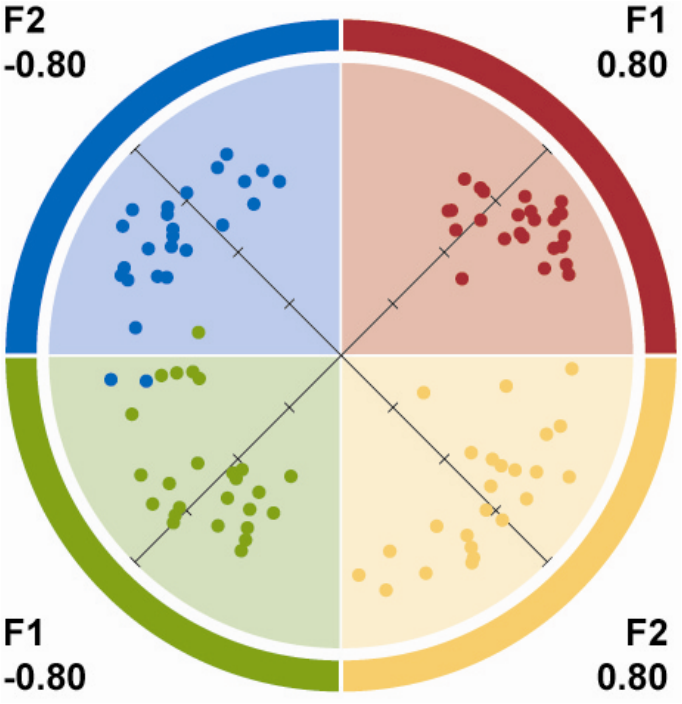


Figure 2 – Dutch S1.2 (beta) IDE – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Cronbach Alpha and Split-Half analysis) and construct validity (using Factor Analysis) of the Dutch S1.2 (beta) IDE.

Version Specific Addendum: The Development, Validity and Reliability of the English - Young Person of the Insights Discovery Evaluator (version 2).

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the English - Young Person version of the Insights Discovery Evaluator (IDE). It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability - Cronbach Alpha coefficients
- Validity - Factor Analysis

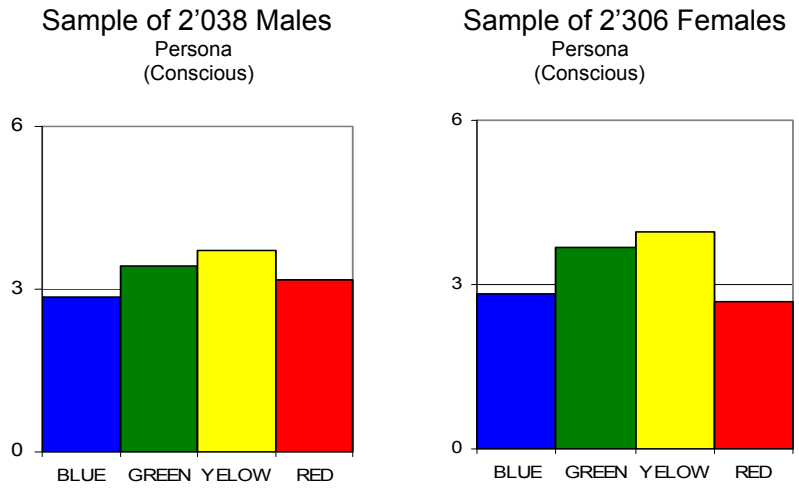
Data on Norms

Presented in Figure 1, is an example specific to the English - Young Person IDE, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and French organisations. This sample is not intended to be representative of English - Young Person overall; it is however, a very useful overview of the French clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[10]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[11].



2.85	3.42	3.71	3.18
47%	57%	62%	53%

2.83	3.67	3.95	2.69
47%	61%	66%	45%

Figure 1 - A graphical view for the norms of Males vs. Females for the French S2.0 IDE

[10] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[11] Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

Reliability

Cronbach-Alpha Coefficients

This 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 inferior limit[12]. Analysing 4'344 completed English (Young Person) IDE reported in Table 1 shows the four colours to have a satisfactory Cronbach-Alpha coefficients between 0.80 and 0.84, providing evidence of excellent reliability.

N=4'344 English (Young Person) IDE	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	0.81	0.80	0.81	0.84

Table 1 – Cronbach-Alpha coefficients

Validity

Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

- The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent
- The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the English (Young Person) IDE.

A two factors solution supporting the bi-polarity of the four IDE constructs accounts for 19% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[13]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background and there bi-polar opposites in yellowed background in the Table 2 below.

English (Young Person) N=4'344	Two factor solution Average Factor Loadings	
	Factor 1	Factor 2
Earth Green	0.36	-0.04
Sunshine Yellow	0.00	0.39
Cool Blue	0.08	-0.35
Fiery Red	-0.40	0.09

Table 2 – Factor Analysis

Conclusion

When read in conjunction with the detailed analysis of the English version of the IDE, this short paper offers good evidence of the internal reliability (using Cronbach Alpha) and construct validity (using Factor Analysis) of the English (Young Person) IDE.

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

Country Specific Addendum: The Development, Validity and Reliability of the English Version S3.0 of the Insights Discovery Evaluator completed in Australia

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the English S3.0 version of the Insights Discovery Evaluator (IDE) completed in Australia. It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability
    - Cronbach Alpha coefficients
    - Split-half correlations
- Validity - Factor Analysis

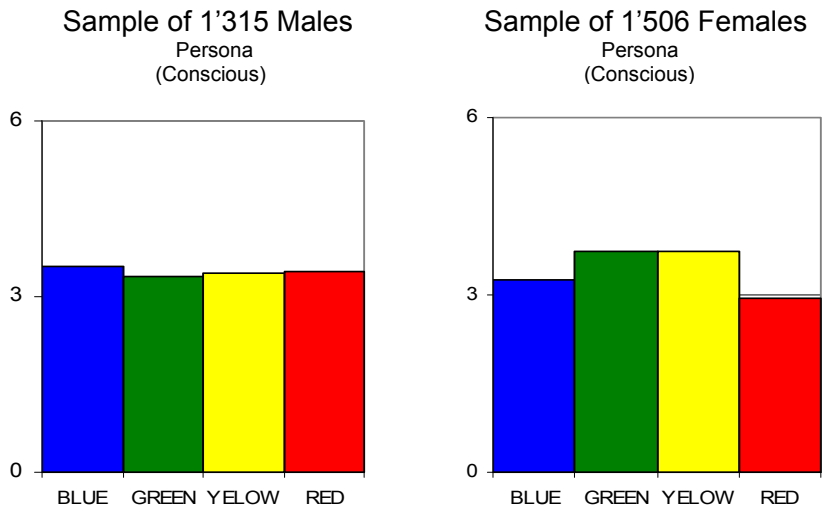
Data on Norms

Presented in Figure 1, is an example specific to the English S3.0 IDE completed in Australia, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and Australian organisations. This sample is not intended to be representative of Australian people overall; it is however, a very useful overview of the Australian clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[14]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[15].



3.48	3.47	3.34	3.32
58%	58%	56%	55%

3.31	3.65	3.66	2.91
55%	61%	61%	49%

Figure 1 - A graphical view for the norms of Males vs. Females for the English S3.0 IDE completed in Australia.

[14] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[15] Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

Reliability: Cronbach-Alpha Coefficients

This 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 inferior limit[16]. The analysis of 2'821 English S3.0 IDE completed in Australia and reported in Table 1 shows that the four colours have high Cronbach-Alpha coefficients between 0.90 and 0.92, providing evidence of excellent reliability.

N=2'821 English S3.0 IDE (Australia)	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	0.91	0.91	0.92	0.90

Table 1 – Cronbach-Alpha coefficients

Reliability: Split-Half Coefficients

The final measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly 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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The analysis of the English S3.0 IDE completed in Australia and reported in Table 2 shows high coefficients for the IDE, with coefficients being:

- Cronbach-Alpha Coefficients between 0.81 and 0.87 for each half
- Pearson Correlation Coefficients between 0.78 and 0.84 i.e. the 2 halves correlate highly

	N=2'821 English S3.0 IDE completed in Australia	Colour preference			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Part 1	Cronbach-Alpha coefficient	0.82	0.81	0.81	0.87
	N of Items	13	13	13	13
Part 2	Cronbach-Alpha coefficient	0.84	0.87	0.83	0.84
	N of Items	12	12	12	12
	Pearson Correlation coefficient between halves	0.81	0.80	0.78	0.84
	Tot. N of Items	25	25	25	25

Table 2 – Split-Half coefficients

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

- The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent
- The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the English S3.0 IDE completed in Australia.

A four factors solution accounts for 37% of the variance, while a two factors solution accounts for 30% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[17]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 3 below.

The four factor solution highlights the presence of the four constructs (translating into the four colour preferences), while the two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

	Four factor solution				Two factor solution	
English S3.0 IDE completed in Australia	Average Factor Loadings				Average Factor Loadings	
N=2'821	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2
Earth Green	0.44	-0.12	-0.08	-0.30	0.52	-0.03
Sunshine Yellow	-0.07	0.46	-0.25	-0.01	-0.03	0.50
Cool Blue	0.04	-0.25	0.48	-0.08	0.08	-0.51
Fiery Red	-0.38	-0.07	-0.14	0.40	-0.55	0.05

Table 3 – Factor Analysis

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

While Table 3 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the English S3.0 IDE completed in Australia, most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

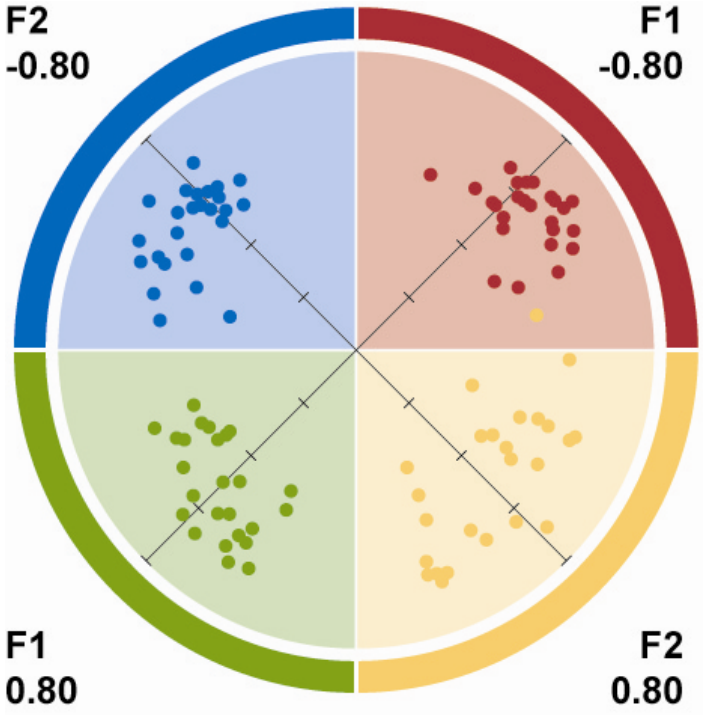


Figure 2 – English S3.0 IDE completed in Australia – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Cronbach Alpha and Split-Half analysis) and construct validity (using Factor Analysis) of the English S3.0 IDE completed in Australia.

University of Westminster

bpc

Country Specific Addendum: The Development, Validity and Reliability of the English Version S3.0 of the Insights Discovery Evaluator completed in Canada

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the English S3.0 version of the Insights Discovery Evaluator (IDE) completed in Canada. It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability
    - Inter-Item correlations
    - Cronbach Alpha coefficients
    - Split-half correlations
    - Test-retest analysis
- Validity - Factor Analysis

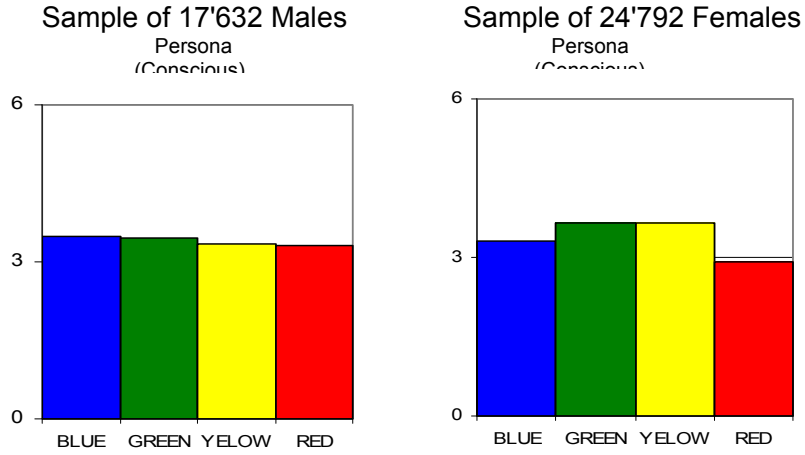
Data on Norms

Presented in Figure 1, is an example specific to the English S3.0 IDE completed in Canada, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and Canadian organisations. This sample is not intended to be representative of Canadian people overall; it is however, a very useful overview of the Canadian (English speaking) clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[18]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[19].



3.48	3.47	3.34	3.32
58%	58%	56%	55%

3.31	3.65	3.66	2.91
55%	61%	61%	49%

Figure 1 - A graphical view for the norms of Males vs. Females for the English S3.0 IDE completed in Canada.

Reliability: 'Inter-item' Correlations

[18] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[19] Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

'Inter-item' correlation coefficients have been calculated using the Pearson Product-Moment Correlation. This involved creating four colour based '25 by 25' matrices showing the correlation between the 25 colour items. In 1991, Robinson et al.[20] concluded that the mean 'inter-item' correlation coefficient should equal or exceed 0.30.

The analysis of 42'424 evaluators performed on the English S3.0 IDE completed in Canada is reported in Table 1 and shows that for the four colours in the evaluator, the average 'inter-item' correlation coefficient is above 0.30. This provides good evidence of the case for reliability.

N=42'424 English S3.0 IDE completed in Canada	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Mean	**0.32**	**0.32**	**0.32**	**0.34**
Minimum	0.09	0.11	0.04	0.12
Maximum	0.60	0.61	0.61	0.54

Table 1 – Inter-item correlations

Reliability: Cronbach-Alpha Coefficients

This 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 inferior limit[21]. The analysis of 42,424 English S3.0 IDE completed in Canada and reported in Table 2 shows that the four colours have high Cronbach-Alpha coefficients between 0.92 and 0.93, providing evidence of excellent reliability.

N=42'424 English S3.0 IDE completed in Canada	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	**0.92**	**0.92**	**0.92**	**0.93**

Table 2 – Cronbach-Alpha coefficients

Reliability: Split-Half Coefficients

[20] 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
Robinson, J.P., Shaver, P.R. (1973) 'Measure of Psychological Attitudes' MI: Survey Research Centre Institute for Social Research, University of Michigan
DeVellis, R. F. (1991) 'Scale Development: Theory and Applications', Sage Publications, Newbury Park, CA
Swailes, S., & McIntyre-Bhatty, T. (2002) 'The "Belbin" team role inventory: reinterpreting reliability estimates', Journal of Managerial Psychology, 17, 6, 529 – 536

The final measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly 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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The analysis of the English S3.0 IDE completed in Canada and reported in Table 3 shows high coefficients for the IDE, with coefficients being:

- Cronbach-Alpha Coefficients between 0.84 and 0.88 for each half
- Pearson Correlation Coefficients between 0.80 and 0.85 i.e. the 2 halves correlate highly

	N=42'424 English S3.0 IDE completed in Canada	Colour preference			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Part 1	Cronbach-Alpha coefficient	0.85	0.84	0.85	0.88
	N of Items	13	13	13	13
Part 2	Cronbach-Alpha coefficient	0.87	0.88	0.86	0.86
	N of Items	12	12	12	12
	Pearson Correlation coefficient between halves	0.83	0.80	0.83	0.85
	Tot. N of Items	25	25	25	25

Table 3 – Split-Half coefficients

Reliability: Temporal Stability – Test / Re-test Correlations

The analysis of 42'424 evaluators performed on the English S3.0 IDE completed in Canada is reported 'Temporal stability' or 'test/re-test' reliability is determined through the administration of the same evaluator across time and it helps us gauge how robust the items are. A sample of 189 people who completed the English IDE twice (with at least 3 months gap) had their original and re-tested colour scores assessed through a Pearson correlation analysis. Reliability is expressed as correlation coefficients, ranging from 1 to 0. Temporal stability tests are generally expected to yield reliability coefficients ranging between 0.70 and 0.90.

The results of the Test / Re-test analysis performed on the four colour scores of the English IDE completed in Canada, and reported in Table 4, show a good reliability, translating into Pearsons' correlation coefficients ranging from 0.80 to 0.86.

N = 189	RETEST Cool Blue	RETEST Earth Green	RETEST Sunshine Yellow	RETEST Fiery Red
TEST Cool Blue	**0.86**	0.14	-0.71	-0.35
TEST Earth Green	0.15	**0.80**	-0.21	-0.70
TEST Sunshine Yellow	-0.69	-0.12	**0.84**	0.15
TEST Fiery Red	-0.35	-0.68	0.23	**0.84**

All correlations in this table are significant at the 0.01 level (2-tailed).

Table 4 – Test re-test correlation

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

• The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent

• The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the English S3.0 IDE completed in Canada.

A two factors solution accounts for 34% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[22]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 5 below.

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

The two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

English S3.0 IDE completed in Canada N=42'424	Two factor solution Average Factor Loadings	
	Factor 1	Factor 2
Earth Green	0.55	-0.03
Sunshine Yellow	-0.03	0.55
Cool Blue	0.08	-0.55
Fiery Red	-0.57	0.04

Table 5 – Factor Analysis

While Table 5 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the English S3.0 IDE completed in Canada, most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

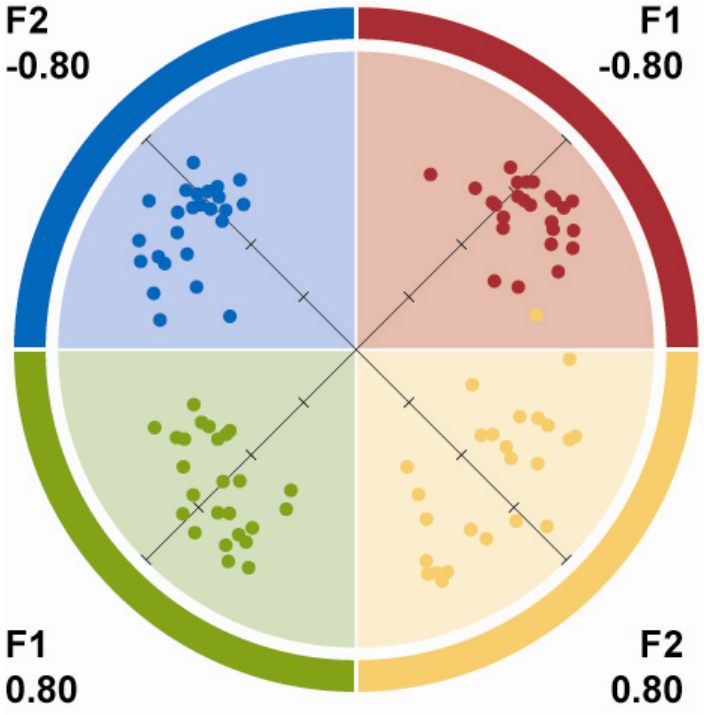


Figure 2 – English S3.0 IDE completed in Canada – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Inter-Item correlations, Cronbach Alpha, Split-Half, and Test-retest analysis) and construct validity (using Factor Analysis) of the English S3.0 IDE completed in Canada.

Country Specific Addendum: The Development, Validity and Reliability of the English Version S3.0 of the Insights Discovery Evaluator completed in Ireland

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the English S3.0 version of the Insights Discovery Evaluator (IDE) completed in Ireland. It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability
    - Cronbach Alpha coefficients
    - Split-half correlations
- Validity - Factor Analysis

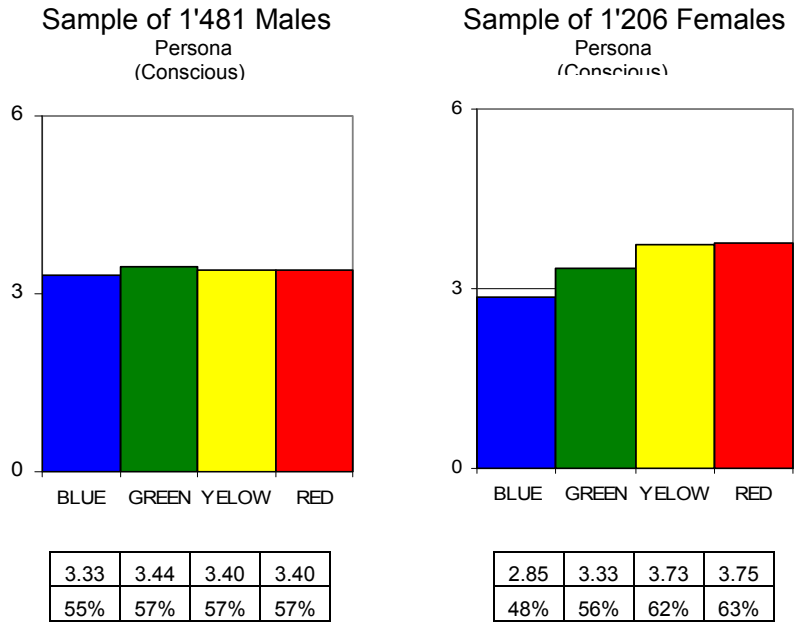
Data on Norms

Presented in Figure 1, is an example specific to the English S3.0 IDE completed in Ireland, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and Irish organisations. This sample is not intended to be representative of Irish people overall; it is however, a very useful overview of the Irish clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[23]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[24].



3.33	3.44	3.40	3.40
55%	57%	57%	57%

2.85	3.33	3.73	3.75
48%	56%	62%	63%

Figure 1 - A graphical view for the norms of Males vs. Females for the English S3.0 IDE completed in Ireland.

Reliability: Cronbach-Alpha Coefficients

[23] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[24]Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

This 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 inferior limit[25]. The analysis of 2'687 English S3.0 IDE completed in Ireland and reported in Table 1 shows that the four colours have high Cronbach-Alpha coefficients between 0.90 and 0.93, providing evidence of excellent reliability.

N=2'687 English S3.0 IDE completed in Ireland	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	0.92	0.91	0.90	0.93

Table 1 – Cronbach-Alpha coefficients

Reliability: Split-Half Coefficients

The final measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly 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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The analysis of the English S3.0 IDE completed in Ireland and reported in Table 2 shows high coefficients for the IDE, with coefficients being:

- Cronbach-Alpha Coefficients between 0.82 and 0.88 for each half
- Pearson Correlation Coefficients between 0.77 and 0.84 i.e. the 2 halves correlate highly

	N=2'687 English S3.0 IDE completed in Ireland	Colour preference			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Part 1	Cronbach-Alpha coefficient	0.85	0.83	0.82	0.88
	N of Items	13	13	13	13
Part 2	Cronbach-Alpha coefficient	0.86	0.86	0.83	0.86
	N of Items	12	12	12	12
	Pearson Correlation coefficient between halves	0.82	0.77	0.78	0.84
	Tot. N of Items	25	25	25	25

Table 2 – Split-Half coefficients

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

• The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent

• The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the English S3.0 IDE completed in Ireland.

A two factors solution accounts for 32% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[26]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 3 below.

The two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

	Two factor solution	
English S3.0 IDE completed in Ireland N=2'687	Average Factor Loadings	
	Factor 1	Factor 2
Earth Green	0.52	-0.01
Sunshine Yellow	-0.02	0.50
Cool Blue	0.08	-0.54
Fiery Red	-0.58	0.07

Table 3 – Factor Analysis

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

While Table 3 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the English S3.0 IDE completed in Ireland, most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

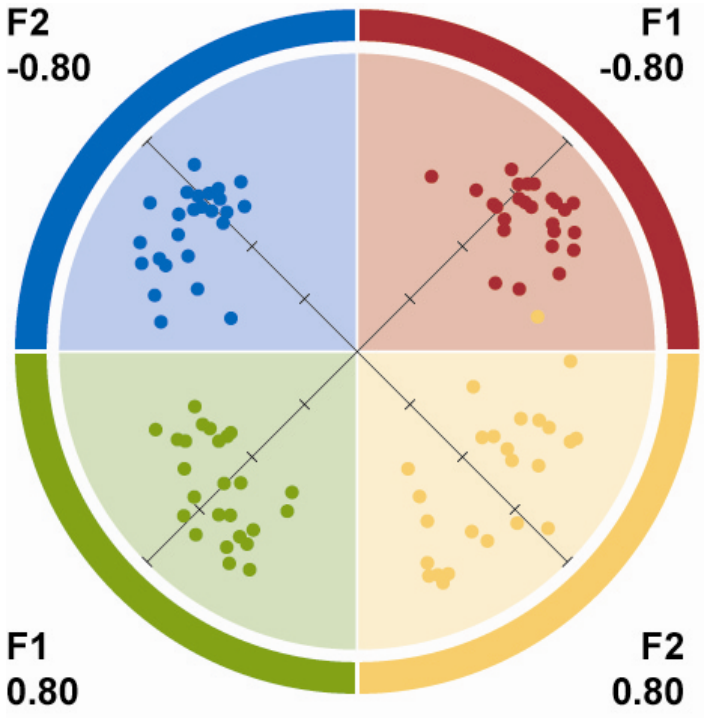


Figure 2 – English S3.0 IDE completed in Ireland – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Cronbach Alpha and Split-Half, analysis) and construct validity (using Factor Analysis) of the English S3.0 IDE completed in Ireland.

Country Specific Addendum: The Development, Validity and Reliability of the English Version S3.0 of the Insights Discovery Evaluator completed in South Africa

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the English S3.0 version of the Insights Discovery Evaluator (IDE) completed in South Africa. It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability
    - Inter-Item correlations
    - Cronbach Alpha coefficients
    - Split-half correlations
- Validity - Factor Analysis

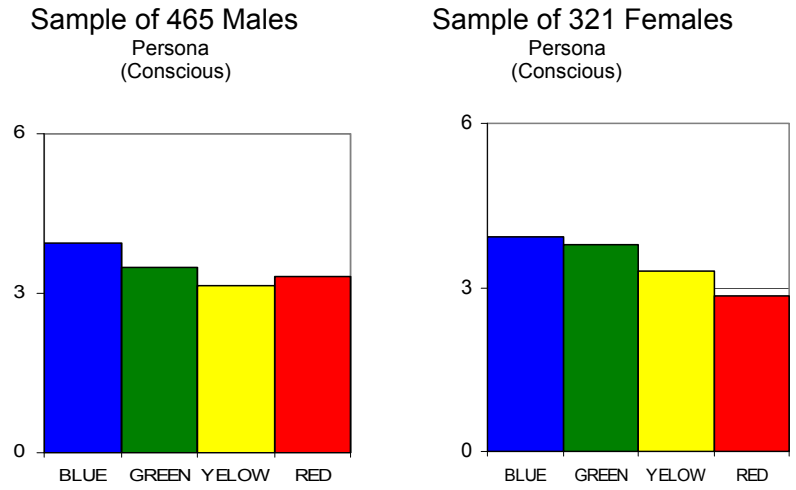
Data on Norms

Presented in Figure 1, is an example specific to the English S3.0 IDE completed in South Africa, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and South African organisations. This sample is not intended to be representative of South African people overall; it is however, a very useful overview of the South African clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[27]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[28].



3.93	3.48	3.14	3.31
65%	58%	52%	55%

3.93	3.77	3.30	2.84
65%	63%	55%	47%

Figure 1 - A graphical view for the norms of Males vs. Females for the English S3.0 IDE completed in South Africa.

[27] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[28] Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

Reliability: 'Inter-item' Correlations

'Inter-item' correlation coefficients have been calculated using the Pearson Product-Moment Correlation. This involved creating four colour based '25 by 25' matrices showing the correlation between the 25 colour items. In 1991, Robinson et al.[29] concluded that the mean 'inter-item' correlation coefficient should equal or exceed 0.30.

The analysis of 789 evaluators performed on the English S3.0 IDE completed in South Africa is reported in Table 1 and shows that for the four colours in the evaluator, the average 'inter-item' correlation coefficient is equal or above 0.30. This provides evidence of the case for reliability.

N = 786 English S3.0 IDE (South Africa)	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Mean	**0.30**	**0.36**	**0.33**	**0.30**
Minimum	0.08	0.14	0.10	0.08
Maximum	0.55	0.61	0.52	0.59

Table 1 – Inter-item correlations

Reliability: Cronbach-Alpha Coefficients

This 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 inferior limit[30]. The analysis of 789 English S3.0 IDE completed in South Africa and reported in Table 2 shows that the four colours have high Cronbach-Alpha coefficients between 0.92 and 0.93, providing evidence of excellent reliability.

N = 786 English S3.0 IDE (South Africa)	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	**0.91**	**0.93**	**0.92**	**0.91**

Table 2 – Cronbach-Alpha coefficients

[29] 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
Robinson, J.P., Shaver, P.R. (1973) 'Measure of Psychological Attitudes' MI: Survey Research Centre Institute for Social Research, University of Michigan
DeVellis, R. F. (1991) 'Scale Development: Theory and Applications', Sage Publications, Newbury Park, CA
Swailes, S., & McIntyre-Bhatty, T. (2002) 'The "Belbin" team role inventory: reinterpreting reliability estimates', Journal of Managerial Psychology, 17, 6, 529 – 536

Reliability: Split-Half Coefficients

The final measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly 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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The analysis of the English S3.0 IDE completed in South Africa and reported in Table 3 shows high coefficients for the IDE, with coefficients being:

- Cronbach-Alpha Coefficients between 0.83 and 0.90 for each half
- Pearson Correlation Coefficients between 0.82 and 0.84 i.e. the 2 halves correlate highly

	N = 786 English S3.0 IDE (South Africa)	Colour preference			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Part 1	Cronbach-Alpha coefficient N of Items	0.83 13	0.86 13	0.84 13	0.87 13
Part 2	Cronbach-Alpha coefficient N of Items	0.86 12	0.90 12	0.85 12	0.85 12
	Pearson Correlation coefficient between halves	0.84	0.83	0.82	0.84
	Tot. N of Items	25	25	25	25

Table 3 – Split-Half coefficients

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

• The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent

• The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the English S3.0 IDE completed in South Africa.

A two factors solution accounts for 32% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[31]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 4 below.

