



REVIEW

The Power of Tests
A Critical Perspective on the Uses of
Language Texts
Elana Shohamy (Longman, 2001)

BOB BROAD

Illinois State University

Most of us interested in assessment would agree that our field's thinking about what educational evaluation is and does has developed through at least a couple of distinct historical-conceptual phases. Initially, we understood evaluation as a way of gathering information about what people had learned, and we focused on getting the most accurate and reliable information possible. More recently, our conception of evaluation was broadened and enhanced as we came to understand the rich cluster of assessment concepts that point out to us how assessment not only collects data, but also teaches people (produces knowledge) and transforms educational systems and processes. We, therefore, now also attend to validity issues related to consequences, washback, and assessment's "educative" nature (Wiggins, 1998). In this second historical phase, we still seek valid and valuable information about learning and learners, but we also turn our attention to the impact of assessment decisions on the broader educational project.

In *The Power of Tests*, Elana Shohamy grapples with the nature and implications of a third, relatively new phase or mode of consciousness related to educational evaluation. Policymakers and bureaucrats, having fully grasped the transformative

Bob Broad is Professor of English at Illinois State University, where he teaches graduate courses in writing assessment and undergraduate courses in composition pedagogy. His book *What We Really Value: Beyond Rubrics in Teaching and Assessing Writing* was published in 2003. His articles have appeared in the journals *Research in the Teaching of English*, *Journal of Writing Assessment*, *Assessing Writing*, *Works and Days*, *African American Review*, and *Illinois English Bulletin*.

Direct all correspondence to: Bob Broad, Department of English, Illinois State University, Normal, IL 61790-4240.

power of assessment (the second phase just noted), have taken to implementing tests explicitly as a way of bringing about educational change. In other words, tests have come to function as policy tools and, in many cases, as substitutes for meaningful education reform. Shohamy's experiences as a student, test-taker, citizen, researcher, educator, and parent of a child in schools all inform her exploration and critique of this impulse by educational leaders to use tests to achieve policy goals, in other words to wield political and educational power.

Shohamy divides her book into four main sections: Part I: "The Power of Tests" sets a theoretical, political, historical, and multicultural context for her inquiry into testing. Part II: "Uses of Tests: Studies and Cases," presents three studies of specific tests in her home nation of Israel. Part III: "Uses of Tests: Conclusions and Interpretations" pulls together her critical reflections on the cases she presents in Part II. The final section, Part IV: "Democratic Perspectives of Testing," struggles admirably to weave together critical education theory with a concrete and detailed plan of action by which to "democratize" the enterprise of educational testing.

Each of the book's four sections accomplishes something valuable on the topic of testing, and moves Shohamy's project forward. Part I lays out issues that caught Shohamy's attention when she awakened from her traditional psychometric slumber and began investigating the "uses" of tests, their "meanings," the "feelings" they evoke in test-takers and others, and—most compelling in my reading—the voices of those affected by tests. In Chapter 2, "Voices of Test Takers," we hear, for example, about a school in Israel in which the principal excluded from a national reading test a group of students he expected would perform poorly on the test, and instead sent them to watch a movie in the gym. "The article [from which the account is drawn] provides a close insight into [students'] feelings of humility [*sic*] and shame as the non-participation perpetuated their feelings of failure and contributed to their low self-esteem." Meanwhile, the same principal refused the request of one high-performing student to miss school that day for personal reasons; the principal insisted that the student participate in the test because he feared the student's absence would lower the school's average test performance.

Chapter 2 also offers us the tale of a young driver testing for a license. With the examiner sitting in the passenger seat, the driver waits at a red traffic light. As the light turns green, a pedestrian enters the street against the light, crossing the path of the vehicle. The examiner scolds the driver for not proceeding since the light has turned green. The driver explains that the obligation to follow a traffic rule is not sufficient reason to run over a pedestrian. The tester fails the driver, noting that the driver "Stops at green light." With detailed and personal accounts like these and others, Shohamy evokes a vividness and personal tone often absent from assessment discussions.

As a qualitative researcher, I found myself particularly drawn to Part II, in which Shohamy presents summaries of three qualitative studies she conducted with colleagues to document and understand the intentions, uses, and consequences of three specific assessments: a reading comprehension test, a test of Arabic proficiency, and a test of oral English proficiency. To the credit of Shohamy and her fellow researchers, the findings of these three studies are appealingly and convincingly complex; the social and educational results of each test are different from what

anyone (including the researchers) anticipated as well as different from the impact of the other tests. Furthermore, Shohamy connects the theoretical work of the other sections with the complex and detailed findings of the three studies. Thus, the book as a whole gains persuasiveness by being firmly grounded in data collected and analyzed in three substantive studies. Shohamy brings to her discussion of educational testing what anthropologist Robert W. Connell (1982) called “a good dose of awkward facts.”

