Leonardo 3. 4. 5.

Scientific accuracy of the methodology

Dr. Heiko Hansjosten

August 2015

Table of Contents

Executive summary

The Leonardo 3.4.5 is a research-based tool that helps identify the potentials and the talents of the individuals inside their working environment. The Leonardo 3.4.5 tool is categorizing the personal attributes of an individual within eight categories, which are formed within four major axes: i. Introvert- Extrovert (E-I), ii. Conceptual- Practical (C-P), iii. Feeling- Analytical (F-A), iv. Open- Structured (O-S).

In this paper the accuracy and reliability of the tool was tested on a big database of users (8058) compiled the past five years. The question was redesigned based on the results of the Leonardo 3.4.5 statistical report of 2009. Applying the Cronbachs alpha methodology (split-half) for each individual axis category tested internal reliability. The highest internal reliability was achieved within the Conceptual-Practical dimension and the lowest on the Open- Structured dimension. However for all dimensions with the exception of a few individual questions of the O-S category were acceptable and can be considered reliable.

The construct validity of the Leonardo 3.4.5 tool was tested on the basis of two criteria: the convergent and discriminant validity. The analysis proved the construct validity of the Leonardo 3.4.5 methodology in general and for the four axis categories and pointed to the need of rethinking of a very limited number of questions (6 in total).

Overall the results on the reliability and internal construct of the test were *positive* indicating that the scientific approach used by the Leonardo 3.4.5 is yielding reliable results on the domain of personal assessment.

The Leonardo 3.4.5 was also evaluated as part of a customer post market research analysis. The assessment was performed in two highly different groups namely an academic community and a set of top managers of a big company. In both cases the reception of the methodology was highly positive with 75% and 82% participants of the respective groups judging the methodology positive and giving a rating eight or more on a scale of zero to ten.

1. Demographics analysis

The reliability analysis was carried out on data of 8058 participants that participated on the Leonardo 3.4.5 assesment. The data were collected for a time period of five years from 2010 to 2015 after a review of the questionnaire in 2009. The review was based on the statistical analysis report of 2009. The demographics of the participants are presented in the following sections.

1.1 Age

On average the users were 34.5 years old, with the youngest user being 18 and the oldest 75. The standard deviation of the five-year sample was 10.8 and the distribution was highly skew to the left to the age of 25 (Fig.1).