The two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

	Two factor solution	
English S3.0 IDE completed in South Africa N=786	Average Factor Loadings	
	Factor 1	Factor 2
Earth Green	-0.59	-0.01
Sunshine Yellow	0.00	-0.52
Cool Blue	0.05	0.51
Fiery Red	0.55	0.03

Table 4 – Factor Analysis

While Table 4 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the English S3.0 IDE completed in South Africa, most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

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

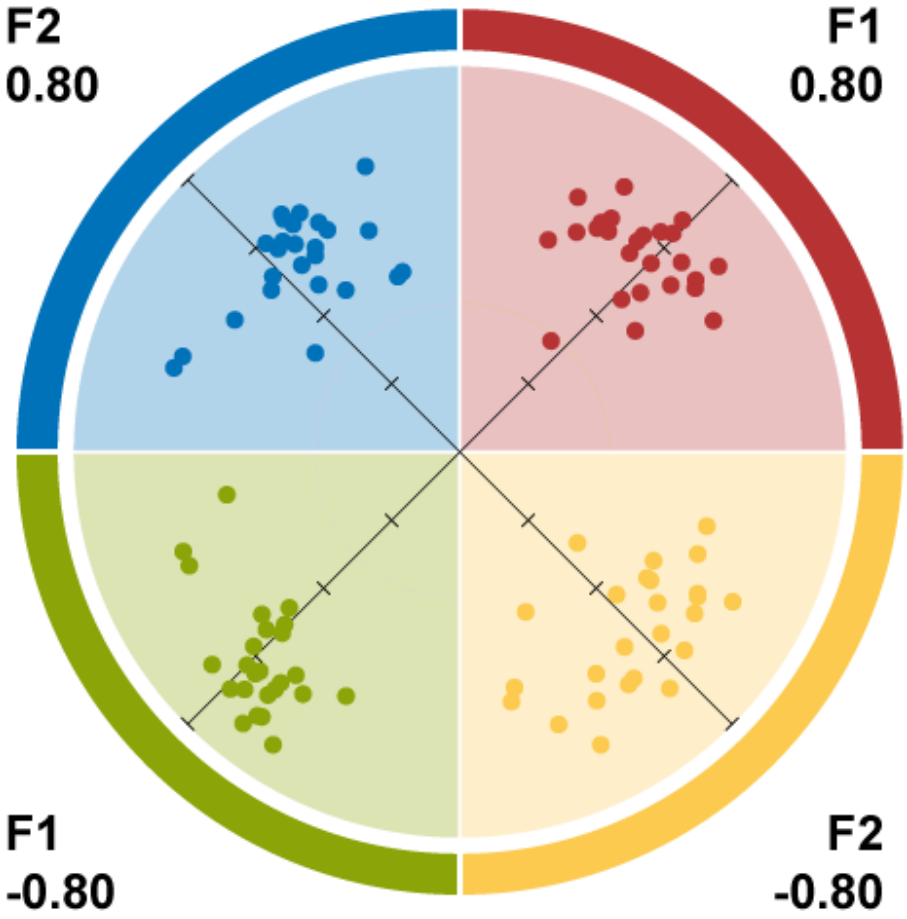


Figure 2 – English S3.0 IDE completed in South Africa – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Inter-Item correlations, Cronbach Alpha, and Split-Half analysis) and construct validity (using Factor Analysis) of the English S3.0 IDE completed in South Africa.

Country Specific Addendum: The Development, Validity and Reliability of the English Version S3.0 completed in the USA of the Insights Discovery Evaluator

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the English S3.0 version of the Insights Discovery Evaluator (IDE) completed in the USA. It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability
    - Inter-item correlations
    - Cronbach Alpha coefficients
    - Split-half correlations
    - Test – re-test correlations
- Validity - Factor Analysis

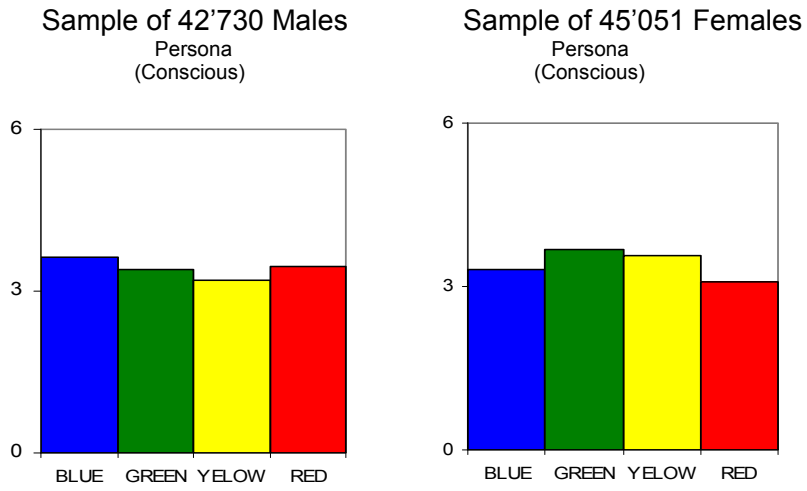
Data on Norms

Presented in Figure 1, is an example specific to the English S3.0 IDE completed in the USA, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and American organisations. This sample is not intended to be representative of American people overall; it is however, a very useful overview of the American clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[32]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[33].



3.63	3.40	3.20	3.46
60%	57%	53%	58%

3.32	3.69	3.55	3.07
55%	62%	59%	51%

Figure 1 - A graphical view for the norms of Males vs. Females for the English S3.0 IDE completed in the USA

Reliability: 'Inter-item' Correlations

[32] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[33] Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

'Inter-item' correlation coefficients have been calculated using the Pearson Product-Moment Correlation. This involved creating four colour based '25 by 25' matrices showing the correlation between the 25 colour items. In 1991, Robinson et al.[34] concluded that the mean 'inter-item' correlation coefficient should equal or exceed 0.30.

The analysis of 87'781 evaluators performed on the English S3.0 IDE completed in the USA is reported in Table 1 and shows that for the four colours in the evaluator, the average 'inter-item' correlation coefficient is above 0.30. This provides good evidence of the case for reliability.

N=87'781 English S3.0 IDE completed in the USA	**Colour preference**			
	Cool Blue	**Earth Green**	**Sunshine Yellow**	**Fiery Red**
Mean	**0.32**	**0.32**	**0.32**	**0.34**
Minimum	0.07	0.10	0.03	0.10
Maximum	0.60	0.60	0.62	0.56

Table 1 – Inter-item correlations

Reliability: Cronbach-Alpha Coefficients

This 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 inferior limit[35]. The analysis of 87'781 English S3.0 IDE completed in the USA and reported in Table 2 shows that the four colours have high Cronbach-Alpha coefficients between 0.92 and 0.93, providing evidence of excellent reliability.

N=87'781 English S3.0 IDE completed in the USA	**Colour preference**			
	Cool Blue	**Earth Green**	**Sunshine Yellow**	**Fiery Red**
Cronbach-Alpha Coefficients	**0.92**	**0.93**	**0.92**	**0.93**

Table 2 – Cronbach-Alpha coefficients

Reliability: Split-Half Coefficients

[34] 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
Robinson, J.P., Shaver, P.R. (1973) 'Measure of Psychological Attitudes' MI: Survey Research Centre Institute for Social Research, University of Michigan
DeVellis, R. F. (1991) 'Scale Development: Theory and Applications', Sage Publications, Newbury Park, CA
Swailes, S., & McIntyre-Bhatty, T. (2002) 'The "Belbin" team role inventory: reinterpreting reliability estimates', Journal of Managerial Psychology, 17, 6, 529 – 536

The final measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly 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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The analysis of the English S3.0 IDE completed in the USA and reported in Table 3 shows high coefficients for the IDE, with coefficients being:

- Cronbach-Alpha Coefficients between 0.84 and 0.89 for each half
- Pearson Correlation Coefficients between 0.79 and 0.85 i.e. the 2 halves correlate highly

	N=87'781 English S3.0 (USA)	Colour preference			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Part 1	Cronbach-Alpha coefficient	0.84	0.84	0.85	0.88
	N of Items	13	13	13	13
Part 2	Cronbach-Alpha coefficient	0.86	0.89	0.87	0.86
	N of Items	12	12	12	12
	Pearson Correlation coefficient between halves	0.83	0.79	0.83	0.85
	Tot. N of Items	25	25	25	25

Table 3 – Split-Half coefficients

Reliability: Temporal Stability – Test / Re-test Correlations

The analysis of 87'781 evaluators performed on the English S3.0 IDE completed in the USA is reported 'Temporal stability' or 'test/re-test' reliability is determined through the administration of the same evaluator across time and it helps us gauge how robust the items are. A sample of 394 people who completed the English IDE twice (with at least 3 months gap) had their original and re-tested colour scores assessed through a Pearson correlation analysis. Reliability is expressed as correlation coefficients, ranging from 1 to 0. Temporal stability tests are generally expected to yield reliability coefficients ranging between 0.70 and 0.90.

The results of the Test / Re-test analysis performed on the four colour scores of the English IDE completed in the USA, and reported in Table 4, show a good reliability, translating into Pearsons' correlation coefficients ranging from 0.79 to 0.85.

N = 394	RETEST Cool Blue	RETEST Earth Green	RETEST Sunshine Yellow	RETEST Fiery Red
TEST Cool Blue	**0.83**	**0.07**	**-0.69**	**-0.16**
TEST Earth Green	0.09	**0.79**	-0.09	-0.67
TEST Sunshine Yellow	-0.71	-0.06	**0.85**	0.01
TEST Fiery Red	-0.16	-0.67	0.01	**0.81**

All correlations in this table are significant at the 0.01 level (2-tailed).

Table 4 – Test re-test correlation

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

• The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent

• The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the English S3.0 IDE completed in the USA.

A two factors solution accounts for 34% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[36]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 5 below.

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

The two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

N=87'781 English S3.0 (USA)	Two factor solution Average Factor Loadings	
	Factor 1	Factor 2
Earth Green	0.55	-0.03
Sunshine Yellow	-0.02	0.56
Cool Blue	0.06	-0.54
Fiery Red	-0.58	0.03

Table 5 – Factor Analysis

While Table 5 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the English S3.0 IDE completed in the USA, most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

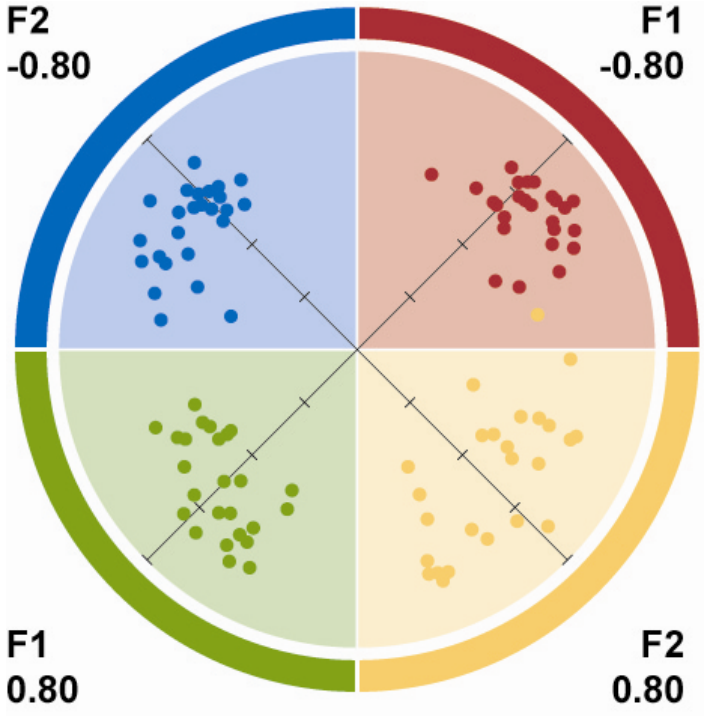


Figure 2 – English S3.0 IDE completed in the USA – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Inter-Item correlations, Cronbach Alpha, Split-Half, and Test-retest analysis) and construct validity (using Factor Analysis) of the English S3.0 IDE completed in the USA.

Country Specific Addendum: The Development, Validity and Reliability of the Finnish Version of the Insights Discovery Evaluator

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the Finnish version of the Insights Discovery Evaluator (IDE). It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability
    - Inter-Item correlations
    - Cronbach Alpha coefficients
    - Split-half correlations
- Validity - Factor Analysis

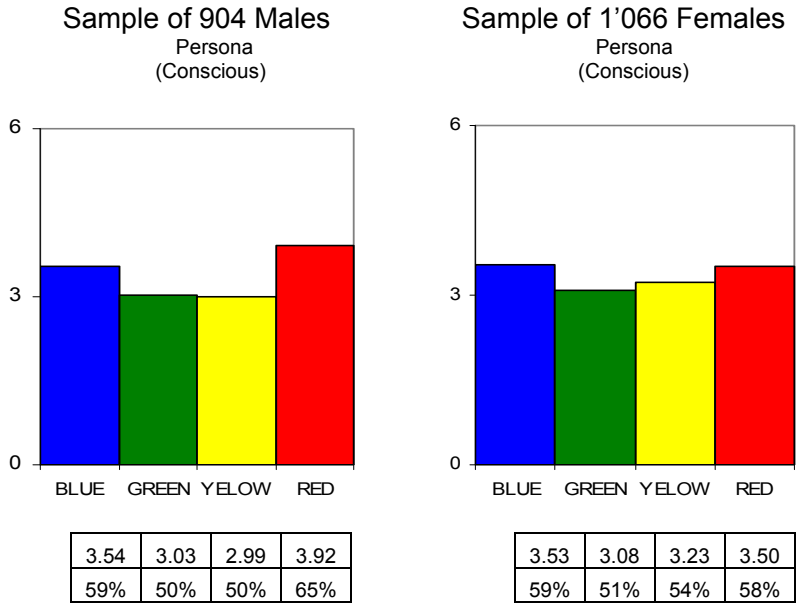
Data on Norms

Presented in Figure 1, is an example specific to the Finnish IDE, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and Finnish organisations. This sample is not intended to be representative of Finnish people overall; it is however, a very useful overview of the Finnish clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[37]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[38].



3.54	3.03	2.99	3.92
59%	50%	50%	65%

3.53	3.08	3.23	3.50
59%	51%	54%	58%

Figure 1 - A graphical view for the norms of Males vs. Females for the Finnish IDE

Reliability: 'Inter-item' Correlations

'Inter-item' correlation coefficients have been calculated using the Pearson Product-Moment Correlation. This involved creating four colour based '25 by 25' matrices showing the correlation between the 25

[37] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[38] Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

colour items. In 1991, Robinson et al.[39] concluded that the mean 'inter-item' correlation coefficient should equal or exceed 0.30.

The analysis of 1'970 evaluators performed on the Finnish IDE is reported in Table 1 and shows that for the four colours in the evaluator, the average 'inter-item' correlation coefficient is equal or above 0.30. This provides good evidence of the case for reliability.

N=1'970 Finnish IDE	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Mean	**0.31**	**0.34**	**0.38**	**0.38**
Minimum	0.02	0.10	0.02	0.09
Maximum	0.55	0.70	0.65	0.68

Table 1 – Inter-item correlations

Reliability: Cronbach-Alpha Coefficients

This 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 inferior limit[40]. Analysing 1'970 completed Finnish IDE reported in Table 2 shows the four colours to have high Cronbach-Alpha coefficients between 0.92 and 0.94, providing evidence of excellent reliability.

N=1'970 Finnish IDE	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	**0.92**	**0.93**	**0.94**	**0.94**

Table 2 – Cronbach-Alpha coefficients

Reliability: Split-Half Coefficients

The final measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly divide all items that are thought to measure the same construct into

[39] 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
Robinson, J.P., Shaver, P.R. (1973) 'Measure of Psychological Attitudes' MI: Survey Research Centre Institute for Social Research, University of Michigan
DeVellis, R. F. (1991) 'Scale Development: Theory and Applications', Sage Publications, Newbury Park, CA
Swailes, S., & McIntyre-Bhatty, T. (2002) 'The "Belbin" team role inventory: reinterpreting reliability estimates', Journal of Managerial Psychology, 17, 6, 529 – 536

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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The analysis of the Finnish IDE reported in Table 3 shows high coefficients for the IDE, with coefficients being:

- Cronbach-Alpha Coefficients between 0.84 and 0.90 for each half
- Pearson Correlation Coefficients between 0.82 and 0.88 i.e. the 2 halves correlate highly

	N=1'970 Finnish IDE	Colour preference			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Part 1	Cronbach-Alpha coefficient	0.84	0.87	0.87	0.89
	N of Items	13	13	13	13
Part 2	Cronbach-Alpha coefficient	0.86	0.87	0.90	0.87
	N of Items	12	12	12	12
	Pearson Correlation coefficient between halves	0.82	0.85	0.87	0.88
	Tot. N of Items	25	25	25	25

Table 3 – Split-Half coefficients

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

- The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent
- The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the Finnish IDE.

A four factors solution accounts for 44% of the variance, while a two factors solution accounts for 37% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[41]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 4 below.

The four factor solution highlights the presence of the four constructs (translating into the four colour preferences), while the two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

	Four factor solution				Two factor solution	
Finnish IDE	Average Factor Loadings				Average Factor Loadings	
N=1'970	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2
Earth Green	0.56	-0.16	-0.01	-0.11	-0.55	0.11
Sunshine Yellow	-0.08	0.53	-0.32	0.01	0.08	-0.60
Cool Blue	0.04	-0.21	0.52	-0.07	-0.06	0.52
Fiery Red	-0.48	-0.05	-0.08	0.38	0.60	-0.02

Table 4 – Factor Analysis

While Table 4 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the Finnish IDE most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

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

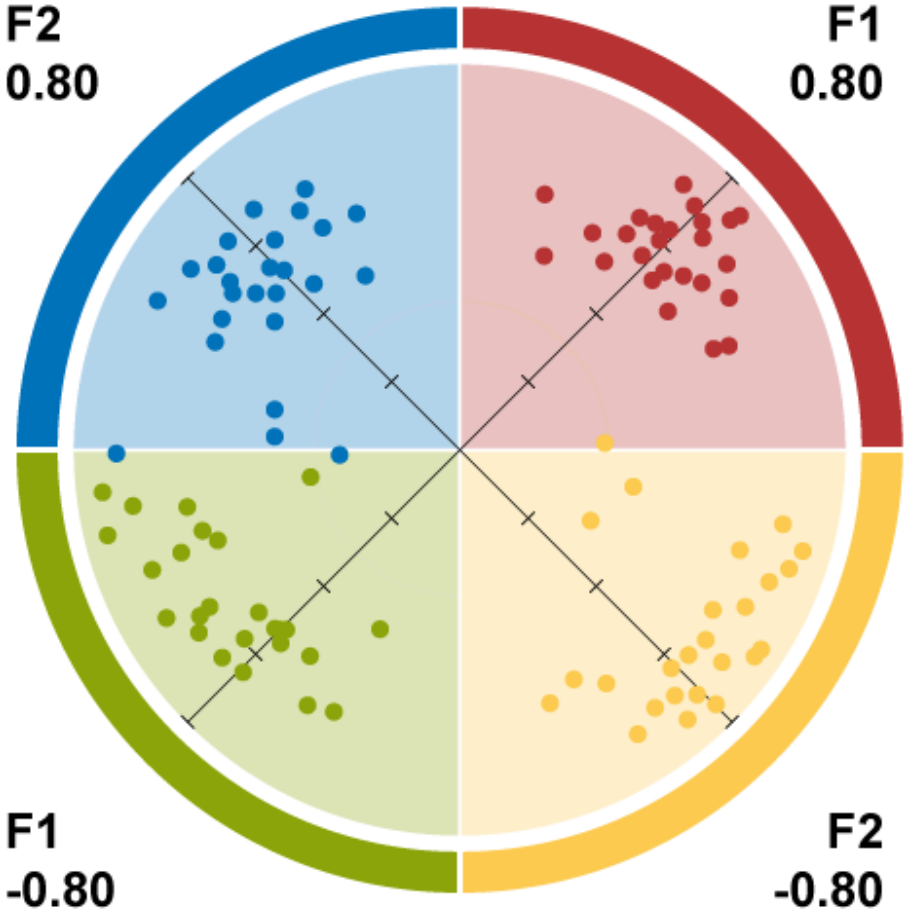


Figure 2 – Finnish IDE – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Inter-Item correlations, Cronbach Alpha and Split-Half analysis) and construct validity (using Factor Analysis) of the Finnish IDE.

Country Specific Addendum: The Development, Validity and Reliability of the French Canadian Version S2.1 of the Insights Discovery Evaluator

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the French Canadian version of the Insights Discovery Evaluator (IDE). It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability
    - Split-half correlations
    - Cronbach Alpha coefficients
- Validity - Factor Analysis

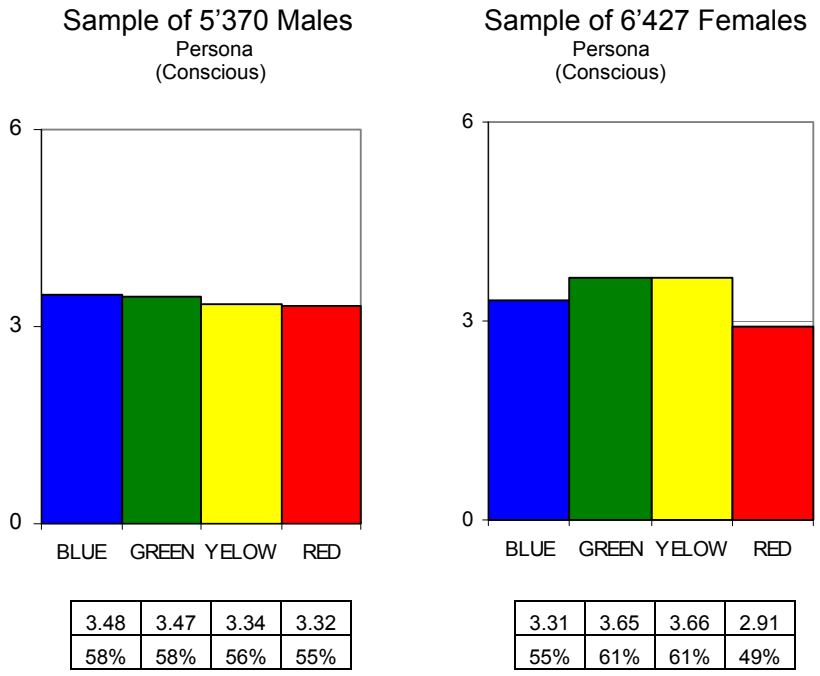
Data on Norms

Presented in Figure 1, is an example specific to the French Canadian IDE, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and French Canadian organisations. This sample is not intended to be representative of French Canadian people overall; it is however, a very useful overview of the French Canadian clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[42]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[43].



3.48	3.47	3.34	3.32
58%	58%	56%	55%

3.31	3.65	3.66	2.91
55%	61%	61%	49%

Figure 1 - A graphical view for the norms of Males vs. Females for the French Canadian S2.1 IDE

Reliability: Cronbach-Alpha Coefficients

This 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

[42] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[43] Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

commonly accepted inferior limit[44]. Analysing 11,797 completed French Canadian S2.1 IDE reported in Table 1 shows the four colours to have high Cronbach-Alpha coefficients between 0.90 and 0.92, providing evidence of excellent reliability.

N=11'797 French Canadian S2.1 IDE	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	0.91	0.90	0.92	0.92

Table 1 – Cronbach-Alpha coefficients

Reliability: Split-Half Coefficients

The final measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly 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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The analysis of the French Canadian IDE reported in Table 2 shows high coefficients for the IDE, with coefficients being:

- Cronbach-Alpha Coefficients between 0.82 and 0.85 for each half
- Pearson Correlation Coefficients between 0.80 and 0.86 i.e. the 2 halves correlate highly

	N=11'797 French Canadian S2.1 IDE	Colour preference			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Part 1	Cronbach-Alpha coefficient	0.82	0.82	0.85	0.85
	N of Items	13	13	13	13
Part 2	Cronbach-Alpha coefficient	0.84	0.83	0.85	0.84
	N of Items	12	12	12	12
	Pearson Correlation coefficient between halves	0.82	0.80	0.86	0.84
	Tot. N of Items	25	25	25	25

Table 2 – Split-Half coefficients

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

- The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent
- The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the French Canadian S2.1 IDE.

A four factors solution accounts for 39% of the variance, while a two factors solution accounts for 32% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[45]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 3 below.

The four factor solution highlights the presence of the four constructs (translating into the four colour preferences), while the two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

French Canadian S2.1 IDE N=11'797	Four factor solution Average Factor Loadings				Two factor solution Average Factor Loadings	
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2
Earth Green	-0.27	-0.04	0.42	-0.14	0.48	-0.09
Sunshine Yellow	-0.06	-0.29	-0.13	0.48	-0.04	0.56
Cool Blue	-0.10	0.49	0.07	-0.21	0.12	-0.49
Fiery Red	0.48	-0.10	-0.27	-0.03	-0.54	0.04

Table 3 – Factor Analysis

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

While Table 3 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the French Canadian IDE most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

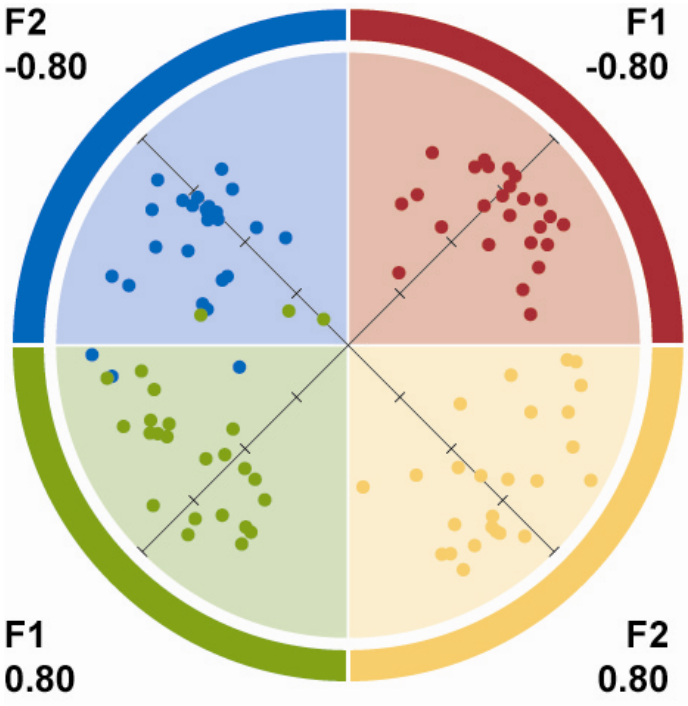


Figure 2 – French Canadian S2.1 IDE – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Cronbach Alpha and Split-Half analysis) and construct validity (using Factor Analysis) of the French Canadian S2.1 IDE.

Country Specific Addendum: The Development, Validity and Reliability of the French Version S2.0 of the Insights Discovery Evaluator

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the French S2.0 version of the Insights Discovery Evaluator (IDE). It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability
    - Split-half correlations
    - Cronbach Alpha coefficients
- Validity - Factor Analysis

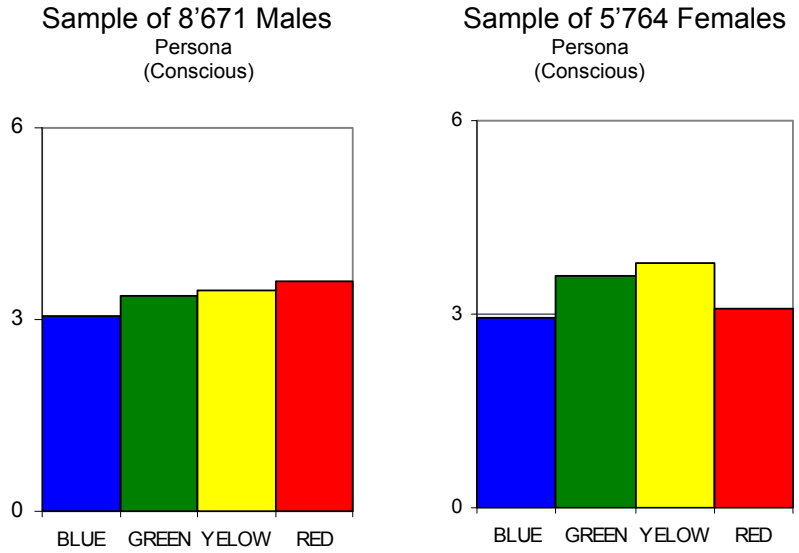
Data on Norms

Presented in Figure 1, is an example specific to the French S2.0 IDE, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and French organisations. This sample is not intended to be representative of French people overall; it is however, a very useful overview of the French clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[46]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[47].



3.06	3.37	3.47	3.59
51%	56%	58%	60%

2.93	3.59	3.79	3.09
49%	60%	63%	51%

Figure 1 - A graphical view for the norms of Males vs. Females for the French S2.0 IDE

[46] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[47] Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

Reliability: Cronbach-Alpha Coefficients

This 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 inferior limit[48]. Analysing 14,435 completed French S2.0 IDE reported in Table 1 shows the four colours to have high Cronbach-Alpha coefficients between 0.87 and 0.92, providing evidence of excellent reliability.

N=14'435 French S2.0 IDE	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	0.88	0.87	0.91	0.92

Table 1 – Cronbach-Alpha coefficients

Reliability: Split-Half Coefficients

The final measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly 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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The analysis of the French IDE reported in Table 2 shows good coefficients for the IDE, with coefficients being:

- Cronbach-Alpha Coefficients between 0.75 and 0.87 for each half
- Pearson Correlation Coefficients between 0.74 and 0.84 i.e. the 2 halves correlate highly

	French S2.0(FRE) N=14'435	Colour preference			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Part 1	Cronbach-Alpha coefficient N of Items	0.75 13	0.76 13	0.81 13	0.87 13
Part 2	Cronbach-Alpha coefficient N of Items	0.83 12	0.80 12	0.85 12	0.84 12
	Pearson Correlation coefficient between halves	0.74	0.75	0.83	0.84
	Tot. N of Items	25	25	25	25

Table 2 – Split-Half coefficients

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

- The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent
- The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the French S2.0 IDE.

A two factors solution accounts for 29% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[49]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 3 below.

