

## A mixed methods approach to validating a test

Mária Vargová  
Testing team  
Language Institute  
Slovakia

### **SUMMARY**

The presentation attempted to illustrate the great potential of utilizing a mixed methods approach when validating a test (Slides 1 – 2).

As most of the national testing teams collect numerical data and analyse them using quantitative methods, the presentation first focused on discussing the advantages of taking a quantitative approach to test validation (Slides 3 – 4).

After that, the results of a statistical analysis run on the data Slovak national team collected during pretesting a SLP2 listening test (presented in a separate document and not shared for security reasons) were shown to the audience and interpreted. The main reason for including this activity was to demonstrate that although the results of a quantitative analysis of test data are very useful in telling us WHAT is good about our test/items and what needs some improvement, their ability to help us reaching a better understanding of WHY our test/items did not work the way we had expected them to is rather limited. At this point, the limitations of quantitative analysis methods were mentioned (Slide 5) and the advantages of utilizing qualitative analysis methods when establishing test construct validity were highlighted (Slide 6). Different tools of qualitative data collection and analysis were then presented (Slide 7).

The second part of the presentation was devoted to sharing the experience Slovak national team has with using verbal protocols in test validation. The team uses verbal protocols to gather qualitative information about test-takers' thinking processes when taking a test to check the alignment between the observed strategies and the expectations of the test designers. Before describing the process of collecting a verbal protocol (Slide 9), its definition and different types were listed (Slide 8).

In order to demonstrate how data gathered through think aloud protocols can help test designers understand WHY some of test items do not work as expected, extracts from two verbal reports (transcribed, translated and coded) were shown to the audience (presented in a separate document and not shared for security reasons). The extracts contained the comments made by a good listener (targeted level - SLP2) and a less able listener (SLP1+) when attempting the item which, based on the interpretation of the statistical analysis results, seemed the most problematic item on the test. This activity was included to show the potential of verbal protocols in getting insights into why individual items are faulty. The particular verbal protocols for example revealed that instead of measuring the intended ability, the problematic item was measuring a construct-irrelevant factor, namely real-life knowledge.

Besides exploring the match between the test-designers' predictions and the actual skills and processes test-takers use during the test, another two ways in which verbal protocols can be used to get insights into test quality were mentioned (Slide 10).

The presentation was concluded with a reference to a great potential of combining quantitative and qualitative methods when collecting evidence for a validity argument (Slide 11).