

BOOK REVIEW

Anya Kamenetz

The Test

***Why Our Schools Are Obsessed with Standardized
Testing—but You Don't Have to Be***

Public Affairs, 2015

reviewed by Richard P. Phelps

PERHAPS IT IS BECAUSE I avoid most tabloid journalism that I found journalist Anya Kamenetz's loose-cannon introduction to *The Test: Why Our Schools Are Obsessed with Standardized Testing—but You Don't Have to Be* so jarring. In the space of seven pages, she employs the pejoratives "test obsession," "test score obsession," "testing obsession," "insidious . . . test creep," "testing mania," "endless measurement," "testing arms race," "high-stakes madness," "obsession with metrics," and "test-obsessed culture."

Those unmeasured words fit tightly alongside assertions that education testing, standardized testing, or high-stakes testing is responsible for numerous harms, ranging from stomachaches, stunted spirits, family stress, "undermined" schools, demoralized teachers, and paralyzed public debate, to the Great Recession (pp. 1, 6, 7), which was initially sparked by problems with mortgage-backed financial securities (and parents choosing home locations in part based on school

average test scores). Oh, and tests are "gutting our country's future competitiveness," too (p. 1).

Kamenetz tells us, "[T]here's lots of evidence that these tests are doing harm, and very little in their favor" (p. 13), but she has made almost no effort to search for counter-evidence.¹ Her sources for information in the relevant research literature include some of the country's most prolific proponents of her claim.² Ergo, why bother to look for it?

Had a journalist covering the legendary Hatfield-McCoy feud talked only to Hatfields, one might expect a surplus of reportage favoring the Hatfields over the McCoy's and a deficit of reportage favoring the McCoy's over the Hatfields.

Looking at tests from any angle, Kamenetz sees only evil. Tests are bad because they were used to enforce Jim Crow discrimination (p. 63). Tests are bad because some of the first scientists to use intelligence tests were racists (pp. 40–43).

Tests are bad because they employ the statistical tools of latent trait theory and factor analysis (the same tools, incidentally, currently used by tens of thousands of social scientists worldwide), but the "eminent paleontologist" Stephen J. Gould doesn't like them (pp. 46–48). (Gould argued that if you cannot measure something directly, it doesn't really exist.) And by the way, did you know that some of the early-twentieth-century scientists of intelligence testing were racists (pp. 48–57)?

Tests are bad because of Campbell's Law: "When a measure becomes a target, it ceases to be a good measure" (p. 5). Such a criticism, if valid, could be used to condemn any measure used to evaluate anything in society's realm. Forget health and medical studies, sports statistics, Department of Agriculture food-monitoring protocols, *Consumer Reports* ratings, Angie's List, the Food and Drug Administration. None of them are "good measures" because they are all targets.

Tests are bad because they are "controlled by a handful of companies" (pp. 5, 81); testing companies "determine . . . the quality of teachers' performance" (p. 20); and "tests shift control and authority into the hands of the unregulated testing industry" (p. 75). Criticisms such as Kamenetz's, if valid, could justify nationalizing all businesses in industries with high-scale economies (e.g., there are only four big national wireless telephone companies, so perhaps the federal

government should take over) and outlaw all government contracting. Most of our country's roads and bridges, for example, are built by private construction firms under contract to local, state, and national government agencies to the latter's specifications, just like most standardized tests: but who believes that those firms control our roads?

Kamenetz swallows anti-testing dogma whole. She claims that multiple-choice items can test only recall and basic skills (p. 35), that students learn nothing while they are taking tests (p. 15), and that U.S. students are tested more than any others (pp. 15–17, 75). That's true if you make calculations the way her information sources do—counting at minimum an entire class period for each test administration, even a one-minute DIBELS test; counting all students in all a school's grades as taking a test whenever any students in any grade are taking a test; counting all subtests in the United States independently (e.g., making each ACT count as five because it has five subtests) but only the whole tests in other countries; etc.

Standardized testing absorbs way too much money and time, according to Kamenetz. Later in the book, however, she recommends an alternative education universe of fuzzy assessments that, if enacted, would absorb far more time and money.

What are the author's solutions to the insidious, obsessive mania of testing? She engages in some Rousseauian fantasizing: all schools should be like her daughter's happy pre-school, where each student learns at his or her own pace (pp. 3–4) and the school's job is "customizing learning to each student" (p. 8).