The two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

French S2.0(FRE) N=14'435	Two factor solution Average Factor Loadings	
	Factor 1	Factor 2
Earth Green	-0.44	-0.04
Sunshine Yellow	-0.04	0.52
Cool Blue	-0.07	-0.46
Fiery Red	0.56	0.02

Table 3 – Factor Analysis

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

While Table 3 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the French IDE most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

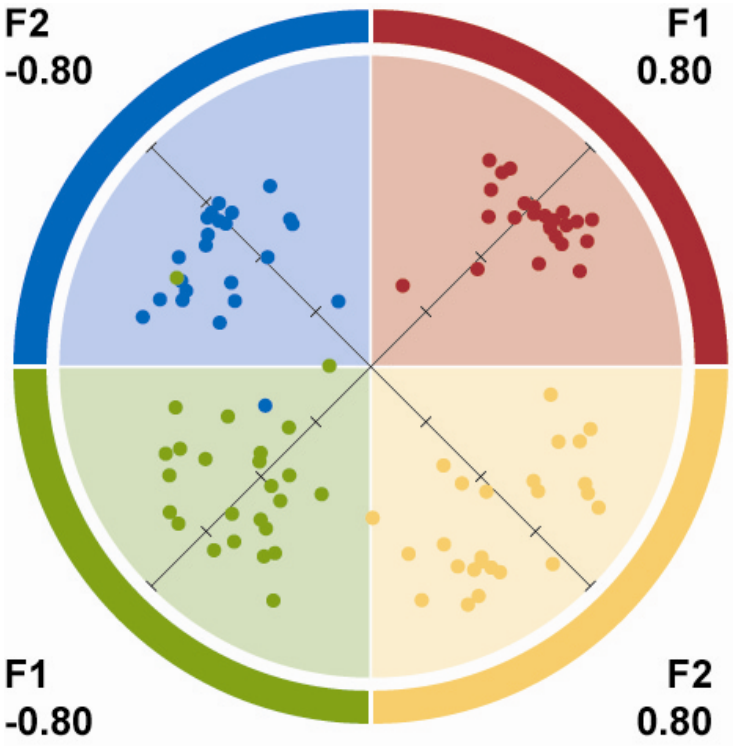


Figure 2 –French S2.0 IDE – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Cronbach Alpha and Split-Half analysis) and construct validity (using Factor Analysis) of the French S2.0 IDE.

Country Specific Addendum: The Development, Validity and Reliability of the German Version S3.1 of the Insights Discovery Evaluator

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the German version of the Insights Discovery Evaluator (IDE). It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability
    - Split-half correlations
    - Cronbach Alpha coefficients
    - Test-retest analysis
- Validity - Factor Analysis

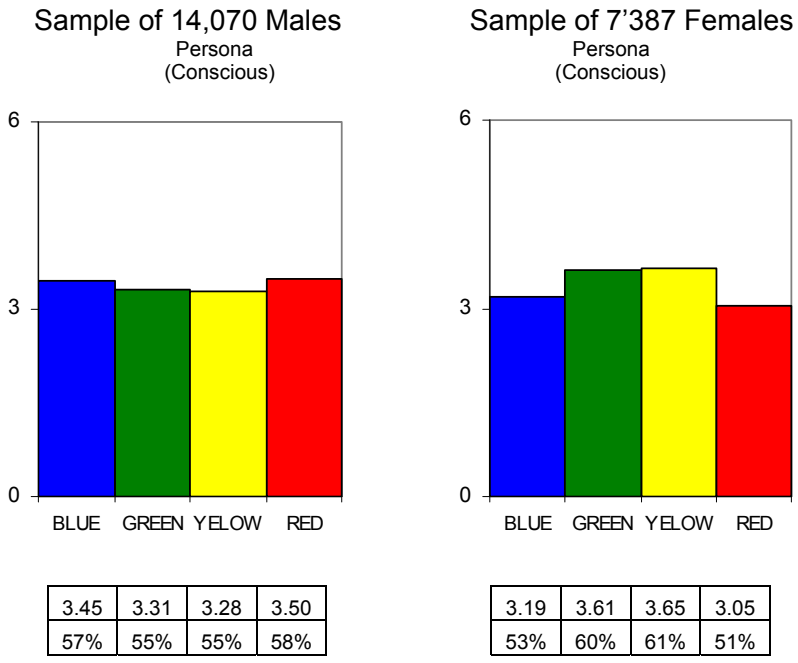
Data on Norms

Presented in Figure 1, is an example specific to the German S3.1 IDE, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and German organisations. This sample is not intended to be representative of German people overall; it is however, a very useful overview of the German clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[50]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[51].



3.45	3.31	3.28	3.50
57%	55%	55%	58%

3.19	3.61	3.65	3.05
53%	60%	61%	51%

Figure 1 - A graphical view for the norms of Males vs. Females for the German S3.1 IDE

Reliability: Cronbach-Alpha Coefficients

This 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

[50] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[51] Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

commonly accepted inferior limit[52]. Analysing 21,417 completed German S3.1 IDE reported in Table 1 shows the four colours to have high Cronbach-Alpha coefficients between 0.90 and 0.91, providing evidence of excellent reliability.

N=21,417 German S3.1 IDE	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	0.908	0.909	0.906	0.910

Table 1 – Cronbach-Alpha coefficients

Reliability: Split-Half Coefficients

The final measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly 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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The analysis of the German IDE reported in Table 2 shows high coefficients for the IDE, with coefficients being:

- Cronbach-Alpha Coefficients between 0.80 and 0.85 for each half
- Pearson Correlation Coefficients between 0.82 and 0.84 i.e. the 2 halves correlate highly

	N=21'417 German S3.1 IDE	Colour preference			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Part 1	Cronbach-Alpha coefficient	0.811	0.801	0.847	0.844
	N of Items	13	13	13	13
Part 2	Cronbach-Alpha coefficient	0.849	0.865	0.806	0.828
	N of Items	12	12	12	12
	Pearson Correlation coefficient between halves	0.84	0.82	0.83	0.82
	Tot. N of Items	25	25	25	25

Table 2 – Split-Half coefficients

Reliability: Temporal Stability – Test / Re-test Correlations

'Temporal stability' or 'test/re-test' reliability is determined through the administration of the same evaluator across time and it helps us gauge how robust the items are. A sample of 437 people who completed the German IDE twice (with at least 3 months gap) had their original and re-tested colour scores assessed through a Pearson correlation analysis. Reliability is expressed as correlation coefficients, ranging from 1 to 0. Temporal stability tests are generally expected to yield reliability coefficients ranging between 0.70 and 0.90.

The results of the Test / Re-test analysis performed on the four colour scores of the German IDE, and reported in Table 3 show a high reliability, translating into Pearsons' correlation coefficients ranging from 0.81 to 0.84.

N = 437	RETEST Cool Blue	RETEST Earth Green	RETEST Sunshine Yellow	RETEST Fiery Red
TEST Cool Blue	**0.82**	-0.33	-0.67	0.24
TEST Earth Green	-0.28	**0.84**	0.09	-0.68
TEST Sunshine Yellow	-0.66	0.11	**0.81**	-0.19
TEST Fiery Red	0.22	-0.66	-0.21	**0.84**

All correlations in this table are significant at the 0.01 level (2-tailed).

Table 3 – Test re-test correlation

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

- The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent
- The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the German S3.1 IDE.

A four factors solution accounts for 39% of the variance, while a two factors solution accounts for 31% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[53]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 4 below.

The four factor solution highlights the presence of the four constructs (translating into the four colour preferences), while the two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

	Four factor solution				Two factor solution	
N=21'417 German S3.1 IDE	Average Factor Loadings				Average Factor Loadings	
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2
Earth Green	0.53	-0.10	-0.03	-0.07	-0.52	-0.03
Sunshine Yellow	-0.04	0.48	-0.23	0.05	0.07	0.51
Cool Blue	0.00	-0.26	0.48	-0.07	-0.04	-0.52
Fiery Red	-0.44	-0.01	-0.14	0.30	0.52	0.09

Table 4 – Factor Analysis

While Table 4 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the German IDE most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

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

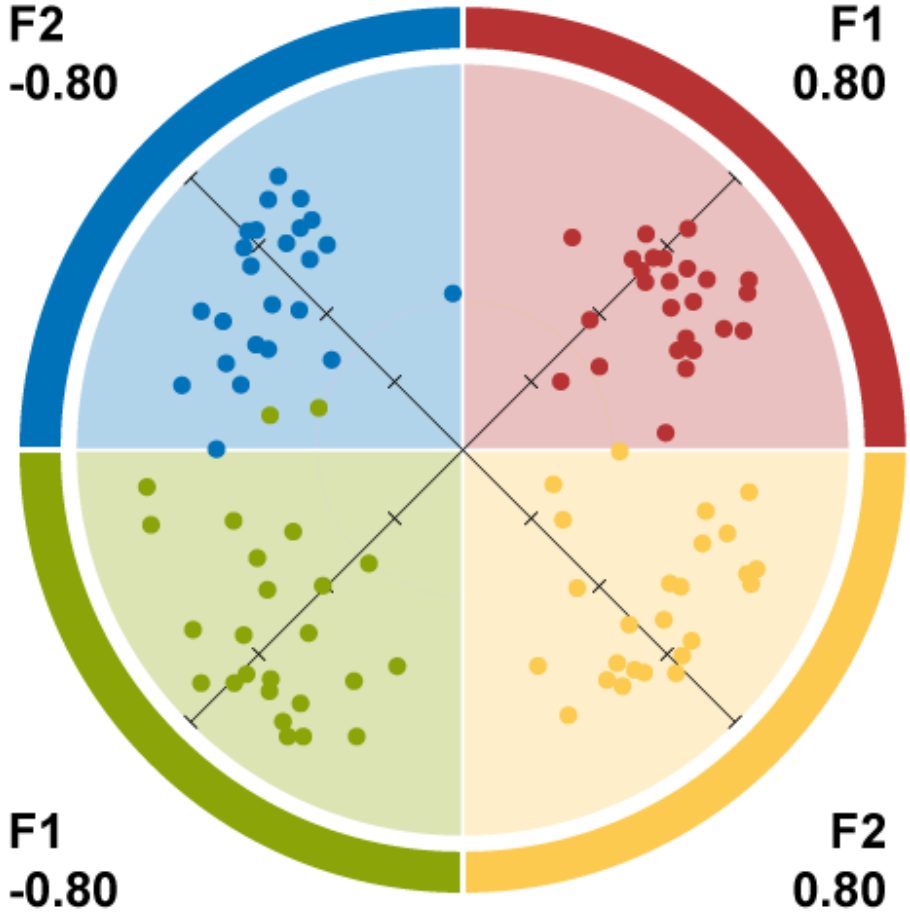


Figure 2 – German S3.1 IDE – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Cronbach Alpha, Split-Half analysis and Test-retest analysis) and construct validity (using Factor Analysis) of the German S3.1 IDE.

Country Specific Addendum: The Development, Validity and Reliability of the Italian Version R2 of the Insights Discovery Evaluator

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the Italian version of the Insights Discovery Evaluator (IDE). It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability
    - Split-half correlations
    - Cronbach Alpha coefficients
- Validity - Factor Analysis

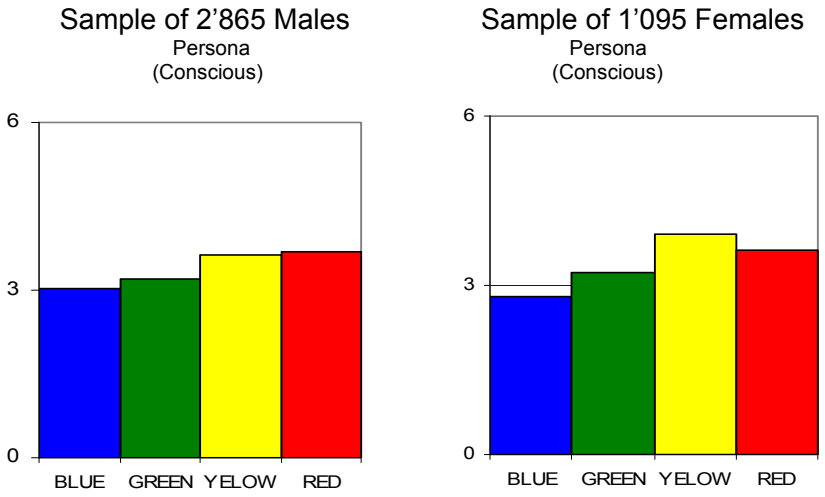
Data on Norms

Presented in Figure 1, is an example specific to the Italian Version R2 IDE, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and Italian organisations. This sample is not intended to be representative of Italian people overall; it is however, a very useful overview of the Italian clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[54]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[55].



3.04	3.21	3.64	3.70
51%	54%	61%	62%

2.81	3.22	3.90	3.61
47%	54%	65%	60%

Figure 1 - A graphical view for the norms of Males vs. Females for the Italian R2 IDE

Reliability: Cronbach-Alpha Coefficients

[54] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[55]Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

This 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 inferior limit[56]. Analysing 3'960 completed Italian R2 IDE reported in Table 1 shows the four colours to have high Cronbach-Alpha coefficients between 0.80 and 0.90, providing evidence of excellent reliability.

N=3'960 Italian R2 IDE	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	0.85	0.80	0.89	0.90

Table 1 – Cronbach-Alpha coefficients

Reliability: Split-Half Coefficients

The final measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly 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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The analysis of the Italian IDE reported in Table 2 shows acceptable coefficients for the IDE (although the 'Earth Green' colour preference is border line), with coefficients being:

- Cronbach-Alpha Coefficients between 0.67 and 0.84 for each half
- Pearson Correlation Coefficients between 0.70 and 0.82 i.e. the 2 halves correlate highly

	N=3'960 Italian R2 (ITA)	Colour preference			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Part 1	Cronbach-Alpha coefficient N of Items	0.72 13	0.67 13	0.79 13	0.84 13
Part 2	Cronbach-Alpha coefficient N of Items	0.76 12	0.69 12	0.82 12	0.80 12
	Pearson Correlation coefficient between halves	0.75	0.70	0.77	0.82
	Tot. N of Items	25	25	25	25

Table 2 – Split-Half coefficients

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

- The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent
- The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the Italian R2 IDE.

A four factors solution accounts for 33% of the variance, while a two factors solution accounts for 25% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[57]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 3 below.

The four factor solution highlights the presence of the four constructs (translating into the four colour preferences), while the two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

	Four factor solution				Two factor solution	
Italian R2 (ITA)	Average Factor Loadings				Average Factor Loadings	
N=3'960	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2
Earth Green	-0.06	-0.17	0.32	-0.10	-0.34	-0.04
Sunshine Yellow	0.49	-0.01	-0.08	-0.12	0.02	0.49
Cool Blue	-0.28	-0.05	0.12	0.31	-0.09	-0.40
Fiery Red	-0.03	0.42	-0.30	-0.06	0.52	0.02

Table 3 – Factor Analysis

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

While Table 3 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the Italian IDE most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

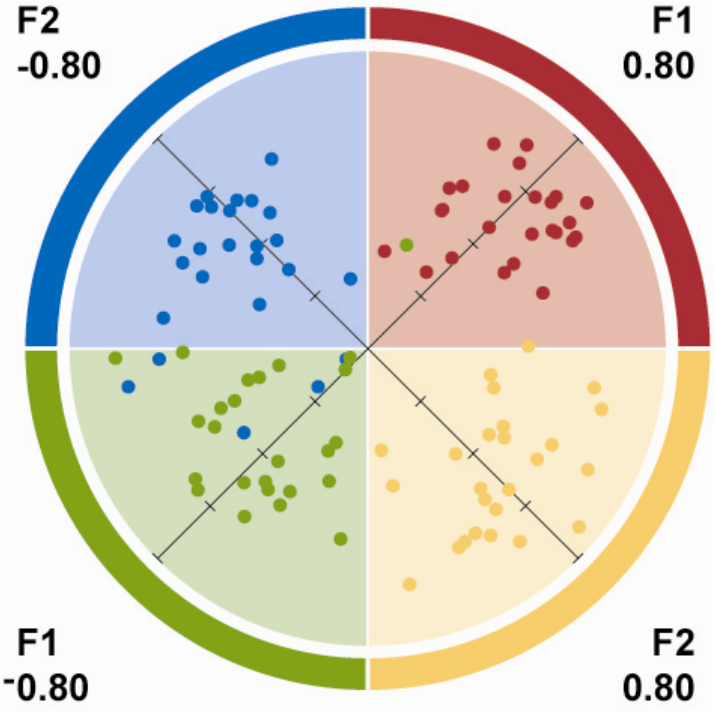


Figure 2 – Italian R2 IDE – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Cronbach Alpha and Split-Half analysis) and construct validity (using Factor Analysis) of the Italian R2 IDE.

Country Specific Addendum: The Development, Validity and Reliability of the Japanese Version of the Insights Discovery Evaluator

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the Japanese version of the Insights Discovery Evaluator (IDE). It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability
    - Cronbach Alpha coefficients
    - Split-half correlations
- Validity - Factor Analysis

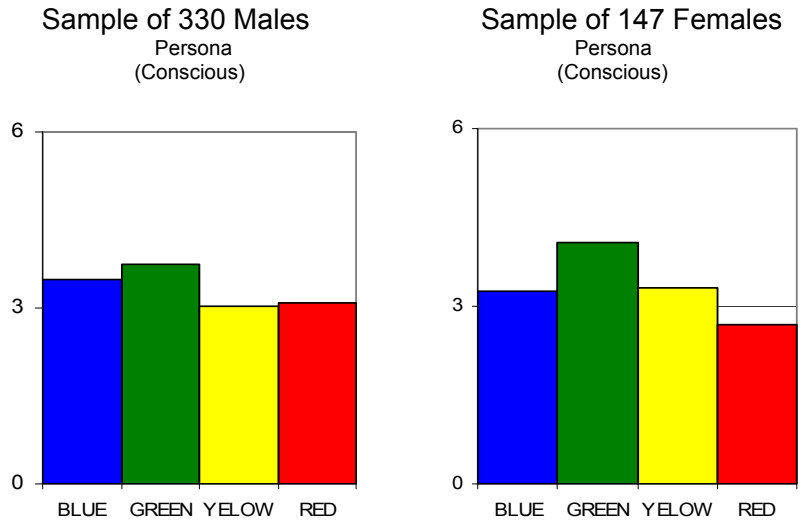
Data on Norms

Presented in Figure 1, is an example specific to the Japanese IDE, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and Japanese organisations. This sample is not intended to be representative of Japanese people overall; it is however, a very useful overview of the Japanese clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[58]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[59].



3.47	3.74	3.02	3.08
58%	62%	50%	51%

3.26	4.08	3.30	2.69
54%	68%	55%	45%

Figure 1 - A graphical view for the norms of Males vs. Females for the Japanese IDE

Reliability: Cronbach-Alpha Coefficients

[58] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[59]Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

This 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 inferior limit[60]. Analysing 477 completed Japanese IDE reported in Table 1 shows the four colours to have high Cronbach-Alpha coefficients between 0.85 and 0.91, providing evidence of excellent reliability.

N=477 Japanese IDE	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	0.90	0.85	0.90	0.91

Table 1 – Cronbach-Alpha coefficients

Reliability: Split-Half Coefficients

The final measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly 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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The analysis of the Japanese IDE reported in Table 2 shows reliable coefficients for the IDE, with coefficients being:

- Cronbach-Alpha Coefficients between 0.73 and 0.85 for each half
- Pearson Correlation Coefficients between 0.72 and 0.83 i.e. the 2 halves correlate highly

	N=477 Japanese IDE	Colour preference			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Part 1	Cronbach-Alpha coefficient	0.81	0.73	0.85	0.84
	N of Items	13	13	13	13
Part 2	Cronbach-Alpha coefficient	0.83	0.73	0.83	0.80
	N of Items	12	12	12	12
	Pearson Correlation coefficient between halves	0.78	0.72	0.83	0.80
	Tot. N of Items	25	25	25	25

Table 2 – Split-Half coefficients

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

- The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent
- The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the Japanese IDE.

A four factors solution accounts for 38% of the variance, while a two factors solution accounts for 30% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[61]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 3 below.

The four factor solution highlights the presence of the four constructs (translating into the four colour preferences), while the two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

	Four factor solution				Two factor solution	
Japanese IDE N=477	Average Factor Loadings				Average Factor Loadings	
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2
Earth Green	-0.06	-0.27	-0.03	0.33	-0.01	-0.41
Sunshine Yellow	0.49	-0.02	-0.27	0.00	0.54	-0.01
Cool Blue	-0.25	-0.08	0.45	-0.02	-0.48	-0.07
Fiery Red	0.00	0.50	-0.06	-0.16	0.05	0.50

Table 3 – Factor Analysis

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

While Table 3 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the Japanese IDE most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

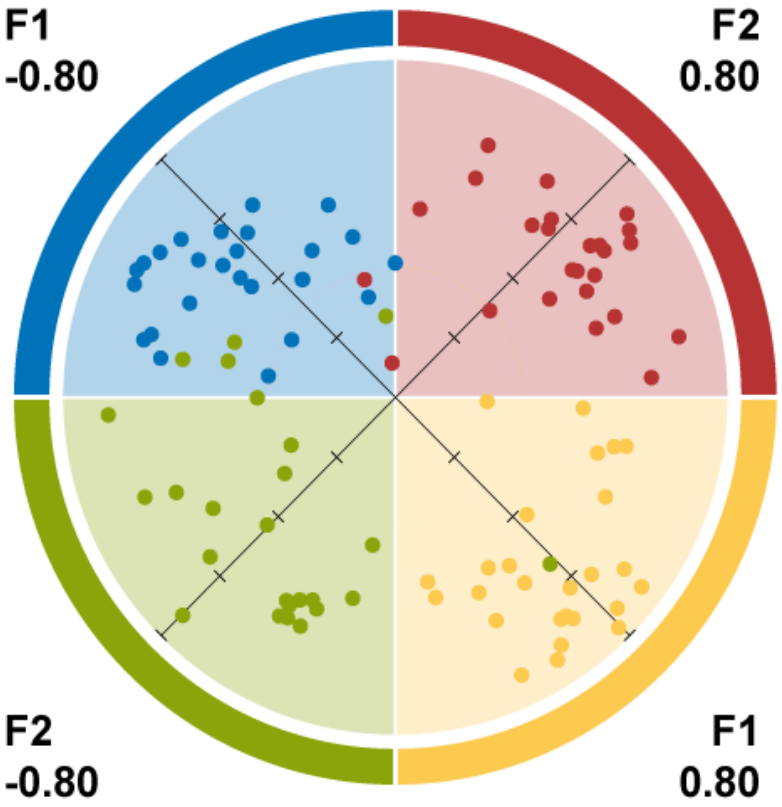


Figure 2 – Japanese IDE – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Cronbach Alpha and Split-Half analysis) and construct validity (using Factor Analysis) of the Japanese IDE.

Country Specific Addendum: The Development, Validity and Reliability of the Norwegian (Bokmal) Version R22 of the Insights Discovery Evaluator

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the Norwegian (Bokmal) R22 version of the Insights Discovery Evaluator (IDE). It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability
    - Inter-Item correlations
    - Cronbach Alpha coefficients
    - Split-half correlations
- Validity - Factor Analysis

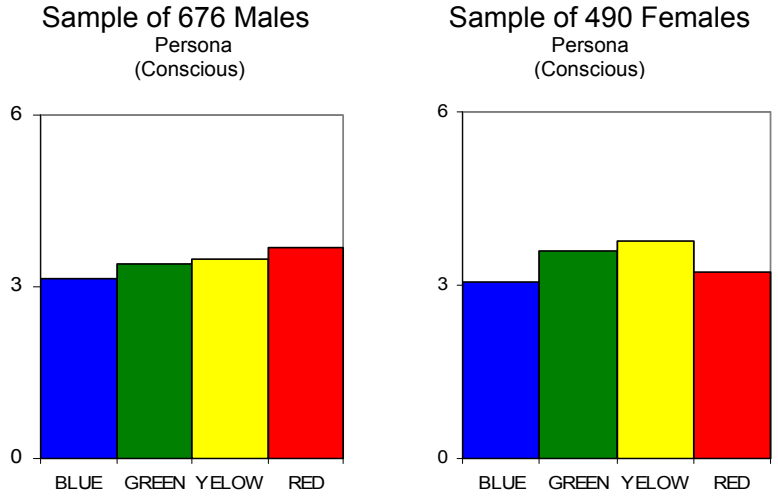
Data on Norms

Presented in Figure 1, is an example specific to the Norwegian (Bokmal) R22 IDE, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and Norwegian organisations. This sample is not intended to be representative of Norwegian people overall; it is however, a very useful overview of the Norwegian clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[62]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[63].



3.15	3.40	3.49	3.67
52%	57%	58%	61%

3.06	3.60	3.78	3.23
51%	60%	63%	54%

Figure 1 - A graphical view for the norms of Males vs. Females for the Norwegian (Bokmal) R22 IDE

Reliability: 'Inter-item' Correlations

[62] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[63]Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

'Inter-item' correlation coefficients have been calculated using the Pearson Product-Moment Correlation. This involved creating four colour based '25 by 25' matrices showing the correlation between the 25 colour items. In 1991, Robinson et al.[64] concluded that the mean 'inter-item' correlation coefficient should equal or exceed 0.30.

The analysis of 1'166 evaluators performed on the Norwegian (Bokmal) R22 IDE is reported in Table 1 and shows that for the four colours in the evaluator, the average 'inter-item' correlation coefficient is equal or above 0.30. This provides evidence of the case for reliability.

N=1'166 Norwegian R22 IDE	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Mean	**0.30**	**0.30**	**0.30**	**0.36**
Minimum	0.03	0.10	0.03	0.14
Maximum	0.54	0.55	0.60	0.72

Table 1 – Inter-item correlations

Reliability: Cronbach-Alpha Coefficients

This 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 inferior limit[65]. Analysing 1'166 completed Norwegian (Bokmal) R22 IDE reported in Table 2 shows the four colours to have high Cronbach-Alpha coefficients between 0.88 and 0.91, providing evidence of excellent reliability.

N=1'166 Norwegian R22 IDE	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	**0.91**	**0.91**	**0.91**	**0.93**

Table 2 – Cronbach-Alpha coefficients

Reliability: Split-Half Coefficients

[64] 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
Robinson, J.P., Shaver, P.R. (1973) 'Measure of Psychological Attitudes' MI: Survey Research Centre Institute for Social Research, University of Michigan
DeVellis, R. F. (1991) 'Scale Development: Theory and Applications', Sage Publications, Newbury Park, CA
Swailes, S., & McIntyre-Bhatty, T. (2002) 'The "Belbin" team role inventory: reinterpreting reliability estimates', Journal of Managerial Psychology, 17, 6, 529 – 536

The final measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly 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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The analysis of the Norwegian (Bokmal) R22 IDE reported in Table 3 shows high coefficients for the IDE, with coefficients being:

- Cronbach-Alpha Coefficients between 0.82 and 0.87 for each half
- Pearson Correlation Coefficients between 0.77 and 0.83 i.e. the 2 halves correlate highly

	N=1'166 Norwegian R22 IDE	Colour preference			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Part 1	Cronbach-Alpha coefficient N of Items	0.83 13	0.84 13	0.82 13	0.89 13
Part 2	Cronbach-Alpha coefficient N of Items	0.86 12	0.85 12	0.87 12	0.86 12
	Pearson Correlation coefficient between halves	0.80	0.80	0.77	0.83
	Tot. N of Items	25	25	25	25

Table 3 – Split-Half coefficients

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

- The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent
- The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the Norwegian (Bokmal) R22 IDE.

A four factors solution accounts for 34% of the variance, while a two factors solution accounts for 27% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[66]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 4 below.

The four factor solution highlights the presence of the four constructs (translating into the four colour preferences), while the two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

	Four factor solution				Two factor solution	
Norwegian R22 IDE	Average Factor Loadings				Average Factor Loadings	
N=1'166	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2
Earth Green	-0.30	-0.02	0.47	-0.06	0.52	-0.05
Sunshine Yellow	0.02	-0.28	-0.10	0.47	-0.04	0.53
Cool Blue	-0.14	0.48	0.04	-0.24	0.13	-0.52
Fiery Red	0.50	-0.17	-0.30	-0.05	-0.58	0.11

Table 4 – Factor Analysis

While Table 4 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the Norwegian (Bokmal) R22 IDE most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

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

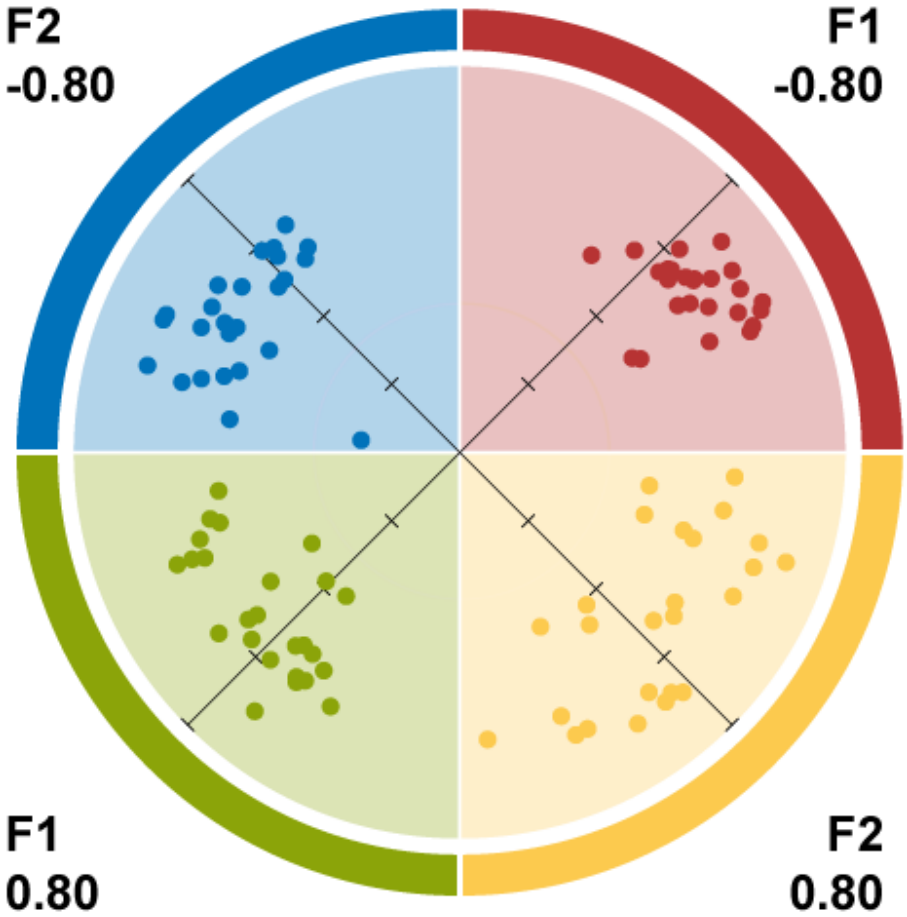


Figure 2 – Norwegian (Bokmal) R22 IDE – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Inter-Item correlations, Cronbach Alpha and Split-Half analysis) and construct validity (using Factor Analysis) of the Norwegian (Bokmal) R22 IDE.

