"SOPHIMETRICS": A REINTERPRETATION OF MENO'S PARADOX APPLIED TO STANDARDIZED TESTING IN SCHOOLS

Darren Pascavage
Georgia State University

> But the sophist. . .was generally proud of his name, and proud of providing services for which the [people] were so obviously hungry
>
> James L. Jarrett
> *The Educational Theories of the Sophists*

In an editorial in the June 3, 2001 edition of the New York Daily News, New York state education commissioner Richard Mills defends standardized tests. The headline of this editorial was "High-stakes School Tests: They Measure Knowledge." This was not the first time I had seen or heard the claim that standardized tests measured knowledge. Importantly, however, I cannot recall having ever seen or heard the claim made by a philosopher of education. This is so, I suspect, because the elusive nature of knowledge makes the idea of a standardized test that measures knowledge seem foolish, if only because it would prove so difficult to settle on a definition of knowledge before writing any questions. Nonetheless, there are many who claim that standardized tests do, in fact, measure knowledge, and I cannot help but ask, "Who has convinced them of this fact, and how did they manage to do so?"

This paper attempts to address these questions by arguing that claims regarding standardized tests as measures of knowledge should be indicted as sophistry. Specifically, it will be shown that the operationalized, behavior-oriented definitions of knowledge employed by measurement specialists should be interpreted as a kind of "measurement sophistry"—what I will call sophimetrics. On this view, psychometricians first assume that there must be a way to measure knowledge (a "right answer"), then define knowledge in such a way as to permit the identification of specific skills that are presumed to be evidence of knowledge. Tests are constructed which ask questions leading to knowledge being measured, and the measurement process is defended with carefully constructed, technical arguments (rhetoric) specifically intended to respond to criticism. Corporations profit from the widespread use of standardized tests, and employ the psychometricians whose technical skill makes tests marketable. Taken together, one sees evidence of the kind of sophistic stance that Plato argued against—the pursuit of wisdom, or truth, replaced with a kind of ends-oriented approach to inquiry that presumes the existence of both a result and a means to achieve it, coupled with the provision of services for a fee and, thus, a profit motive.

Meno's Paradox and Sophistry

In Plato's Meno, Meno asks Socrates if inquiry is possible. Socrates begins his reply by restating Meno's question as follows:

> It is thus impossible for a man to inquire either into what he knows, or into what he does not know. He cannot inquire into what he knows, for he knows it, and there is no need for inquiry into a thing like that. Nor would he inquire into what he does not know; for he does not know what he is to inquire into.[1]

This quandary (known as Meno's Paradox) is resolved by Socrates via the Theory of Recollection. The paradox is often understood to be an example of Meno's sophistry: eristical in nature, Meno's intent is not to pursue wisdom but, instead, merely to stump Socrates through the use of an apparent paradox.

Scholars have characterized Sophists and sophistry in a variety of ways: some favorable, some not.[2] Of particular relevance to this paper is the link between sophistry and rhetoric. The Sophists were recognized as skilled rhetoriticians, offering to teach others the art of persuasion for a fee. Of note is the observation, made by Roger Chance, that Sophistic rhetoric was "the art of strengthening the weaker statement, of turning 'bad' into 'good' and 'black' into 'white,' of drowning reason and experience in a flood of words."[3] It is this notion of rhetoric that gives rise to the term "sophistry."[4]

There are many ways in which Meno's Paradox, and indeed the Meno itself, can be applied to contemporary schooling. One could hold up the paradox as an example of sophistry and then illustrate the ways in which sophistry has taken hold in schools. On this view, the focus on getting the right answers to questions, and in developing specific teaching practices that will enable students to learn how to respond to, for example, test questions, all at the expense of pursuing truth for the sake of truth or learning for the sake of learning, exemplify the manner by which sophistry is present in schools.[5]

Differently, one could interpret the paradox more literally and then consider its implications on the practice of testing in schools. In fact, a liberal reinterpretation of Meno's Paradox illustrates well how the problem of testing for knowledge emerges. To wit: if students know something (call it X) then there really is no point in testing to see if they know X since, alas, they know it. Presumably, the teacher would know that the student knows X via class work, conversations with the student, student projects, and familiarity with the student's written work, such that the use of a standardized test would be redundant in light of what the teacher already knows about what the student knows. Given this, it can be concluded that teachers justify the belief that a student knows (to themselves, colleagues, parents, as well as the student herself) via a kind of

epistemological coherence. Standardized tests, likewise represent, at best, a kind of fallible foundationalist attempt at justification.

