

Testing Today in Context: History, Impact, and Alternatives

While almost everyone who lives outside of classrooms—from US presidents to local PTA chairs—celebrates the presumed gains from testing, they fail to see what is being lost in the process. What we mean here by “testing” is the imposition on students of tests that are designed and graded by individuals, state agencies, or companies outside the classroom and the use of these tests to rank teachers, students, and schools.¹

Why has the test—throughout history, and perhaps most pervasively today—come to define our relationship to questions of truth, knowledge and even reality?²

In light of the very real ways that standardized testing has come to dominate educational reform in Chicago and across the nation, and how test results have come to define success and failure,³ it is important to explore a number of misconceptions that are central to understanding the answer to the above question.

Testing Misconceptions

These *misconceptions* include, but are not limited to:

- *Testing is the way we've always measured learning.*
- *How one performs on tests is an accurate assessment of one's learning.*
- *Testing is the best way to ensure that teachers are teaching well.*
- *Standardized high-stakes testing is a necessary evil.*
- *Testing is here to stay, and there is nothing we can do about it.*

Embedded in each misconception are taken-for-granted assumptions that need to be understood and questioned in order to engage in meaningful, substantive, and informed public dialogue about educational reform.

Toward this end, **CR_eATE** (Chicagoland Researchers and Advocates for Transformative Education) has produced this policy brief that uses research in the field of education to demystify how testing—particularly high-stakes testing—has come to serve as not only the mechanism of assessment, but also the foundation for teaching and learning. The way that tests are currently being used is so problematic that even test makers and proponents have raised concerns.⁴ This brief will draw on research to address the history of our current approach to educational testing; the reasons why even testing experts are concerned; the impact of this singular focus on how assessment is defined and practiced; and the range of assessment alternatives that exist, but have been ignored in current debates.

A Brief History of Testing

Misconception: *Testing is the way we've always measured learning.*

Facts:

- *The modern era of testing began at the turn of last century.*
- *Testing is based on that era's faith in science and the needs of industry at that time.*
- *Tests were used to sort and differentiate more than to assess learning.*
- *Tests are now being used as evidence of schools' 'failure.'*

To understand the role of testing today, we need to look back in time and examine the social, political, and economic factors that influenced school reform over the past century. We might begin by asking a basic question: Why do we believe in the value of testing and the significance of test results as a basis for educational decision making? Some adults may respond that “it worked for me” or that this is “what we’ve always done,” but history tells us that testing arose in particular forms, for particular reasons, and with particular goals. Over the past century, at least two events played pivotal roles in promoting the value of testing: one that took place at the start of the 20th century and one that emerged toward the end of that century.

The industrial and scientific revolutions of the late 19th and early 20th centuries promoted a faith in the scientific method that, when applied to the field of education, resulted in a growing admiration for what researcher Peter Taubman⁵ terms the “science of learning.” This paradigm shift was one response to the visible achievements within the natural sciences in measurement and in quantifying data to make claims of objectivity and truth, and soon extended to the social sciences as well.⁶ That is, one of the key aspects of the scientific method was to remove the scientist from the science, thereby reinforcing claims of objectivity, whether the purpose was to study the natural world or people and society.

One of the first and most significant developments in this area was the birth of intelligence testing in France at the start of the last century. Constructed by Alfred Binet as an instrument for identifying students whose cognitive abilities were too limited to benefit from mainstream schooling, his creation eventually evolved into the Stanford-Binet Intelligence Scales,⁷ a test claiming to be a scientific measurement of innate intelligence. Edward Thorndike, building on the work of Binet in the emergent field of experimental psychology, expanded the idea of mental testing in education and developed the standardized achievement test, a measurement not of innate intelligence, but of what a student has learned. These tests, consisting of a larger number of questions with multiple-choice answers, resulted in a single numerical score that could easily be compared with the scores of others taking the test. Here began what Taubman⁸ refers to as the unwavering “faith in the science of testing” in education.