Country Specific Addendum: The Development, Validity and Reliability of the Polish Version of the Insights Discovery Evaluator

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the Polish version of the Insights Discovery Evaluator (IDE). It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability
    - Inter-Item correlations
    - Cronbach Alpha coefficients
    - Split-half correlations
- Validity - Factor Analysis

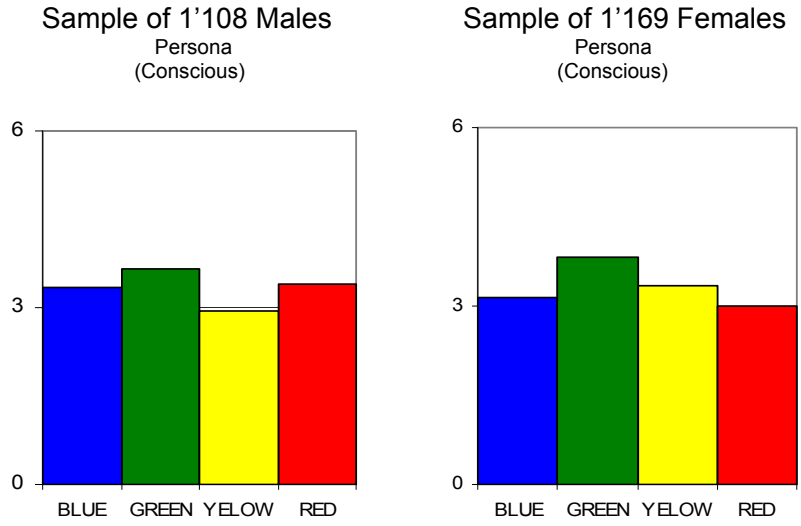
Data on Norms

Presented in Figure 1, is an example specific to the Polish IDE, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and Polish organisations. This sample is not intended to be representative of Polish people overall; it is however, a very useful overview of the Polish clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[67]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[68].



3.35	3.66	2.95	3.41
56%	61%	49%	57%

3.13	3.83	3.34	2.99
52%	64%	56%	50%

Figure 1 - A graphical view for the norms of Males vs. Females for the Polish IDE

Reliability: 'Inter-item' Correlations

'Inter-item' correlation coefficients have been calculated using the Pearson Product-Moment Correlation. This involved creating four colour based '25 by 25' matrices showing the correlation between the 25

[67] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[68] Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

colour items. In 1991, Robinson et al.[69] concluded that the mean 'inter-item' correlation coefficient should equal or exceed 0.30.

The analysis of 2'277 evaluators performed on the Polish IDE is reported in Table 1 and shows that for the four colours in the evaluator, the average 'inter-item' correlation coefficient is equal or above 0.30. This provides evidence of the case for reliability.

N=2'277 Polish IDE	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Mean	**0.34**	**0.30**	**0.36**	**0.32**
Minimum	0.07	0.07	0.06	0.01
Maximum	0.55	0.54	0.63	0.55

Table 1 – Inter-item correlations

Reliability: Cronbach-Alpha Coefficients

This 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 inferior limit[70]. Analysing 2'277 completed Polish IDE reported in Table 2 shows the four colours to have high Cronbach-Alpha coefficients between 0.91 and 0.93, providing evidence of excellent reliability.

N=2'277 Polish IDE	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	**0.93**	**0.91**	**0.91**	**0.93**

Table 2 – Cronbach-Alpha coefficients

Reliability: Split-Half Coefficients

The final measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly divide all items that are thought to measure the same construct into

[69] 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
Robinson, J.P., Shaver, P.R. (1973) 'Measure of Psychological Attitudes' MI: Survey Research Centre Institute for Social Research, University of Michigan
DeVellis, R. F. (1991) 'Scale Development: Theory and Applications', Sage Publications, Newbury Park, CA
Swailes, S., & McIntyre-Bhatty, T. (2002) 'The "Belbin" team role inventory: reinterpreting reliability estimates', Journal of Managerial Psychology, 17, 6, 529 – 536

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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The analysis of the Polish IDE reported in Table 3 shows high coefficients for the IDE, with coefficients being:

- Cronbach-Alpha Coefficients between 0.83 and 0.88 for each half
- Pearson Correlation Coefficients between 0.81 and 0.86 i.e. the 2 halves correlate highly

	N=2'277 Polish IDE	Colour preference			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Part 1	Cronbach-Alpha coefficient N of Items	0.86 13	0.82 13	0.88 13	0.88 13
Part 2	Cronbach-Alpha coefficient N of Items	0.87 12	0.84 12	0.87 12	0.83 12
	Pearson Correlation coefficient between halves	0.86	0.81	0.85	0.85
	Tot. N of Items	25	25	25	25

Table 3 – Split-Half coefficients

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

- The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent
- The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the Polish IDE.

A four factors solution accounts for 41% of the variance, while a two factors solution accounts for 34% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[71]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 4 below.

The four factor solution highlights the presence of the four constructs (translating into the four colour preferences), while the two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

	Four factor solution				Two factor solution	
Polish IDE N=2'277	Average Factor Loadings				Average Factor Loadings	
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2
Earth Green	-0.10	0.51	-0.13	-0.06	-0.06	-0.50
Sunshine Yellow	0.60	-0.02	0.05	-0.10	0.59	0.01
Cool Blue	-0.45	0.01	-0.08	0.38	-0.56	-0.06
Fiery Red	-0.04	-0.43	0.35	-0.17	0.04	0.56

Table 4 – Factor Analysis

While Table 4 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the Polish IDE most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

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

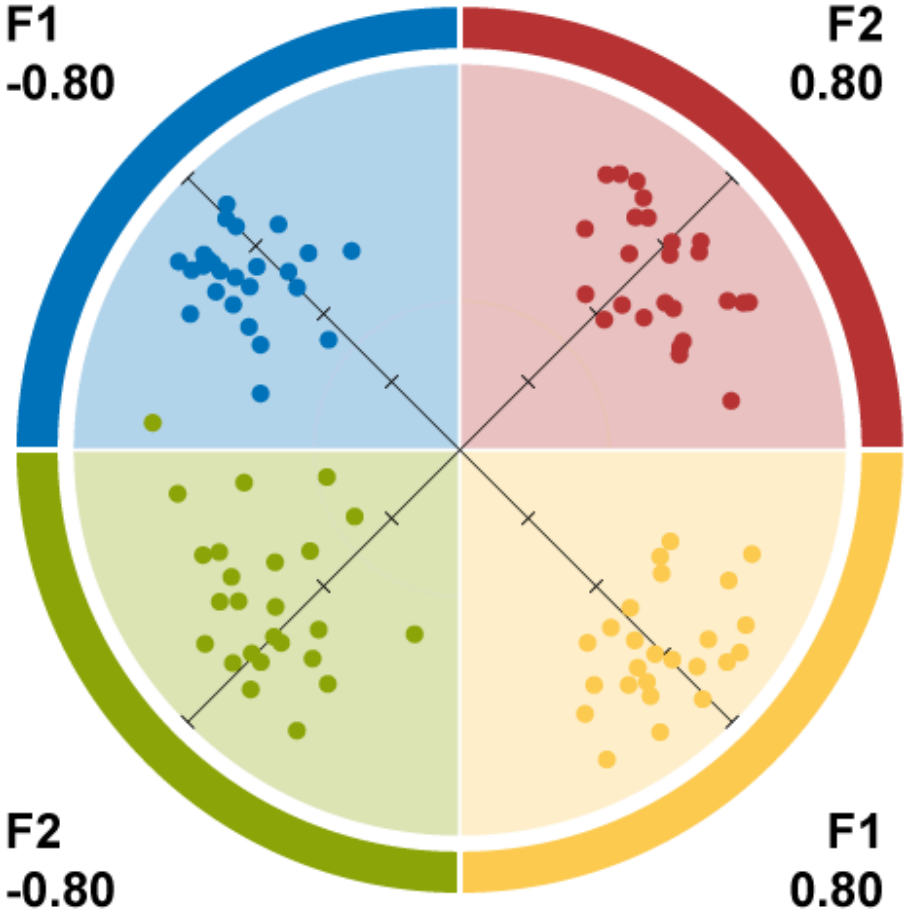


Figure 2 – Polish IDE – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Inter-Item correlations, Cronbach Alpha and Split-Half analysis) and construct validity (using Factor Analysis) of the Polish IDE.

Country Specific Addendum: The Development, Validity and Reliability of the Portuguese 'Brazilian' Version POR2 of the Insights Discovery Evaluator

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the Portuguese 'Brazilian' POR2 version of the Insights Discovery Evaluator (IDE). It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability - Cronbach Alpha coefficients
- Validity - Factor Analysis

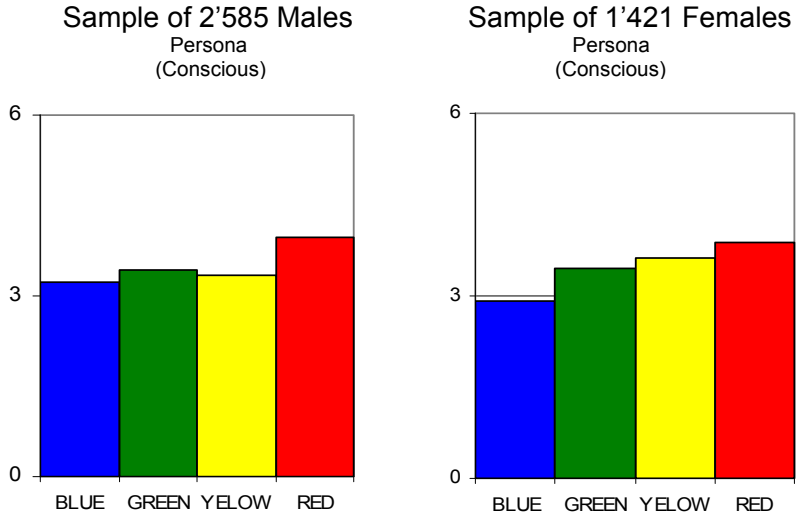
Data on Norms

Presented in Figure 1, is an example specific to the Portuguese 'Brazilian' POR2 IDE, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and Brazilian organisations. This sample is not intended to be representative of Brazilian people overall; it is however, a very useful overview of the Brazilian clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[72]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[73].



3.22	3.42	3.33	3.97
54%	57%	55%	66%

2.90	3.46	3.61	3.87
48%	58%	60%	65%

Figure 1 - A graphical view for the norms of Males vs. Females for the Portuguese 'Brazilian' POR2 IDE

Reliability: Cronbach-Alpha Coefficients

[72] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[73]Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

This 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 inferior limit[74]. Analysing 4'011 completed Portuguese 'Brazilian' POR2 IDE reported in Table 1 shows the four colours to have high Cronbach-Alpha coefficients between 0.75 and 0.88, providing evidence of good reliability.

N=4'011 Portuguese 'Brazilian' POR2 IDE	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	0.83	0.75	0.88	0.87

Table 1 – Cronbach-Alpha coefficients

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

- The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent
- The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the Portuguese 'Brazilian' POR2 IDE.

A two factors solution accounts for 22% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[75]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 2 below.

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

The two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

	Two factor solution	
Portuguese 'Brazilian' POR2	Average Factor Loadings	
N=4'011	Factor 1	Factor 2
Earth Green	-0.04	-0.29
Sunshine Yellow	0.47	-0.04
Cool Blue	-0.39	-0.07
Fiery Red	-0.02	0.47

Table 2 – Factor Analysis

While Table 2 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the Portuguese 'Brazilian' IDE most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

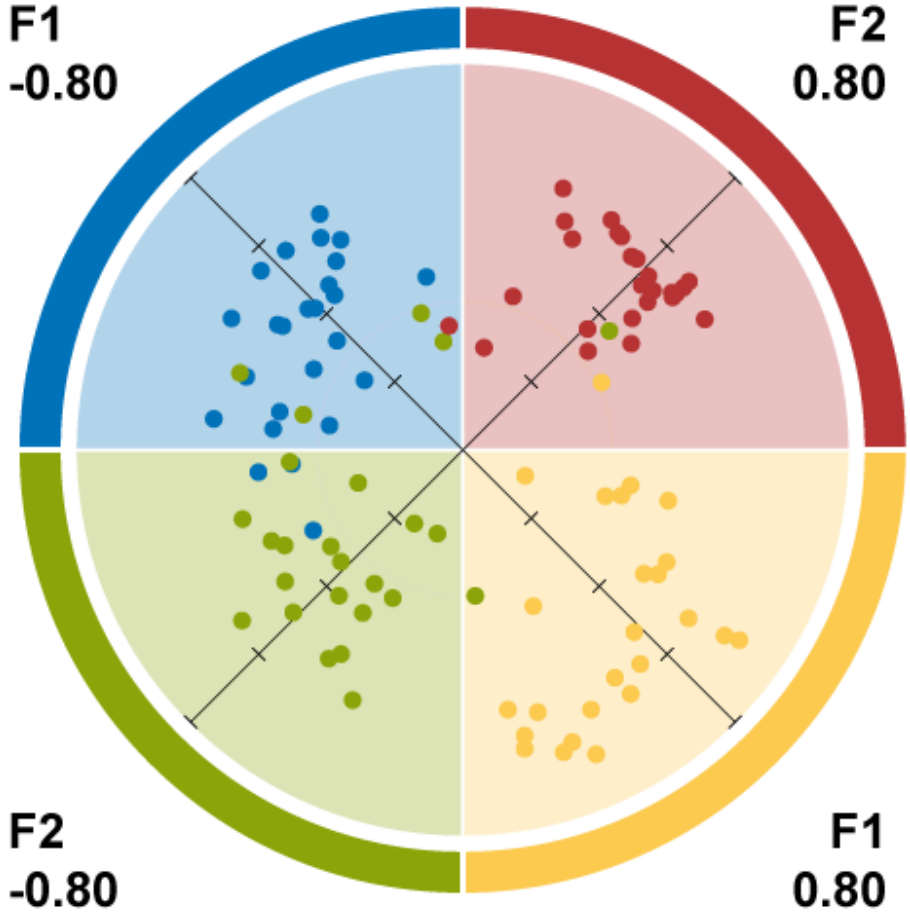


Figure 2 – Portuguese 'Brazilian' POR2 IDE – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Cronbach Alpha analysis) and construct validity (using Factor Analysis) of the Portuguese 'Brazilian' POR2 IDE.

Country Specific Addendum: The Development, Validity and Reliability of the Portuguese Version PIR2 of the Insights Discovery Evaluator

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the Portuguese PIR2 version of the Insights Discovery Evaluator (IDE). It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability
    - Split-half correlations
    - Cronbach Alpha coefficients
- Validity - Factor Analysis

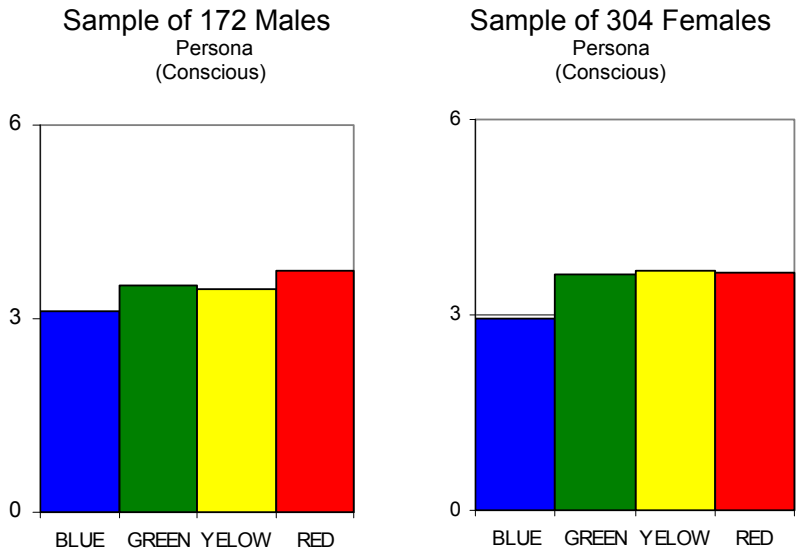
Data on Norms

Presented in Figure 1, is an example specific to the Portuguese PIR2 IDE, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and Portuguese organisations. This sample is not intended to be representative of Portuguese people overall; it is however, a very useful overview of the Portuguese clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[76]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[77].



3.12	3.53	3.47	3.76
52%	59%	58%	63%

2.94	3.63	3.67	3.65
49%	61%	61%	61%

Figure 1 - A graphical view for the norms of Males vs. Females for the Portuguese PIR2 IDE

Reliability: Cronbach-Alpha Coefficients

This 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

[76] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[77] Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

commonly accepted inferior limit[78]. Analysing 476 completed Portuguese PIR2 IDE reported in Table 1 shows the four colours to have high Cronbach-Alpha coefficients between 0.84 and 0.91, providing evidence of excellent reliability.

N=476 Portuguese PIR IDE	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	0.86	0.84	0.89	0.91

Table 1 – Cronbach-Alpha coefficients

Reliability: Split-Half Coefficients

The final measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly 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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The analysis of the Portuguese IDE reported in Table 2 shows reliable coefficients for the IDE, with coefficients being:

- Cronbach-Alpha Coefficients between 0.72 and 0.86 for each half
- Pearson Correlation Coefficients between 0.71 and 0.81 i.e. the 2 halves correlate highly

	N=476 Portuguese PIR2 IDE	Colour preference			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Part 1	Cronbach-Alpha coefficient	0.72	0.73	0.79	0.80
	N of Items	13	13	13	13
Part 2	Cronbach-Alpha coefficient	0.80	0.72	0.86	0.81
	N of Items	12	12	12	12
	Pearson Correlation coefficient between halves	0.75	0.71	0.81	0.74
	Tot. N of Items	25	25	25	25

Table 2 – Split-Half coefficients

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

- The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent
- The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the Portuguese PIR2 IDE.

A four factors solution accounts for 32% of the variance, while a two factors solution accounts for 25% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[79]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 3 below.

The four factor solution highlights the presence of the four constructs (translating into the four colour preferences), while the two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

	Four factor solution				Two factor solution	
Portuguese PIR2 IDE	Average Factor Loadings				Average Factor Loadings	
N=476	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2
Earth Green	-0.26	-0.02	-0.01	0.35	-0.04	-0.39
Sunshine Yellow	-0.05	0.47	-0.25	-0.08	0.53	-0.02
Cool Blue	-0.03	-0.28	0.36	0.01	-0.44	-0.03
Fiery Red	0.48	-0.04	-0.05	-0.11	-0.01	0.48

Table 3 – Factor Analysis

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

While Table 3 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the Portuguese IDE most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

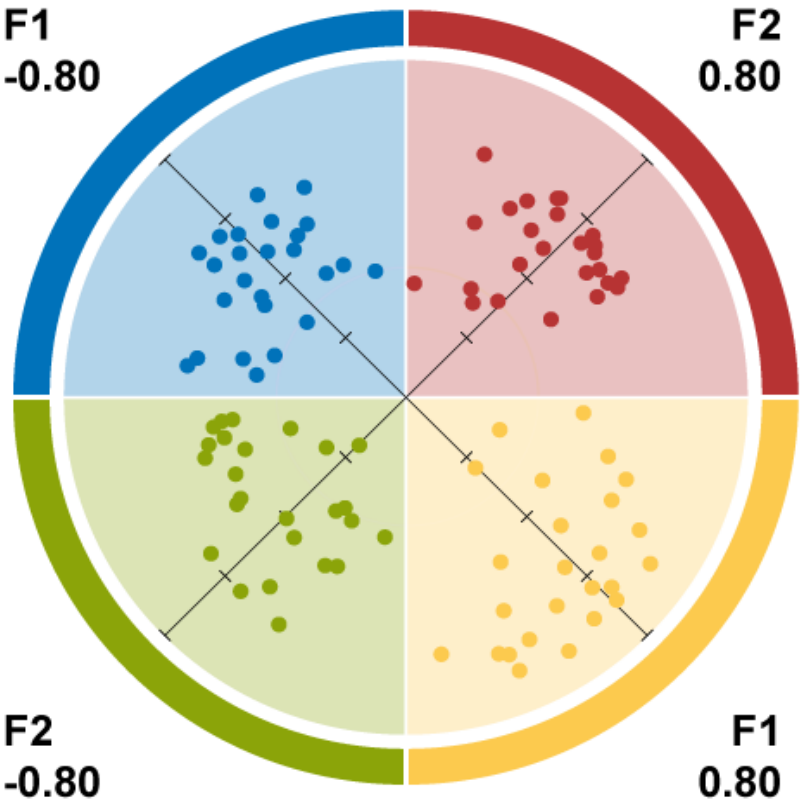


Figure 2 – Portuguese PIR2 IDE – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Cronbach Alpha and Split-Half analysis) and construct validity (using Factor Analysis) of the Portuguese PIR2 IDE.

Country Specific Addendum: The Development, Validity and Reliability of the Spanish 'Mexican' Version S1.0 of the Insights Discovery Evaluator

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the Spanish 'Mexican' version of the Insights Discovery Evaluator (IDE). It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability
    - Split-half correlations
    - Cronbach Alpha coefficients
- Validity - Factor Analysis

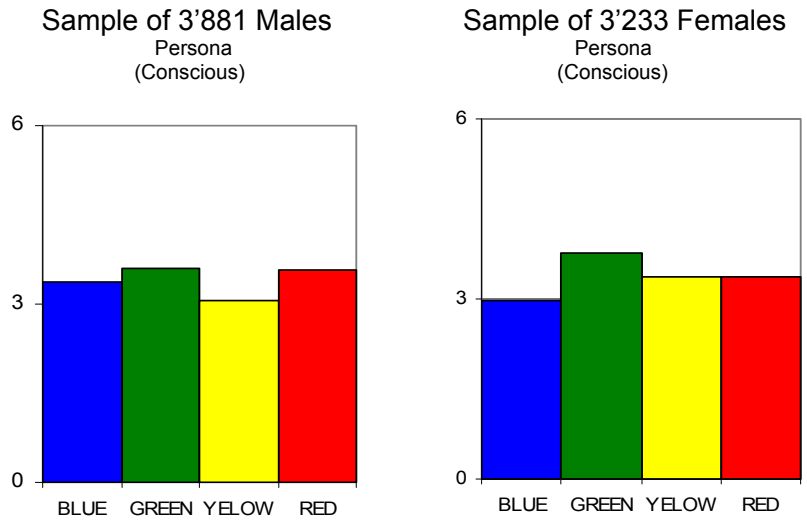
Data on Norms

Presented in Figure 1, is an example specific to the Spanish 'Mexican' IDE, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and Mexican organisations. This sample is not intended to be representative of Mexican people overall; it is however, a very useful overview of the Mexican clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[80]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[81].



3.40	3.61	3.33	3.50
57%	60%	56%	58%

3.15	3.73	3.56	3.31
52%	62%	59%	55%

Figure 1 - A graphical view for the norms of Males vs. Females for the Spanish 'Mexican' S1.0 IDE

Reliability: Cronbach-Alpha Coefficients

[80] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[81] Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

This 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 inferior limit[82]. Analysing 7'114 completed Spanish 'Mexican' S1.0 IDE reported in Table 1 shows the four colours to have high Cronbach-Alpha coefficients between 0.87 and 0.89, providing evidence of excellent reliability.

N=7'114 Spanish (Mexico) S1.0 IDE	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	0.89	0.87	0.89	0.88

Table 1 – Cronbach-Alpha coefficients

Reliability: Split-Half Coefficients

The final measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly 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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The analysis of the Spanish 'Mexican' IDE reported in Table 2 shows reliable coefficients for the IDE, with coefficients being:

- Cronbach-Alpha Coefficients between 0.76 and 0.82 for each half
- Pearson Correlation Coefficients between 0.74 and 0.81 i.e. the 2 halves correlate highly

	N=7'114 Spanish (Mexico) S1.0 IDE	Colour preference			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Part 1	Cronbach-Alpha coefficient	0.79	0.76	0.80	0.81
	N of Items	13	13	13	13
Part 2	Cronbach-Alpha coefficient	0.81	0.81	0.82	0.76
	N of Items	12	12	12	12
	Pearson Correlation coefficient between halves	0.80	0.74	0.81	0.76
	Tot. N of Items	25	25	25	25

Table 2 – Split-Half coefficients

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

- The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent
- The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the Spanish 'Mexican' S1.0 IDE.

A four factors solution accounts for 32% of the variance, while a two factors solution accounts for 25% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[83]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 3 below.

The four factor solution highlights the presence of the four constructs (translating into the four colour preferences), while the two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

Spanish (Mexico) S1.0 IDE N=7'114	Four factor solution Average Factor Loadings				Two factor solution Average Factor Loadings	
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2
Earth Green	0.44	-0.08	-0.03	-0.18	-0.02	-0.46
Sunshine Yellow	-0.05	-0.24	0.45	-0.05	-0.49	0.00
Cool Blue	-0.03	0.47	-0.20	0.03	0.47	0.04
Fiery Red	-0.28	-0.05	-0.10	0.39	0.05	0.46

Table 3 – Factor Analysis

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

While Table 3 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the Spanish 'Mexican' IDE most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

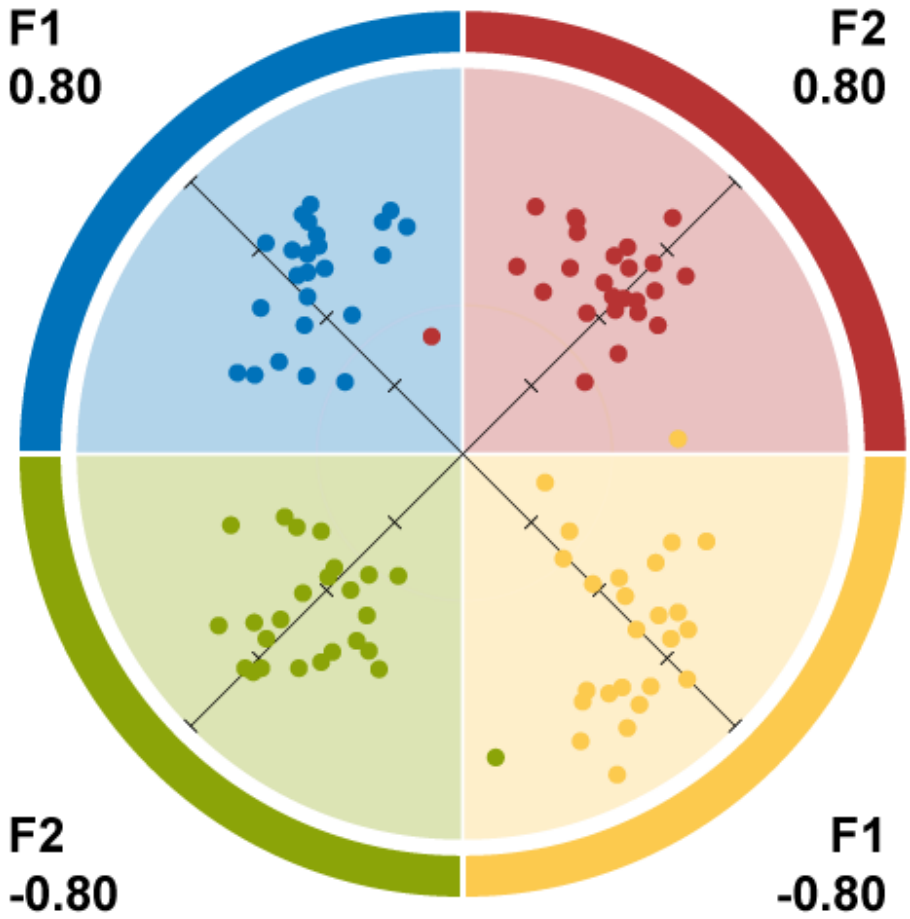


Figure 2 – Spanish 'Mexican' S1.0 IDE – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Cronbach Alpha and Split-Half analysis) and construct validity (using Factor Analysis) of the Spanish 'Mexican' S1.0 IDE.

Country Specific Addendum: The Development, Validity and Reliability of the Spanish Version S1.3 of the Insights Discovery Evaluator

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the Spanish S1.3 version of the Insights Discovery Evaluator (IDE). It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability
    - Split-half correlations
    - Cronbach Alpha coefficients
- Validity - Factor Analysis

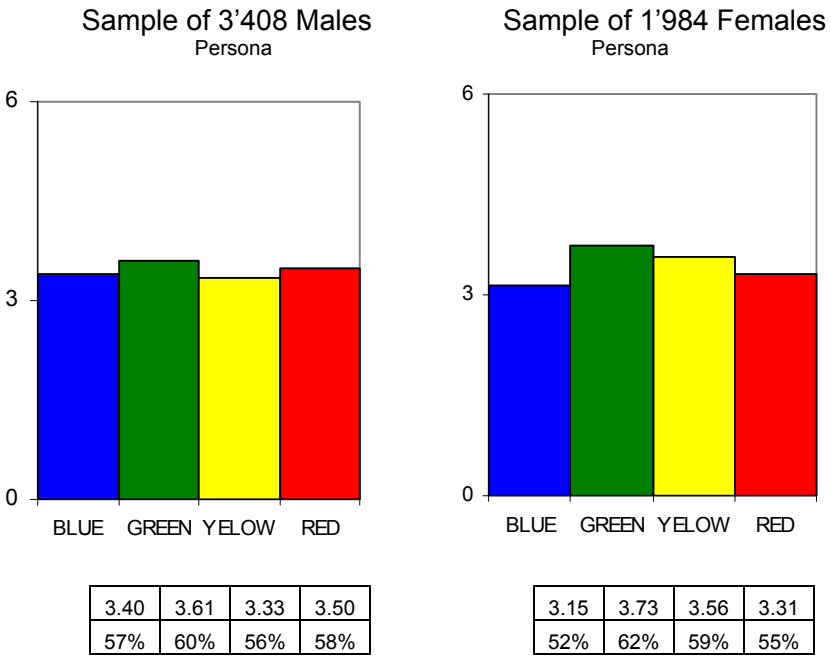
Data on Norms

Presented in Figure 1, is an example specific to the Spanish S1.3 IDE, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and Spanish organisations. This sample is not intended to be representative of Spanish people overall; it is however, a very useful overview of the Spanish clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[84]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[85].



3.40	3.61	3.33	3.50
57%	60%	56%	58%

3.15	3.73	3.56	3.31
52%	62%	59%	55%

Figure 1 - A graphical view for the norms of Males vs. Females for the Spanish S1.3 IDE

Reliability: Cronbach-Alpha Coefficients

This 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

[84] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[85] Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

commonly accepted inferior limit[86]. Analysing 5'392 completed Spanish S1.3 IDE reported in Table 1 shows the four colours to have high Cronbach-Alpha coefficients between 0.88 and 0.91, providing evidence of excellent reliability.