This conclusion demands clarification. I assert that teachers employ a kind of coherent justification when, in an attempt to justify their belief that the student does/does not know, they evaluate a wide range of different school work completed over the course of the year. This may include classroom tests and quizzes, projects, reports, presentations, and student participation in class. The collective weight of this evidence either does or does not stand to justify the belief of student as knower.

Standardized tests, by contrast, yield a fallible score which is taken as evidence, almost exclusively, of the student's status as a knower. In cases where the teacher and the test score differ with respect to the student's status as a knower (for example, where the teacher believes that a student does know, but the test score suggests that the student does not,) it is the test score which is assumed (possibly by virtue of its privileged status as a part of state legislation, possibly because of its quantitative and supposedly objective nature) to be correct. This should be obvious given the structure of accountability programs, whereby teachers and administrators are expected to be able to account for poor student performance on tests independent of the performance of students in the classroom.[6] If a teacher's recommendation to pass or retain a student based on classroom performance conflicts with the test score that a student makes on a test of promotion, it is the teacher, and not the psychometricians who designed the test, who is expected to account for this disagreement.

Likewise, if the student does not know X, the teacher would know this in light of the same kinds of evidence mentioned previously—inadequate classwork, unsatisfactory conversations with the student, poorly done projects, and familiarity with the deficient quality of a student's written work. Naturally, such an interpretation assumes that teachers can recognize that the student knows, and recognize that they (the teachers) know that the student knows.[7]

The possible use of standardized tests becomes viable in two instances. The first instance occurs when the teacher has no evidence that the student knows. The second instance occurs when the teacher has evidence of student knowledge, but that evidence is not accepted by policy makers or administrators as sufficient. We might say that the former instance is the result of poorly considered classroom policies, where students have been neither expected nor required to demonstrate knowledge of subject matter. In such instances, school administrators may have some responsibility to attempt to ascertain if, and to what extent, student learning has occurred. Seen in this light, the motives behind standardized testing might possibly be justified.[8] Likewise, we might say that the latter instance results from a lack of faith in teachers, generally, and their

ability to assess student knowledge, specifically.[9] The standardized test, in such an instance, replaces de facto the teacher's professional evaluation of both the student's work and the student's status as knower of subject matter.

To clarify this point further, consider a student in a classroom located in a state where a standards-based curriculum determines what the student is supposed to learn. At the end of the academic year, the status of the student as knower can be rendered in the simplest possible terms as one of the following:

1S. The student knows the required elements of the curriculum.

2S. The student does not know the required elements of the curriculum.

In most circumstances, a teacher is present throughout the year and has assigned work, graded projects, administered tests, held conversations with the student, and so forth, so as to form some basis of evidence to justify the belief that the student does or does not know. Presumably the teacher's status as knower, relative to what the student knows, is one of the following:

1T. The teacher knows the student's status as knower with respect to the curriculum.

2T. The teacher does not know the student's status as knower with respect to the curriculum.

The same legislative mandate that establishes the priority of such curricula often includes an assessment component, so that states administer standardized tests aimed at determining what students know. Importantly, this is seen as necessary even in cases where the teacher may already know the student's status, again because the professional judgment of teachers is seen as inadequate with respect to evaluating what students know. Were it otherwise, such programs would be unnecessary. Therefore, we can conclude that states, as a matter of legislation, accept 2T. In so doing, the need for objective, standardized tests is manufactured.

But using a standardized test to measure knowledge presumes that a standardized test could be designed to measure knowledge. This is hardly a trivial activity, because it is unclear how one would adequately define knowledge or, for that matter, would know that the test was measuring knowledge once knowledge was defined. In other words, and again recalling Meno's Paradox, if student knowledge is the object of inquiry, how would one be able to identify (or construct) a test that measured knowledge, and then recognize that fact? Psychometricians, by defining knowledge operationally and linking that definition to the concept of construct validity, have a ready response to such a challenge.[10]

Measuring Knowledge in Schools

Demands for accountability in education have made the (presumed) measurement of student knowledge an important goal. To accomplish the task of measuring knowledge, itself, and knowing that knowledge is being measured, test designers accept, as a definition of knowledge, an operational definition. Operational in this context means, following Robert Ennis, that there is "an important relationship between the meaning of a term and the instruments and procedures that one would use to see whether the term applies to a particular situation and, if so, how."[11] Such a definition carries the advantage of rendering knowledge as measurable, thus making viable the possibility of constructing a test which has knowledge as its construct.

