

Assessment Literacy Self-Study Quiz #11

by Tim Newfields

This column about testing, statistics, and assessment is in a quiz format and designed to promote greater assessment literacy. Suggested answers to all of the problems below are online at <http://jalt.org/test/SSA11.htm>.

Part I: Open Questions

1. At most universities in Japan, how do the AO entrance examinations tend to differ from regular entrance exams and the National Center Test?
2. What is a "false consensus effect" (FCE)? How might it skew some social science data and how should it be avoided?
3. What unstated assumptions are inherent in the TOEIC "Can Do" statement chart below? What issues need to be considered when interpreting this chart? How could this information be more ethically presented?

What do the scores mean?

TOEIC Scores	English Ability Expectations	Position/Activities
250-350 points	Basic English	<ul style="list-style-type: none"> • Office Clerk
350-450 points	Employees who do not need to do business outside of the country.	<ul style="list-style-type: none"> • Technician, Receptionist
450-550 points	Ability to communicate at a minimum level with people from other countries.	<ul style="list-style-type: none"> • Computer operator, Tech Coordinator • Fairs, Technical Course
550-630 points	Ability to make business trips to other countries with some technical experience.	<ul style="list-style-type: none"> • Assistant Manager, Tech Manager • Conference, Convention
630-700 points	Total capacity of managing business travel in English without any assistance.	<ul style="list-style-type: none"> • Customer Service, Sales Representative, Quality Control
700-750 points	Ability to negotiate with people from other countries and make presentations.	<ul style="list-style-type: none"> • Purchasing Manager, Financial Analyst, Auditor • Go to university, promotional tours.
750-990 points	Has enough skill for any activity that is assigned.	<ul style="list-style-type: none"> • HR Manager, Executive Manager, Marketing Manager, Director • Get a masters

Source: Ingles Personalizado. (2011). *TOEIC*. Retrieved from <http://ingles-personalizado.com/toEIC>

4. Illustrate how a study with a *regression discontinuity design* could be structured to ascertain the effects of a specific intervention on English language reading fluency among a group of EFL high school students.

5. What additional information should have been mentioned in the chart below to help readers to interpret the data more meaningfully? Also, how would you interpret this data distribution?

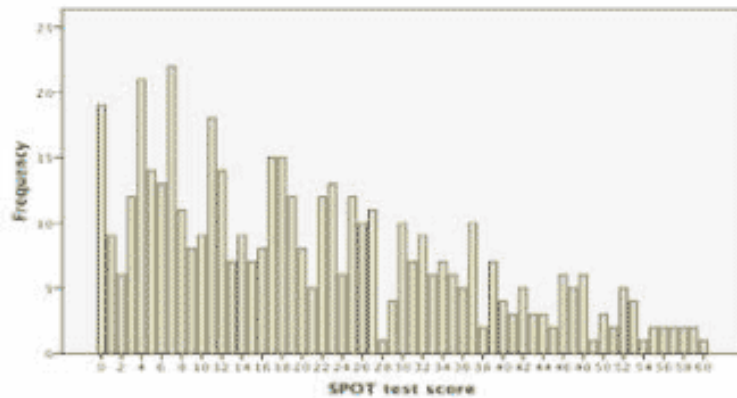


Figure 1. SPOT test scores
N=464; Mean=20.87; Min=0; Max=60

Source: Northwood, B. & Thompson, C. K. (2010). Why stop studying Japanese? A case in Australia. In A. M. Stoke (Ed.) *JALT 2009 Conference Proceedings*. (p. 821). Retrieved from <http://jalt-publications.org/archive/proceedings/2009/E040.pdf>

6. In factor analysis, how does a *communality* differ from a *commonality*?

Suggested answers online at <http://jalt.org/test/SSA11.htm>

Part II: Multiple Choice Questions

1. A tendency of respondents to concur with the opinions implicitly expressed in a given rating scale (and avoid offering counter-opinions) is known as _____.
- (a) contamination bias (b) acquiescence bias (c) social desirability bias
(d) central tendency bias (e) nonresponse bias

NOTE: One of these choices is most relevant; two are somewhat relevant, and two are irrelevant.