N=5'392 Spanish S1.3 IDE	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	0.91	0.88	0.91	0.91

Table 1 – Cronbach-Alpha coefficients

Reliability: Split-Half Coefficients

The final measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly 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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The analysis of the Spanish IDE reported in Table 2 shows high coefficients for the IDE, with coefficients being:

- Cronbach-Alpha Coefficients between 0.77 and 0.85 for each half
- Pearson Correlation Coefficients between 0.77 and 0.82 i.e. the 2 halves correlate highly

	N=5'392 Spain S1.3 IDE	Colour preference			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Part 1	Cronbach-Alpha coefficient	0.80	0.77	0.82	0.85
	N of Items	13	13	13	13
Part 2	Cronbach-Alpha coefficient	0.83	0.82	0.83	0.81
	N of Items	12	12	12	12
	Pearson Correlation coefficient between halves	0.81	0.77	0.82	0.81
	Tot. N of Items	25	25	25	25

Table 2 – Split-Half coefficients

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

- The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent
- The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the Spanish S1.3 IDE.

A four factors solution accounts for 34% of the variance, while a two factors solution accounts for 27% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[87]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 3 below.

The four factor solution highlights the presence of the four constructs (translating into the four colour preferences), while the two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

Spain S1.3 IDE N=5'392	Four factor solution Average Factor Loadings				Two factor solution Average Factor Loadings	
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2
Earth Green	0.45	-0.05	-0.03	-0.20	-0.02	0.46
Sunshine Yellow	-0.04	0.49	-0.22	-0.06	0.52	0.01
Cool Blue	0.02	-0.27	0.45	0.00	-0.49	0.05
Fiery Red	-0.33	-0.06	-0.14	0.40	0.02	-0.52

Table 3 – Factor Analysis

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

While Table 3 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the Spanish IDE most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

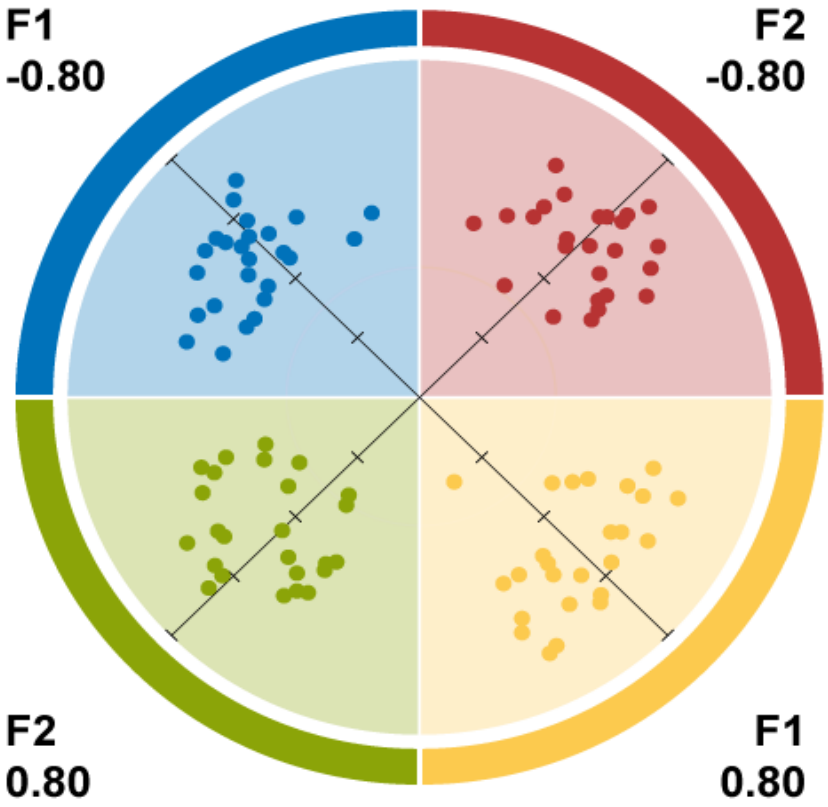


Figure 2 – Spanish S1.3 IDE – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Cronbach Alpha and Split-Half analysis) and construct validity (using Factor Analysis) of the Spanish S1.3 IDE.

Country Specific Addendum: The Development, Validity and Reliability of the Swedish Version R 2 (Sverige) Beta of the Insights Discovery Evaluator

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the Danish version of the Insights Discovery Evaluator (IDE). It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability
    - Cronbach Alpha coefficients
    - Split-half correlations
- Validity - Factor Analysis

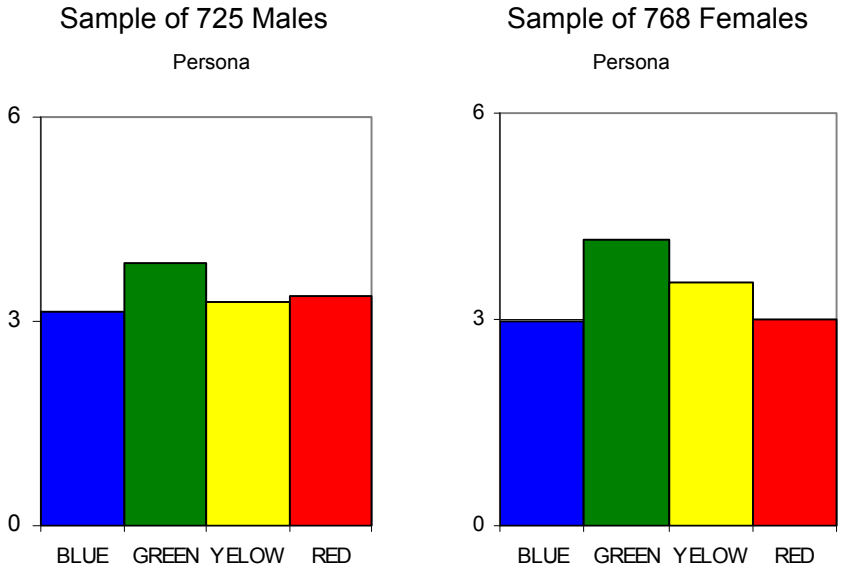
Data on Norms

Presented in Figure 1, is an example specific to the Danish IDE, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and Danish organisations. This sample is not intended to be representative of Swedish people overall; it is however, a very useful overview of the Swedish clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[88]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[89].



3.15	3.86	3.28	3.37
53%	64%	55%	56%

2.98	4.16	3.53	3.01
50%	69%	59%	50%

Figure 1 - A graphical view for the norms of Males vs. Females for the Swedish IDE

Reliability: Cronbach-Alpha Coefficients

This 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

[88] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[89] Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

commonly accepted inferior limit[90]. Analysing 1'493 completed Swedish IDE reported in Table 1 shows the four colours to have high Cronbach-Alpha coefficients between 0.90 and 0.91, providing evidence of excellent reliability.

N=1'493 Swedish IDE	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	0.91	0.90	0.90	0.90

Table 1 – Cronbach-Alpha coefficients

Reliability: Split-Half Coefficients

The final measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly 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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The analysis of the Swedish IDE reported in Table 2 shows very reliable coefficients for the IDE, with coefficients being:

- Cronbach-Alpha Coefficients between 0.81 and 0.85 for each half
- Pearson Correlation Coefficients between 0.78 and 0.83 i.e. the 2 halves correlate highly

	N=1'493 Swedish IDE	Colour preference			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Part 1	Cronbach-Alpha coefficient	0.84	0.81	0.83	0.83
	N of Items	13	13	13	13
Part 2	Cronbach-Alpha coefficient	0.85	0.84	0.82	0.81
	N of Items	12	12	12	12
	Pearson Correlation coefficient between halves	0.83	0.78	0.82	0.83
	Tot. N of Items	25	25	25	25

Table 2 – Split-Half coefficients

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

- The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent
- The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the Swedish IDE.

A four factors solution accounts for 37% of the variance, while a two factors solution accounts for 30% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[91]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 3 below.

The four factor solution highlights the presence of the four constructs (translating into the four colour preferences), while the two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

	Four factor solution				Two factor solution	
N=1'493	Average Factor Loadings				Average Factor Loadings	
Swedish IDE	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2
Earth Green	-0.03	0.51	-0.16	0.01	-0.01	0.51
Sunshine Yellow	0.53	-0.06	0.08	-0.04	0.50	-0.07
Cool Blue	-0.45	-0.01	-0.12	0.31	-0.54	0.07
Fiery Red	0.04	-0.35	0.32	-0.22	0.12	-0.49

Table 3 – Factor Analysis

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

While Table 3 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the Swedish IDE most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

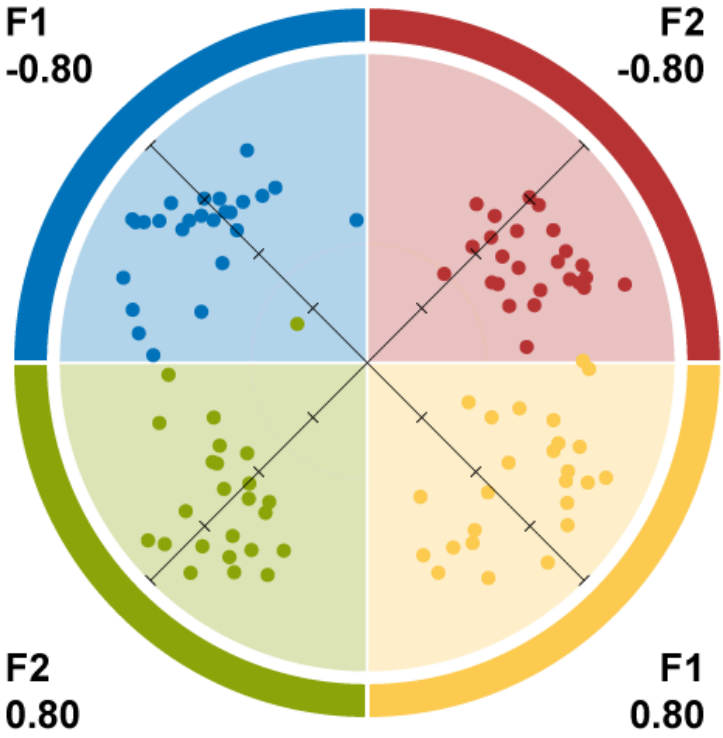


Figure 2 – Swedish IDE – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Cronbach Alpha and Split-Half analysis) and construct validity (using Factor Analysis) of the Swedish IDE.

Country Specific Addendum: The Development, Validity and Reliability of the Turkish Version R2 of the Insights Discovery Evaluator

Dr. Stephen Benton

Director of The Business Psychology Centre (bpc),

The University of Westminster, 309 Regent Street, London, W1B 2UW, UK

Dr. Corine van Erkom Schurink

Director of The Analytical Research Bureau (Pty) Ltd,

16 Central Avenue, Pinelands 7405, South Africa

Stewart Desson

Head of Learning, Insights Learning & Development Ltd,

Jack Martin Way, Dundee, DD4 9FF, Scotland, UK

This addendum presents key data on the psychometric properties of the Turkish R2 version of the Insights Discovery Evaluator (IDE). It draws upon an extensive research and development programme undertaken between Insights Learning & Development Ltd. and the University of Westminster, aimed at the development of a psychometrically robust evaluator. This paper assumes the reader has first read "An Overview of the Development, Validity and Reliability of the English Version 3.0 of the Insights Discovery Evaluator" produced at the University of Westminster's Business Psychology Centre (bpc).

This report presents information covering reliability and validity. Key statistics have been computed on:

- Norms data by gender
- Reliability
    - Split-half correlations
    - Cronbach Alpha coefficients
- Validity - Factor Analysis

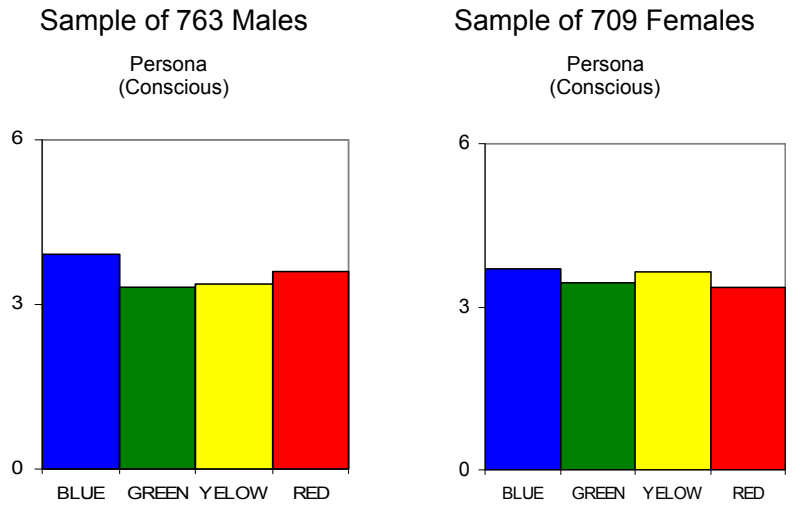
Data on Norms

Presented in Figure 1, is an example specific to the Turkish R2 IDE, providing norms data by gender.

The samples are all taken from the management communities of large multi-national and Turkish organisations. This sample is not intended to be representative of Turk people overall; it is however, a very useful overview of the Turkish clients Insights Learning and development work with.

Please note that while this data indicates that people of a certain gender tend to have a preference for certain colours, it does not correlate or necessarily relate to how well they do their job or how capable they are of fulfilling a particular role.

On average, females have lower scores for blue and red, and higher scores for green and yellow. This data is consistent with other instruments that suggest gender differences in personality. Hammer and Mitchell's research done on the distribution of MBTI types by ethnicity and gender in the US show that women indicated a higher preference for extraverted feeling (f:31.1%, m:16.2%) and introverted feeling (f:29.9%, m:15.2%). In a similar vein men showed a preference for extraverted thinking (f:15.9%, m:29.2%) and introverted thinking (f:22.9%, 39.4%)[92]. Rubinstein and Strul found similar gender differences in their study carried out with the NEO-PI-R, women had a significantly higher preference for agreeableness and neuroticism attributes that have both been linked to the Jungian attitudinal function of feeling[93].



3.91	3.32	3.37	3.60
65%	55%	56%	60%

3.68	3.44	3.65	3.36
61%	57%	61%	56%

Figure 1 - A graphical view for the norms of Males vs. Females for the Turkish R2 IDE

Reliability: Cronbach-Alpha Coefficients

This 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

[92] Hammer, A. L., Mitchell, W.D. (1996) *The Distribution of MBTI Types In the US by Gender and Ethnic Group*, Journal of Psychological Type, Vol.37, 2 – 15.

[93] Rubinstein, G., Strul, S. (2006) *The Five Factor Model (FFM) Among Four Groups of Male and Female Professionals*, Journal of Research in Personality, Vol. 41, 931-937

commonly accepted inferior limit[94]. Analysing 1,472 completed Turkish R2 IDE reported in Table 1 shows the four colours to have high Cronbach-Alpha coefficients between 0.86 and 0.89, providing evidence of excellent reliability.

N = 1,472 Turkish R2 IDE	Colour preference			
	Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Cronbach-Alpha Coefficients	0.86	0.86	0.89	0.87

Table 1 – Cronbach-Alpha coefficients

Reliability: Split-Half Coefficients

The final measure of internal consistency that supports the case for reliability is the 'split-half' measure. In split-half reliability we randomly 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 IDE were achieved by splitting the 25 frames into two groups of 12 and 13. 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. The analysis of the Turkish IDE reported in Table 2 shows good coefficients for the IDE, with coefficients being:

- Cronbach-Alpha Coefficients between 0.74 and 0.81 for each half
- Pearson Correlation Coefficients between 0.75 and 0.79 i.e. the 2 halves correlate highly

	N = 1'472 Turkish R 2 IDE	Colour preference			
		Cool Blue	Earth Green	Sunshine Yellow	Fiery Red
Part 1	Cronbach-Alpha coefficient N of Items	0.80 13	0.77 13	0.80 13	0.81 13
Part 2	Cronbach-Alpha coefficient N of Items	0.81 12	0.74 12	0.81 12	0.74 12
	Pearson Correlation coefficient between halves	0.79	0.75	0.79	0.76
	Tot. N of Items	25	25	25	25

Table 2 – Split-Half coefficients

Validity - Confirmatory Factor Analysis

Confirmatory Factor Analysis was used to test the hypothesized factor structure of the Insights Discovery model. Specifically, it is hypothesized that the four sets of 25 colour based items in the IDE, should load onto the factors such that:

- The polar opposite nature of the 'Fiery Red' vs. 'Earth Green' items is apparent
- The polar opposite nature of the 'Sunshine Yellow' vs. 'Cool Blue' items is 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 hypothesized structure and offer evidence for the construct validity of the Turkish R2 IDE.

A four factors solution accounts for 34% of the variance, while a two factors solution accounts for 26% of the variance. Generally factor loadings equal or greater than 0.30 are considered to meet the minimal level for significance required[95]. Using these criteria the statistically significant factor loadings have been highlighted with greyed background in the Table 3 below.

The four factor solution highlights the presence of the four constructs (translating into the four colour preferences), while the two factor solution highlights the bipolarity of those four constructs, as the opposite signs of these loadings (negative values highlighted with yellowed background) support the theoretical construct that 'Fiery Red' and 'Earth Green' are polar opposites. The same holds true for 'Cool Blue' and 'Sunshine Yellow'.

	Four factor solution				Two factor solution	
N=1,472	Average Factor Loadings				Average Factor Loadings	
Turkish R 2 IDE	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2
Earth Green	-0.12	-0.27	0.34	-0.11	-0.59	-0.01
Sunshine Yellow	0.48	0.10	0.01	-0.15	0.00	-0.52
Cool Blue	-0.26	-0.02	0.02	0.38	0.05	0.51
Fiery Red	-0.10	0.37	-0.28	-0.05	0.55	0.03

Table 3 – Factor Analysis

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

While Table 3 reports an average of the factor loadings, it is also possible to report the factor loadings for each of the 100 items in the IDE in the form of a scatter plot (both diagonal scales range from 0.8 to minus 0.8, with the axis crossing at zero). Based on the two factor solution Figure 2 shows the factor loadings onto all 100 items. The data has been superimposed onto the Insights Discovery Wheel. In the case of the Turkish IDE most of the 100 items appear in the 'correct' quadrant.

This item level data provides further evidence of the bi-polar nature of the colour scores and the construct validity of the model.

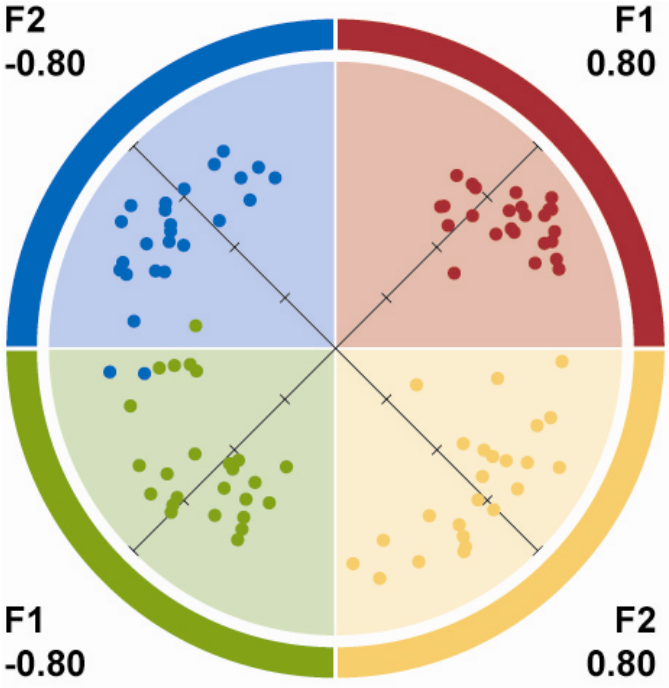


Figure 2 – Turkish S2.1 IDE – Graph of the 100 items (25 x 4 colours) plotted against the factors

Conclusion

In conjunction with the detailed analysis of the English version of the IDE, this brief paper offers good evidence of the internal reliability (using Cronbach Alpha and Split-Half analysis) and construct validity (using Factor Analysis) of the Turkish R2 IDE.

The Sport Psychologist, 2005, 19, 203-220

Communication Within Sport Teams: Jungian Preferences and Group Dynamics

Mark R. Beauchamp
University of Leeds, UK

Alan Maclachlan
Insights Learning and Development, Toronto, Canada

Andrew M. Lothian
Insights Learning and Development, Dundee, UK

Contemporary group dynamics theorists and practitioners consistently highlight the importance of effective communication in facilitating successful team functioning (cf. Carron & Hausenblas, 1998). In this review paper, we explore how an understanding of Jungian preferences (cf. Jung, 1921/1971a) can provide an important theory-driven framework for those concerned with group dynamics in sport. As a basis for improved interaction, this model suggests that in order to effectively "adapt and connect" with other team members, one must first develop an acute understanding of *self* as well as the patterns of preferences that characterize those with whom one interacts. In this paper, we discuss the theoretical structure of this model and explain how the model can inform group dynamics interventions in sport.

Groups are often highly complex entities, with each group member contributing a unique set of psychological attributes. When members are acutely aware of their own and other's individual differences, and can adapt their behaviors accordingly to meet the group's needs, it is likely that effective group functioning will follow (cf. Carron & Hausenblas, 1998; Forsyth, 1999). As Shaw (1981) so

Mark R. Beauchamp is with the School of Sport and Exercise Sciences at the University of Leeds, Leeds, LS2 9J2, United Kingdom. E-mail: m.r.beauchamp@leeds.ac.uk. Alan Maclachlan and Andrew M. Lothian are with Insights Learning and Development in Toronto, Canada and Dundee, UK, respectively.

eloquently put it, "personality characteristics of group members play an important role in determining their behavior in groups. The magnitude of the effect of any given characteristic is small but taken together the consequences for group processes are of major significance" (p. 208). In this paper, we present a framework for understanding personality that is grounded in Jung's (1921/1971a) theory of *Personality Types*. Within the context of sport teams, athletes typically spend much time together, through training, competition, and on social occasions. As it relates to sport, the basic premise of this model is that if members of sport teams can begin to understand their own preferences (as a reflection of their personalities) within these varied contexts, as well as the preferences of others with whom they interact, this may provide an effective basis for enhanced group functioning.

Despite a departure in recent years from the study of personality in sport (cf. Van den Auweele, Nys, Rzewnicki, & Van Mele, 2001), interest in personality research has continued to flourish in other domains of psychology, including human resources (Hogan, Hogan, & Roberts, 1996), project management (Barry & Stewart, 1997), work motivation (Judge & Ilies, 2002), leadership (Hogan, Curphy, & Hogan, 1994; Judge, Bono, Ilies, Gerhardt, 2002), organizational consulting (Camara, Nathan, & Puente, 2000), and medicine (Piedmont & Ciarrocchi, 1999; Shope, Frohna, & Frohna, 2000). The purpose of this review is not to lament the considerable conceptual, methodological, and interpretive limitations (Gill, 2000) that have beset personality research in sport over the past few decades. Substantive reviews on the subject have been provided elsewhere (e.g., Van den Auweele et al., 2001; Vealey, 2002). Rather, the purpose of this review is to present a conceptual model for understanding personality that has been employed with success in other fields of *performance psychology*, including organizational communication (Allen & Brock, 2000; Hirsh & Kummerow, 1990), decision making (Nutt, 1990), leadership (Atwater & Yammarino, 1993; Roush & Atwater, 1991), management (Gardner & Martinko, 1996; Walck, 1992a), as well as team building in business (McCaulley, 2000). Given the emphasis within this model on understanding *self* (cf. Rogers, 1951, 1977) and others as a means toward facilitating effective interpersonal communication, it would appear to have particular relevance for those concerned with enhancing group dynamics in sport.

In his seminal work, Carl Jung (1921/1971a) developed a typology for understanding human cognition, affect, and behavior. As an astute psychologist and psychiatrist, Jung recognized that each individual's personality is unique and it would be impossible to *reduce* people down to *pure types* that fit neatly within a simple classification. As he was clear to point out, "one can never give a description of a type, no matter how complete, that would apply to more than one individual, despite the fact that in some ways it aptly characterizes thousands of others. Conformity is the one side of man, uniqueness is the other." (Jung, 1923/1971b, p. 516). However, Jung (1936/1971d) also recognized that an understanding of typology may act as a compass for psychological orientation, remarking that

> It is not the purpose of a psychological typology to classify human beings into categories—this in itself would be pretty pointless. Its purpose is to provide a critical psychology which will make a methodological investigation and presentation of the empirical material possible. First and foremost it is a critical tool for the research worker, who needs definite points of view and guidelines

> if he is to reduce the chaotic profusion of individual experiences to any kind of order. . . . Secondly, a typology is a great help in understanding the wide variations that occur among individuals, and it also furnishes a clue to the fundamental difference in the psychological theories now present. Last but not least, it is an essential means for determining the "personal equation" for the practicing psychologist, who, armed with an exact knowledge of his differentiated and inferior functions, can avoid many serious blunders when dealing with his patients. (pp. 554 - 555)

In the past, a primary motive for studying personality in sport was to determine whether certain traits could predict performance, or even whether a generic personality profile existed for "the elite athlete" (cf. Morgan, 1979, 1980; Vanek & Cratty, 1970). Not surprisingly, subsequent research failed to provide any clear findings in this regard (cf. Bakker, Whiting, & Van der Brug, 1990; Vealey, 2002). We do not propose a return to this reductionist and outmoded approach to personality research in sport psychology. Instead, we suggest that if members of sport teams (through the intervention of a sport psychologist) can begin to understand themselves (i.e., their personalities) and others with whom they interact, this can enable them to more effectively adapt their communication behaviors to successfully "connect" with each other. This, in turn, can facilitate improved group functioning.

The Basic Model

Given that sport psychologists remain largely unaware of Jung's (1921/1971a) conceptual framework, a brief overview of this model is presented. Jung described one's personality as consisting of both conscious and unconscious structures. Although Jung recognized the influence of genetic or biological factors in shaping one's personality, he moved beyond Freud's (1900/1953) conception that behaviors are essentially *determined* by unconscious instinctual motives. Jung viewed one's personality as a self-regulating system where individuals are active agents in their own development. Jung not only theorized that the structure of personality is dynamic (shaped by various environmental experiences), but also embraced an active conception of human agency (cf. Bandura, 1997). Indeed, Jung's (1921/1971a) concept of *individuation* "to become as whole or complete a human being as one's personal circumstances would allow" (Stevens, 1998, p.53) is directly comparable to what Maslow (1968) was later to term "self-actualization."

In his conceptual model, Jung (1921/1971a) theorized that individuals differ in the way they orient themselves to the world, the way they make decisions, and the way they perceive information. According to Jung, these differences emanate from the interaction of four *functions* (or mental processes) and two personality *attitudes*. The four functions refer to *thinking*, *feeling, sensing* and *intuition*. The two attitudes refer to *extraversion* and *introversion*.

The Attitudes

Jung's (1921/1971a) distinction between introversion and extraversion as psychological modes of adaptation is perhaps the one aspect of his typology that academic psychologists (and lay-people alike) are most familiar. Although empirical support for the presence of these dimensions has emerged from a variety of sources

(e.g., Cattell, 1965; Cattell, Ebber, & Tatsuoka, 1970; Eysenck & Eysenck 1975; McCrae & Costa, 1987, 1990), these terms are often misunderstood. Jung differentiated between both attitudes by the *direction* of psychic energy, with each attitude being dichotomously opposite to the other. With introversion, this direction is toward the internal world (or the subject), and as such, the introvert may be perceived as reserved, reflective, quiet, and even cautious. For extraversion, on the other hand, this psychic energy is directed outward (toward the object) and as such the extrovert may be perceived as flamboyant, action oriented, talkative, or bold. Jung was also clear that while some people have a preference for introversion and others for extraversion, some people have the capacity to switch back and forth between both attitudes and are "influenced as much from within as from without" (Jung, 1923/1971b, p. 516). Whether one is predominantly introverted or extraverted, or is able to switch between the two, the attitudes of introversion and extraversion operate as a "switchboard from which on the one hand external behavior is regulated and, on the other, specific experiences are formed" (Jung, 1931/1971c, p. 534). In isolation, the attitudes of introversion and extraversion are insufficient in determining our behavioral, emotive, or cognitive preferences. It is in *combination* with the four functions of thinking, feeling, sensing, and intuition that our personality preferences are shaped.

The Four Functions

Thinking and feeling represent what Jung (1921/1971a) referred to as *rational*, or judging, functions. These are described as rational because they characterize how people make decisions. Within sport teams, athletes are continually required to make decisions related to issues such strategy implementation, developing tactics, or even how to approach a coach or teammate for feedback. With a thinking preference, decisions are based on cognitions associated with logic, whereas with a feeling preference, decisions are based on values and individual worth. In the former instance decisions could be said to originate "from the head," whereas in the latter case decisions stem "from the heart." For example, a key decision maker (e.g., team captain, pivotal player) may choose to provide performance-related feedback to another athlete (e.g., rookie); however, the manner with which this is done will differ depending on whether a thinking or feeling preference is engaged. Characteristics of a thinking preference might include being analytical, detached, impersonal, task-focused, and assertive. Conversely, a feeling preference might reflect being accommodating, considerate, relationship-orientation, and receptive. One should note, however, that one should not confuse Jung's (1921/1971a) concept of feeling with general affect or emotion. As Sharp (1987) reminds us, "in Jung's model, the term feeling refers strictly to the way in which we subjectively evaluate what something, or someone, is worth to us. This is the sense in which it is rational; in fact, to the extent to which it is not colored by emotion" (Sharp, 1987, p.17).

Sensing and intuition represent what Jung (1921/1971a) termed *irrational*, or perceiving, functions. In Jungian terminology, irrational does not mean unreasonable or groundless as in common parlance (cf. Stevens, 2001), rather that sensation and intuition are non-decision-making functions that refer to the way in which people *perceive* the world. Sensing (or sensation) reflects a preference for practical or real experiences, derived from our five senses, which are grounded in the "here and now." Intuition, on the other hand, is more concerned with the inferences related to a particular situation such as "what was" or "what might be." For example, two

athletes (one with a sensing preference and another with an intuition preference) may approach training and competition very differently. Although the athlete with a sensing preference may be highly adept at focusing on the practice (or task) at hand, the intuitive athlete may be more interested with how that task or practice might be adapted for future competitions.