An example of this kind of definition of knowledge as a measurable psychometric objective is offered by Tom Kubiszyn and Gary Borich in clarifying Bloom's Taxonomy of educational objectives in the cognitive domain. "Objectives at the knowledge level" write Kubiszyn and Borich, "require the students to remember."[12] A list of "action verbs" that are synonymous with knowledge include, according to Kubiszyn and Borich, "list," "match," "recall," "recite," and "state."[13]

Were one to object that such a definition of knowledge was too narrow, given its obvious emphasis on *techne* at the expense of *arête*, Kubiszyn and Borich would counter that any other definition of knowledge might not be measurable and, therefore, not useful. Indeed, they explicitly warn against "being too 'sophisticated' in measuring learning outcomes," cautioning readers to write items that are "simple" and "straightforward."[14]

Such an account also links measurable knowledge directly to the test, so that without the test there is no knowledge. The question, "What is knowledge?" can thus be answered, "That which the tests measures," while the question, "What does the test measure?" is easily answered as "Knowledge."[15] Criticism of this reasoning as circular has a practical response: that there is no way to measure knowledge unless knowledge is defined in a measurable way. On this view, such an account of knowledge is justified on practical grounds, since the definition (as means) eventually permits the calculation of test scores (as ends).

Conclusion

This paper has argued that the claim that standardized tests measure knowledge should be condemned as sophistic. Inasmuch as behavioral notions of observability and measurability are the primary terms by which knowledge is defined, psychometricians have carefully crafted a notion of knowledge which resists disputation. Increasing calls for accountability and student testing have created an environment within which claims regarding the measurement of

knowledge are liable to be accepted without careful scrutiny. That there is a significant profit behind such persuasion (estimated to be on the order of $400 million for high-stakes tests alone last year) raises questions about the degree to which psychometricians are willing to forgo the pursuit of wisdom in favor of providing a willing and eager public with a service they desire.[16] Viewed in this way, psychometrics becomes "sophimetrics" as firms that employ measurement specialists seek to profit from the desire of the public to "know" if, and to what extent, students "know" the required curriculum. Indeed, knowledge of what students know might possibly require just a bit more than No. #2 pencils and bubble-sheets. Thoughtful interlocution of students by teachers, the drawing of careful distinctions between propositional and non-propositional knowledge types as exemplars of what students should know, and the elimination of cart-before-the-horse definitions of knowledge that yield to measurement but fall to close examination, are aspects of contemporary schooling where philosophers of education have much to offer. The haste with which philosophers of education rise to these challenges may make the difference between student assessments that, through thoughtful design, posses epistemic virtue and those which are bereft of any measure of epistemic legitimacy. The latter already exist, in droves –we should help to develop the former.

Notes

1. Reginald Allen, trans., *Euthyphro, Apology, Crito, Meno, Gorgias, Menexenus: The Dialogues of Plato*, Volume 1 (New Haven, Yale University Press, 1984), 163.

2. For example, see Thomas R. Martin, *Ancient Greece: From Prehistoric to Hellenistic Times* (New Haven: Yale University Press, 1996); Werner Jaeger, *Paideia: The Ideals of Greek Culture* (New York: Oxford University Press, 1945); Peter A. Redpath, *Wisdom's Odyssey: From Philosophy to Transcendental Sophistry* (Amsterdam: Rodopi Press, 1984); Ch. Perelman and L. Olbrechts-Tyteca, *The New Rhetoric: A Treatise on Argumentation*, John Wilkinson and Purcell Weaver, trans. (Notre Dame: University of Notre Dame Press, 1969).

3. Roger Chance, *Until Philosophers are Kings: A Study of the Political Theory of Plato and Aristotle in Relation to the Modern State* (Port Washington, New York: Kennikat Press, Inc., 1968), 29.

4. Ibid.

5. Deron R. Boyles, "Sophistry, Dialectic, and Teacher Education: A Reinterpretation of Plato's *Meno*," *Philosophy of Education* (1996): 102–109.