While standardized achievement testing did not immediately replace other forms of assessment (including the tests and other measurements created by individual teachers for their own courses and students), it was one of the first examples of the norm-referenced testing that would soon become standard practice in schools across the nation. A central feature of such tests is the comparison of a student to all students who took the test. This norm referencing involves giving the same test to large numbers of students in different classrooms and schools throughout a state or the nation. The interest in comparing students

only increased as the student population rapidly diversified in subsequent decades due to increased immigration, urbanization and the resulting integration of geographic regions, and legislation to make school attendance free and compulsory. Standardized, norm-referenced testing inserted a presumably objective tool into efforts to sort students into academic tracks and otherwise differentiate students. Two important assumptions emerged from this focus: Teaching and learning can (and should) be standardized; and standardized tests are an accurate and unbiased measure of what has been learned (and taught). At the time, educational scholarship was already emerging that questioned the validity of both of these assumptions, but the practices continued to spread.⁹

Starting in the 1970s and 80s and continuing today, a second paradigm shift occurred that would frame standardized testing, particularly high-stakes testing, as the educational panacea. A primary catalyst was the Cold War and growing concerns about how the United States could maintain its preeminent position globally as an economic, scientific, and political “superpower.” A crisis was brewing that the nation’s schools were failing, and that the nation itself was at risk for decline, as argued in the Reagan Administration’s 1983 report, *A Nation at Risk*. According to this logic, schools needed to be held accountable, and the primary means to assess their progress was to be the most objective of measures: the standardized, norm-referenced tests that could reliably carry high stakes for student promotion, teacher retention, school autonomy, and school funding. This emphasis on standardization, accountability, and high-stakes testing continued to permeate the educational visions of subsequent administrations, from Clinton’s Goals 2000 to Bush’s No Child Left Behind to Obama’s Race to the Top, despite that the science of testing, particularly the insights from the field of psychometrics, does not support such assumptions about what these tests tell us or how the results of such tests should be used. The most fundamental principles, prescriptions, and proponents within the psychometrics field tell us a very different story.¹⁰

Psychometrics 101

Misconception: *How one performs on tests is an accurate assessment of one’s learning.*

Facts:

- *Tests designed to compare student performance are being misused to assess learning.*
- *All test scores contain error.*
- *Tests vary in their validity, or how well they measure what they are supposed to.*
- *No big decision should ever be made on the basis of a single test score.*

The first important consideration of testing is purpose. Test developers and users must be absolutely clear about what is being measured and how the results will be applied in making decisions. The process of test construction is so specialized that an instrument designed for one purpose cannot be effectively used for another.

For example, as noted earlier in this brief, norm referencing involves comparing a test taker's performance to that of all other test takers. Norm referencing serves certain measurement purposes very well. Some would argue that college entrance exam results, for instance, are appropriately reported this way, as there are limited slots and it is important to know which students are most deserving. Norm-referenced scores are often reported as a percentile, where an 87th percentile ranking means that 87% of test takers received a lower score (a strong result), and a 10th percentile ranking means that only 10% of others performed worse (a weak showing).

Individual items on exams whose scores are to be norm referenced must be painstakingly written and piloted. One reason these questions are given a trial run is to ensure that about half of test-takers answer them correctly and about half get them wrong. Questions that too many people miss, or that too many people get right, do not contribute to the spread of scores that will allow performance to be ranked.

When the purpose is not to rank, such as when determining the degree to which certain standards are met, it is not difficult to see why norm referencing is highly inappropriate. Standards-referenced assessments should be measuring the performance of a student in relation to curriculum standards, not in relation to the performance of other students. Using norm-referenced scores for these purposes would render it mathematically impossible for all students to achieve success, and a large proportion of our students would be guaranteed to be held back at each testing threshold. After all, 50% of students will *always* score below the 50th percentile, even if all students achieve unprecedented mastery of the material.