Jung (1921/1971a) theorized that those functions that are available to us in consciousness are complemented by a balancing set of functions in our unconscious. He described the function that we tend to employ most (in consciousness) as our *primary* or *dominant* function, whereas the function that we employ least corresponds to our *inferior* function. Figure 1 illustrates the psyche as a circle containing each of the four functions (cf. Jacobi, 1973). In this case, Figure 1 is oriented to represent an individual where thinking is the primary function, and given its place in conscious use is represented by a "light" background. In this person, the opposite or inferior function would be feeling, and is presented in the "dark" or shadow background as a reflection of its position in the unconscious. In Jung's conceptual model, it is possible for an auxiliary (secondary) and sometimes a tertiary (third) function to support the primary function. For example, if an athlete's thinking preference is supported by intuition (as the auxiliary), a preference for logical analysis might be supplemented by a desire to envisage future possibilities. In Figure 1, sensation and intuition are presented as supportive (i.e., auxiliary and tertiary) functions, whereby they reside partly in consciousness and partly in the unconscious.

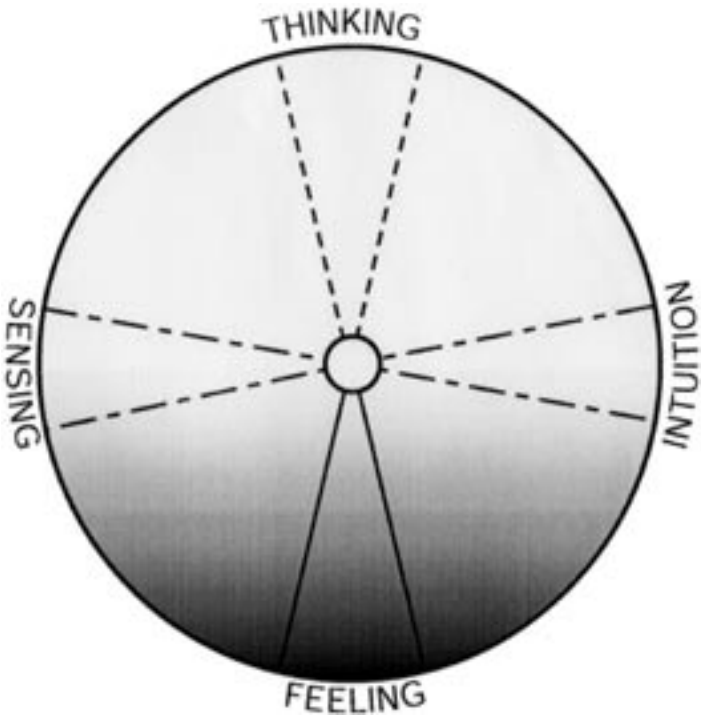


Figure 1—Graphical representation of Jung's rational and irrational functions. From Jacobi (1973), reprinted with permission.

Each of the four functions presented in Figure 1 (i.e., thinking, feeling, sensing, and intuition) can manifest themselves in either introverted or extraverted form resulting in eight possible attitudinal-functions (extraverted thinking, introverted thinking, extraverted feeling, etc.). When each function is aligned with either attitude, this gives rise to a disparate range of cognitive and behavioral preferences. Jung's (1921/1971a) typology assumes that every human (either consciously or unconsciously) uses each of these mental processes to some extent. However, some of these attitudinal-functions will be more developed (or differentiated) in each of us than others. This means that while we may employ some of these preferences with relative ease (i.e., in consciousness), others are less easily employed, and others reside within the unconscious (i.e., untrained and reluctantly used). The following section provides an overview of the preferences that result when each of the two attitudes combine with the four functions in conscious use.

Eight Jungian Attitudinal-Functions[1]

Extraverted Thinking—Seeking Order and Taking Action

Extraverted thinking may be characterized by a preference for bringing order and structure to training, team sessions, and even social interactions. Such an athlete might be keen to ensure that training programs are comprehensively thought through by both athlete and coach well in advance of competition, and that s/he plays an active role in decisions (e.g., game plans, developing his/her own personal training programs). An athlete with an extraverted thinking preference will typically try to seek out the objective facts to inform his or her ability to resolve problems (e.g., confronting a teammate for breaking team "rules") or make decisions (e.g., moving to another team/club at the end of a season). Those with an extraverted thinking preference like to organize, categorize, structure, and clarify before taking decisive action. They may also be perceived as bold, assertive, and commanding.

Introverted Thinking—Developing Ideas and Strategies With Analytical Precision

Introverted thinking is characterized by a preference for developing new ideas and discovering solutions. Within a team sport setting, athletes with such a preference may become particularly interested in creating novel performance strategies or systems, in order to give them an advantage over their competitors. An introverted thinking preference typically involves processes of investigation, observation, and thinking things through carefully before acting. Athletes with an introverted thinking preference may sometimes be viewed as distant or removed, but with a strong capacity for critical analysis.

Extraverted Feeling—Sociable and Selfless

A preference for extraverted feeling is characterized by being outgoing, sociable, and generally considerate of others' (e.g., teammates) needs. In both practice and social settings, athletes with this type of preference may choose to seek out the company of other teammates rather than train or relax on their own. They may feel particularly at home in team situations (e.g., group training, on tour, social settings) and may value the opportunity to share their interests with others. When athletes

with an extraverted feeling preference engage in decision-making processes within team settings (e.g., establishing team goals, setting out rules for team etiquette), they will often ensure that the needs of others are considered (e.g., teammates, coaches, support staff). They may sometimes feel uncomfortable when required to "keep their own company" or train by themselves (i.e., would rather be with others).

Introverted Feeling—Personal Reflection and One-on-One Connections

Interactions are typically made by an introverted feeling preference after a process of considered reflection that is underpinned by one's deeply held personal beliefs. Athletes with an introverted feeling preference may choose to communicate with a coach or teammate based on these values and may be seen by others as supportive, considerate, and reliable. For example, such an athlete may be in direct competition with a teammate for a place on a starting line-up, but will still ensure that pleasant relations are maintained and that support is provided to that athlete when necessary. Athletes with an introverted feeling preference typically avoid the limelight (e.g., media attention, public adulation) craved by those with an extraverted feeling preference and are generally not social animals. However, this does not mean they dislike social interaction and indeed may feel particularly comfortable when surrounded by a small group of friends or teammates.

Extraverted Sensing—Down to Earth and Practical

People with an extraverted sensing preference are typically down to earth and practical and are more concerned with the here and now rather than what might happen in the future. In the context of sport teams, they act in ways that are aimed at the short-term and may not fully consider the implications of their actions. For example, this might involve making a hasty decision to leave a coach after a particular poor performance rather than thinking things through first. With less concern for the deeper meaning behind relationships or future possibilities, individuals with such a preference are more concerned with what is immediately in front of them. Those with an extraverted sensing preference may be highly motivated by practical hands-on tasks or experiences. Indeed, such athletes might prefer to physically practice various strategies or team systems on the field of play, rather than deliberate over the use of those systems away from the training environment (e.g., chalk and talk sessions, team discussions).

Introverted Sensing—Reflecting on Fine Details

Introverted sensing is characterized by a preference for employing all of the five senses for reflection and consideration. This means that, like a highly sensitive photographic plate, those with an introverted sensing preference seek to take in and study every aspect of their environment. In the competitive sport environment, they may be particularly adept at technical analysis of performance, paying particular attention to the detail underpinning both individual actions and team systems. They may also be particularly proficient at clueing into the actions of their competitors and use this information to their advantage (e.g., spot weaknesses in body language, self-talk). Introverted sensing enables people to record the reality of the here and now, although they may be less adept at imagining alternative strategies for action (e.g., game plans).

Extraverted Intuition—Creative Vision and Strategy Development

Extraverted intuition is characterized by a preference to move beyond what might immediately be apparent and examine issues in greater depth. Individuals with such a preference are typically interested in what might be going on behind the scenes or under the surface. In the sport team environment, this may enable individuals to challenge existing processes (e.g., training protocols) and be innovative in applying enterprising new strategies or tactics for action. They have a particular interest in unexplored possibilities. Although athletes with a preference for extraverted intuition may be particularly motivated by a long-term vision, if they are not careful, they may fail to consider pressing needs of the present.

Introverted Intuition—Introspection and Innovation

Introverted intuition is characterized by introspection and considered innovation. A preference for challenging the mental processes used during training/competition may stimulate an athlete to work on his or her performance-related cognitive processes. With a capacity to consider issues with insight, originality, and depth, the person with introverted intuition typically has a preference for working independently and often seeks to understand the real meaning behind concepts (e.g., technique, strategy). A person with a preference for introverted intuition may be viewed by others as a daydreamer who shows little interest in the real world, only "what might be." They may have fantastic insights into various training, performance, or even social issues but may be reluctant to communicate these with their teammates or coaches.

Organization of Attitudinal-Functions and Preferences

It is important at this stage to distinguish between Jungian preferences and *traits*. Trait theorists (e.g., McCrae & Costa, 1989) assume that behavior is caused by relevant underlying traits and that traits are largely independent of each other. Preferences, on the other hand, are theorized to result from both dispositional and situational factors and arise from the complex interaction of *attitudes* and *functions* (Quenk, 2000). Behavior in this framework is considered to be an expression of one's (combined) preferences and further dependent on the degree to which one's preferences are brought into consciousness.

In Jung's (1921/1971a) typology, the primary attitudinal-function in consciousness will be balanced by an opposing attitudinal-function in the unconscious. For example, if a person's primary attitudinal-function is represented by extraverted intuition, the balancing attitudinal-function in the unconscious will be introverted sensing. The inferior attitudinal-function relates to that aspect of our personality to which we have little or no access. In the vernacular, this might be described as our "blind spot" or less conscious negative characteristics (i.e., those aspects that others might be aware of but we may not). Positioned in between the primary and inferior attitudinal-functions, other attitudinal-functions may be found, which may vary in their development from conscious use, through to the unconscious. By considering the degree to which athletes' attitudinal-functions are consciously or unconsciously expressed, one may gain an important insight into how they like to (a) communicate and be communicated with (e.g., one-on-one versus team meetings), (b) interact with teammates in both practice and social

situations (e.g., outgoing and sociable versus quiet and reserved), (c) resolve interpersonal conflicts (e.g., direct and assertive versus calm and considerate), and (d) deal with stressful situations (e.g., seek out the company of others or keep things to one's self).

Assessment of Jungian Preferences

A number of Jungian preference instruments exist, perhaps the most recognizable of which is the Myers-Briggs Type Indicator (MBTI; Myers, 1962; Myers, McCaulley, Quenk, & Hammer, 1998). This is also one of the most extensively used personality instruments with nonclinical populations, with an estimated two million people completing the MBTI annually (Quenk, 2000).

The MBTI identifies a person's primary attitudinal-function and by inference the inferior attitudinal-function as well as an auxiliary (or supporting) attitudinal-function. The result is the identification of 16 MBTI "Types" that represent possible combinations of dominant and auxiliary attitudinal-functions (i.e., extraverted thinking with introverted intuition as auxiliary, introverted thinking with extraverted sensing as auxiliary, etc). An extensive body of research has sought to examine the construct, convergent, and discriminant validity of the MBTI. From a *within-network* perspective (cf. Marsh, 1997), acceptable internal consistencies (cf. Nunnally, 1978) have been found for the separate subscales, as well as interscale correlations as predicted by theory (Myers et al., 1998). Support has also been provided for the factor structure of the MBTI (Johnson & Saunders, 1990), and from a face validity perspective, respondents report generally high levels of agreement with the results of their MBTI profiles (Kummerow, 1988; Walck, 1992b).

Other Jungian instruments also exist that not only identify the dominant and auxiliary attitudinal functions, but also the degree (or intensity) with which a person's preferences are employed (cf. Lothian, 1996). Regardless of whether one wishes to use the MBTI or any other personality instrument with a client (i.e., the athlete, coach), applied practitioners must obtain the necessary professional training and requisite accreditation (cf. American Psychological Association; APA, 2002) that enable the use of that instrument.

Applied Issues: Preferences and Interactions

Ultimately, the applied sport psychologist will be interested in how an understanding of personality can inform his or her work with athletes. Once the client has completed an assessment of his or her Jungian preferences, the practitioner will typically discuss the results with the client and have the client reflect on the *meaning* this assessment has for him or her. This process should involve the reinforcement that there are no good or bad types, only a different set of gifts (or different roads to excellence) that accompany different types (Jung, 1921/1971a; Myers & Myers, 1995). In their work with athletes, sport psychologists should also ensure that when using this conceptual model, the language used is appropriate for the athlete in question. For example, a consultant might choose to employ terms such as "preferred behaviors" over "dominant attitudinal-functions," or "decision-making styles" over "Jung's rational-functions." In short, we recognize that Jungian terminology might differ from common parlance, but we encourage the consulting sport psychologist to adapt his/her own language to meet with the needs of the client.

Having identified in discussion with the client what types of thoughts and behaviors the athlete typically employs when making decisions, interacting with others, and so forth, the client might be encouraged to reflect on his/her preferences and the implications these might have for team functioning. In addition to recognizing likely strengths, this may also involve trying to recognize potential blind spots (i.e., those behaviors that the athlete may tend to employ but may not be aware of). Given the very nature of an athlete's blind spots (i.e., difficult to access and recognize), the sport psychologist may encourage the client to involve those close to him or her (e.g., spouse, a carefully selected teammate) in this appraisal process. This is directly comparable to what is often called 360 degree appraisal (London & Smither, 1995; Sulsky & Keown, 1998) within the organizational literature, whereby one obtains a number of independent appraisals to provide a better picture of a person's strengths and weaknesses.

Once members of a team have (individually) completed this process and are at a stage whereby they are able to recognize their own preferences for interaction as well as potential blind spots, the consulting sport psychologist would be able to take the intervention to the *group-level*. At this stage, the consultant might first choose to explain that we are all different in the way we see the world (cf. "the world exists not merely in itself, but also as it appears to me," Jung, 1921/1971a, p. 374) and that we have different preferences for interaction. The consultant might, through an educational workshop, explain some of the behaviors that are characteristic of different types (being explicit that there are no good or bad types, just *differences*). Once athletes are able to recognize the disparate range of preferences for interaction (or collective gifts within the team), the consultant sport psychologist might seek to ensure that athletes become acutely aware of (a) their own and others' communications styles, (b) potential barriers to effective communication, (c) their own (and the team's) possible blind spots, as well as (d) strategies for personal and collective (i.e., team) development. We will discuss each of these considerations in turn.

In terms of effective communication, it is important that athletes can approach interactions from the point of view of the *other* person. One highly effective strategy to achieve this goal of empathy is to engage in a process of role-playing (McCann, 2000; Nelson-Jones, 2003), whereby athletes are asked to act out different scenarios or case studies and seek to connect with other team members using preferences that are very different from one's own. In actual sport settings, especially when under pressure (e.g., critical points in competition, debriefing following painful losses), this "goal of empathy" may be particularly challenging, especially when requiring a person to communicate with another whose preferences are, psychologically speaking, opposite to one's own. However, if through effective role-playing one can first seek to understand (others), one can then endeavor to become understood.

Although role-playing may be particularly effective when the intervention is presented at the group level, it may also be helpful when a consultant is working one-on-one with an athlete. For example, consider the case of Peter (a 26 year-old club-level soccer player and team captain; case names are fictitious), who is having difficulty in getting Simon, a talented 19-year old teammate, to contribute to team discussions. Peter could be described as strong willed and particularly demanding of his teammates (i.e., extraverted thinking) but after talking with his coach believes that if he approaches Simon in a very direct fashion, he might not get the best out of his teammate. Simon is a fairly quiet, conscientious, and reflective (i.e.,

introverted feeling preference) member of the team. In a one-on-one session with his sport psychologist, Peter could be encouraged to recognize Simon's preferences for interaction and consider using strategies, such as the following: allow time for reflection before asking Simon to contribute his views, approach Simon in private, and consider his opinions before acting. In this session, Peter could role-play the behaviors required to adapt to Simon's (as played by the sport psychologist) preferences. Once Peter is able to effectively implement these behaviors in his interactions with Simon, this would enable Simon to feel more at ease in group discussions and contribute more fully to the team's efforts. Alternatively, if the roles were reversed and Simon wanted to communicate effectively with Peter, Simon could be encouraged to (a) clearly and succinctly specify the objectives for the discussion, (b) be direct and to the point whenever possible, and (c) recognize Peter's need to feel in control of the discussion and be aware that he may come across as forceful.

As another example, consider the case of Susan, a university-level basketball player, who often gets frustrated with the way Kate (another team member) provides her with performance-related feedback following competition. Although Kate may be particularly sociable, outgoing, and considerate (i.e., extraverted feeling), Susan often believes that her observations are imprecise and vague. When Susan's need for comprehensive and detailed feedback (i.e., introverted thinking) are not met, she gets particularly irritated, which in turn puts a strain on their relationship. In this instance, Kate could be encouraged to communicate with Susan by (a) listing the criteria she has used to underpin her feedback to Susan, (b) being thorough with her explanations, and (c) support her performance appraisals with evidence. For example, rather than tell Susan that she could have been a bit more assertive on offense, Kate could highlight those key moments that she has used to inform her judgments and provide specific feedback related to those facets of her game that could be improved.

The above two cases illustrate how a sport psychologist might employ *direct intervention* (Terry, 1998) strategies to ensure that effective communication occurs between two members of the same team. A sport psychology consultant may also draw from his/her understanding of Jungian preferences to inform *indirect interventions* (cf. Terry, 1998), whereby the sport psychology consultant works with the coach and the coach then delivers the intervention to the team. One group-related setting that lends itself well to this type of intervention is in the delivery of *performance analysis and review* sessions that coaches typically run with their players prior to or following competition. Consider the case of Bob, a highly promising 18-year old rugby player who is fairly reserved and astute in his understanding of the game but often feels uncomfortable in these group sessions (i.e., introverted sensing). This is especially noticeable when the coach puts Bob on the spot with very direct and pointed questions about his recent role performances. Once the team's sport psychologist is aware of Bob's behavioral preferences, s/he could encourage the coach to provide Bob with information to reflect on, in advance of the team session. By considering Bob's preference to reflect on issues in detail before acting, this strategy would enable him to more fully contribute to the ensuing team appraisal/review process.

Although this intervention might be appropriate for an introverted sensing preference (i.e., reflects on fine details), in a similar regard, different (indirect) intervention strategies could be applied that take into account other cognitive and behavioral preferences. For example, if certain athletes were to have a preference

for explicit structure, order, and logic (i.e., extraverted thinking) in their interaction with others, the sport psychologist could encourage the coach to ensure that the objectives for any meeting with those athletes are explicitly laid out at the outset and that the criteria used to inform the group process (e.g., performance review) are explicitly structured and detailed. Alternatively, if a team were comprised of highly sociable and gregarious athletes (e.g., extraverted feeling), the coach could be encouraged to implement team sessions that are player-directed and involve small-group activities in which the players can debate their observations with others (i.e., satisfies their need for self-expression).

Although these examples highlight some of the implications that arise when team members have a particular preference for interaction, from an applied perspective, a sport psychology consultant should also consider the contributions of athletes' supporting (i.e., auxiliary) conscious preferences. A cursory examination of the primary attitudinal-functions of a sport team's constituent members may initially suggest a fairly homogeneous distribution of a group's resources. For example, half of a team might have a primary preference for introspection and independence (i.e., introverted intuition) and the other half may have a primary preference for being outgoing and considerate (i.e., extraverted feeling). By considering the auxiliary functions of the team's members, it is possible that additional skills will also be available to the group that otherwise may have been overlooked, such as being highly organized and decisive (i.e., extraverted thinking) or the ability to analyze performances in particular detail (i.e., introverted sensing). In short, in order to bring about effective intragroup communication the sport psychologist (as facilitator) and, in turn, the team as a whole should consider the degree to which each attitudinal-function (or preference) is brought to consciousness within the group as a whole. Furthermore, if a team (through the guidance of the sport psychologist) can become aware of any unused preferences or blind spots, this may serve as a basis for further group development. For example, if an intuitive preference is rarely used within a team, the coach and psychologist could structure practices to enable the team to consider how various systems might differentially be implemented in future competitions.

In terms of potential barriers to communication, one only has to examine the contrasting nature of each of the rational and irrational functions to see how potential barriers to effective communication might arise. Those with a thinking preference will primarily be guided by notions of logic, whereas those with a feeling preference will be guided by personal notions of value, morality, or a heightened sense of loyalty. Similarly, while a preference for intuition may motivate a person toward future possibilities, these motives might potentially conflict with a teammate's need to attend to the present (i.e., sensing preference). By considering how an athlete's pattern of preferences might conflict with those of another, those concerned with intervention (e.g., coach, psychologist) could develop and implement a set of communication strategies amenable to both parties. For example, this might involve the provision of a set of ground rules that include considerations such as the following: everyone has a chance to express his/her preferences; feedback should be provided in constructive manner; athletes should not make hasty assumptions about what others are thinking; even if athletes think they are right, they should try to see things from the other side of the fence (Anshel, 1994; Sullivan, 1993).

In terms of potential "blind spots," group members could be encouraged to recognize that their own perceptions of self may differ somewhat from the way

others perceive them. Indeed, at times we may be distinctly unaware of certain behaviors that we engage in and the impact they have on those around us. For example, if an athlete is motivated by seeking out innovative new training techniques (i.e., extraverted intuition), once those techniques have been found, the athlete may be more interested in finding even newer techniques, rather than consolidating on the information that has already been acquired. In this instance, the athlete may appear brash and quick to apply future directions when, in fact, it may be more appropriate to get to grips with the detail (i.e., introverted sensing) of one particular training regimen. Given the difficulty one may have in recognizing and working on one's weaknesses, a consultant might first encourage athletes to work on their listening skills (cf. Nelson-Jones, 2003) and then, under the consultant's guidance, reflect on possible blind spots with other trusted teammates or peers. From an interventionist perspective, one should note that it is just as important to understand the weaknesses represented by one's inferior attitudinal-function as it is to understand the strengths expressed by one's dominant attitudinal-function.

Within the context of sport teams, each member will bring a unique collection of gifts or pattern of preferences to the group's collective resources. By mapping out the different patterns of preferences exhibited by each member, the group could become more acutely aware of where its strengths and weaknesses lay. For example, if a group is heterogeneous in its diversity of resources (i.e., patterns of preferences), this will bring a different set of challenges for the team and consulting psychologist compared to a group that is relatively homogeneous in nature (Forsyth, 1999; Johnson & Johnson, 1997). For a homogeneous group, this structural property may enable group members to come to decisions more quickly and agree that the chosen solution is the most appropriate. However, invariability of a group's personal resources may also lead to *group-think* "when members' strivings for unanimity over-ride their motivation to realistically appraise alternative courses of action" (Janis, 1972, p.9). Janis coined the term group-think after analyzing the decision-making procedures that were employed prior to and during the Bay of Pigs invasion, as well as the Korean and Vietnam Wars. If a group is relatively homogenous in its distribution of preferences, it may be useful to see where potential weaknesses in the group may be found and employ strategies to ensure that necessary processes are not overlooked. These may include encouraging team members to play devil's advocate or have consultants (e.g., other coaches, independent experts) from outside the team environment challenge its strategies for action.

A heterogeneous group, on the other hand, may potentially have a greater diversity of personal resources (i.e., patterns of preferences) available. However, when faced with such potentially conflicting preferences for action, this will provide a very different set of challenges for the sport psychology consultant, requiring very different strategies for intervention. Lothian (1997) suggested that when faced with competing interests and priorities, a group could be encouraged to employ a "Four-Sight" approach to facilitate effective intra-group interaction (see Figure 2). This process involves first, brainstorming for ideas using the imagination of group members (intuition); second, gathering the relevant factual data (sensing); third, analyzing the effects of adopting alternative outcomes (thinking); and finally, considering how different solutions affect the people involved (feeling). Such an approach could encourage different members to contribute their strengths to the group solution and at the same time enable them to value the contributions of others with alternative perspectives.

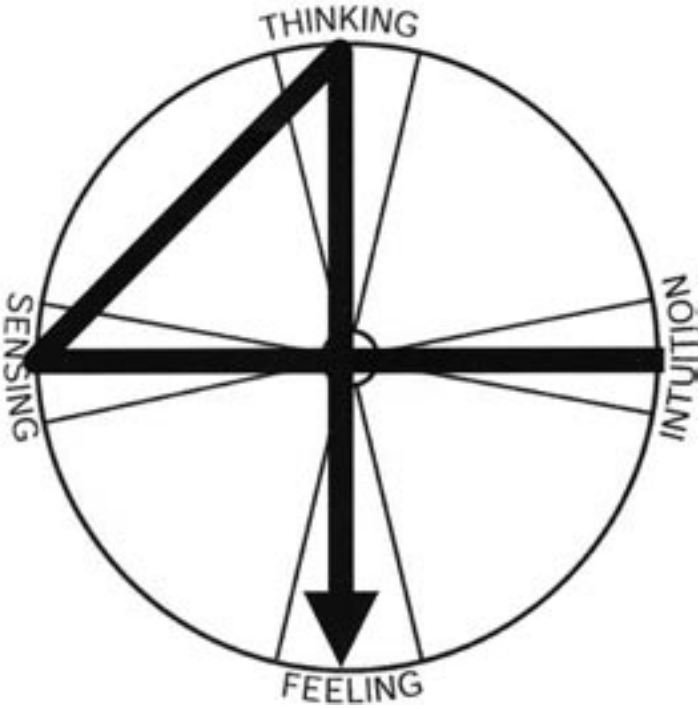


Figure 2—The "Four-Sight" approach to intra-group communication. Adapted from Lothian (1997).

Note. The goal of the "Four-Sight" approach is to effectively employ all four functions (thinking, feeling, sensing, and intuition) in conscious use. This would explain why the shadow evident in Figure 1 is not present in Figure 2.

In his writing, Jung (1921/1971a) was particularly interested in the process of individuation, to become "rounded," by raising to consciousness those preferences typically resident in the unconscious. In much the same way that individuation relates to individual growth and development, it may also be possible for the group to collectively individuate and seek to develop those preferences (or blind spots) that are undeveloped and untrained. Indeed, if all members of a team can effectively adapt their communication to connect more effectively with others, not only will successful interdependence become more likely, but also the team (as a whole) will be a happier, healthier, and more effective environment in which to perform.

As with any psychological intervention, the applied sport psychologist should also be acutely aware of the potential for misuse and misinterpretation arising from any Jungian preference (or any other) personality assessment. Although the application of this model within sport has much potential, consultants must be aware of, and act within, the ethical principles guiding the use of personality instruments with clients (cf. APA, 2002). For example, psychologists must take necessary steps to explain to the client the results of any assessment (APA, 2002, principle 9.10) and ensure that conflicts of interest between ethical and organizational demands

(e.g., in professional sport this may be the employing organization) are avoided (APA, 2002, principle 1.03). Psychologists should also be aware of their boundaries of competence based on training and education and only employ intervention strategies that are commensurate with that education and training (APA, 2002, principle, 2.01).

In summary, we do not believe that certain attitudinal-functions or preferences are better than others or that certain personalities are more likely to succeed in sport. However, we do suggest that if group members, under the guidance of a suitably qualified sport psychologist, can begin to understand how their preferences might complement or conflict with those of others, they might function more effectively, both individually and as a team.

References

Allen, J., & Brock, S.A. (2000). *Healthcare communication using personality type: Patients are different.* London: Routledge.

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist*, **57**, 1060-1073.

Anshel, M.H. (1994). *Sport psychology: From theory to practice.* Scottsdale, AZ: Gorsuch Scarisbrick.

Atwater, L., & Yammarino, F.J. (1993). Personal attributes as predictors of superiors' and subordinates' perceptions of military academy leadership. *Human Relations*, **46**, 645-668.

Bakker, F.C., Whiting, H.T.A., & Van der Brug, H. (1990). *Sport psychology: Concepts and applications*. Chichester, UK: Wiley.

Bandura, A. (1997*). Self-efficacy: The exercise of control.* New York: Freeman and Company.

Barry, B., & Stewart, G.L. (1997). Composition, process, and performance in self-managed groups: The role of personality. *Journal of Applied Psychology*, **82**, 62-78.

Camara, W.J., & Nathan, J.S., & Puente, A.E. (2000). Psychological test usage: Implications in professional psychology. *Professional Psychology*: *Research and Practice*, **31**, 141-154.

Carron, A.V., & Hausenblas, H. (1998). *Group dynamics in sport* (2nd ed). Morgantown, WV: Fitness Information Technology.

Cattell, R.B. (1965). *The scientific analysis of personality.* Baltimore: Penguin.

Cattell, R.B., Ebber, H.W., & Tatsuoka, M.M. (1970). *Handbook for the Sixteen Personality Factor Questionnaire (16 PF).* Champaign, IL: Institute for Personality and Ability Testing.

Eysenck, H.J., & Eysenck, S.B.G. (1975). *Manual of the Eysenck Personality Questionnaire.* London: Hodder & Stoughton.

Forsyth, D.R. (1999). *Group dynamics* (3rd ed.). Belmont, CA: Wadsworth.

Freud, S. (1953). The interpretation of dreams. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 4 & 5). London: Hogarth Press. (Originally published 1900)

Gardner, W.L., & Martinko, M.J. (1996). Using the Myers-Briggs Type Indicator to study managers: A literature review and research agenda. *Journal of Management*, **22**, 45-83.

Gill, D.L. (2000). *Psychological dynamics of sport and exercise* (2nd ed.). Champaign, IL: Human Kinetics.

Hirsh, S.K., & Kummerow, J.M. (1990) *Introduction to type in organizations* (2nd ed.). Palo Alto, CA: Consulting Psychologists Press.

Hogan, R., Curphy, G.L., & Hogan, J. (1994). What we know about leadership: Effectiveness and personality. *American Psychologist*, **49**, 493-504.

Hogan, R., Hogan, J., & Roberts, B. (1996). Personality measurement and employment decisions: Questions and answers. *American Psychologist*, **51**, 469-477.

Jacobi, J. (1973). *The psychology of C.G. Jung.* New Haven, CT: Yale University Press.

Janis, I. (1972). *Victims of groupthink.* Boston, MA: Houghton Mifflin.

Johnson, D.W., & Johnson, F.P. (1997). *Joining together: Group theory and group skills* (6th ed.). Boston, MA: Allyn & Bacon.

Johnson, D.A., & Saunders. D.R. (1990). Confirmatory factor analysis of the Myers-Briggs Type Indicator: Expanded analysis report. *Educational and Psychological Measurement*, **50**, 561-571.

Judge, T.A., Bono, J.E., Ilies, R., & Gerhardt, M.W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, **87**, 767-780.

Judge, T.A., & Ilies, R. (2002) Relationship of personality to performance motivation: A meta-analytic review. *Journal of Applied Psychology*, **87**, 797-807.

