
TITLE

Looking into listening: Using eye-tracking to establish the cognitive validity of the Aptis Listening Test

This is a summary of a report by Franz Holzknacht, Kathrin Eberharter, Benjamin Kremmel, Matthias Zehentner, Gareth McCray, Eva Konrad, Carol Spöttl as part of the ARAGs Research Online Series. For a copy of the full report, see www.britishcouncil.org/exam/aptis/research/publications/

WHAT WE LOOKED AT:

We were interested in learning more about how test takers process information during an Aptis Listening Test. We investigated two aspects:

- 1) What kind of information from the sound file do test takers use in order to answer test items
- 2) How do they use the information from the sound file.

The Aptis Listening Test specifications list the test developers' expectations regarding these two aspects. However, it is important to continuously examine whether a language test actually measures the kind of skills it claims to test so that test users feel they can trust the test results.

HOW WE DID IT:

Researching how information is processed during a test poses a methodological challenge, as thought processes cannot be observed directly. To overcome this challenge we opted to combine eye-tracking technology with stimulated recall interviews. Eye-tracking devices recorded, via an infrared camera, what participants looked at while doing a certain task on a computer screen. The stimulated recall interviews took place after the eye-tracking. Participants were shown the recordings of their eye movements while they were solving the task and were asked to report in as much detail as they could what they were thinking while completing the task.

We invited 30 German speakers to participate and first asked them to complete the Aptis Vocabulary and Grammar, Listening and Reading test components in order to get information on their general English language proficiency level. Next, all participants completed one Aptis Listening Test on an eye-tracking device. We selected 16 participants to also tell us what they were thinking while they had answered the test items. To help them create a more accurate account of their thoughts, we stopped the Listening Test after two to three items and asked them to remember how they had answered the listening items. Participants were shown a recording of their eye-movements to stimulate recall. These interviews were recorded and transcribed before analysis.

Through this procedure we collected interview data from 16 selected participants and eye-tracking data from all 30 participants. The interviews were analysed to see how test takers answered the items. The eye-tracking data was analysed using statistical methods to show whether there were any significant patterns.

WHAT WE FOUND:

Stimulated recall interviews:

The analysis of the stimulated recall interviews showed that while taking the Aptis Listening Test, test-takers generally behaved as expected by the test developers. In terms of what kind of information the test-takers used to answer the items, we found a clear pattern. More challenging test items also required the test-takers to handle more complex information units. However, we found a less clear relationship between item difficulty and how test-takers approached the item, i.e. the cognitive processes they employed. Also, one higher-level listening process labelled “discourse construction” was not used enough by test-takers to answer test items designed to target this process.

Eye-traces:

The analysis of the eye-tracking data also revealed interesting results. First, we found that test-takers looked at answer options of more difficult items for a longer period of time compared to easier items. Second, answer options that are situated higher up on the screen were looked at longer than answer options lower down. Also, better readers of English spent more time looking at the answer options than poorer readers.