

Dynamic Placement Test: Validity of the test

Anchor items and achieving validity

Validity in assessment refers to the degree to which a test measures the knowledge or skills it claims to measure, while reliability refers to whether the test provides consistent results each time it is taken.

Steps are taken at different points in the lifecycle of a test to ensure that a score given to a candidate is accurate. This means that if a test taker's score is B1, he or she really is a B1, and another test would show the same thing. Some of the measures used include adding anchor items, pilot testing and data analysis.

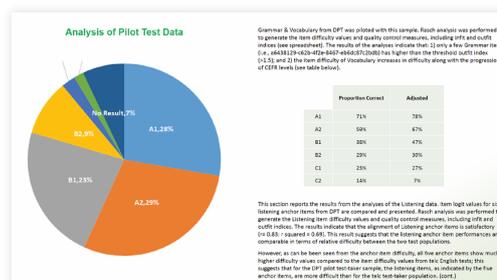
To check the validity of a test, anchor items are used. These are test questions that have been taken from existing tests and are proven, through the use of data analysis, to correspond to a specific level. For the Dynamic Placement Test, anchor items were taken from telc language tests at all levels from A1 to C2 of the Common European Framework of Reference.

The anchor items are placed into the test alongside the newly created items. Piloting provides the data to show which items candidates get right and which they get wrong. If a large number of candidates answer the B1 anchor items correctly, but have difficulty with a new 'B1' item, this indicates that the new item is not performing as it should. The non-performing item will then be evaluated by test experts who decide to adapt it or replace it. Further analysis will be then carried out to check the performance of the new or adapted items. The process continues for the duration of a test's life.

This shows the fundamental role of data analysis in creating and maintaining a test.

Regular item analysis

The Dynamic Placement Test team regularly runs data analysis on tests taken by users, allowing us to refine and refresh the questions. With the constant influx of data from thousands of test takers we are also able to look at the validity of every item, ensuring it helps determine a test taker's CEFR level within the 30 minute time-span. This means the test will become ever more relevant, reliable and valid. The chart is a snapshot of the result of a recent item analysis carried out by the telc team in Frankfurt, Germany.



DPT item analysis June 2023

As part of the ongoing quality control of the Dynamic Placement Test (DPT), an analysis of all the live items in the item bank was carried out in June 2023, using data collected from every instance of the DPT taken between May 2022 and April 2023. Statistical analysis showed that the overall test reliability was very high.

- Using Cronbach's alpha, the reliability of DPT as a whole was measured at 0.95 (in other words, a 95% level of reliability). Cronbach's alpha is a standard measure used by many test providers and a value above 0.90 is generally considered excellent.

- Analysis of individual items, using measures of item facility and discrimination, showed that the items distinguish well between stronger and weaker candidates. Any items falling outside agreed thresholds will be subjected to further investigation and then rewritten or replaced as necessary.

As the analysis was based on 12 months of testing, the size of the data set allows a high level of confidence in the findings.

Further refinements

The Dynamic Placement test is committed to being culturally fair so that the test is discrimination-free and reliable amongst test takers from different cultural and language backgrounds. Feedback from test-takers and administrators is constantly reviewed – that allows us to remove or further refine questions that are deemed problematic from a fairness point of view. Options for L1 instructions and practice questions are also available (e.g. Arabic and Spanish) to ensure technology literacy does not get in the way of language testing.

For more detailed explanations on how the test works, and interpreting scores, please visit www.clarityenglish.com/dpt/