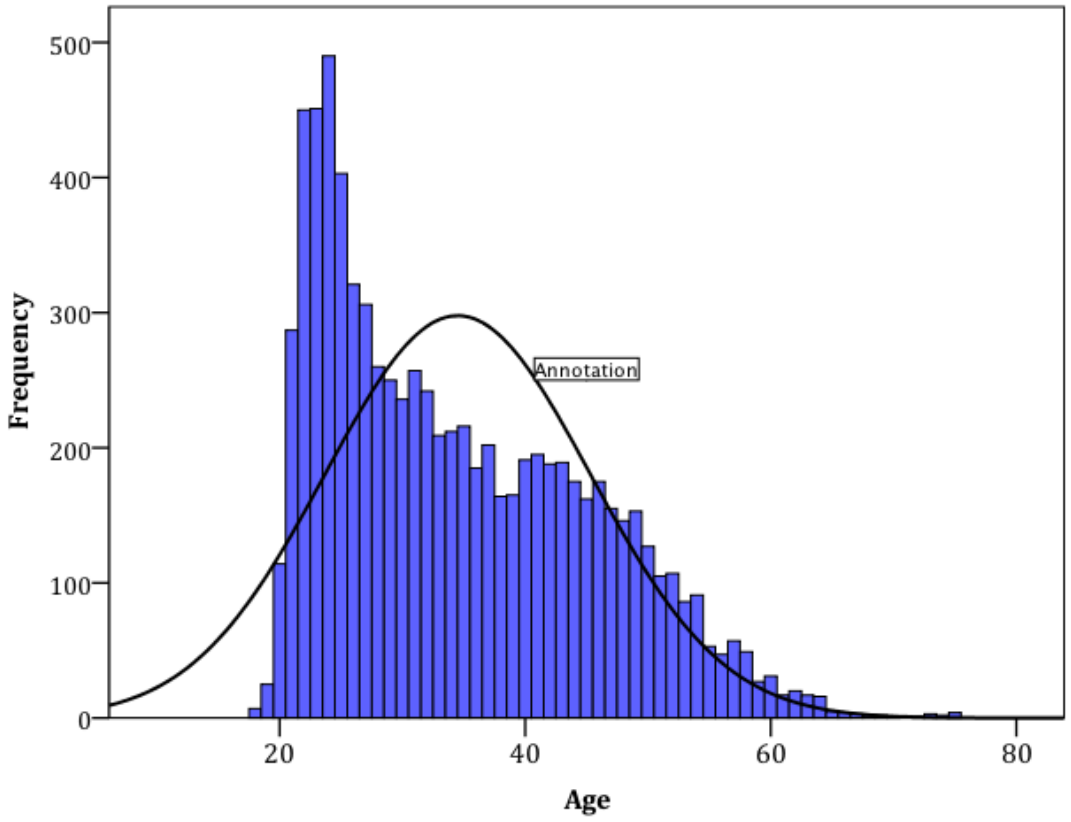

Fig 1. Age distribution of 8058 participants

1.2 Language profile

Main speaking language was not proportionally distributed between the respondents survey. Out of the 8058 participants, 67,4% was using German as

their main language, 23,9% French, 3,1% English and 5,6% Italian (Table 1). This disproportional distribution of languages was due to the fact that the survey was contacted mainly on Switzerland and Germany where although most people speak more than one language the frequency of German speakers is significantly higher. However based on the analysis performed and presented in earlier reports no significant differences were found between the spoken language of the participants and their test results.

Table 1. Language distribution of Leonardo 3.4.5 users

Language	Total number of users	% of Users
German	5428	67.4
French	1927	23.9
English	249	3.1
Italian	454	5.6

2. Leonardo 3.4.5 reliability analysis

2.1 Reliability measures and the Cronbachs Alpha test

Reliability is defined as the total consistency of a measure. If a measure produces similar results under the same conditions then the measure is considered to be highly reliable. A good example of such a measure is the human height, as it is expected to be the same every time it is measured within the same time period.

The reliability analysis allows identifying the critical items on a measure that affect the consistency of the test and thus should no longer be used within a tool. The reliability of a tool can be measured with a number of empirical tools, where the temporal stability and equivalent measurements are tested. One of the simplest reliability tests is the test- retest method. In this method the subjects are presented with the same measure twice and the correlation between these two values gives an estimate of the reliability. As this method works under the assumption that the true values between the measures remain unchanged, it can lead to significant estimation errors and thus it is rarely used.

During the administration of the Leonardo 3.4.5. test to our subjects retesting was not possible for practical reasons. Individual behavior within a working group is subjected to many influences and cannot be expected to be stable on the long term. Therefore any test of reliability that would retest the subjects within a

short time period would lead to a high reliability coefficient that could be proven wrong in the long term and therefore should be avoided.

A more relevant way to test reliability would be to take many repetitive measurements within the same time period using different measuring instruments. This method is difficult to reproduce in practice but it could be applied in this particular database if several of the questions that were under the same category produced similar scores. This method is called internal consistency test. Internal consistency is usually measured by using Cronbach's alpha, which is calculated based on the pairwise correlations between items within the same test. When using Cronbach's alpha to check for internal consistency of the data the following rule applies on the calculated alpha values.

Table 2. Internal consistency description based on Cronbachs alpha test

Cronbach's alpha	**Internal consistency**
$\alpha \geq 0.9$	Excellent (High-Stakes testing)
$0.7 \leq \alpha < 0.9$	Good (Low-Stakes testing)
$0.5 \leq \alpha < 0.7$	Acceptable
$\alpha < 0.5$	Unacceptable

Alpha can take values from zero to one. Values under 0.5 are considered unacceptable indicating a high unreliable test. In practice values between 0.5 and 0.7 are acceptable with values between 0.7 and 0.9 considered the best indicators of the internal stability of a tool.

2.2 Reliability within the Extrovert- Introvert (E-I) dimension

The reliability coefficient alpha within the subtests of the Extrovert- Introvert axis was 0.597. The test sample was 6904 individuals with a full item scale of 20.

In Table 3 the alpha values for each of the E-I parameters are shown. In general the alpha values varied from a value of 0.532 the lowest to a value of 0.685 the highest. Although the alpha values in this category are in general lower than 0.8, they are still within the acceptable range for consistency of the test. Further investigation therefore is needed within the subcategories of the users in order to understand more on the nature of these correlations. However the removal or rethinking of some parameters (i.e. EI42, EI54, EI68) would improve the internal consistency of the category.