2. What is one difference between the *standard error of measurement* (SEM) and *standard error of difference* (SED)?
- (a) The former is a test of significance between two individual test scores.
(b) The latter indicates how widely an individual's given test score is likely to differ from his/her true score.
(c) The latter is the critical value of a *t-test* of the difference between two means.
(d) To calculate the latter, the former must be known.

NOTE: Just one answer is considered fully correct.

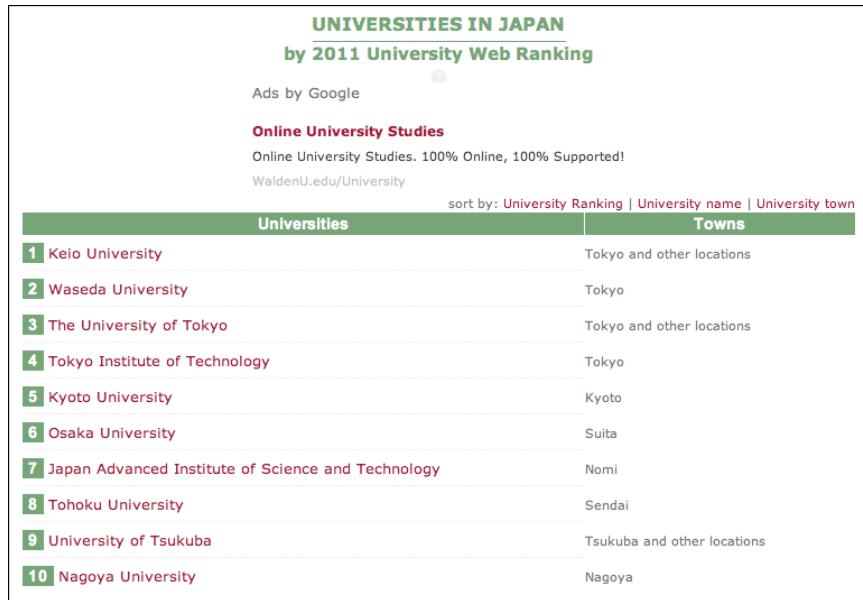
3. What sort of test would be most appropriate to ascertain whether a large sample of male and female university students differed significantly in their attitudes towards study abroad

according to a Likert scale with 4 or 5 options??

- (a) An unpaired t-test (b) An ANCOVA (c) An ANOVA
 (d) A MANOVA (e) A chi-square test

NOTE: Two answers are viable, but one answer will yield more information than the other.

4. Which of the following points is not problematic about the avowed ranking of these universities in the following chart?



UNIVERSITIES IN JAPAN by 2011 University Web Ranking	
Universities	Towns
1 Keio University	Tokyo and other locations
2 Waseda University	Tokyo
3 The University of Tokyo	Tokyo and other locations
4 Tokyo Institute of Technology	Tokyo
5 Kyoto University	Kyoto
6 Osaka University	Suita
7 Japan Advanced Institute of Science and Technology	Nomi
8 Tohoku University	Sendai
9 University of Tsukuba	Tsukuba and other locations
10 Nagoya University	Nagoya

Source: 2011 University Web Ranking. (2011). *Universities in Japan*. Retrieved from <http://www.4icu.org/jp/>

- (A) The criteria for the ranking is not explained in the title of the chart or anywhere on the same page as the chart.
 (B) This ranking was made entirely on the basis of a secret web metric formula – presumably, universities with lots of web traffic according to some search engines will receive higher ratings than those with low web traffic.
 (C) The fact that the different faculties at the same university may have different rankings is ignored.
 (D) The point that the lesson quality might vary widely from teacher to teacher at the same school is also ignored.
 (E) The location of some of the schools appears to be incorrect.
 (F) There is no mention of the measurement error of this rubric or indication of how widely apart the various school rankings are. Moreover, the possibility that two schools may have essentially equivalent rankings according to a given criterion does not appear to be considered.

Suggested answers are online at <http://jalt.org/test/SSA11.htm>