Jung, C.G. (1971a). Psychological types (H.G. Baynes, Trans., revised by R.F.C. Hull). In *The Collected Works of C. G. Jung* (Vol. 6, pp. 1-495). Princeton, NJ: Princeton University Press. (Original work published 1921)

Jung, C.G. (1971b). Psychological types (H.G. Baynes, Trans., revised by R.F.C. Hull). In *The Collected Works of C. G. Jung* (Vol. 6, pp. 510-523). Princeton, NJ: Princeton University Press. (Original work published 1923)

Jung, C.G. (1971c). A psychological theory of types (H.G. Baynes, Trans., revised by R.F.C. Hull). In *The Collected Works of C.G. Jung* (Vol. 6, pp. 524-541). Princeton, NJ: Princeton University Press. (Original work published 1931)

Jung, C.G. (1971d). Psychological typology (H.G. Baynes, Trans., revised by R.F.C. Hull). In *The Collected Works of C.G. Jung* (Vol. 6, pp. 542-555). Princeton, NJ: Princeton University Press. (Original work published 1936)

Kummerow, J. (1988). A methodology for verifying type: Research results. *Journal of Psychological Type*, **15**, 20-25.

London, M., & Smither, J.W. (1995). Can multi-source feedback change perceptions of goal accomplishment, self-evaluations, and performance-related outcomes? Theory-based applications and directions for research. *Personnel Psychology*, **48**, 803–839.

Lothian, A.M. (1996). *Insights Discovery Preference Evaluator.* Dundee, UK: Insights Learning and Development.

Lothian, A.M. (1997). *Insights into personal effectiveness: Workbook.* Dundee, UK: Insights Learning and Development.

Lothian, A.M. (2003). *The birth of analytical typology in coaching.* Dundee, UK: Insights Learning and Development.

Marsh, H.W. (1997). The measurement of physical self-concept: A construct validation approach. In K.R. Fox (Ed.), *The physical self: From motivation to well-being* (pp. 27-58). Champaign, IL: Human Kinetics.

Maslow, A.H. (1968). *Towards a psychology of being* (2nd ed.). Princeton, NJ: Van Nostrand.

McCann, S.C. (2000). Doing sport psychology at the really big show. In M.B. Andersen (Ed.), *Doing sport psychology* (pp. 209-222). Champaign, IL: Human Kinetics.

McCrae, R.R., & Costa, P.T., Jr. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, **52**, 81-90.

McCrae, R.R., & Costa, P.T., Jr. (1989). Reinterpreting the Myers-Briggs Type Indicator from the perspective of the five-factor model of personality. *Journal of Personality*, **57**, 17-40.

McCrae, R.R., & Costa, P.T., Jr. (1990). *Personality in adulthood.* New York: Guilford Press.

McCaulley, M.H. (2000). Myers-Briggs Type Indicator: A bridge between counseling and consulting. *Consulting Psychology Journal: Practice and Research*, **52**, 117-132.

Morgan, W.P. (1979). Prediction of performance in athletics. In P. Klavora, & J.V. Daniel (Eds.), *Coach, athlete, and the sport psychologist* (pp. 172-186). Champaign, IL: Human Kinetics.

Morgan, W.P. (1980). Sport personology: The credulous-skeptical argument in perspective. In W.F. Straub (Ed.), *Sport psychology: An analysis of athlete behavior* (2nd ed., pp 330-339). Ithaca, NY: Mouvement Publications.

Myers. I.B. (1962). *Manual: The Myers-Briggs Type Indicator.* Princeton, NJ: Educational Testing Service.

Myers, I.B., McCaulley, M.H., Quenk, N.L., Hammer, A.L. (1998). *MBTI manual: A guide to the development and use of the Myers-Briggs Type Indicator.* Palo Alto, CA: Consulting Psychologists Press.

Myers, I.B., & Myers, P.B. (1995). *Gifts differing: Understanding personality type.* Palo Alto, CA: Davis-Black.

Nelson-Jones, R. (2003). *Practical counseling and helping skills: Text and exercises for the lifeskills counseling model* (4th ed.). London: Sage.

Nunnally, J.C. (1978). *Psychometric theory.* New York: McGraw-Hill.

Nutt, P.C. (1990). Strategic decisions made by top executives and middle managers with data and process dominant styles. *Journal of Management Studies*, **27**, 173-194.

Piedmont, R.L., & Ciarrocchi, J.W. (1999). The utility of the revised NEO personality inventory in an outpatient, drug rehabilitation context. *Psychology of Addictive Behaviors*, **13**, 213-226.

Quenk, N.L. (2000). *Essentials of Myers-Briggs Type Indicator assessment*. New York: Wiley.

Rogers, C.R. (1951). *Client-centered therapy.* Boston: Houghton Mifflin.

Rogers, C.R. (1977). *Carl Rogers on personal power.* New York: Delacorte Press.

Roush, P.E., & Atwater, L. (1991). Using the MBTI to understand transformational leadership and self-perception accuracy. *Military Psychology*, **4**, 17-34.

Sharp, D. (1987). *Personality types: Jung's model of typology.* Toronto: Inner City Books.

Shaw, M.E. (1981). *Group dynamics: The psychology of small group behavior* (3rd ed.). New York: McGraw-Hill.

Shope, T.C., Frohna, J.G., & Frohna, A.Z. (2000). Using the Myers-Briggs Type Indicator (MBTI) in the teaching of leadership skills. *Medical Education*, **34**, 956.

Stevens, A. (1998). *An intelligent person's guide to psychotherapy.* London: Duckworth.

Stevens, A. (2001). *Jung: A very short introduction.* Oxford: Oxford University Press.

Sullivan, P. (1993). Communication skills training for interactive sports. *The Sport Psychologist*, **7**, 79-91.

Sulsky, L.M., & Keown, J.L. (1998) Performance appraisal in the changing world of work: Implications for the meaning and measurement of work performance. *Canadian Psychology*, **39**, 52-59.

Terry, P.C. (1998, September). *Delivery of applied sport psychology: Lessons learned over 15 years.* Paper presented at the annual convention of the Association for the Advancement of Applied Sport Psychology (AAASP), Cape Cod, MA, USA.

Walck, C.L. (1992a). Psychological type and management research: A review. *Journal of Psychological Type*, **24**, 13-23.

Walck, C.L. (1992b). The relationship between indicator type and "true type": Slight preferences and verification process. *Journal of Psychological Type*, **23**, 17-21.

Van den Auweele, Y., Nys, K., Rzewnicki, R., & Van Mele, V. (2001). Personality and the athlete. In R.N. Singer, H.A. Hausenblas, & C.M. Janelle (Eds.), *Handbook of research on sport psychology* (2nd ed., pp. 239-268), New York: Wiley.

Vanek, M., & Cratty, B.J. (1970). *Psychology of the superior athlete*. London: Macmillan.

Vealey, R.S. (2002). Personality and sport behavior. In T. Horn (Ed.), *Advances in sport psychology* (2nd ed., pp. 43-82). Champaign, IL: Human Kinetics.

End Note

[1]Adapted from Lothian, A.M. (2003).

Manuscript submitted: October 14, 2003
Revision received: May 5, 2004

Coping Strategies for Low Frequency Noise

by

Geoff Leventhall, Stephen Benton and Donald Robertson

reprinted from

Journal of LOW FREQUENCY NOISE, VIBRATION AND ACTIVE CONTROL

VOLUME 27 NUMBER 1 2008

MULTI-SCIENCE PUBLISHING COMPANY LTD.

Coping Strategies for Low Frequency Noise

Geoff Leventhall[*], Stephen Benton[**] and Donald Robertson[***]
[*]*150 Craddocks Avenue, Ashtead, Surrey. KT21 1NL geoff@activenoise.co.uk*
[**]*Department of Psychology, University of Westminster, Regent St. London.*
[***]*10 Harley Street, London*

Received 15th January 2008

ABSTRACT

A small group (n=9), whose complaints of low frequency noise had not been resolved by Environmental Health Officers and related care professionals, were invited to attend a series of intervention sessions led by a psychotherapist. The aims of the sessions were to improve the participants' coping strategies and their quality of life, in order to relieve them from some of the distress caused by their belief that they were exposed to low frequency noise. Prior to the psychotherapy sessions the group was evaluated on a number of self report questionnaires, which measured individual responses for reaction to low frequency noise, quality of life, quality of coping and a personality questionnaire (Insights^{tm} Discovery Preference Evaluator). A before and after, within group analysis of responses was based upon repeat measures of the three behavioural response questionnaires for noise reactivity, quality of life and coping. A general reduction in the subjects' stress levels was shown, suggesting positive effects of psychotherapy upon symptoms that had, in this group's case, proved resistant to improvement by conventional local authority and specialist interventions. This 'therapeutic' approach to LFN interventions could lead to improved health and effectiveness and fewer demands on local services. Although the techniques of tinnitus management were informative, analogy between the problems of low frequency noise sufferers and those of tinnitus sufferers fails at the point where low frequency noise sufferers believe that an external agency is the cause of their problems.

1. EXPERIMENTAL AIMS AND DESIGN

The aim of the study was to examine how long-term LFN sufferers' perception of their quality of life and coping quality responded to group and self-help therapeutic interventions.

To this aim, parameters were selected which supported a before and after, within-subjects comparison. All subjects had agreed to participate in the assessments, both before and after the therapeutic intervention. The sequence of measurements and interventions included:

1. Administration of Questionnaires
2. Therapeutic intervention by psychotherapist
3. Repeat administration of questionnaires

2. QUESTIONNAIRES AND PSYCHOLOGICAL MEASURES

The questionnaires were designed to register the before-and-after perceptions of quality of life and coping ability. A further measure was made of personality. The personality measure, using the Insights Discovery Preference Evaluator^{tm}, was introduced for two principal reasons: firstly to be able to place individuals within a typological system in order to examine possible groupings within particular types (e.g. Introverted thinking/Extraverted feeling), as personality has been indicated as an independent variable in sensitivity to noise (Belojevic et al. 2003). Secondly,

measurement of personality within this system offered the potential for highlighting communication blind spots (a mismatch of communication styles is likely to increase anxiety) associated with personality type.

3. CLINICAL GROUP-WORK

The aim of the clinical component of the study was to establish whether group therapy combined with self-help processes could be of benefit to sufferers from low frequency noise related stress (LFNRS). The initial combination of the techniques used was based on a preliminary review of the literature and analogies with treatment of phobic conditions and stress management.

Initially, emphasis was placed upon a combination of multi-modal relaxation and imaginal exposure techniques based upon Cognitive Behavioural Therapy and Group Hypnotherapy. Different techniques were explained and demonstrated to participants and a written handout prepared for them containing a "menu" of possible therapeutic techniques, which could be used as self-help coping skills. In addition, participants were given a relaxation CD Rom based upon the multi-modal relaxation processes used in the group sessions. As the work progressed, individual interventions were combined, and adapted, according to discussion and feedback received from the group.

The techniques discussed with the group included the Neural Linguistic Programming (NLP) rewind (or "fast phobia") technique, changing the sound to a visual image and altering its sensory sub-modalities, different forms of visual-kinaesthetic dissociation (VKD), anchoring, different forms of multi-modal relaxation, affirmations, and imaginal exposure. (See Appendix 2). Six group Sessions were held for the project.

4. SUBJECT SELECTION AND INITIAL EVALUATION

A database of subjects was available from a survey which had been carried out as part of earlier work (Leventhall, et al. 2003). Those who lived within access of London were telephoned and the project explained to them. Responses varied from "I don't hear the noise anymore" to "The noise has made me too ill to travel to London". Some subjects would have liked to take part but were constrained by their work. However, there was a good positive response and subjects were selected as in Table I, including some new contacts. The subjects are typical of low frequency noise sufferers. (Leventhall, et al 2003)

Table I.

Subjects who took part in the relaxation sessions. Subject E was self-employed and F was in employment Subject H dropped out early in the project

Subject	Age	Sex	Hearing problems
A	76	F	Y
B	65	M	N
C	69	F	N
D	69	F	Y
E	56	F	N
F	59	F	N
G	71	M	N
(H	75	M	N)
I	72	M	N

5. LOW FREQUENCY NOISE REACTION QUESTIONNAIRE (LFNRQ).

As an initial assessment, each subject was sent a questionnaire to complete before subjects met the researchers. The questionnaire registered subjects' responses with minimal influence from the project while requiring a preliminary effort from the subjects, as an indication of the seriousness of their feelings towards the project. An existing Tinnitus Reaction Questionnaire (Wilson, et al. 1991) was used as the basis for the questionnaire.

The original questionnaire was modified in the following ways:

Each question originally commenced with "My tinnitus has...." All questions were changed to commence with "The noise has..." and the questionnaire was titled "Low Frequency Noise Questionnaire".

An additional question, No. 27, was added at the end to reflect some low frequency noise sufferers' statements that the noise drives them from their homes.

Three further general questions permitted subjects to write down comments on their noise and, finally, subjects were asked to list prescription drugs that they were taking. Four of the original nine subjects were taking anti-psychotic drugs, none of which was listed in pharmacopia as having auditory illusional effects.

The Low Frequency Noise Questionnaire is shown in Appendix 1 and the comparison of outcomes before and after the relaxation sessions is given in section 12.1.

6. PSYCHOLOGICAL TESTS.

At their first meeting with the researchers, each of the subjects (n=9), completed questionnaires, which assessed personality, perceived Quality of Life (current), Quality of Coping (current) and their earlier Quality of Life (as estimated by the subjects before the onset of the noise). The primary aim was to build a profile of their individual subjective experience of, and reactions to, the perceived LFN.

Personality profiles were elicited using the Insights ^{tm} Discovery Preference Evaluator. This is a Jungian based system and provides for an in-depth and easy to read report on personality preferences, interpersonal and communication styles. Subjects each received a copy of their report.

7. QUALITY OF LIFE AND COPING

7.1 Current Quality of Life

The Current Quality of Life questionnaire addressed key areas of physical well-being, coping, anxiety, emotional support and emotional well-being. The Quality of Coping Questionnaire registered specific responses to an individual's perceived capacity to manage the noise and other competing demands. The 'Before Noise' Quality of Life Questionnaire provided the subjects with an opportunity to describe how they might have responded to the questionnaire before the onset of the noise problem.

The Quality of Life questionnaire required responses to 15 questions within the categories of Strongly Agree, Agree, Neutral, Disagree and Strongly Disagree for each item. The Quality of Coping questionnaire contained eight questions in the same response format.

Table II.Shows key Quality of Life results from the Questionnaires before the therapeutic intervention (number of subjects out of 8)

Table II.

Quality of Life Questionnaire - key results

Current Quality of Life and Anxiety		Perceived Coping	
Feel sad	4	Dissatisfied with coping	3
Feel Anxious	5	Cannot accept the noise	6
Unable to have fun (relax)	4	Losing hope	3
Discontent with Q of L	7	In a state of worry	5
Unable to sleep well	5	Worried will get worse	4

The general reported Quality of Life was shown to be of real concern to the group and of a generally low level, while anxiety was high and sleeping disturbed. These responses, combined with a high degree of worry about the noise and its continued effects, all served to characterise subjects' Quality of Life as poor and point to a likely ongoing and confounding influence on the perceived effectiveness of existing interventions. Clearly, any support for subjects' condition from existing medical and assessment interventions was highly likely to be compromised by the strength of associated anxiety and worry.

7.2 Quality of Coping Questionnaire:

The second questionnaire completed by the subject group comprised items related to personal coping. Some key results show the responses to items that addressed individuals' assessment of their present state of coping and the quality of their ability to manage living with the noise, as shown in Table III.

Table III.

Quality of coping questionnaire - key results

	Agree	neutral	disagree
Having a hard time adjusting to the noise	5	1	2
Feelings of inadequacy	5	1	2
Will never cope with the noise and be happy	3	2	3
People around them are uncomfortable because of the noise problem	4	1	3

Although the general level of coping was not good, the group still displayed a resilient attitude to the situation; indicative of a strong need to re-assert control over their personal environment. The partner/main support relationship was viewed as central to most subjects' coping, probably reflecting a dependency rather like that initiated in response to an illness. However, the broader social network seems to pose further and unwanted demands upon at least half of the group as shown in Table III.

7.3 Quality of Life: Comparison - Before and After the Noise

All group members completed a questionnaire that allowed them to assess items against their impression of how they would have scored that item before the onset of the noise as in Fig 1. In the 15 Quality of Life (Q of L) questions, the first eight questions relate to negative aspects of their life whilst the remaining questions relate to positive aspects. The questions were:

NEGATIVE	Q1	I have lack of energy
	Q2	I have nausea
	Q3	Because of my feelings of fatigue I have trouble meeting the needs of my family
	Q4	I experience bad headaches
	Q5	I feel ill
	Q6	I am forced to spend time in bed
	Q7	I feel sad
	Q8	I feel anxious
POSITIVE	Q9	I feel close to my friends
	Q10	I am able to concentrate at home
	Q11	My work at home is fulfilling
	Q12	I am able to enjoy life
	Q13	I am sleeping well
	Q14	I am enjoying the things I usually do for fun
	Q15	I am content with the quality of my life right now

Figure 1 summarises the subjects' responses to these questions and displays them as a comparison of scores 'Before and After' onset of the noise.

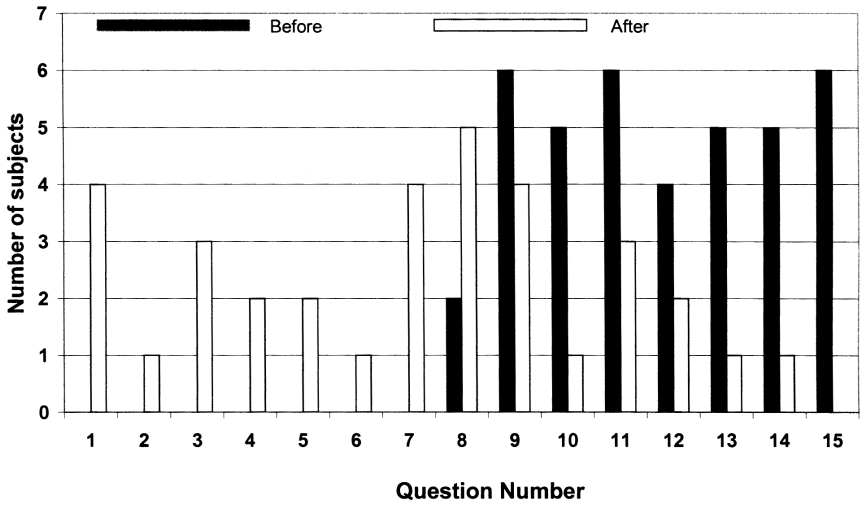


Fig.1 Quality of Life questions before and after onset of noise

Fig. 1 shows the number of subjects who selected 'agree' or 'strongly agree' to the items. For example, when asked to consider their previous level of energy, before the noise onset, -Q1 I have lack of energy - seven subjects selected Strongly Disagree and one selected Disagree giving a zero score for the question before onset of the noise. The first seven questions, relating to negative indicators in their life, all gave zero scores for their memory of their condition before the onset of the noise. Question 8, on anxiety, showed an increase after the onset of the noise. Similarly, the remaining questions, referring to indicators of positive aspects of their life showed a diminished agreement after the onset of the noise. In order to increase the robustness of level of data, any neutral scores were not included as evidence for an item. This means that, in Fig 1, the number of responses to a question is less than the number of subjects.

Whilst their memories of their situation before the noise may be idealised, the differences illustrate the subjects' belief in the way that the noise has affected them, which is a major contributor to their current levels of stress and perceived levels of 'Quality of Life', e.g. "I was so happy before the noise started". Indeed, the comparison of ratings shows a picture of a home environment being stripped of control, relaxation and enjoyment. The loss of sleep and degraded capacity to concentrate are just two items which, when combined, create an interactive and corrosive process that inhibits recuperation, whilst elevating anxiety and thereby further undermining the individual's sense of well-being. Clearly the quality of mood-state was perceived as deteriorating as both the sadness and anxiety items (Q 7 and 8 in Fig 1) display marked increase across the before and after noise onset conditions.

Perhaps the strongest aspect of these findings can be found in the recognition of the chronic nature of the situation within which individuals find themselves. As the key Quality of Life elements are distorted and fail, the individual may start to lose hope, which exacerbates the already eroded quality of emotionality. The general trend is pronounced, with the key indicators for self-rated Quality of Life showing a marked decline when compared with the Quality of Life as remembered by subjects before the noise onset. Even allowing for an element of idealisation in subjects' perception of their lives before the onset of noise, the noise produces a clear shift towards the less favourable indicators in their lives, which lie at the left of Figure 1.

8. PERSONALITY QUESTIONNAIRE. INSIGHTS DISCOVERY EVALUATOR

For the purpose of this study, the Insights Discovery Preference Evaluator [tm] (IDPE) was used to locate individuals within one of the Jungian Types. The Jungian system is comprised of two attitudes (introversion and extraversion) and two rational functions (thinking and feeling) and two irrational functions (sensing and intuition). The IDPE provides a bi-polar two factor space mapped as a wheel around which individuals are located as shown in Fig. 2, which shows the placing of the subjects on the wheel of the Insights bi-polar quadrants.

All but two of the subject's scores placed them in the introverted quadrants.

It can be noted that subjects C and D, who appear in the extroverted section, showed high scores on the LFN Reaction Questionnaire. Subject C was sensitive to all noise and very distressed at the start of the sessions. Subject D, who said that her normal personality was outgoing and ebullient, had been driven to tears by the noise and had become prone to panic attacks. The remaining subjects, in the introverted section, were more subdued in their responses, although felt just as keenly.

The Jungian model, which generates this classification, posits how individuals prefer to organise and cope with incoming information. The dimension along which this preference is measured is that of 'sensing'. According to Jung, individuals who score highly along this dimension, tend to be comfortable organising information around specific examples and on a 'here and now' basis. Moreover, these individuals would tend to perceive what is going on or happening to them, based upon the concrete experiences available to them from the senses. There is a predicted dependence upon the senses for perception, rather than internalised abstraction. The Jungian model predicts that such individuals are heavily inclined to build models of the world based upon a need for 'trustworthy' information, from which coping follows. For these individuals, trustworthy information is strongly influenced by the quality of the sensory information available. Any decline in the level of 'trustworthy' sense-based information (for whatever reason), undermines their fundamental reliance upon their preference to verify, based upon access to and use of concrete information.

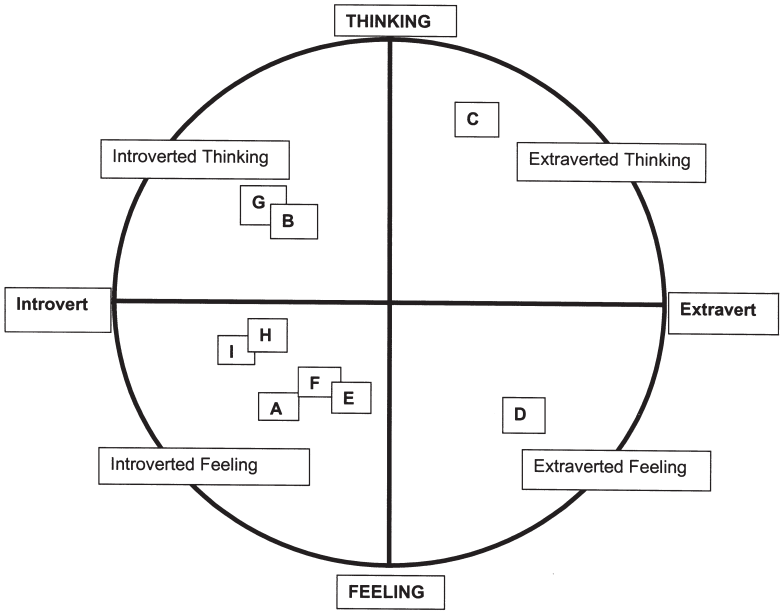


Fig. 2 Illustrating the location of subjects in terms of Introversion and Extraversion

The majority of subjects (seven from nine) are located within the introverted sensing and feeling quadrants. Failure to be able to control sensory information is probably

disturbing for most types of individuals. However, for this group the Jungian model suggests that failure to control such sense-based events tends to undermine the personal strategies preferred by these individuals when seeking to cope with environmental demands.

The inability to establish control over the sensory environment inhibits any accommodation to changes in the composition of the noise situation within which they found themselves; the ability to "simply get used to it" would be inhibited and remote. For these types, accommodation to changing sense data is conditional upon re-establishing control and their sensitivity to the noise impact will likely increase over time, as duration prompts further evidence of unwanted intrusion, almost irrespective of the sound level.

9. RATIONALE FOR TREATMENT PLAN

It was hypothesised that the stress reaction to the sound could be treated by a combination of three basic therapeutic (personal coping) strategies.

1) Reassurance, explanation, and support.
2) Relaxation therapy techniques.
3) General stress management advice and exercise (coping skills).

It was also hypothesised, based on the clinical analogy with phobia treatment, that some form of imaginal exposure therapy might help participants to desensitise, i.e. habituate to the sound. Hence, a fourth strategy is:

4) Imaginal exposure using "anchoring" of relaxation.

The final intervention constituted a later phase of treatment, developing out of the subjects' acquisition of basic relaxation skills.

The intervention methods are described in Appendix 2.

10 GROUP STRUCTURE AND ATTENDANCE.

The initial group of participants was composed of nine sufferers from low frequency noise related stress (LFNRS). One subject was accompanied by her husband. She requested his presence for emotional support, as she was subject to panic attacks, making a tenth participant. He took part in the exercise and provided useful feedback and comments from the perspective of someone observing a sufferer at close quarters. However, as he was not himself a sufferer, he was not included in the formal assessment of the group.

Most group members attended each workshop session. Some members missed sessions, however we anticipated that this would happen in the design of the project and had integrated much into the sessions that reconsolidated materials and practices, such that participant should still have been in a position to benefit sufficiently from their cumulative attendance at the other workshops, provided they had maintained their commitment to the homework assigned. One subject dropped out at an early stage, leaving eight subjects for before and after intervention comparisons.

The therapeutic interventions used in the project were delivered by means of a series of six two-hour group workshops. During the workshops, participants were assigned time to discuss their feelings about LFNRS and related issues, were given information and advice about heightening understanding of their symptoms and taught coping skills. A range of coping skills were discussed and practised, but the key intervention was a version of the Benson Relaxation Response Method (Benson 1975; Benson and Stuart 1996). This is an evidence-based relaxation therapy technique, widely used in stress management and psychotherapy. The Benson Method was taught and rehearsed in each session, supported and reinforced with other relaxation techniques derived from psychotherapy and self-hypnosis. Participants were given handouts explaining it as a protocol and assigned the homework of practising it twice daily for twenty minutes, and then feeding back on their experiences at the start of each subsequent session. Participants were also given a generic stress management CD, recorded by Donald Robertson, containing a number of common relaxation exercises to supplement their use of the Benson Method. Participants were asked to use those which they felt helpful, as regular "home work".

11 GROUP DYNAMIC AND CHARACTERISTICS

One of the most interesting aspects of the project has been the opportunity to informally observe how members of the group interacted with each other during the exercises they were given, and with the workshop facilitator. Discussions related to issues such as difficulties dealing with noisy neighbours, local authorities, as well as health issues. It was necessary for the group to feel that they had been listened to and given an opportunity to air their concerns, as to prohibit such discussion may have caused them to feel undervalued or ignored.

The number and variety of potentially stress-related symptoms among the group was notable. For example, group members complained of, or exhibited behaviour, which might be interpreted as paranoia, sleep problems, anger management problems, panic attacks, tremors, and headaches. We also noted that several group members reported a range of other hearing-related issues, e.g., tinnitus, hyperacusis and hearing impairment. The group also contained a number of members who had either resorted to, or considered, litigation in relation to their low frequency noise problem. Many of the group had also had acoustic measurements taken to locate the low frequency noise, but without a positive outcome.

Another interesting development was that, when group members were encouraged to discuss and compare their symptoms, it became clear that they experienced LFN in a number of ways. Of note was the fact that three group members said that they were more distressed by feelings of vibration, which they associated with LFN and other sounds, than by the sound itself. It is highly unlikely that LFN of the levels encountered could directly cause a physical vibration of this kind. Sometimes, not surprisingly, participants found it quite difficult to explain the stimulus or its effects and expressed frustration with the inability of non-sufferers to comprehend their experiences.

Other specific comments, made by participants on feedback forms, included the following. Three people said they found the Neural Linguistic Programming (NLP) technique known as Visual-Kinaesthetic Dissociation (VKD) to be helpful. Group members were introduced to this technique as part of a "menu" of possible coping skills for their evaluation. Four specifically mentioned that they found the Benson method helpful. Two noted that the stress management CD had been helpful. Three commented on the importance they placed on actually identifying the external source of the sound. Four mentioned that they found it helpful to meet and speak with other sufferers. Two mentioned that they found the use of white noise or pleasant background sounds useful for masking the LFN.

The small number of subjects limits the reliability of statistical deductions, but anecdotal evidence, based upon subjects' comments and responses to questions during the sessions, suggest that the majority of group members, felt significant improvement in their levels of stress associated with LFN. The high levels of adherence to the programme and attendance at sessions was surprising, and indicated that the participants placed considerable value upon the treatment they were receiving. Only one group member dropped out from the project, at a very early stage, due to illness (subject H).

12. COMPARISON OF BEFORE AND AFTER OUTCOMES

12.1 Low Frequency Noise Reaction Questionnaire

The questionnaire, which is shown in Appendix 1, contains questions which all relate to negative effects of the noise. It was scored in the following way. Each question answered was given a score from 0 to 4, where:

Not at all = 0
A little of the time = 1
A good deal of the time = 3
Most of the time = 4

The maximum score for the 27 questions is 108 and actual scores were expressed as a percentage of this.

Most subjects answered all of the LFNRQ questionnaire statements, but one very stressed subject, H, did not respond to 11 of the statements and he subsequently

dropped out of the study. Two others missed either one or two statements. The results are summarised in Figs. 3 and 4. A question not answered was left blank in the scoring.

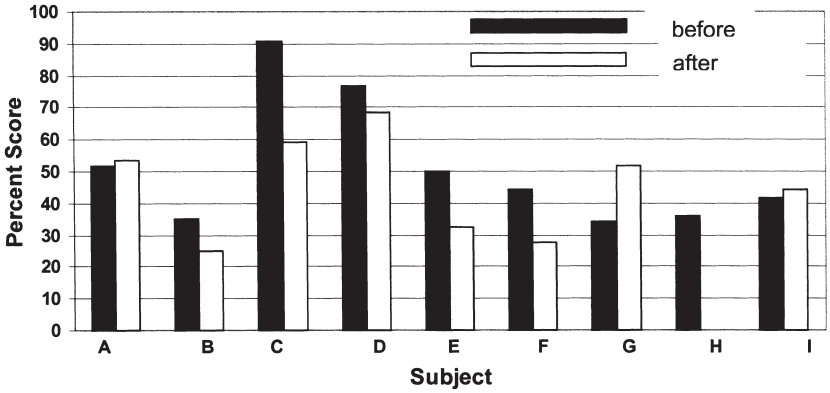


Fig 3. Subject's overall LFNRQ scores before and after the therapeutic sessions

Figure 3 commentary:
The bar score in Figure 3 indicates the overall level of impact experienced by each of the subjects, the higher the score the greater the adverse effects of the noise. The level of impact is shown in terms of a percentage of the 'amount' of impact. The results show that Subjects C and D displayed the highest overall unwanted reaction to their LFN experience. The remaining subjects are more moderately affected, as they typically score 30% to 50%, which is, on average, between "A little of the time" and "Some of the Time".