Table 3. Statistical analysis for the Extrovert- Introvert dimension

	Corrected itemized scale correlations	Cronbach-Alpha,
EI1	-,336	,642
EI5	-,594	,685
EI7	-,181	,636
EI12	-,251	,639
EI15	,296	,570
EI16	,379	,557
EI20	,251	,578
EI22	,170	,588
EI24	,416	,559
EI38	,147	,591
EI42	,518	,532
EI44	,286	,572
EI46	,219	,582
EI54	,447	,546
EI58	,405	,554
EI61	,322	,567
EI68	,464	,542
EI72	,436	,548
EI76	,378	,556
EI79	,327	,567

2.3 Reliability within the Conceptual- Practical (C-P) dimension

The reliability coefficient alpha within the subgroup of the Conceptual- Practical dimension was equal to 0.747, for a sample size of 6904 individuals and a full scale of 20 items. The overall alpha value although not higher than 0.8 it still within the good internal consistency range so no changes should be made.

The alpha value varied between 0.713 and 0.787 as shown in Table 4. No significant increase of the overall alpha can be achieved by the removal of one of the parameters, so all of them should be kept within the tool.

Table 4. Statistical analysis for the Conceptual- Practical dimension

	Corrected itemized scale correlations	Cronbach-Alpha,
NT6	-,274	,778
NT8	-,390	,787
NT13	,437	,723
NT18	,241	,739
NT27	,302	,735
NT32	,297	,735
NT37	,436	,724
NT39	,480	,720
NT48	,405	,726
NT50	,438	,724
NT53	,336	,732
NT57	,276	,737
NT59	,491	,720
NT62	,229	,740
NT65	,358	,730
NT67	,571	,713
NT69	,494	,719
NT71	,323	,733
NT75	,418	,725
NT78	,414	,726

2.4 Reliability within the Feeling- Analytical (F-A) dimension

The reliability coefficient alpha within the subgroup of the Feeling- Analytical dimension was equal to 0.695, for a sample size of 6904 individuals and a full scale of 20 items. Although the overall alpha value was lower than the ideal 0.8 or the threshold of 0.7, being so close to the later can be considered as acceptable. This is especially true if the variety of the participants, age and language is considered.

The alpha value varied between 0.663 and 0.747 as shown in Table 5. No significant increase of the overall alpha can be achieved by the removal of one of the parameters, so all of them should be kept within the tool.

Table 5. Statistical analysis for the Feeling-Analytical dimension

AF2	-,390	,747
AF9	,199	,690
AF10	,001	,711
AF17	,415	,668
AF19	,471	,663
AF21	,253	,685
AF25	,475	,664
AF30	,454	,665
AF33	,059	,705
AF35	,114	,697
AF40	,489	,662
AF43	,271	,684
AF47	,451	,666
AF52	,287	,683
AF56	,179	,693
AF60	,417	,670
AF64	,221	,689
AF66	,481	,661
AF74	,354	,675
AF80	,287	,682

2.5 Reliability within the Open- Structured (O-S) dimension

The reliability coefficient alpha within the subgroup of the Open- Structured dimension was equal to 0, 542, for a sample size of 6904 individuals and a full scale of 20 items. The overall alpha value is within the range of acceptable but a number of variables are not and they should be altered in order to increase the internal consistency of the test. Specifically the items OS14, OS28, OS34, OS 41, OS 49, OS 51, OS 55, OS 70 and OS 77 have an alpha coefficient of less than 0.5.

Table 6. Statistical analysis for the Open- Structured dimension

	Corrected	itemized	scale	Cronbach-

	correlations	Alpha,
OS3	-,243	,586
OS4	-,405	,613
OS11	-,055	,558
OS14	,327	,493
OS23	-,016	,547
OS26	-,058	,557
OS28	,473	,460
OS29	,276	,501
OS31	,203	,512
OS34	,320	,493
OS36	,053	,539
OS41	,401	,479
OS45	,171	,518
OS49	,499	,461
OS51	,423	,474
OS55	,282	,497
OS63	,070	,533
OS70	,345	,486
OS73	,138	,523
OS77	,383	,475

3. Construct validity

The construct validity of the Leonardo 3.4.5 tool was tested on the basis of two criteria: the convergent and discriminant validity. The convergent validity refers to the degree to which two measures of constructs that theoretically should be related are in fact related. In contrast discriminant validity tests whether the measures that are supposed to be unrelated are in fact unrelated.

A factor analysis was performed in order to assess the construct validity of the test. For the purpose of the analysis the four dimensions defined on the previous chapters (E-I, F-A, N-T, O-S) were used as the underlying factors. The ensuing from the analysis component matrix shows the relationships between the individual questions of each category with each one of the four factors. According to the discrimination criterion an indicator should be associated only with the related factor, and have no or very weak correlation with the other dimensions.

