

basicprofile

Fundamental People Business

Construction & Reliability

Forced Choice

By obliging people to choose between two equally valued alternatives, they will remain genuine in answering. The great advantage of forced choices, is that the social desirability will be largely ruled out. Searching for the `right` answers, one of the big contaminants testing personality aspects becomes very difficult this way.

72 choices

Making choices can be energy consuming and the mental exhaustion can contaminate the reliability of the results. BPS contains 72 choices and should not take more than fifteen minutes. This amount of time ensures that also the very last answers will be given with focus and not just randomly chosen.

9 Primary, neutral variables

Fear and anger are two primary emotions. Willingness to listen and the need for exploring are primary personality traits. They are elementary in contrast to for example, motivation to perform, ambition or perfectionism that cover a whole range of urges, attitudes and needs.

The advantage of primary variables is that people will be less consciously aware of their relative position towards others. Because of this, people will be inclined much less to adjust or steer their answers to certain directions.

Furthermore, there are no good or bad answers. All variables contain both good and bad sides. Of course each score on a variable does say something but what really counts, is this score in relation to the eight other scores. A high score can be good or bad (suitable, less suitable), depending on the specificity of the question/job AND on how this score relates to the remaining other scores.

The social desirability (already diminished by the ipsative format) is hereby even further eliminated because the variables themselves are not manifest socially desirable.

The elementary and neutral character of the variables does not only make the test more reliably measurable, but also and even more, it allows us to uncover a great range of more complex variables. The primary variables are, in different words, also valuable keystones for PsychoLogica or the logic framework around psychological concepts such as attitudes, affinities, needs and urges.

Independent Variables (1700 tests)

	Receive	Send	Dictate	Think	Do	Note	Explore	Brake	Steer
Receive	1	-0,025	-0,227	0,054	-0,204	-0,003	-0,08	-0,022	-0,153
Send	-0,025	1	0,132	0,012	-0,306	-0,232	0,102	-0,255	-0,027
Dictate	-0,227	0,132	1	-0,116	-0,298	-0,109	-0,004	-0,357	0,111
Think	0,054	0,012	-0,116	1	-0,037	0,039	-0,014	-0,134	-0,248
Do	-0,204	-0,306	-0,298	-0,037	1	-0,197	-0,12	0,063	-0,204
Note	-0,003	-0,232	-0,109	0,039	-0,197	1	-0,358	0,126	-0,025
Explore	-0,08	0,102	-0,004	-0,014	-0,12	-0,358	1	-0,346	-0,169
Brake	-0,022	-0,255	-0,357	-0,134	0,063	0,126	-0,346	1	-0,08
Steer	-0,153	-0,027	0,111	-0,248	-0,204	-0,025	-0,169	-0,08	1

These correlations reflect the fluctuations between the different variables. The closer these correlation values are to zero, the stronger the independency between them. The correlation between Receive and Send is -0.025. The highest positive correlation is the one between Note and Brake and is 0.126. The strongest negative correlation is the one between Explore and Note and is -0.358.

The common measured variability is the square of this correlation and reaches in the worst case a minor 10%. These results justify the combining of the variables to more complex constructs and the (psycho)logics behind it.

This result is even further confirmed by the results of a Principal Component Analyses where at least six components are needed to explain more than 80% of the observed connection between the variables. Even the 8th component still explains 7% of the correlation!!

Results Principal Component Analyses (on correlation matrix above)



School of Public Health
 U.Z. Sint-Rafaël
 Kapucijnenvoer 35
 B-3000 Leuven
 Belgium

Steffen Fieuw

Tel: +32/16/33.68.97

Fax: +32/16/33.70.15

E-mail: steffen.fieuw@med.kuleuven.be

Eigenvalues of the Correlation Matrix

	Eigenvalue	Difference	Proportion	Cumulative
1	1.91528444	0.42538939	0.2128	0.2128
2	1.48989506	0.15030794	0.1655	0.3784
3	1.33958711	0.28737011	0.1488	0.5272
4	1.05221701	0.18808375	0.1169	0.6441
5	0.86413326	0.05661885	0.0960	0.7401
6	0.80751441	0.04293783	0.0897	0.8298
7	0.76457658	0.11802378	0.0850	0.9148
8	0.64655280	0.52631347	0.0718	0.9866
9	0.12023933		0.0134	1.0000

Proportion=Proportion of explained variation by each of the nine components (factors)

Reminder; each component is a linear combination of the nine original variables.