Consider Subject C, who showed a percentage reduction of about 30%. Overall, this subject's score, before and after intervention, dropped by more than 30 steps in the 0 to 4 scoring range of the answers, which is a good result.

For most other subjects, the overall trend suggests that the unwanted reactions to noise had been reduced relative to pre-intervention scores.

Results from the LFNRQ are shown in Figure 4, where the percentage of the total possible scores across subjects for the before and after therapy conditions are shown in question- specific scores. A lower score after the intervention indicates an improvement in the subjects' conditions.

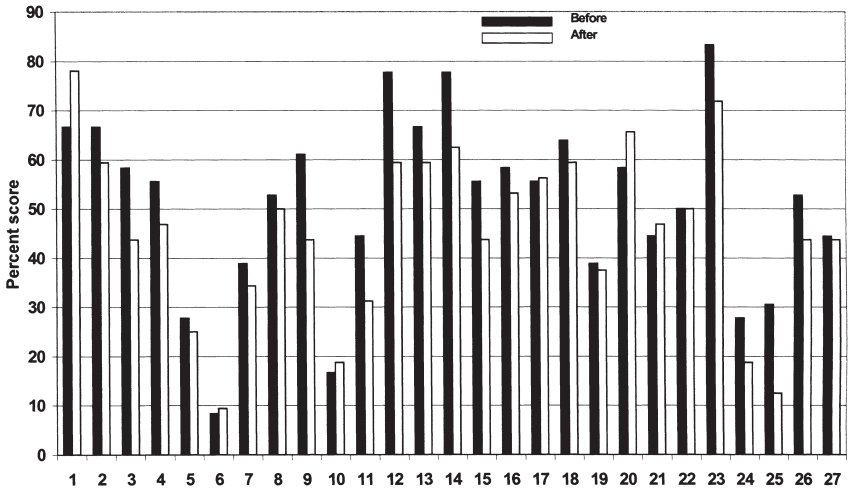


Figure 4 Total scores for the before and after therapeutic intervention shown for each question

Figure 4 commentary:
The strongest indicators of impact can be found from the questions;

Q. 23 (sleep disturbance) 83% reducing to 72% after therapy,
Q. 12 (interfere with enjoyment of life) 78% reducing to 59% and
Q.14 (made hard to relax) 78% reducing to 62%.

These are followed by:

Q. 1 (made worry) 67% increasing to 78%
Q. 9 (annoyed) 61% reducing to 44%
Q. 2 (made tense) 67% reducing to 59%
Q. 16 (made feel helpless) 58% reducing to 53%
Q. 13 (hard to concentrate) 67% reducing to 59%
Q. 20 (made avoid noisy situations) 58% increasing to 67%
Q. 3 (made irritable) 58% reducing to 44%
Q. 18 (interfered with work) 64% reducing to 59%
Q. 4 (made angry) 56% reducing to 47%

The least adverse effects in terms of percentage scores were

Q. 6 (led to avoid quiet situations) 8% increasing to 9%
Q. 10 (made feel confused) 17% increasing to 19%
Q. 5 (made cry) 28% reducing to 25%
Q. 24 (made think of suicide) 28% reducing to 19%
Q. 25 (made feel panic) 28% reducing to 13%

Consequently, the strongest effects relate to sleep disturbance, interference with enjoyment, relaxation, concentration and work, whilst leading to annoyance, anger, irritation, helplessness and avoidance of noise.

Following the therapeutic intervention, subjects' scores indicate improvement in the quality of their reactions to LFN, and this is not incompatible with elevated awareness of the noise.

12.2 Quality of Life Questionnaire

Subjects were again asked to complete the Q of L questionnaire after the series of therapeutic sessions. Scoring was the same as in Fig. 1, using the five point scale Strongly Disagree: Disagree: Neutral: Agree: Strongly Agree. The numbers for Strongly Disagree and Disagree are combined into one group, as are the numbers for Agree and Strongly Agree. The numbers who responded as Neutral are a single group.

The answers to the questions are shown in Figure 5, which gives subjects' responses under the three headings:

1. Agreement with the Q of L questions
2. Disagreement with the Q of L questions
3. Neutrality - unable to make a decision either way.

Figure 5 displays subjects' responses to questions which reflect the quality of their emotional and physical well-being, before and after relaxation therapy. Reading vertically down the three bar charts shows how agreement, disagreement and neutrality changed from before and after the relaxation sessions. The 15 Q of L questions are shown above, in section 7.3.

In Figure 5 questions 1 to 8 cover negative aspects of the subject's life, so that a reduction in subject numbers after the therapy is a positive outcome. Similarly, questions 9 to 15, cover positive aspects of the subject's life, so that an increase in the subject numbers after therapy is a positive outcome. There are eight replies to each question.

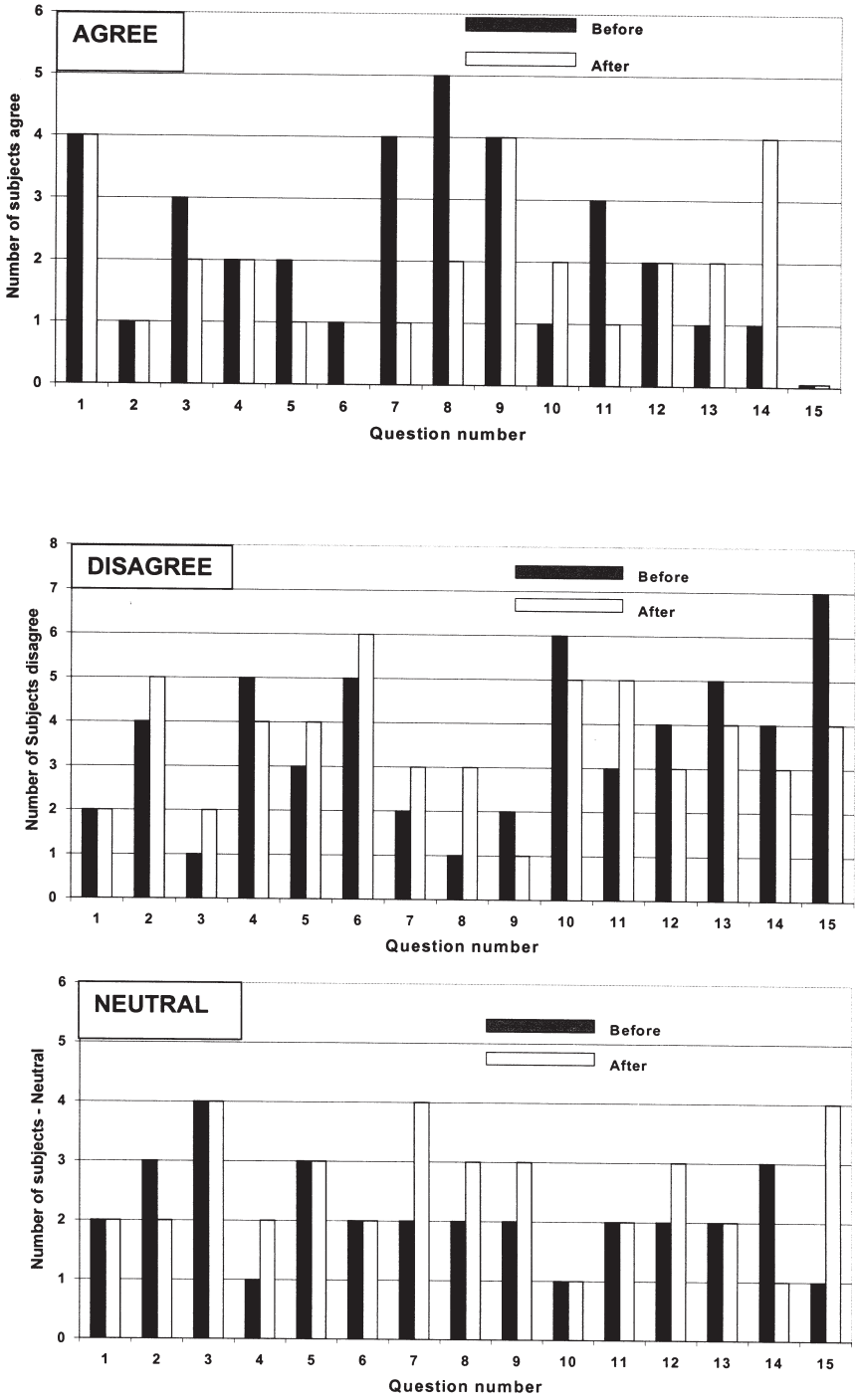


Figure 5 Quality of Life Scores before and after therapeutic intervention

The subjects' pre and post therapy Quality of Life scores show that a number of questions have elicited stronger positive scores. If the ratings can be taken as a commentary on subjects' quality of life, it can be seen that a number of aspects of their lives have improved. Subjects report that, in general, their physical well being (Q3, Q5, Q6, Q7 and Q8) showed signs of improvement, increasing or decreasing as appropriate. For example, consider Q8 - I feel anxious - see section 7.3. Prior to intervention, five of the group agreed with this statement, one disagreed and two were neutral. After intervention two agreed, three disagreed and three were neutral. This showed a trend away from the negativity in their lives, also shown by responses to other questions. For example, when ratings for questions 1-8 are examined

together, there is a trend towards decreasing negatives, although some of the improvement is in neutrality rather than disagreement e.g. Q7 (I feel sad). An increase in neutral response coupled with a decrease in negative response is a positive outcome.

Subjects report a continual battle against the noise and its intrusions. Many of the subjects have a long-term history with the noise and so it is against this well-established tendency that encouraging gains have been recorded.

The impact upon subjects' quality of emotional life, as recorded by the scores to questions 7 and 8, indicate improvement. The capacity to reassert control over the impact of unwanted stimuli is central to effective coping. The more effective the individual's coping the more likely that their mood will improve and confidence increase, leading to a lowering of anxiety.

The responses to question 11 (My work at home is fulfilling) were not consistent with the general trend towards improvement as, prior to intervention, three subjects agreed, compared with one subject after intervention. It is possible that the therapy roused increased expectations in this area.

Subjects' reports suggest evidence for an increased capacity to relax and to replenish energy levels - questions 12 - 14. This is a move towards breaking the downward pressure upon subjects' quality of life, a pressure that characterizes LFN complaints and accentuates stress. However, none of the subjects were experiencing a general level of quality of life with which they were content. The general level of well-being seems to have responded favourably as illustrated by scores of Question 14 (I am enjoying the things I usually do for fun) which shows particular improvement, where agreement rose from 1 subject to 4 subjects, caused by a drop of one in disagreement and a drop of two in neutral. None of the subjects are content with the quality of their life before and after intervention, but there is a move of three from disagreement to neutral.

Given the relatively short duration of the relaxation sessions in relation to long-term problems, it does seem that subjects were able to build an improved quality of life as a result of the therapeutic intervention. Overall, subjects' ratings indicate decreased sensitivity to the noise, and improved coping. The quality of home life appears to have also benefited, given the trend for improvement found in questions 7, 8, 13 and 14.

It is possible that benefits may have occurred as an artefact, a consequence of being able to share their feelings and experiences with other co-sufferers within a supportive group. However, results from the quality of coping questionnaire (shown below) suggest that subjects were able to apply control techniques, drawn from the group sessions, as scores on active coping questions also indicated improvement.

12.3 Quality of Coping Questionnaire

Overall, as the quality of life measures were showing improvement, the underlying mechanism supporting this was likely to be that of improved coping. Subjects' scores on the coping questionnaire, shown for each question in Figure 6, indicated stronger positive perceptions. (Answers were given on the same five-point scale as for the Quality of Life Questionnaire). The techniques acquired by subjects from the therapy sessions seem to have initiated an awareness of how to manage 'unwanted' responses, which supported them in countering the negative effects of failed coping. When asked to consider their longer term capacity to cope with the noise(Q.7) seven subjects, compared with three subjects pre-session, thought that they would now be able to regain happiness.

Figure 6 shows the Subjects' scores under three headings

1. Agreement with the coping questions
2. Disagreement with the coping questions
3. Neutrality - unable to make a decision either way

The coping questionnaire required response to the following eight questions, all of which, except Q.6, refer to inadequacy in coping.

Coping Questions

1. I have a hard time adjusting to the noise
2. Because of the noise I miss the things I like to do most
3. The noise makes me feel useless at times
4. The noise has made me more dependent on others than I want to be
5. The noise has made me a burden on my family and friends
6. The noise does not make me feel inadequate
7. I will never cope with the noise well enough to make me happy
8. I think that people are uncomfortable around me because of my problem with the noise

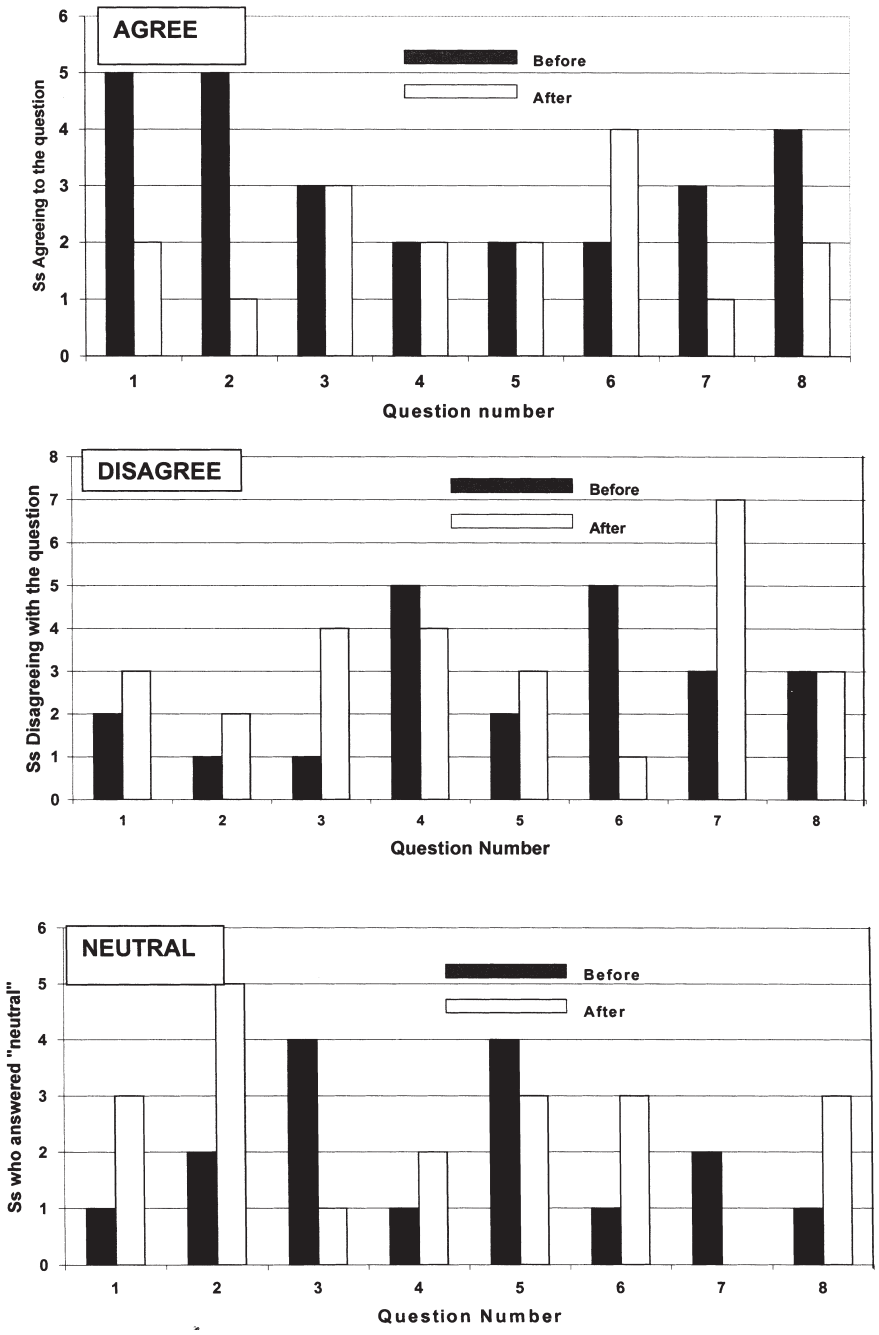


Figure. 6 Subjects who scored agreed, disagreed or were neutral on the coping questionnaire

Consider Figure 6. Reading vertically down the three bar charts shows how agreement, disagreement and neutrality changed from before and after the therapy sessions.

For example, Q.1 - I have a hard time adjusting to the noise - changed from 5 subjects agreeing to 2 subjects agreeing. Disagreement rose from 2 to 3 subjects, but there was a rise in Neutral from 1 subject to 3 subjects, showing a move towards disagreement, even though some subjects were not able to go all the way.

Q2 - Because of the noise I miss the things I like to do most - showed a marked change from agreement to neutrality.

Q3 - The noise makes me feel useless at times - showed no change in the number agreeing, but a reduction in neutrality leading to a rise in disagreement.

Q4 - The noise has made me more dependent on others than I want to be - showed little change.

Q5 - The noise has made me a burden on my family and friends - also showed only small changes.

Q6 - The noise does not make me feel inadequate - here there is a clear change towards agreeing with the question or to neutrality.

Q7 - I will never cope with the noise well enough to make me happy - Again a clear change towards disagreement and with no neutral answers after intervention.

Q8 - I think that people are uncomfortable around me because of my problem with the noise - therapy led to a reduction in agreement and increase in neutrality.

The scores show an overall improvement in the Quality of Coping. There is an improvement in the scores on questions of a positive and forward looking nature, for example, (Q.7) which is a key element in building positive behaviours, and which can provide a future - oriented source of goal directed reinforcement, rather than that commonly described by sufferers, which is firmly anchored in the past. In this past context, failed coping is a self-fulfilling and self perpetuating process, the cycle repeats and learning only serves to reinforce the original view, leading to a restricted and distorted range of response options. It is possible that the therapeutic interventions were able to provide subjects with a wider range of response options, breaking with learned behaviours and thereby countering previously acquired responses and associated negative emotions.

The responses to Q. 6 (inadequacy) are evidence of an increased capacity to meet the emotional demands placed upon them by LFN. The more effective an individual's strategies are for managing the impact of noise, the more likely they are to experience the noise as less intrusive and annoying. This would support reduced anxiety levels; associated reduction in anxiety and negative mood states.

One of the most telling aspects of individual's experience of trying to cope with LFN is the gradual yet apparently remorseless growth of disconnection between them and others who do not have their noise experience.

The post therapy ratings for Q.8 indicate that subjects have reduced concerns about how their responses to the noise may impact upon their relationships with others. Given the general trend towards improved coping, this may be taken as evidence that the noise is less of a concern to the subject and therefore figures less in their lives with others. If subjects were beginning to feel the benefits of improved coping and associated quality of life, they may perceive themselves as placing fewer demands upon their partners/others and consequently be more at ease with the relationship. Again, this would be a significant development, as many LFN sufferers describe a focus that frequently dominates relationships with others.

Figure 6 illustrates that the relaxation therapy has led to an improvement in coping capacity for subjects whose experience is that of living with LFN at home.

The group had also been invited to make written comments on their perceptions of the therapy. The following comment from one member summarises what the group generally found to be helpful:

> "Focusing awareness on the *choices* one has to assist in coping. Greater detachment from the noises and from one's own stress reactions. Letting go of negative feelings towards the people making the noise and trying

to deal with only the noise.
Having "somewhere to go" via the visualisations, which gives a sense of power and control.
Sharing the group with fellow sufferers."

13. CONCLUSIONS

A group of subjects, long term complaints of noise, and in particular of low frequency noise, were introduced to relaxation and related therapies as a means of relieving the worst symptoms consequent on the noise. Nine subjects commenced the therapy, but one dropped out due to illness, so that deductions have been based on the same eight subjects before and after the therapy sessions.

Subjects responses to 'their noise problem' were elicited by questionnaires administered before and after the therapeutic sessions. There is a clear trend in the overall results that suggests an improved Quality of Life and improved Coping with Life. The number of subjects is insufficient for meaningful statistical deductions, but a review of the comments of the subjects on changes in their personal situations showed that, whilst all reported benefits, some felt that they had derived considerable benefit.

Relaxation and other psychotherapeutic techniques have been shown to be useful interventions in reducing the stress caused by noise problems, such as those from some low frequency noises for which a technical noise control solution was not available.

At the time when this work was in progress, it was not known to the authors that the UK National Health Service was investigating two computerised Cognitive Behavioural Therapy (CCBT) programs, one for use in cases of mild to moderate depression and the other for anxiety/phobia problems. See
http://guidance.nice.org.uk/TA97.

Primary Care Trusts were required to make these packages available to patients from 31 March 2007.

This supports the case that the CCBT has promised as a means of helping those who suffer from a variety of problems, including noise which cannot be solved technically, and who exhibit symptoms of both depression and anxiety. These people are often widely dispersed and it may not be possible to give them individual or group attention. The next requirement is to develop computerised Cognitive Behaviour Therapy for use by noise sufferers.

ACKNOWLEDGEMENT.

This work was carried out under contract for Defra and the Devolved Administrations.

14. APPENDIX 1 LFN REACTION QUESTIONNAIRE.

The following questionnaire was completed by the subjects before the start of the relaxation sessions to be returned by post. It was completed again at the end of the sessions with modification to the final three questions. Each question was answered by ticking off one of the following choices

Not at all A little of the time Some of the time
A good deal of the time Most of the time

1 *The noise has made me worry about it*
2 *The noise has made me feel tense*
3 *The noise has made me feel irritable*
4 *The noise has made me feel angry*
5 *The noise has made me cry*
6 *The noise has led me to avoid quiet situations*
7 *The noise has made me feel less interested in going out*
8 *The noise has made me feel depressed*
9 *The noise has made me feel annoyed*

10 The noise has made me feel confused
11 The noise "drives me crazy"
12 The noise has interfered with my enjoyment of life
13 The noise has made it hard for me to concentrate
14 The noise has made it hard for me to relax
15 The noise has made me feel distressed
16 The noise has made me feel helpless
17 The noise has made me feel frustrated with things
18 The noise has interfered with my ability to work
19 The noise has led me to despair
20 The noise has led me to avoid noisy situations
21 The noise has led me to avoid social situations
22 The noise has made me feel helpless about the future
23 The noise has interfered with my sleep
24 The noise has led me to think about suicide
25 The noise has made me feel panicky
26 The noise has made me feel tormented
27 The noise has forced me out of my home

Please write a few words in response to the following questions. Use a separate sheet if necessary.

28 Do you believe that you know the source of the noise?
29 Have you tried to have it reduced, and with what result?
30 Give any other comments which you wish.
And finally, are you taking prescription drugs? Some of the standard prescription drugs may have side effects on hearing. If you are taking any, could you have a look at the pack and write the name below.

15 APPENDIX 2. BASIS OF THE THERAPEUTIC INTERVENTIONS

15.1 The Benson Relaxation Response

In 1960, at Harvard Medical School in the US, the cardiologist Herbert Benson developed the "Relaxation Response" approach to stress therapy. Benson's research into human physiology showed that, in addition to the established 'fight or flight' response, the body possessed the ability to deliberately induce a counterbalancing state of physical rest and emotional calm which he labelled the 'relaxation response'.

> When the mind is focused, whether through meditation or other repetitive mental activities, the body responds with a dramatic decrease in heart rate, breathing rate, blood pressure (if elevated to begin with), and metabolic rate - the exact opposite effects of the fight-or-flight response. (Benson, 1975: 9).

Benson discovered that the relaxation response could be elicited in a number of ways, using established relaxation techniques from yogic meditation to progressive muscle relaxation. Once he had established the existence of a measurable and clinically significant relaxation mechanism, Benson proceeded to search for the simplest possible protocol capable of inducing it. He concluded that the relaxation response could be elicited by a combination of two essential factors:

1. A monotonous mental stimulus. That is, a sound, word, phrase, or prayer repeated silently or aloud, or a fixed gaze at an object.
2. A passive mental attitude. Not worrying about how well one is performing the technique and simply putting aside distracting thoughts to return to one's focus. (Benson, 1975: 10).

Consequently, Benson developed his own protocol, which generally involves the subject sitting in a comfortable chair, with eyes closed, repeating a simple word, such as "one" or "peace", on each exhalation of breath. This is usually done for 10-20 min. twice per day, on an ongoing basis.

The group of low frequency noise sufferers were taught how to use the Benson Method by means of group exercises facilitated by Donald Robertson. They practised the technique at home and discussed their experiences at the start of each session, where time was allocated to coach them through any difficulties in technique and to answer their questions and offer emotional support, reassurance and encouragement.

The Relaxation Response protocol, sometimes known as the "Benson Method", is currently one of the most popular relaxation techniques in modern stress management and psychological therapy. It is often used in conjunction with cognitive-behavioural therapy (CBT) (Beck 1976) and other solution-focused and evidence-based interventions in psychological therapy.

A recent clinical literature review published in the British Medical Journal (BMJ) outlines some of the most reliable research evidence on the effects of relaxation techniques like the Benson Method and self-hypnosis.

> There is good evidence from randomised controlled trials that both hypnosis and relaxation techniques can reduce anxiety, particularly that related to stressful situations [...]. They are also effective for panic disorders and insomnia, particularly when integrated into a package of cognitive therapy [...] (Vickers and Zollman 1999)

Anxiety, panic attack, and insomnia are typical stress-related symptoms, which were found to be particularly common among the sample group of LFNRS sufferers, and which are believed to be frequently encountered among LFNRS sufferers in general.

15.2 Principles of Relaxation Therapy

"Relaxation therapy" is a broad term that encompasses a range of different therapy interventions. However, most of these techniques are essentially highly directive and systematic processes which exhibit measurable physiological results, and can therefore be considered as a form of behaviour therapy. Most forms of relaxation therapy are evidence-based and reasonably well accepted - though that does not mean widely practised - within mainstream medicine. In that respect they can legitimately be considered as a branch of "orthodox" therapy rather than "complementary and alternative medicine" (CAM). That is, "Relaxation" and "Stress Management" are found to a certain extent within conventional medicine (Vickers and Zollman 1999). The BMA define 'relaxation techniques' as follows:

> Methods of reducing muscle tension to achieve mental calm. Can assist people with anxiety, help reduce hypertension, and relieve stress. (BMA 2002)

Most forms of relaxation therapy entail teaching specific coping skills designed to induce the relaxation response in the body. There are a number of different ways of achieving this, e.g., progressive muscle relaxation, guided visualisation, meditation, self-hypnosis, breathing exercises, contemplative meditation, etc.

When relaxation techniques are taught by a facilitator or therapist, the results tend to be more pronounced. This is partly because simple processes, such as the Benson Method, can be combined with longer and more sophisticated exercises facilitated by the group leader. For instance, the LFN group were "talked through" lengthy progressive muscle relaxation exercises and visualisation techniques which led into the practice of the Benson Method, which they were instructed to use at home. This meant that they were already in a fairly relaxed state before commencing the part of the exercise which they were to repeat at home. When they then practised the Benson Method between sessions, they could recall the relatively deep levels of relaxation which they were coached into during the sessions, making it easier for them to make progress in developing their coping skills for relaxation.

The effects of deep relaxation are numerous. For example, in their advice book on stress, the British Medical Association asserts that systematic relaxation "Improves sleep, increases mental and physical performance, combats tiredness, decreases anxiety and tension" (Wilkinson 2004). Essentially, it evokes a physiological and psychological state which is the opposite of, and mutually

exclusive with, the state of stress, including the kind of stress that appears to result, in certain cases, from exposure to LFN.

One advantage of this approach is that it is well suited to group work as well as individual therapy. This obviously makes it easier to carry out research on efficacy, and renders the therapy more cost-effective to deliver.

In addition to dealing with the symptoms of low frequency noise stress, an attempt was made to induce the process of habituation to the low frequency noise stimulus by means of a process of "imaginal exposure" otherwise known as systematic desensitisation, described as:

> A technique of behaviour therapy, developed in the 1950s by the South African-born US psychiatrist Joseph Wolpe (1915-97) for treating phobias in particular, in which each member of a hierarchy of increasingly anxiety-provoking imaginary situations involving the phobic stimulus is repeatedly paired with a response that is physiologically incompatible with fear and anxiety, such as deep muscular relaxation [...] (Colman 2001).

The clinical analogy with phobia treatment was assumed, as many of the sufferers described their reactions to LFN in terms which resembled phobia. Once group members had mastered the basics of relaxation therapy, an element of imaginal exposure was introduced by guiding them through the process of imagining themselves to be in the place where the sound typically occurs and remembering the sound and any accompanying sensations (often described as "vibrations"), while maintaining and reinforcing their sense of emotional calm and physical relaxation. This was done repeatedly, in an attempt to neutralise the anxious and stressful feelings associated with LFN by pairing it with the relaxation response.

REFERENCES

Beck, A. T. (1976). "Cognitive Therapy and Emotional Disorders." *New York: International Universities Press*.

Belojevic, G., et al. (2003). "Noise and Mental Performance: Personality Attributes and Noise Sensitivity." *Noise and Health* **6**(21): 67-89

Benson, H. (1975). "The Relaxation Response." *New York: Harper Collins*

Benson H. and E. Stuart (1996). "The Wellness Book: The Comprehensive Guide to Maintaining Health and Treating Stress-related Illness." *Citadel Press*.

BMA (2002). "Relaxation Techniques." *In: Illustrated Medical Dictionary London: Dorling Kindersley.*

Colman, A. (2001). "The Oxford Dictionary of Psychology. Oxford: OUP."

Leventhall, H. G. et al., (2003). "A review of published research on low frequency noise and its effects. Report for Defra." http://www.defra.gov.uk/environment/noise/research/lowfrequency/pdf/lowfreqnoise.pdf.

Vickers, A. and C. Zollman (1999). "Clinical review: Hypnosis and relaxation therapies." *BMJ* **319**: 1346-1349.

Wilkinson, G. (2004). "Understanding Stress." *Dorset: Family Doctor Publications*.

Wilson, P. H. et al. (1991). "Tinnitus reaction Questionnaire: Psychometric Properties of a Measure of Distress Associated with Tinnitus." *Jnl. Speech and Hearing Research* Vol.**34**: 197 - 201.