The construct analysis was based on four factors that were the four dimensions of the model. Each of these factors had 20 subcategories, giving 80 possible correlations. The results of the analysis are shown on Table 6.

Table 7. Factors component analysis

Categories	EI	NT	OS	AF	Categories	EI	NT	OS	AF
AF10	0.011	0.072	0.084	0.154	**EI1**	-0.221	-0.068	-0.009	0.078
AF17	-0.059	-0.074	-0.121	0.528	**EI12**	-0.146	-0.183	-0.087	0.002
AF19	-0.005	-0.031	-0.066	0.541	**EI15**	0.444	0.063	0.057	-0.068
AF2	0.004	0.009	0.088	-0.244	**EI16**	0.496	0.225	0.118	-0.018
AF21	-0.047	-0.061	-0.116	0.339	**EI20**	0.379	-0.035	0.037	-0.116
AF25	-0.047	-0.042	-0.152	0.54	**EI22**	0.308	-0.024	0.025	-0.074
AF30	-0.128	-0.109	-0.111	0.535	**EI24**	0.494	0.089	0.054	-0.131
AF33	-0.068	-0.032	0.075	0.227	**EI38**	0.274	0.009	-0.204	0.072
AF35	-0.107	-0.026	0.027	0.252	**EI42**	0.617	0.155	0.066	-0.129
AF40	-0.058	-0.048	-0.131	0.575	**EI44**	0.409	0.137	0.157	-0.128
AF43	0.012	0.11	0.074	0.406	**EI46**	0.363	0.477	0.258	-0.035
AF47	-0.072	-0.054	-0.101	0.533	**EI5**	-0.479	-0.12	-0.019	0.109
AF52	-0.021	-0.157	-0.175	0.273	**EI54**	0.55	0.228	0.107	-0.156
AF56	-0.045	0.101	0.003	0.315	**EI58**	0.528	0.116	-0.04	0.007
AF60	0.015	-0.035	-0.125	0.487	**EI61**	0.46	0.059	0.065	-0.118
AF64	0.009	-0.03	0.015	0.35	**EI68**	0.566	0.209	0.064	-0.02
AF66	-0.063	-0.08	-0.146	0.571	**EI7**	0.004	-0.04	-0.144	0.093
AF74	-0.114	-0.115	-0.14	0.468	**EI72**	0.56	0.187	0.045	-0.092
AF80	-0.099	-0.068	-0.13	0.334	**EI76**	0.52	0.117	0.033	-0.086
AF9	-0.042	-0.058	-0.044	0.318	**EI79**	0.431	0.159	0.021	0.007
NT13	0.236	0.528	0.212	-0.052	**OS11**	0.105	0.062	0.11	0.195
NT18	0.139	0.342	0.118	-0.001	**OS14**	-0.003	0.186	0.435	-0.065
NT27	0.006	0.429	0.105	0.043	**OS23**	0.003	0.011	0.155	0.001
NT32	0.077	0.421	0.144	-0.05	**OS26**	-0.249	-0.083	0.166	-0.008
NT37	0.144	0.534	0.294	-0.082	**OS28**	0.043	0.258	0.563	-0.071
NT39	0.059	0.584	0.175	0.034	**OS29**	0.2	0.246	0.388	-0.054
NT48	0.015	0.507	0.188	0.009	**OS3**	-0.027	-0.167	-0.087	0.077
NT50	0.121	0.536	0.229	-0.117	**OS31**	0.025	0.09	0.355	-0.08
NT53	0.006	0.449	0.136	0.028	**OS34**	0.28	0.416	0.381	-0.096
NT57	0.227	0.394	0.076	0.079	**OS36**	0.025	0.012	0.227	-0.103
NT59	0.254	0.586	0.248	-0.092	**OS4**	-0.086	-0.151	-0.201	0.114
NT6	0.013	-0.147	-0.064	0.005	**OS41**	0.011	0.181	0.474	-0.106
NT62	0.271	0.353	0.138	-0.035	**OS45**	-0.065	-0.071	0.292	-0.14
NT65	0.246	0.472	0.299	-0.214	**OS49**	0.127	0.247	0.558	-0.182
NT67	0.107	0.651	0.207	-0.047	**OS51**	0.101	0.252	0.474	-0.086
NT69	0.119	0.599	0.233	-0.063	**OS55**	0.094	0.235	0.36	-0.055
NT71	0.006	0.435	0.306	-0.153	**OS63**	0.031	-0.002	0.197	-0.116
NT75	0.264	0.537	0.203	0.03	**OS70**	0.123	0.264	0.457	-0.09
NT78	0.199	0.516	0.217	-0.1	**OS73**	-0.109	0.176	0.304	0.004
NT8	-0.143	-0.292	-0.218	0.1	**OS77**	-0.002	0.171	0.478	-0.025