Some of the book's latter half addresses "innovative" (of course) solutions that are not quite as innovative as National Public Radio's "lead digital education reporter" seems to believe. True, some interesting recent technologies suffuse Kamenetz's recommendations. But even jazzing up the context, format, and delivery mechanisms with the latest whiz-bang gizmos will not eliminate the problems inherent in her old-new solutions: performance testing, simulations, demonstrations, portfolios, and the like. Like so many Common Core Standards boosters advocating the same "innovations," she seems unaware that they have been tried in the past, with disastrous results.³

Lacking personal acquaintance with Ms. Kamenetz, I must assume the sincerity of her beliefs and her decisions about what to write. Nonetheless, if she had naively allowed herself to be wholly misled by those with a vested interest in education-establishment doctrine, the result would have been no different.

The book is basically a slapped-together rant, unworthy of an established journalist. Ironically, however, I agree with Kamenetz on many issues. Like her, I do not much like the assessment components of the old No Child Left Behind Act or the new Common Core Standards. (My solution would be to repeal both programs, not eliminate standardized testing.) Like her, I oppose the U.S. practice of relying on single proficiency standard for all students (pp. 5, 36). (My solution would be to employ multiple targets, as most countries do. Kamenetz would dump the tests.)

Again like Kamenetz, I believe it unproductive to devote more than a smidgen of time (at most half a day) to test preparation, with test forms and item formats, that is separate from subject-matter learning. And like her (p. 194), I am convinced that most test prep does more harm than good. Kamenetz, however, blames the tests and the testing companies for the abomination; in fact, the testing companies prominently and frequently discourage the practice. The advocates of test prep are actually the same testing opponents Kamenetz has chosen to trust. Trying to establish the legitimacy of non-subject-matter-related test preparation serves the argument of testing opponents because, if true, it would expose tests as invalid measurement instruments that can be gamed with tricks.

Like Kamenetz, I oppose firing teachers based on student test scores, as current value-added measurement (VAM) systems do, while the students suffer no consequences. I believe the VAM systems wrong because they rely on too-few data points and because student effort in such conditions is unreliable, varying by age, gender, socio-economic level, and more. I would eliminate VAM programs, or drastically revise them; Kamenetz, by contrast, would eliminate the tests.

Like Kamenetz, I believe that educators' cheating on tests is unacceptable, far more common than is publicly known, and should be stopped. I say, stop the cheating. She says, dump the tests.

It defies common sense to have teachers administering high-stakes tests in their own classrooms. Rotating test-administration assignments so that teachers do not proctor their own students is not particularly difficult, nor is rotating assignments further so that every testing room is proctored by at least two adults. So why aren't these and other remarkably simple fixes for test-security problems implemented? (Note that the education professionals responsible for managing test administrations are often the same individuals who complain that testing is impossibly unfair.)

The sensible solution is to take test-administration management out of the hands of those who may welcome test-administration fiascos and to hire independent professionals with no conflict of interest. Like many education insiders, though, Kamenetz would ban the testing and thereby reward those who have mismanaged test administrations, sometimes deliberately, by giving them a vacation from reliable external evaluation.

If Kamenetz were correct on all these issues—that the testing is the problem in each case—shouldn't we also eliminate examinations for doctors, lawyers, nurses, and pharmacists (many of which rely on the multiple-choice format, by the way)?

Our country has a problem. More than in most other countries, our public education system is independent, self-contained, and self-renewing. The education professionals who staff school districts make the hiring, purchasing, and school catchment-area boundary-line decisions. School district boundaries often differ from those of other governmental jurisdictions, confusing the electorate. In many jurisdictions, school officials set the dates for votes on bond issues or school board elections and can do so to their advantage. Those school officials are trained, and socialized, in graduate schools of education.

A half century ago, many faculty members in graduate schools of education may have received their own professional training in such core disciplines as psychology, sociology, or business management. Today, by contrast, most members of education school faculties are themselves education school graduates, socialized in the prevailing culture. The dominant expertise in schools of education can maintain that dominance with faculties that support the conventional wisdom and deny tenure to those who stray. The dominant expertise in education journals can control education knowledge when article submissions with agreeable results are accepted and those without are rejected.