In an attempt to address this mismatch in purpose, some large-scale tests, such as the one used to report Adequate Yearly Progress (AYP) by Illinois elementary schools (the ISAT), also offer reporting of performance in relation to the material, a scoring system termed "criterion referencing." However, the fact that the test was developed to be norm referenced means that each item is tailored for half correct/half incorrect responses, and countless items that better assessed the standards were rejected. Some scholars posit that the variance of responses in norm-referenced test items is only achieved by requiring knowledge that is not part of the standards, information that students gain through their out-of-school experience and is not learned in the classroom—no matter how hard the teacher and pupils work.¹¹

Another fundamental principle in psychometrics is reliability, which is most simply defined as "the consistency with which a test measures whatever it's measuring."¹² Test developers work very hard to ensure that their products are reliable. However, it is an undisputed fact in the measurement field that no test is completely reliable. In other words, "some degree of inconsistency is present in all measurement procedures."¹³ Psychometricians conceptualize this as the difference between a test taker's true score (a perfectly accurate reflection of ability, necessarily hypothetical given the limitations of measurement) and their attained score on a test. The difference is called "error of measurement," and it is a component of *every* test score. This is why results are often accompanied by a range representing the standard error of measurement, or an estimate of how much a particular test taker's score is likely to vary. A decision would be psychometrically unsound—not to mention very unfair and a misrepresentation—unless it was supported by a score anywhere within this range.

Even more important than reliability for a test is validity, or how well it actually measures what it was designed to measure. According to Popham, this is "hands down, the most significant concept in assessment."¹⁴ An analogy, taken from Thorndike,¹⁵ might help to illustrate the difference between reliability and validity, and explain why the latter is more important. Imagine that the instrument in question is a bathroom scale. If you were to get up on the scale several times in immediate succession to be met with widely varying numbers each time, you'd quickly conclude that the scale was inconsistent (unreliable) and therefore broken (invalid). If, however, the number on the scale was consistently 5 pounds below or 5 pounds above your actual weight, the instrument would have reliability, but not validity. The scale is still broken, even if discovering this fact involves comparison with some other scale (and a disappointing or pleasant surprise). A test that is invalid can be reliable (though still of no use for its purpose), but a test that is not reliable can never be valid.

Reliability is just one of many factors that go into a determination of test validity. Traditionally, measurement experts think about validity in three ways (technically, three types of evidence): content, criterion, and construct. Content validity is concerned with how well a test reflects the curricular content in question. In other words, does the test match the course standards and objectives? Criterion validity involves comparison of a test's results with the outcomes from some other measure of the same ability. Positive criterion-related evidence would show a strong correlation between the test scores and the outcomes from the other measure, which is called the criterion. Finally, construct validity is concerned with how well the test measures the "construct of interest." A construct is an idea, or something that is not concrete. Psychometric instruments, as mental (psycho) measures (metrics), are by

definition tests of constructs. Here, we ask the big question of whether or not this is a good test. The above types of validity evidence, reliability, as well as other factors all have an influence on this determination. This is the most challenging aspect of validity to conceptualize, but also the most comprehensive and significant.

Because measurement experts understand the limitations of tests that have just been described, there is universal agreement in the field on a caveat for their use: No big decision should ever be made based on single test score. This principle is espoused by test developers, academic bodies and professional associations alike, including the National Research Council and the National Council on Measurement in Education.¹⁶ Attaching high stakes to any single test breaks one of the most fundamental rules of psychometrics.

Currently, most measurement experts continue to hold tests to strict and high standards of purpose, reliability, validity and limited use. Unfortunately, some users and publishers do not. It is up to us to hold these institutions accountable to the standards of the field of measurement.

The Impact of Testing in Classrooms and Schools

Misconception: *Testing is the best way to ensure that teachers are teaching well.*

Facts:

- *Testing and test preparation have narrowed the curriculum.*
- *Pressure to raise test scores is driving good teachers from the profession.*
- *Reading scores have not risen in Chicago, and the racial achievement gap has grown.*
- *Most high-scoring Chicago schools have low rates of student poverty.*

Testing in K-12 schools is not a new phenomenon. However, until the mid-1990s, standardized tests taken by students in Chicago (and in most of the United States, for that matter) were decidedly low stakes. That changed in 1995, when then-Mayor Richard Daley was given control of the school district by the Illinois State Legislature. Daley quickly ushered in a new CEO of Chicago Public Schools (CPS)—his former budget director, Paul Vallas—along with a mandate to bring strict “accountability” and higher test scores to the state’s largest district.¹⁷ Seemingly overnight, standardized test scores were being held up as the new—and sole—measure of progress for students and schools in the city, and moves toward a more structured, prescriptive curriculum were underway.

