

Point to Point

A Reaction to MacGregor's "The *Eiken* Test: An Investigation"

Nigel Henry

International Masters Academy, Okayama

This paper represents a critique of the *Eiken* test investigation by Laura MacGregor (Volume 19, No. 1, May 1997, pp. 24-42). MacGregor provides detailed background information on the *Eiken* test explaining its origins, importance, and contents. In order to investigate what kind of test the *Eiken* is, explanations of the purpose and goals of the two types of tests used in language teaching, criterion-referenced tests (CRTs) and norm-referenced tests (NRTs), are given. Through evidence found in the nature of the *Eiken* test, it is correctly judged to be a hybrid CRT/NRT. However, there are a number of difficulties with her methodology which call into question her conclusions. Discussions that follow will be limited due to word restrictions.

The test used in this study is the pre-second level test, originally developed for second and third year high school students (16- and 17-year-olds). In this study, however, the test subjects were 182 first year junior college students (aged 18 to 20). Because this study was not aimed at the correct target population, the results cannot be compared with those provided by STEP or even utilized in a valid analysis of the test. If a test was developed for a certain group, investigating its validity and reliability should be done using the target group.

The reliability of the test was measured using both descriptive and item statistics. Descriptive statistics revealed that the test performed like a true NRT, though since the author classifies the test as a hybrid CRT/NRT it seems strange to apply purely NRT standards when analysing the results. Item statistics, according to guidelines set out in Brown regarding items on an NRT (1996, p. 69), showed that 60% of the test items needed refinement or improvement. As above, MacGregor failed to ei-

ther defend her use of NRT standards or present an alternative system for the analysis of the items. Four general questions were posed to help determine the validity of the test. Answers revealed that:

- 1) The items were suitable for senior high school students.
- 2) Two items were found to be invalid where content validity was concerned.

The argument against one of the items is presented using anecdotal rather than empirical evidence. It is also difficult to determine which items the author was referring to since items and sections were not clearly and consistently outlined from the beginning.

- 3) STEP claimed successful examinees are able to converse, read and write about everyday topics.

MacGregor challenges this, stating that *Eiken* only tests reading and listening skills. However, neither provides evidence to support their claims.

- 4) There were some poorly constructed items on the test.

However, there is some incongruity as to the problems with items. For example, the first example MacGregor (1997, p. 38) presents is not necessarily problematic because of the structures but, instead, because of the length of the distracters. Though poorly constructed items were found, investigations into their nature were subjective rather than methodical, systematic, and empirical. A framework by which items might be analyzed less subjectively might, for example, be based on Chapter 4 of Henning (1987).

The results of the examination of scoring revealed that passing percentages were actually much lower than those stated by STEP. This again suggests that the test group employed by MacGregor was not representative of the STEP population.

The above discussion has examined the relevance and usefulness of this investigation into the *Eiken* test. It found that though the investigation took on some detail and identified strong and weak areas within the test, it failed to determine the validity and reliability of the test adequately and accurately. A more thorough investigation, using more appropriate tools for measuring and analyzing test components, is clearly required.

References

- Brown, J.D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall Regents.
- Henning, G. (1987). *A guide to language testing*. Boston: Heinle & Heinle.
- MacGregor, L. (1997). The *Eiken* test: An investigation. *JALT Journal*, 19(1), 24-42.