6. Ralph Tyler was among the first to champion the use of tests for such purposes; see George F. Madaus and Daniel Stufflebeam, eds., *Educational Evaluation: Classic Works of Ralph W. Tyler* (Boston: Kluwer Academic Publishers, 1989).Additionally, Herbert Kliebard recounts the links between testing, curriculum, and the social efficiency movement; see Herbert M. Kliebard, *The Struggle for the American*

Curriculum, 1893-1958 (New York: Routledge and Kegan Paul, 1987).See also: Harvey Goldstein and Toby Lewis, "The Scope of Assessment," in Harvey Goldstein and Toby Lewis, eds., *Assessment: Problems, Developments, and Statistical Issues* (Chicago: John S. Wiley and Sons, 1996); John Gray, "The Use of Assessment to Compare Institutions," in Goldstein and Lewis, eds.; Robert Glaser and Anthony J. Nitko, "Measurement in Learning and Instruction," in Robert L. Thorndike, ed., *Educational Measurement*, Second Edition (Washington: American Council on Education, 1971), 625-670;Junius A. Davis, "Use of Measurement in Student Planning and Guidance," in Thorndike, ed., 671–679; Robert Wood, *Measurement and Assessment in Education and Psychology* (London: The Falmer Press, 1987), 233–248; Joy A. Frechtling, "Administrative Uses of School Testing Programs," in Robert L. Linn, ed., *Educational Measurement* (New York: Macmillan Publishing Company, 1989); John R. Frederiksen and Allan Collins, "A Systems Approach to Educational Testing," *Educational Researcher* 18, no. 9 (1989): 27–32.

7. Knowing that one knows, knowing that someone else knows, and knowing what someone else knows should not be thought of as insignificant epistemological activities. For a consideration of problems that arise from such knowledge relationships, see Richard Feldman, "Fallibilism and Knowing That One Knows," in Kenneth G. Lucey, ed., *On Knowing and the Known: Introductory Readings in Epistemology* (Amherst, NY: Prometheus Books, 1996), 257–270, Bertrand Russell, "The Analogy Argument for Other Minds," in Louis P. Pojman, ed., *The Theory of Knowledge: Classic & Contemporary Readings*, Second Edition (Belmont, CA: Wadsworth Publishing Company, 1999), 514-515; and Norman Malcom, "The Behavioral Criterion and the Problem of Other Minds," in Pojman, ed., 527–532.

8. Whether standardized tests are the best means for administrators to accomplish this is a debatable point.

9. See Jeannie Oakes, "Tracking, Inequality, and the Rhetoric of Reform: Why Schools Don't Change," in H. Svi Shapiro and David E. Purpel, eds., *Critical Social Issues in Education: Transformation in a Postmodern World*, Second Edition (Mahwah, NJ: Lawrence Erlbaum Associates, 1998), 142. See also K.A. Sirotnik and J. I. Goodlad, "The Quest for Reason Amidst the Rhetoric of Reform: Improving Instead of Testing Our Schools," in W. J. Johnson, ed., *Education on Trial: A Midterm Report* (San Francisco: Institute for Contemporary Studies, 1985), cited in Oakes.

10. Construct validity refers to the extent to which a test measures what it is designed to measure. See Lee J. Cronbach, "Test Validation," in Robert L. Thorndike, ed., *Educational Measurement* (Washington: American Council on Education, 1971), 446; see also Lee J. Cronbach and Paul E. Meehl, "Construct Validity in Psychological Tests'" in Harry S, Broudy, Robert H. Ennis, and Leonard I. Krimerman, eds., *Philosophy of Educational Research* (New York: John Wiley and Sons, Inc., 1973), 567–594.

11. Robert H. Ennis, "Operational Definitions," in Broudy, et al., 651.

12. Tom Kubiszyn and Gary Borich, *Educational Testing and Measurement: Classroom Application and Practice* (New York: HarperCollins Publishers, 1996), 60.

13. Ibid.

14. Ibid., 56.

15. Robert Ennis characterizes this close relationship between that which is measured and the measurement instrument as operationalism. Ennis goes so far as to identify such a situation as a specific kind of operationalism: "Equating a Phrase or Sentence Containing the Term With a Phrase or Sentence About a Combination of 1. Operations and Observations." His examples use IQ as the construct, and not knowledge, but this makes no substantive difference. Ibid., 657.

16. See Andrew Goldstein, "Making Another Big Score," *Time*, March 12, 2001, 67, and David J. Hoff, "States Spend Nearly Half-a-Billion On Testing," *Education Week*, March 4, 2001.