Part II is also useful in two other ways. First, we get to hear the perspectives of testing bureaucrats, teachers, and student test-takers. They articulate views that are sometimes predictable but other times illuminating. In Part II, however, we get very few direct quotations from participants. Because they are presented almost exclusively in general terms, the voices of these various stakeholders in the testing enterprise are muted and abstracted. I was struck by the contrast between the muted character of the research reporting in Part II with the robustly humorous, poignant, and detailed quotations that appear in Chapter 2, just described.

The final reason I particularly valued Part II is the lesson it teaches us about opinions and research—about testing or practically anything else. Shohamy clearly shows us that people across the spectrum of opinions about educational testing routinely reach conclusions about issues without conducting any actual systematic inquiries. Instead, they reach decisions about the consequences and value of various assessments based on their theories, impressions (often self-interested), or personal experiences.

For the English Inspectorate [the Israeli agency responsible for overseeing English as a Foreign Language instruction], the reactions to the test (*without their ever examining its real impact*) were overwhelming: “The introduction of the oral test was a great success and created a very positive educational impact . . . (Shohamy, 2001 p. 84)

The administrators express this optimism about a test regarding a part of which teachers commented, among other things, “I feel it is all a sham” and “I think this part of the test is really meaningless.” The point is that anyone concerned with developing assessments that support the best teaching and learning needs to conduct an inquiry—or carefully attend to others’ inquiries—to be able to make sound decisions about tests and other assessments.

Part III presents Shohamy’s intense and methodical effort to work from a systematic analysis of the three cases of tests presented in Part II toward “a model that describes the process of using tests for power and control” (p. 95). Among other things, Shohamy lays out the impact on teachers and teaching of imposed tests, in language that is not only harrowing but that also precisely foreshadows the findings of Hillocks (2002) in his study of five statewide writing assessments in the United States:

[Teachers] experience fear and anxiety as students, principals and parents demand preparation for this high-stake test. Since teachers have no explicit knowledge of how to teach the topic they turn to the most immediate pedagogical source, the test itself, to learn how to carry out and comply with the new orders. The test becomes the single most influential pedagogical source, and the *de facto* knowledge. Teachers

are reduced to “following orders”—a frustrating role as their responsibility increases while their authority diminishes. (p. 107)

Finally, in Part IV, the author presents her boldest venture of all: She lays out a plan of action “to limit, minimize, and control such powerful uses of tests as described in this book.” She offers “critical language testing” as an area of inquiry located within critical pedagogy. Drawing on the work of Freire, Giroux, and others, Shohamy provides a series of principles and guidelines that she believes will make testing more democratic because they will “minimize, limit and control the powerful uses of tests” (p. 131).

Near the end of Part IV (and the end of the book), Shohamy offers a series of six examples of the “democratic” approach to testing that she seeks for education. I found the examples encouraging, but I also wished for much more detail in this account of what assessment *should* be following 130 pages of discussion of what testing should not be. It may be too much to ask one book to provide the sort of thorough-going critique Shohamy delivers and also to provide detailed examples enough to give readers a sense of how they might act in their worlds so as to transform assessment in the ways the author recommends. The book concludes with Shohamy’s discussion of test-makers’ responsibilities and test-takers’ rights.

In *The Testing Trap*, George Hillocks (2002) takes George W. Bush to task for the naïve promises he made related to education and assessment during his 2000 presidential campaign and especially for the hollowness of the “No Child Left Behind” (NCLB) legislation Bush championed. Hillocks expresses his frustration that NCLB ended up implementing testing requirements as a relatively cheap substitute for what was really needed: education reform, including new resources for supporting teachers (e.g., by reducing class size) and helping those teachers develop professionally through graduate study and attendance at professional conferences.

It is sobering to see Shohamy, in her book published a year before Hillocks’s, criticizing Bill Clinton for almost exactly the same political strategy around tests. In the 1990s, Clinton proposed a set of national standards and national tests, again with the same goal Bush later pursued: achieving educational improvements without investing significantly in the education system. This is a helpful reminder that the political problems of testing in the United States, although currently dogging educators in the form of Bush’s NCLB Act and its accompanying testing requirements, are rooted more deeply than the ambitions of any one politician or party. They appear to be woven throughout the contemporary American approach to education.

Much of the powerful critical and theoretical analysis Shohamy presents in the book comes from two French social theorists: Michel Foucault and Pierre Bourdieu. Both of these theorists are widely admired for the breadth and intensity of their analysis of the functioning of social power. Both theorists have also been faulted for presenting an account of the vast machinations of power so overwhelming that their readers are likely to surrender any hope of transforming social reality. That is, students of Bourdieu and Foucault often end up believing that the power structure is so massive, elaborate, and entrenched that nothing can be done to change it. Shohamy, however, does not suffer from this lack of agency. Part IV

lays out an energetic and detailed plan for documenting and addressing what Shohamy sees as abuses of the power of tests.

