

DNV certification process

1) How was the certification process of DNV structured/organized?

It is organized in a way that we first deliver our documentation on the instrument we want to certify. Thereafter DNV send the documentation to two independent experts. They read through the documentation, test the tests etc. and then ask for more info from us (the provider). If they are satisfied with the info they begin scoring according to the requirements, and then they write a proposal to DNV – advising them to certify the product (or not).

2) Which requirements had to be met?

Check-list with the following certification criteria:

Quality of the explanation of the rationale, the presentation and the quality of information provided

- Quality of the explanation of the rationale
 - o Theoretical foundations of the constructs
 - o Test development procedure
 - o Thoroughness of the item analyses and item analysis model
 - o Explanation of content validity
 - o Summary of relevant research:
- Adequacy of documentation available to the user (user and technical manuals, norm supplements etc)
 - o Rationale
 - o Development
 - o Standardization
 - o Norms
 - o Reliability
 - o Validity
- Quality of the Procedural instructions provided for the user
 - o For test administration
 - o For test scoring, norming etc
 - o For interpretation and reporting
 - o For providing feedback and debriefing test takers and others
 - o For providing good practice issues on fairness and bias
 - o Restrictions on use
 - o References and supporting materials
- Quality of the materials
 - o General quality of test materials (test booklets, answer sheets, test objects, software, etc)
 - o Test quality of the local adaptation
 - o Ease with which the test taker can understand the task
 - o Ease with which responses or answers can be made by the test taker
 - o Quality of the items

Norms or reference group information

- Appropriateness for local use, whether local or international norms
- Appropriateness for intended applications
- Sample sizes
- Procedures used in sample selection
- Quality of information provided about minority/protected group differences, effects of age, gender etc

Validity

- Construct Validity
 - o Designs used
 - o Sample sizes
 - o Procedure of sample selection
 - o Median and range of the correlations between the test and other similar tests
 - o Quality of instruments as criteria or markers
 - o Differential Item Functioning (DIF) analyses
- Criterion-related Validity
 - o Description of the criteria used and characteristics of the populations
 - o Sample sizes
 - o Procedure of Sample selection
 - o Median and range of the correlations between the test and criteria
 - o Correction for attenuation by measurement error
 - o Corrections for restrictions of range

Reliability

- Data provided about reliability
- *Internal consistency*
 - o Sample size
 - o Median of coefficients
- *Test retest stability*
 - o Sample size
 - o Median of coefficients
 - o Coefficients detailed
- *Equivalence reliability*
 - o Sample size
 - o Median of coefficients
 - o Coefficients detailed

Quality of Computer Generated Reports

- Scope or coverage
 - o *Does the report cover the range of attributes measured by the instrument?*
 - o *Does it do so at a level of specificity justifiable in terms of the level of detail obtainable from the instrument scores?*
 - o *Can the 'granularity' of the report (i.e. the number of distinct score bands on a scale that are used to map onto different text units used in the report) be justified in terms of the scales measurement errors?*
 - o *Is the report used with the same populations of people for who the instrument was designed? (E.g. Groups for whom the norm groups are relevant, or for whom there is relevant criterion data etc).*
- Reliability
 - o *How consistent are reports in their interpretation of similar sets of score data?*
 - o *If report content is varied (e.g. by random selection from equivalent text units) is this done satisfactorily?*
 - o *Is the interpretation of scores and differences between scores justifiable in terms of the scale measurement errors?*
- Relevance or validity
 - o *How strong is the relationship between the content of the report and the scores on the instrument? To what degree does the report go beyond or diverge from the information provided by the instrument scores?*
 - o *Does the report content relate clearly to the characteristics measured by the instrument?*
 - o *Does it provide reasonable inferences about criteria to which we might expect such characteristics to be related?*
 - o *What empirical evidence provided to show that these relationships actually exist?*
- Fairness, or freedom from systematic bias
 - o *Is the content of the report and the language used likely to create impressions of inappropriateness for certain groups?*
 - o *Does the report make clear any areas of possible bias in the results of the instrument?*
 - o *Are alternate language forms available? If so, have adequate steps been taken to ensure their equivalence?*
- Acceptability
 - o *Is the form and content of the report likely to be acceptable to the intended recipients?*
 - o *Is the report written in a language that is appropriate for the likely levels of numeracy and literacy of the intended reader?*
- Practicality
 - o *How much time does each report save the user?*
 - o *How much time does each report take to read and use?*
- Length

The instrument fails, if it got a score below 0 on any of the parameters.

3) Which steps had to be passed through during the certification process?

This was a two-way dialogue; we had to provide them with more info and studies when they asked for it etc. So there aren't any real steps, just a thorough examination of the instruments.

4) Which criteria were considered for the certification of cut-e instruments?

When it comes to criteria we use to gain concurrent and predictive validity we mostly had subjective criteria (360 evaluations against test results), and a few studies with objective criteria (actual sales results)

5) How was the result of the certification process and how was it documented?

The result was very good, we had just two 0's, and the rest was 1's and 2's. Details are documented in the technical reports. We also got a certificate, and we have to recertify the tests after 3 years.