

Modern Metrology and the Revision of our *Standards for Educational and Psychological Testing* — An Open Letter to American Parents

By Robert Oliphant*

Dear Fellow Americans. . . . Just like cigarette smoke, the crisp odor of corruption will hit Americans right away when they read the opening statement of the “Standards for Educational and Psychological Testing” (SEPT) (www.teststandards.net). Certainly its “follow the money” link to our new federal Common Core Standards will put American taxpayers on their guard immediately, along with inviting direct attention to the above web site and its solicitation of comments before the closing date of April 21, 2011.

By way of inviting America parents to visit the SEPT web site and raise questions, let’s start by recognizing that legitimate educational and psychological testing is crucially important in a world economy, as W. Edwards Deming and his followers (ESO-9000, etc.) have been reminding us for years. Hence the mischief that can be wrought by testing ventures whose “standards” have absolutely nothing to do with those of our National Institute on Standards and Technology (NIST) and its state affiliates, e.g., the California Division of Measurement Standards. [Deming spent 12 years with the National Bureau of Standards, incidentally.]

At this point I’m certain that most Americans, along with many of their representatives, would support the evaluation of educational tests by NIST or by similar organizations with the same commitment to modern metrology. So by way of illustrating how today’s measurement science (metrology) works, let’s consider how NIST might handle the measurement of vocabulary size — a nightmarish concern that wallops both first graders and their grandparents, especially those worrying about going blank on words and proper names.

In measuring vocabulary size, as with other measurement techniques, our basic requirement is that of Authority (NIST uses the Fahrenheit Scale, not an invented one). We also require

Calibration via the tabulation and comparison of actual results as opposed to correlation and more speculative measures. Finally we require Transparency (call it a cliché if you wish) that enables the public as a whole to understand and even replicate the testing process.

Consensually considered, American dictionaries meet the vocabulary-authority requirement very well. Their spellings, going back to the first Merriam Webster) were demanded and approved by President Theodore Roosevelt as a declaration of orthographic independence from Great Britain.

Just as important, their consensual pronunciations (listed first in each entry and supported now by audio versions) have the authority of national “platform speech” for both television announcers and offshore learners, many of which now use “Chicago to L.A.” pronunciation as telephone-sales professionals working our of cubicles in India, China, and the Philippines.

Dictionaries also meet the calibration requirement. It’s well known that the relative frequency of any headword can be ranked via its number of letters plus the number of its listed definitions. Even more important, especially with respect to transparency, the relative difficulty of any spelling-bee format question can be determined by its number of letters divided by the numerical position of its definition. Thus we would rank the first question presented below as being 7 points more difficult than the second question.

- 1: Please identify the 7-letter headword (in Dictionary X) whose pronunciation is presented as /keuhntayn"/, and whose 4 definition (out of 8) is ‘to keep under proper control; restraint.’
- 2: Please identify the 4-letter headword (in Dictionary X) whose pronunciation is presented as /run/, and whose 2 first definition (out of 125) is ‘to move with haste; act.’

NOTE: Correct answers are CONTAIN (1) and RUN (2) Dictionary X is Random House Webster’s College Dictionary.

Like other American college-size dictionaries, Dictionary X contains roughly 60,000 headword-pronunciation-definition combinations like those used in constructing the above questions. Hence the results of a random sample comprising 30 such questions can be taken seriously as an authoritative assessment of vocabulary size, e.g., 15 correct answers (50%) would represent a personal vocabulary of 30,000 such combinations — verifiably so via additional testing results.

The above case for the use of modern metrology in educational testing is presented here far more as a basis for queries than as an educational proposal on its own. By way of illustration,

here are some possible questions concerned parents might raise after mulling over what's presented on the Standards for Educational and Psychological Testing (SEPT) web site.

- 1) Given SEPT's stated concern with "friendly" testing, are there any educational or psychological tests in current use that a SEPT consensus would describe as "unfriendly"?
- 2) Does SEPT officially recognize measurement science (metrology) as a discipline relevant to its mission and concerns?
- 3) Does the New York Times daily crossword puzzle meet SEPT standards as a vocabulary test or general knowledge test?
- 4) Does the Scripps National Spelling Bee meet SEPT standards as an oracy test?
- 5) Would SEPT encourage the National Institute for Standards and Technology (NIST) to evaluate specific tests and testing innovations (e.g., online testing)?
- 6) As a tripartite body (AREA, APA, and NCME), SEPT is today America's largest and most powerful testing organization, especially when it comes to government contracts. Should SEPT as a national asset encourage public examination of its procedures and policies?
- 7) Would SEPT consider replacing the term "psychometrics" with "gnosti-metrics" or another coined term that more accurately and felicitously describes what most test-makers do?

Let's concede here that questions like these are bound to come across as somewhat speculative, far more so than fact-based questions that specific parents might ask regarding how American tests are actually working, for good or ill, in their specific neighborhoods. But since SEPT itself has already established participation-friendly reaction procedures, I believe that its measurement-science professionals will welcome thoughtful questions like these from American parents and other concerned citizens.

STEP ANNOUNCEMENT. . . . The Standards for Educational and Psychological Testing is produced through a long-standing collaboration of three associations: [American Educational Research Association](#) (AERA), the [American Psychological Association](#) (APA) and the [National Council on Measurement in Education](#) (NCME). First published in 1966, the Standards have been revised periodically. The collaboration of the three associations has been formalized in a cooperative agreement that creates a management structure and sets procedures for maintaining and revising the Standards.

The agreement designates a publisher, currently AERA, to handle distribution and sale of the Standards; a Management Committee composed of one representative from each association to oversee all aspects of the Standards; and a Joint Committee to prepare the actual revision for approval of the associations. From the outset, all income above costs of printing and distribution of the Standards have been placed in a reserve to finance the work of future revisions

Citation: Oliphant, R. (2011). Modern metrology and the revision of our *Standards for Educational and Psychological Testing*: An open letter to American parents. *Nonpartisan Education Review / Essays*, 7(4). Retrieved [date] from <http://nonpartisaneducation.org/Review/Essays/v7n4.pdf>

*Robert Oliphant, a columnist for [Nonpartisan Education Review](#) has a PhD in English Philology from Stanford (1962), where he studied Old English lexicography under Herbert Dean Meritt. The film version of his "A Piano for Mrs. Cimino" earned a Monte Carlo award for Bette Davis. His recent eBooks (available via the [Nonpartisan Education Review](#)) include *BigVocab*[®], which covers dictionary-based measurements of vocabulary size in detail. He is an emeritus Professor of English at California State University, Northridge.