Even doctoral training programs in testing and measurement now reside mainly in schools of education, inside the same cultural cocoon.

Standardized testing is one of the few remaining independent tools American society has for holding education professionals accountable to the public interest, rather than their own. Without valid, reliable, objective external measurement, education professionals can do largely what they please inside our schools, with our

children and our money. When educators are the only arbiters of the quality of their own work, they tend to rate it consistently well.

A substantial portion of *The Test's* girth is filled with complaints that tests fail to measure most of what students are supposed to or should learn: "It's math and reading skills, history and science facts that kids are tested and graded on. Emotional, social, moral, spiritual, creative, and physical development all become marginal. . . ." (p. 4). Kamenetz quotes Daniel Koretz: "These tests can measure only a subset of the goals of education" (p. 14). Several other testing critics are cited making similar claims.

Yet standards-based tests are developed through a multi-year process that enlists scores of legislators, parents, teachers, and administrators to serve on a variety of decision-making committees. The citizens of a jurisdiction and their representatives choose the content of standards-based tests. They could choose content that Kamenetz and the critics she cites prefer, but they don't.

If the critics are unhappy with test content, they should take their case to the appropriate decision-makers, voice their complaints at tedious standards commission hearings, and contribute their time to the rather monotonous work of test-framework review committees. I sense that such patient effort holds little interest for them; they would instead prefer to wield all decision-making power *ex cathedra*, to do as they think best for us.

Moreover, I find some of the testing critics' assertions about what *should be* studied and tested fraught with dangers. Public schools should teach our children emotions, morals, and spirituality?

Likely that prospect would concern most parents, too. But many parents' first reaction to a proposal allowing schools to teach children *everything* might instead be something like: first show us that you can teach our children to read, write, and compute: *then* we can discuss further responsibilities.

So long as education insiders insist that we must hand over our money and children and leave them alone to determine—and evaluate—what they do with both, calls for "imploding" the public education system will only grow louder, as they should.

It is bad enough that so many education professors write propaganda, call it research, and deliberately mislead journalists by declaring the absence of countervailing research and researchers. Researchers confident in their arguments and evidence should be

unafraid to face opponents and opposing ideas. The researchers Kamenetz trusts do all they can to deny dissenters a hearing.

In addition to testing, another potential independent tool for holding education professionals accountable could be an active, skeptical, and inquiring press knowledgeable about education issues and conflicts of interests. Other countries have it. Why are so many U.S. education reporters gullible sycophants?

Notes

1. Kamenetz did speak with Samuel Casey Carter, the author of *No Excuses: Lessons from 21 High-Performing High-Poverty Schools* (2000) (pp. 81–84), but she chides him for recommending frequent testing without “framing . . . the racist origins of standardized testing.” Kamenetz suggests that test scores are almost completely determined by household wealth and dismisses Carter’s explanations as a “mishmash of anecdotal evidence and conservative faith.”
2. Those sources are Daniel Koretz, Brian Jacob, and the “FairTest” crew. In fact, an enormous research literature revealing large benefits from standardized, high-stakes, and frequent education testing spans a century (Brown, Roediger, and McDaniel, 2014; Larsen and Butler, 2013; Phelps, 2012).
3. The 1990s witnessed the chaos of the New Standards Project, MSPAP (Maryland), CLAS (California) and KIRIS (Kentucky), dysfunctional programs that, when implemented, were overwhelmingly rejected by citizens, politicians, and measurement professionals alike. (Incidentally, some of the same masterminds behind those projects have resurfaced as lead writers for the Common Core Standards.)

References

- Brown, P. C., Roediger, H. L., & McDaniel, M. A. (2014). *Make it stick: The science of successful learning*. Cambridge, Mass.: Belknap Press.
- Larsen, D. P., & Butler, A. C. (2013). Test-enhanced learning. In K. Walsh (Ed.), *Oxford textbook of medical education* (pp. 443–452). Oxford: Oxford University Press. <http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2923.2008.03124.x/full>
- Phelps, R. P. (2012). The effect of testing on student achievement, 1910–2010. *International Journal of Testing*, 12(1), 21–43. <http://www.tandfonline.com/doi/abs/10.1080/15305058.2011.602920>