In the ensuing years, Chicago’s high-stakes testing model

would be reflected in President George W. Bush’s No Child Left Behind (NCLB) Act, and schools across the country would begin to feel its impact. The rhetoric of NCLB sounded promising to many: a renewed national focus on efforts to reach underachieving students, a call for all schools—even those noted for their academic excellence—to examine how well they meet the needs of all children. But for classroom teachers, it was impossible to ignore the collateral damage of the resulting tunnel vision. In many schools, the tests now *were* the curriculum, and the impact on both students and classroom teachers was palpable. For students, it meant a curriculum narrowed to little more than reading and math, and more time devoted to test preparation. For teachers, it meant reduced autonomy, increased stress levels, and decreased job satisfaction.

For example, the following email, written in 2006 by a participant in a teacher certification program for career changers, provides a sense of the increasingly pressurized climate in Chicago schools created, in large part, by the intensified focus on raising test scores. The author of the email had become a teacher with high hopes and seemingly boundless energy, determined to create a meaningful experience for her kids and to be in the classroom for the long haul. Yet just three years after teaching her first CPS class, she was leaving:

This is a difficult email to write. ... After months of weighing options and struggling with the complexity of it all, I have decided to resign from Chicago Public Schools. When I went into teaching, I was not naive to the challenges in CPS; I understood that the adventure I was embarking on would be tough. [But] I have really struggled this year to justify the focus on test scores, rather than student learning. ... Trust me, this was not an easy decision ... but as the year went on, and the testing season approached, I realized how much of a mismatch my priorities are with the CPS administration. My personality will not allow me to just shut my door and teach.¹⁸

Another teacher who left CPS in 2010 and later took a long-term substitute position at a suburban school was amazed at the different atmosphere she encountered there:

I am floored at the significant differences in the two experiences. The main one is the lack of talk, hype, professional development, and even weekly memo mentions of standardized testing and data. I was told by a complete stranger on my first day [at the suburban school] that I was trusted—that I was a professional. I almost cried. Testing, especially in the city, has made me see that I am not trusted or valued.¹⁹

Many classroom teachers whom we have taught or worked with over the last decade have questioned the overemphasis on testing. Some watched as their schools failed to make AYP despite a laser-like focus on testing throughout the year. Even teachers whose students' scores increased sometimes wondered if it was worth all they were sacrificing in the process. In the spring of 2011, a third-grade teacher on Chicago's southwest side described her mixed feelings this way:

We received our ISAT scores last Friday. Twenty-one of my 26 students passed. Chicago Public Schools and the administration would say that mattered. However, I honestly cannot agree. Although I breathed a huge sigh a relief when I saw the results, I still feel like I cheated my students out of a year. They passed a test that allows them to continue on to 4th grade, but I don't feel that they learned anything from that. From January to March all we taught was ISAT prep. It was torture for me and for my students. I'm sure there are some other things that mattered this year to my students, but because of the huge emphasis on the ISATs, most of my students would say passing the ISAT was what mattered. In my short five years in CPS, education has gone downhill, and unfortunately what I think matters cannot even be taught.²⁰

Even the much-celebrated test score gains, when examined more closely, are often minimal, if not imaginary. There's mounting evidence that the testing-centered education policies of the past decade have been ineffective even at their own narrow aims.

In Chicago's public schools, a succession of CEOs since the late 1990s has boasted of impressive yearly bumps in student scores systemwide. But in a 2011 report²¹ that synthesized 20 years of research on school reform in the city, the Consortium on Chicago School Research found that statistics used by the district for year-to-year test score comparisons were seriously flawed. When properly adjusted, reading scores in Chicago elementary schools showed almost no growth over the 20-year period. Even worse, the gap in reading scores between White and African American students actually *increased* over the past two decades.