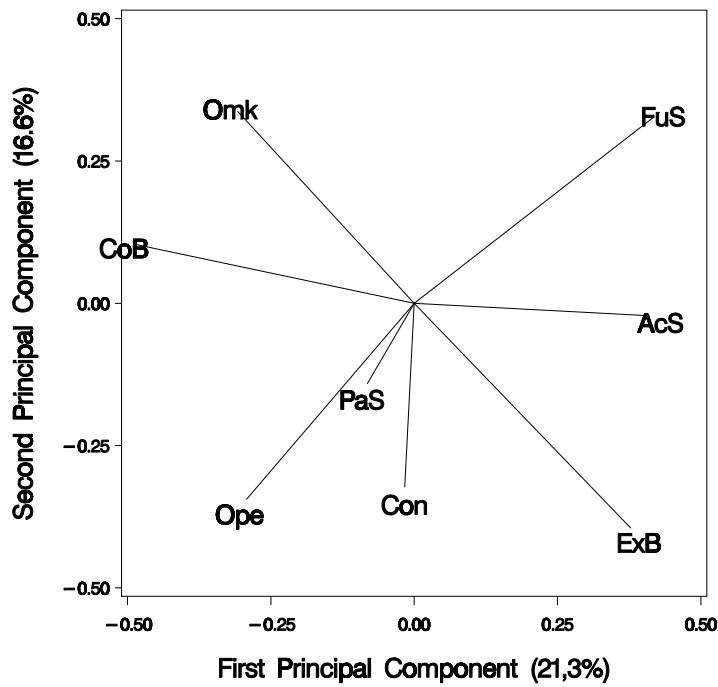
Eg.: $\text{Component1} = b_1 \cdot \text{ReC} + b_2 \cdot \text{Send} + b_3 \cdot \text{Dictate} + \dots$

The coefficients (b_1, b_2, \dots) in this linear combination point to the amount that each variable suits or covers the components. In other words, it is the loading for the variables, the weight they carry to capturing the components. These loadings help interpreting the components.

Variable Prin1 Prin2 Prin3 Prin4 Prin5 Prin6 Prin7 Prin8 Prin9

Receive	-0.09	-0.16	0.55	-0.51	-0.33	0.29	-0.36	0.02	0.29
Send	0.43	-0.02	0.20	-0.18	0.72	0.04	0.03	-0.38	0.26
Dictate	0.43	0.34	-0.08	0.34	-0.06	-0.09	-0.56	0.35	0.37
Think	-0.02	-0.34	0.40	0.55	0.12	0.36	0.33	0.37	0.17
Do	-0.31	-0.36	-0.54	0.16	0.02	0.31	-0.23	-0.31	0.46
Note	-0.32	0.35	0.37	0.37	-0.18	-0.33	0.11	-0.49	0.33
Explore	0.39	-0.41	-0.12	-0.14	-0.36	-0.47	0.38	0.04	0.39
Brake	-0.51	0.11	-0.05	-0.27	0.40	-0.33	0.07	0.50	0.36
Steer	0.11	0.56	-0.21	-0.20	-0.17	0.49	0.49	0.04	0.29

The figure below is the graphical representation of the original variable shaped in space by the two most important principal components. Attention. Because together they only explain 37,84% of all observed variation in the correlations, this image is a very simplified representation.



- Omk= Note
- CoB=Brake
- Ope=Do
- PaS=Receive
- Con=Think
- FuS=Dictate
- AcS=Send
- ExB=Explore

BasicProfileTest Reliability

Reliability of a test makes the test valuable for use. For this it should be reliable and valid. Reliability refers to the stability of measurement. Validity means the measurements are measuring that what it is thought to measure. Further there is the reliability of the norm/reference scores pointing to the active Flemish population as a proper reference group.

STABILITY

1. STABILITY of the basic profiles

Investigating the stability of the BasicProfiles checks to what degree people profile or describe themselves in a stable manner over a period of time.

Because people evolve this seems somewhat contradicting. We assume actually the profile scores would change over time. So we will be actually measuring both the stability of professional profiles as well as the reliability of the test itself.

The BasicProfile of most people actually seems to be stable over time. Even during short but very intense experiences as with joining the TV show “Expedition Robinson”, does not cause fundamental changes. By means of test-retest (3 months up to 2 years) we found a median score of .78*, which is a high score especially for personality tests. Only with intelligence tests this score would be higher, but IQ is much less subject to evolving situations.

*.78 is a value of correlation representing the amount of similarity between first and second test.

2. STABILITY of the variables

Here we will check which variables will show to be more consistent over time. Also here we will have to take into account the natural variability of our professional personality due to changes around us and our adaptive nature.