According to the method only factors that have an eigen value higher of one should be used on the analysis. This condition was met for all four factors of the analysis. According to Bortz if more than 10 factors in Table 7 have a value higher than 0.4 then they can be used for interpretation. This is true for the values of the above table although some factors with negative factoring should be reexamined and probably reformatted (i.e 1, 2, 3,4, 5, 6, 8).

4. Analysis conclusions

The analysis of the five-year users database of Leonardo 3.4.5 yielded positive results that prove the reliability, structural and scientific integrity of the tool.

Based on the analysis presented on the previous chapters, some questions are highlighted as in need of redesign. However these specific questions are considered as highly valuable questions according to the Jung's model on which the Leonardo 3.4.5 assessment tool is based.

This contradiction can be explained by the fact that the questions used in the questionnaire lead the user to be conscious of their answers but this consciousness cannot be captured within the available variables of this specific analysis. Therefore although the test results might indicate that the questions yield statistically poor results, these results can indeed be useful in the correct assessment of ones personality. Therefore it can be concluded that these questions are overall acceptable and should not be changed offhandly but a deeper re-examination of these question should take place before the decision is made.

In conclusion the statistically analysis of a high number of participants responses showed that Leonardo 3.4.5 has a solid scientific base, and that it can calculate a user's profile with accuracy and independent of demographic factors.

5. Case studies

Post customer research was conducted after the administration of the Leonardo 3.4.5 assessment test, within two distinctly different groups.

5.1 Business management case study

In the first case study, the group consisted of 38. All of them were within the top 50 managers of a big international company (ca. 65000 employees). After the administration of the test and their coaching session the following questions were asked to the participants. Firstly how closely did you find yourself relate to the profile predicted by Leonardo 3.4.5 and secondly how would they rate their experience of using the Leonardo 3.4.5 (Table 8).

Table 8. Questions used during the post customer research of the second study. The questions were rated between 0-10 with 0= absolutely not and 10 = totally. Both questions are shown both in English and German as they were posed to the participants

German
Q1. Wie stark finden Sie sich in Ihrem Leonardo 3.4.5-Profil wieder? (0 = überhaupt nicht, 10 = total)
Q2. Wie hilfreich war die Profilanalyse hilfreich für Sie? (0 = überhaupt nicht, 10 = total)
English
Q1. How much did you recognize yourself on the profile predicted by Leonardo 3.4.5) (0= Absolutely not, 10= totally)
Q2. How helpful was the profile analysis for you? (0= Absolutely not, 10= totally)

The questions were rated on scale of 0 to 10 with zero being not at all and ten being totally agree with the question statement. On the first question 82% of the participants gave a rating of 8 and higher (Figure 2a). On the second all participants rated their experience with the Leonardo 3.4.5 tool with an eight and higher (Fig 2b).

