In the field of translation studies, there is an emerging awareness of the need to obtain evidence for the quality of tests used to measure translation competence. To date, translation tests have been informed by practice rather than by empirical research and questions regarding the reliability of the assessment methods have remained unanswered.

According to a recently developed method (Anckaert et al. 2006; Eyckmans et al. 2009) translation performance indicators can be related to the underlying translation competence in a psychometrically controlled way. This norm-referenced method is said to bridge the gap between language testing theory and the epistemological characteristics of translation studies by selecting text segments with discriminating power through a process of pre-testing and item calibration. The dissemination of this method ties in with the European call for test standardisation in the sense that it allows a reliable and valid certification of translation competence across Europe. It also permits an exploration of text robustness so that tests can be validated for different language combinations.

In our paper we will report an experiment in which three methods for assessing translation competence are compared. The pros and cons of the different methods will be discussed. The central research question centers on the reliability and the validity of these methods with reference to each other and homes in on the quintessential issue of text independent measurement of translation competence. The results of the experiment clearly indicate that collaboration between translation teachers and test developers is indispensable if we want to rise to the methodological challenge of achieving equivalent standards across languages.