Of course, these findings run counter to the narrative that central-office administrators have put forth over the years. But they do not come as a surprise to Chicagoans who have read between the lines of district press releases. Even a quick look at the list of schools that score highest on annual tests reveals glaring disparities mapped onto race and social-class differences. At the 15 Chicago elementary schools with the highest test scores in 2010, only 19% of students, on average, were low-income. Contrast that to the district as a whole, where 86% of students are poor.²²

And it is in the schools with the highest poverty rates where teachers, students, and parents feel the pressure to raise test scores most acutely. A seventh-grade teacher in a predominantly Mexican immigrant school says that the emphasis on high-stakes testing has led to a distorted view of teaching and learning.

It has transformed my conversation with parents and community members into a tense double-speak. The scores are looked at as an end, as *the* ends. A misleading valuation is being assigned to my students in which they're defined developmentally, kind of like a cattle brand or a criminal's identification number. You know, *game over* if you don't make this score. Parents see the district's explanation of the scores, and instead of looking at the broader educational landscape and situation for their children, I get asked, "Is there something wrong with my child?"²³

The larger point, though, is this: The obsession with standardized testing in public schools is more than just educationally ineffective. It is undemocratic. The narrow focus on test scores brought about by NCLB and CPS reforms has eroded community control of schools, diminished teacher voice, and—in too many cases—all but extinguished a broad-based, humanistic approach to curriculum. The focus on high-stakes testing actually plays out at ground level as a monumental distraction. It encourages school-based educators to spend all their time focusing on compliance, appearances, and "data" rather than on more essential questions that have gotten lost in the high-stakes shuffle: Education toward what end? What, after all, are the purposes of public schools in a democracy?

Alternatives and Solutions

Misconception: *Standardized high-stakes testing is a necessary evil.*

Facts:

- *Finland, one of the world's highest-scoring nations on international assessments, has almost completely abandoned standardized testing.*
- *Complex thinking and skills are poorly measured by multiple-choice tests.*
- *Portfolio and performance-based tools assess and foster more meaningful learning.*

Some people still ask: Even given the flaws and drawbacks, don't we need some form of standardized testing? Aren't equity and quality ensured by standard, uniform measurements? Aren't tests a necessary feature of a globally competitive, world-class educational system? Research shows compellingly that the answer to these questions

is, no. Assessment is a necessary part of teaching, of curriculum planning, and of educational policymaking, but it does not follow that high-stakes standardized testing is the only tool that can be used for assessment purposes.

While the emphasis on testing can be considered a movement with global dimensions, there are plenty of examples of educational systems around the globe where assessment is structured very differently than what we see in too many of our schools in the United States. A telling example is that of Finland. This country's schools have recently garnered international attention because of high student performances on the Programme for International Student Assessment (PISA) test. While it is important to be cautious about what lessons and features can be "transported" from one setting to another, some of the features that appear to make Finnish education successful at achieving equity, the flourishing of individual students, and national economic competitiveness, include elements that U.S. educational researchers have long documented to be components of successful schools here as well. These include: investment in highly trained teachers whose work days include ample time for planning and preparation; an intentional balance of decentralization and centralization in governance, management, and curriculum design; the provision of resources for those who need them most; high standards and supports for students with special needs; as well as trust and respect within Finnish society for the work of educators.²⁴

What is perhaps most striking about the Finnish education reforms that have been underway since the 1970s is the near-complete abandonment of high-stakes and standardized testing. Instead, Finland uses school-based and student-centered tasks that are embedded into the curriculum. Teachers provide formative and summative reports verbally and in writing, but the major focus is on cultivating students' active learning. Surely, Finland must have some standardized tests? Yes, there is one (1) exam that students may take prior to university, the matriculation exam. This exam is not required for secondary school graduation or for entry into a university, yet it is a common practice for students to take this set of open-ended exams that emphasize problem-solving, analysis, and writing.