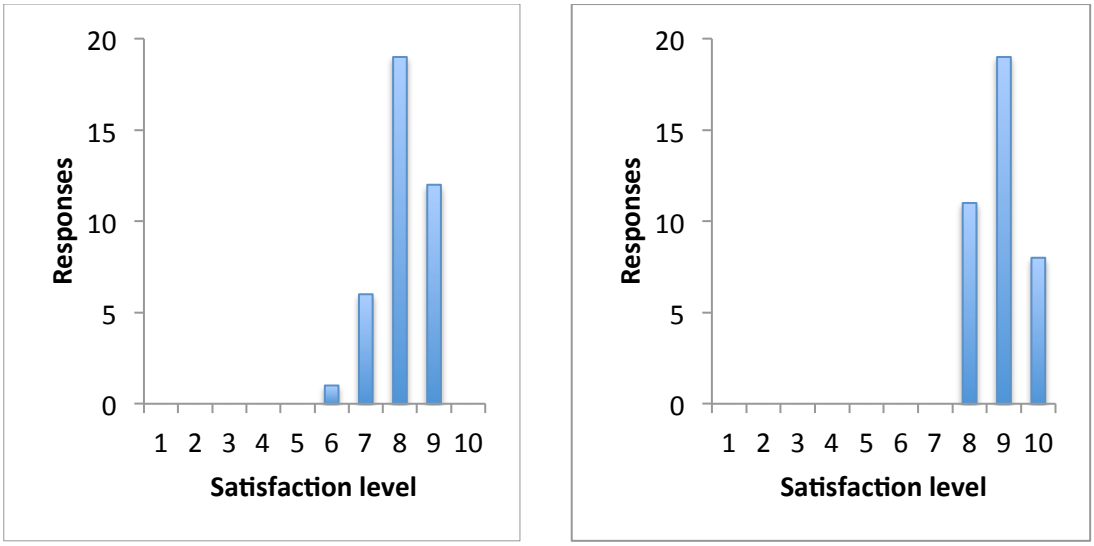

Fig 2. Leonardo 3.4.5 customer satisfaction assessment. 2a (left) depicts the satisfaction rate to the question of how closely the participant relates to the outcome profile. 2b (right) shows the satisfaction rate to the total experience of using Leonardo 3.4.5 assessment tool.

5.2 Academic group case study

In the second case study, the Leonardo 3.4.5 assessment was administrated to a team of university students with an average age of 25 and with 75% of them having no working experience. In this group 75% of the participants were female and 25% male. The customer satisfaction test was conducted within two weeks of finishing the assessment and focused on the perceived improvement of their personal interaction with their group and within their group. At the end of the test they were also asked to assess their total experience with the Leonardo 3.4.5 tool and the effect it had in their everyday interactions within their working environment. In total 45% of the participants perceived an improvement in their interactions with their colleagues while 35% noticed a significant improvement in the interactions between the various team members. In total 75% of the participants described their experience with the Leonardo 3.4.5 as positive or highly positive and 25% as neutral (Fig 3).

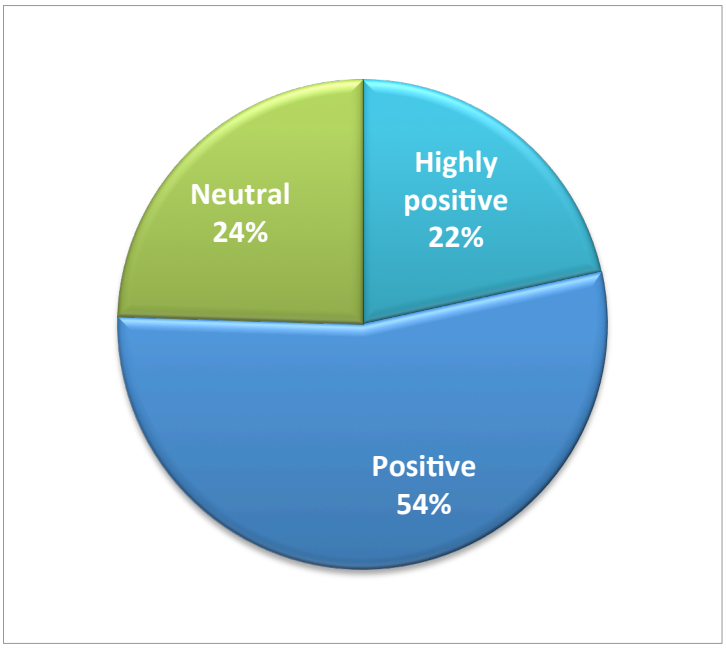

Fig 3. Participant's satisfaction after using the Leonardo 3.4.5

5.3 Case studies conclusions

In general the personal assessment and personality profile provided by the Leonardo 3.4.5 tool was met with a high satisfaction rate by all the participants.

This high satisfaction rate is especially important as the two case studies were conducted in two distinctive and opposing groups. In particular the business group was composed by the top management of a big international company that had

decades of working experience between them and the responsibility of coordinating big groups. On contrast the academic case study included people with minimum working experience younger and with minimal exposure to high intensity working environments. Nevertheless the conclusion after using the Leonardo 3.4.5 from both groups was that the participation on the Leonardo 3.4.5 assessment was very useful as they gained a deeper and better understanding of themselves and their function within their working group and that they could perceive differences in their interactions with others after the assessment. In their own words they found the total experience with Leonardo 3.4.5 as *helpful*, *informative* and *useful* (Fig 4).

These two case studies serve as evidence of the versatility and the wide applicability of the Leonardo 3.4.5 tool that is independent of age, work experience, industry, gender or language.

Fig 4. Word cloud of the most common words used by the Leonardo 3.4.5 users in order to describe their user experience.