The thing that left me scratching my head was that, for all her emphasis on Foucault and Bourdieu, Shohamy seems to work throughout the book with a sense that the power of tests is inherently “bad” and oppressive, and furthermore that it is possible to limit or even eliminate the ways in which tests wield power in the educational system. That is, to have testing without the exercise of power. By contrast, the first and last lesson I learned from Shohamy’s most frequently cited theoretical source, Foucault’s (1977) *Discipline and Punish*, is that it is generally impossible to avoid or limit the exercise of power. Instead, Foucault wants us to understand the pervasiveness (and de-personalized) dynamic and character of power and consider the different ways in which it does and should work.

However, Shohamy resists and resents how tests wield power in educational systems. My readings of Wiggins (1998) and Huot (2002), however, have taught me to accept the power of educational assessment but—crucially—to attach that transformative power to the “right” kind of assessment. As Wiggins puts it, we cannot avoid assessments, and we cannot (and should not) avoid teaching to those assessments. The point of struggle is carefully to develop assessments that transform curriculum and teaching in the direction of best practice, as opposed to curriculum and teaching that trivialize and demean our disciplines.

In the field of writing assessment, the writing portfolio is the favored technology and ideology for delivering authentic and educative assessment. Yet portfolios receive very little attention in Shohamy’s book. Perhaps because she thinks more consistently in terms of tests rather than other assessment approaches, she holds to a view of assessment that is thoroughly cautious and skeptical. I would have been glad to see Shohamy embrace the power of assessment to drive teaching and learning *when it’s the right assessment*.

Interestingly, however, Hillocks shares some of Shohamy’s skepticism regarding whether and how successfully tests can bring about productive educational change. Hillocks’s analysis clearly favors portfolio assessment as the best approach to judging students’ writing ability. However, he also cautions that the two states in which large-scale portfolio assessment have succeeded—Kentucky and Vermont—enjoyed and needed an extensive network of professional development, led mainly by multiple sites of the National Writing Project in each state. In the analysis of interviewees from both these states, large-scale authentic assessment of writing would not have succeeded in the absence of a sustained network of professional development.

So when Shohamy protests against the clumsy or ill-considered political use of tests, she is in good company. A new assessment alone is unlikely to bring about favorable educational change, and Shohamy’s book (especially Part II) provides highly compelling empirical evidence and theoretical arguments in support of a skeptical reception for such assessment efforts. In the opening pages of the book the author recounts her psychometric fall from innocence in a 1985 meeting with the national [Israeli] inspector responsible for instruction in English as a foreign language (EFL). The inspector was highly enthusiastic about the new EFL test because, as he put it, “now teachers will have no choice but to teach students to

speak English.” Shohamy succinctly reflects on what this bureaucrat had overlooked.

It did not matter that teachers were not trained in teaching spoken language, or that there was no curriculum geared for teaching it. He was convinced that the test and its power would take care of everything. (p.xii)

I feel ambivalent about this specific aspect of Shohamy’s book. She rightly challenges the short-sighted and heavy-handed use of tests to drive teaching and learning, but she leaves for another day (or another book) as trenchant an exploration of what a *good* large-scale assessment, backed up with investment in education and other professional development for teachers, looks like. (Remember that Hillocks strongly suggests the blend of good assessment and robust professional development was and is the secret formula that has sustained Kentucky’s and Vermont’s statewide writing portfolio assessments.) As a result, Shohamy risks reinforcing precisely the long-held view that Brian Huot’s (2002) *Re-Articulating Writing Assessment for Teaching and Learning* works to counteract: that assessment is necessarily a threat to good teaching and learning. To the contrary, Huot argues that “we need to reframe assessment for its pedagogical value” (p. 4) and claim assessment as a powerfully positive part of teaching any field of knowledge.

I celebrate Elana Shohamy’s critique of misuses of tests and inadequate tests, and I look forward to a book that undertakes the reframing and claiming called for by Huot with as much intelligence, with as high quality empirical research, and with as much passion for teaching and learning as Shohamy displays in her book.

References

- Connell, R. W. (1982). *Making the difference: Schools, families, and social division*. Sydney: George Allen & Unwin.
- Foucault, M. (1977). *Discipline and punish: The birth of the prison* (A. Sheridan, Trans.). New York: Pantheon.
- Hillocks, G. (2002). *The testing trap: How state writing assessments control learning*. New York: Teachers College Press.
- Huot, B. (2002). *(Re)Articulating Writing assessment for teaching and learning*. Logan: Utah State University Press.
- Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. Harlow, England: Longman.
- Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco: Jossey-Bass.