

Assessment Literacy Self-Study Quiz #4

by Tim Newfields

This column mentions a range of questions about testing, statistics, and assessment in quiz format to help readers become more familiar this evolving field and consider current testing practices. Suggested answers for the items below are online at <http://jalt.org/test/SSA4.htm>.

Part I: Open Questions

1. One Japanese university English entrance examination included this task:

INSTRUCTIONS (translated from Japanese): Read the sentence below, then select one sentence (A-D) in which the term "that" is used in the same manner as in the original sentence:

Sentence: If there is something that you want to see, let me know.

Possible Answers:

- (A) The shoes that I bought did not fit.
- (B) She hid the fact that she spoke French.
- (C) I'm so glad that he could come.
- (D) Such was her anger that her face turned red.

Question: What is this test item likely measuring? What would be the arguments for and against including this item on an EFL test for aspiring college students?

2. On October 9, 2006 a short article in *The Washington Post* mentioned a statistic supporting current testing policies in American schools. Based on an unspecified report by *Public Agenda*, they claimed that "71% [of the students surveyed] felt the number of tests they have to take is 'about right'" and that "79% thought standardized test questions are fair."

Mention at least three problems with this newspaper report and at least four things that should have been done to present the information ethically.

Source: The Washington Post. (2006). *Too much testing?* Retrieved February 10, 2007 from www.washingtonpost.com/wp-dyn/content/article/2006/10/09/AR2006 [inactive URL].

3. Mention one example of each of the following types of hypotheses from the field of applied linguistics: (1) a positive directional hypothesis, (2) a negative directional hypothesis, and (3) a non-directional hypothesis.
4. To estimate the English vocabulary size of a group Japanese examinees, a bilingual test of vocabulary was devised. In this format, six English words appear on one line. Above that, two of those words are defined in Japanese, as in this example:

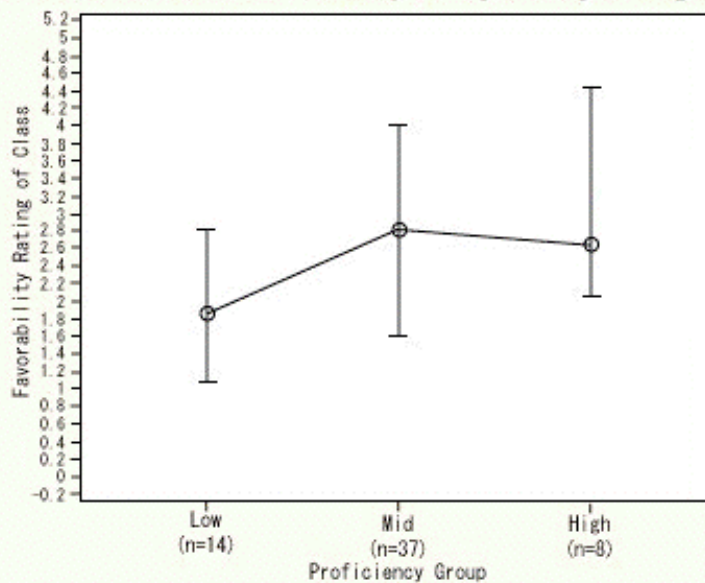
3 信じること・信念 () **4** 盤上で白黒の駒を動かして、勝敗を競うゲーム ()
(1) attention (2) belief (3) chess (4) hook (5) pride (6) union

Source: Mochizuki, M. (1998) . Nihonjin gakusha no tameno goi saizu tesu [A vocabulary size test for Japanese learners of English]. Gogaku Kennkyuujou Kiyou, 12, 27-53. Cited in K. Kasahara. (2006) Producing a revised version of the Mochizuki Vocabulary Size Test. *JLTA Journal*, 9, 55- 72.

Any problems with this type of test design as a measure of English vocabulary?

- Recently an article by Hamada, Oda, and Kito (2007, pp. 42 - 55) explored how classroom performance on a composite 98-item in-house EFL test correlated with a 5-point Likert scale assessment of how students reported liking their class. 59 Japanese high school EFL students participated in their survey. Those who scored 35 points or less (n=14) on the composite tests were placed in a "low" sub-group. Those with 36-65 points (n=37) on the composite tests were placed into a "mid" sub-group. Those with a combined score of 66 points or higher (n=8) were placed in a "high" sub-group. The results are summarized in Figure 1, with small circles indicating the mean scores for each sub-group and anchor bars indicating one S.D. above and below each mean.

Figure 1. Differences in the class favorability rating among three proficiency groups.



First, look at this graph and suggest any ways to make the information clearer. Then, without reading the article, briefly interpret the graph. What can be surmised about the distributions of the sub-groups? Finally, suggest a viable alternative way of comparing the ratings of the high, mid, and low sub-groups.

Source: Hamada, Y., Oda, S., Kito, K. (2007). Correlations between Japanese high school students' language learning beliefs, self-reported attitudes towards group dynamics, and performance in two in-house EFL tests. In T. Newfields, I. Gledall, P. Wanner, & M. Kawate-Mierzejewska. (Eds.) *Second Language Acquisition - Theory and Pedagogy: Proceedings of the 6th Annual JALT Pan-SIG Conference*. May. 12 – 13, 2007. Sendai, Japan: Tohoku Bunka Gakuen University. (pp. 42 - 55) Retrieved January 15, 2008 <http://www.jalt.org/pansig/2007/HTML/HOK.htm>

Part II: Multiple Choice Questions

1. A *t*-distribution is generally resembles a normal Z-distribution except _____.
 - (a) it is a function of the degrees of freedom
 - (b) it is more likely to contain less outlying information than a Z-distribution
 - (c) it is more accurate than a Z-distribution
 - (d) as the sample gets larger, the differences between *t*- and Z-distributions tend to increase

2. Which of the following is generally not considered a test method facet:
 - (a) time allocation
 - (b) test organization
 - (c) the validation procedure
 - (d) test instructions

3. Quasi-experimental research designs (i.e. studies in which the subjects are not randomized into different experimental and control groups) are also often known as _____.
 - (a) correlational designs
 - (b) non-experimental designs
 - (c) classic experimental designs
 - (d) incomplete factorial designs

4. In statistics, the symbol **tc** sometimes refers to _____.
 - (a) t-critical
 - (b) t-confidence
 - (c) Kendall's tau-c
 - (d) t-closure

*If you would like to suggest any questions
or comment on the suggested online answers email
the author at timothy@toyonet.ac.jp.*

HTML: <http://jalt.org/test/SSQ4.htm> PDF: <http://jalt.org/test/PDF/SSQ4.pdf>