Stability: Consistency or reliability score between +1 and -1

Do	.83
Send	.81
Think	.74
Explore	.71
Brake	.71
Dictate	.63
Receive	.56
Steer	.55
Note	.49

Assessment of the Criterion Validity of the Basic Profile Screen ® test

(April 22, 2008)



School of Public Health
U.Z. Sint-Rafaël
Kapucijnenvoer 35
B-3000 Leuven
Belgium

Steffen Fieuids

Tel: +32/16/33.68.97

Fax: +32/16/33.70.15

E-mail: steffen.fieuids@med.kuleuven.be

Purpose

The aim of the present study is to verify the extent to which the Basic Profile Screen ® test discriminates candidates for a specific tv-program from a reference population. More specifically, six hypotheses have been formulated with respect to differences in scale scores between candidates for the tv-program Expedition Robinson (the Netherlands and Flanders, period 2005-2006) and viewers of a tv-program called 'M/V' broadcasted on the channel Talpa during 2006-2007, which focussed on gender differences. Candidates for the tv-program Expedition Robinson (ER) had to fill in the questionnaire as a part of the selection procedure. Information is available from 370 subjects. Viewers of the tv-program 'M/V' could fill in the questionnaire on internet after the tv-program has been broadcasted. Information is available from 4714 viewers, restricted to the age range of the ER-candidates. This set of 4714 viewers, strictly spoken not a random sample, is denoted the "reference sample" in the remainder of the report.

Description of samples

The ER sample consists of 54.9% (203/370) males and 41.1% females (167/370). The median age equals 31 years (range: 18-68 years). The reference sample consists of 31.9% (1502/4714) males and 68.1% females (3212/4714). The median age equals 30 years (range: 20-62 years). Due to the age restriction in the reference sample, no significant age difference exists with the ER sample (Mann-Whitney U test, $p=0.32$). On the contrary, there is a significant imbalance with respect to the gender distribution (χ^2 -test, $p<0.0001$).

Specific Hypotheses

Six specific hypotheses have been formulated pertaining to six scales who should discriminate between both samples, based on the assumed content of the scale and the difference in behavior and/or expectations between both samples.

Subjects in the ER sample are expected to have:

- (1) higher scores on the Send scale,
- (2) higher scores on the Do scale
- (3) higher scores on the Dictate scale
- (4) higher scores on the Explore scale
- (5) lower scores on the Note scale
- (6) lower scores on the Brake scale

Assessment of validity

For each of the considered scales, histograms are constructed to visualise the distribution of the scores in both samples. Due to the imbalance of the gender distribution between both groups, this has been done for males and females separately. Mann-Whitney U tests are used to compare the distribution of the scores between both groups. Using a regression model with the scale score as response, an estimate is obtained for the difference between both samples, after correction for gender. Moreover, it has been verified if the difference between both samples depends on gender and/or age by including the interaction terms (interaction between sample and gender, interaction between sample and age) in the model. Due to the large number of subjects, small negligible differences are expected to be statistically significant. Therefore, focus of interest is not the p-value pertaining to the difference between the ER and reference sample, but the size of the difference (which has been expressed relatively to the standard deviation of the scores). Moreover, the discriminatory ability of the scale score has been quantified by the c-index (which corresponds to the area under the receiver operating curve) obtained with a logistic regression model (with sample as dependent variable and the scale score as predictor). A c-index of 0.5 and 1 correspond to no discrimination and perfect discrimination respectively. Finally, the use of a multiple logistic regression model allows to assess the combined discriminatory ability of the six scales, by considering them simultaneously as predictors.

Results (please let us know if you want to see the tables)

Table 1 in the addendum contains descriptive information in both samples for each of the six scales separately. Tables 1a and 1b present the same information for respectively males and females. In all situations, the Mann-Whitney U test is significant ($p < 0.0001$) confirming a statistically significant difference between males and females.

Table 2 gives an overview of the key results of this validation study. Estimates are given for the difference between the ER and reference sample based on a regression model with age, sample, gender and the interaction between gender and sample as predictors. In all situations, the difference between the ER and the reference sample is significant ($p < 0.0001$), for males as well as for females¹. Moreover, the magnitudes of the differences clearly indicate that the effect is substantial: in all but one situation, the effect sizes exceed 0.50, indicating that the difference between both samples is bigger than half a standard deviation of the distribution of the scores. Further, the c-indices from the logistic regression models using the score as a covariate are relatively high. Especially the 'administrative' and the 'controlling needs' score have strong discriminatory ability. Finally, combination of the six scores to discriminate between the ER and the reference sample yield high c-indices of 0.87 and 0.907 for males and females respectively.

Conclusion

The six scales from the Basic Profile Screen ® test discriminate very well between the ER- and reference sample. The scale differences confirm the formulated hypotheses, are all significant and of substantial size.

¹ Note that for some scales there is evidence that this difference is not of the same magnitude for males and females. However, this is not crucial for the current validation, since none of the hypotheses formulates a specific expectation with respect to this difference.