

Derntl, Victoria. 2009. *Language testing for program evaluation - evaluating the learning outcomes of two language tracks in an Austrian Gymnasium*. MA Thesis, University of Vienna.

## **Abstract**

This thesis presents a language test which was commissioned by an Austrian Gymnasium and which was developed with the purpose of program evaluation. In this particular case the parents' association of the school wanted to compare the two existing language tracks of the school with regard to the learning outcomes of the pupils. Therefore two tests in two different languages (English and French) were designed. The paper features an in-depth discussion of all components of testing theory connecting them to the actual stages in the development, trialling, administration and scoring of the test under consideration.

The results of the tests, which are not the main focus of the paper, show that the test takers from the two language tracks performed almost identically in the English test, whereas one language track exceeded the other one in the performance in the French test. A qualitative and statistical analysis of the tests themselves was carried out in order to ascertain whether the tests developed can be regarded as useful for the purpose of comparison. Results of this analysis demonstrate that although both tests still have potential improvement, they do show a satisfying level of reliability for the purpose of comparison. The results from the analysis of the usefulness of the tests show that the test scores are reliable enough in order to allow for interpretations which are based on them.