Finland serves as a useful reminder that the high-stakes standardized test is not the only way to monitor and improve schools. As noted above, research shows that high-stakes standardized testing is counterproductive. Instead schools, districts, and countries around the globe are increasingly realizing that:

- *The multiple-choice test is poorly suited for measuring—let alone fostering—the complex analytic skills and kinds of critical thinking that children and youth need for meaningful civic*

engagement and economic success in today's world.

- *Instead of testing for the ability to answer discrete pre-set questions, schools should assess students' abilities to synthesize, to collaborate, to deliberate, to manage projects, to solve problems and to innovate creatively.*
- *In place of compulsory end-of-grade tests, students need assessments that foster self-reflection, individual goal setting, and active learning skills.*

Over a century ago, our nation embraced educational testing in the pursuit of uniformity and standardization. While it is questionable whether this was appropriate even for early-to-mid 20th century industrialism, it is very clear that in the contemporary global context, more complex and standards-referenced instruments, like portfolio assessments and performance assessments, are viable and vital alternatives to high-stakes standardized testing.²⁵

Forward to Action!

Misconception: *Testing is here to stay, and there is nothing we can do about it.*

Facts:

- *A variety of alternative assessments can be used in place of high-stakes tests.*
- *High-performing schools are using authentic assessments.*
- *Teachers working together can change testing policy.*
- *The public can, and should, speak out against high-stakes testing.*

The harsh realities of high-stakes testing have permeated our communities and psyches for the last several decades. With the Elementary and Secondary Education Act (ESEA) up for reauthorization, it becomes critically important for communities to take proactive stances against high-stakes standardized testing. CReATE agrees with the Chicago Teachers Union that high-stakes testing continues to be promoted, despite compelling research against its use, because of its profitability by testing-related industries (test makers, scorers, test-prep service providers, and so on), and that it constitutes part of a broad-based attempt by corporate interests to "align the outposts of public education with the needs of an unequal and highly polarized economic system."²⁶ However, all is not lost. Throughout the city and the nation, efforts are underway to interrupt the dehumanizing trend of high-stakes standardized testing as the sole measure of academic achievement.

Organizations like Project Appleseed have begun to operate as a clearinghouse for alternative assessments that school

districts, individual schools, and teachers in classrooms can use in place of high-stakes tests.²⁷ Psychometricians and other researchers have acknowledged that the primary function of standardized testing is to measure how well students take a test; in contrast are performance-based assessments that can provide much more useful feedback to educators about their students' development of knowledge and skills,²⁸ "including the process by which they solve problems."²⁹ By requiring students to be active participants, teachers are able to design projects that allow them to develop the necessary skills to interpret, analyze, and make informed decisions. Performance-based assessments foster the competencies that will be of real-world use for students, for example:

- *Group projects enable a number of students to work together on a complex problem that requires planning, research, internal discussion, and group presentation.*
- *Essays assess students' understanding of a subject through a written description, analysis, explanation, or summary.*
- *Experiments test how well students understand scientific concepts and can carry out scientific processes.*
- *Demonstrations give students opportunities to show their mastery of subject-area content and procedures.*
- *Portfolios allow students to provide a broad portrait of their performance through files that contain collections of students' work, assembled over time.³⁰*

In Chicago, this type of assessment and curriculum is seen more often in high-performing schools. While some schools claim to be centered in project-based learning, the vast majority of schools that have been placed on probation end up limiting or eliminating this type of instruction in order to significantly increase time and resources for test preparation, thereby moving away from the very types of instruction that can increase learning and success.

It is important to note that communities have not taken this lying down. Chicago, with its deep roots in educational activism, has been home to several conscious attempts to interrupt the current testing craze. One example is a group of teachers from Curie High School on the Southwest Side known as the "Curie 12," who refused to administer the district-mandated high school exam called the Chicago Academic Standards Examination (CASE).³¹ They argued that the exam was "a poorly written, content-based test that asked students seemingly random questions unrelated to the curriculum of survey literature."³² They also acted strategically:

- *They compiled hard data to support their assertion that the CASE exam did not improve instruction, but instead actually impaired student achievement.*
- *They let their administration know what they were planning to do ahead of time.*

- *They elected two veteran teachers to be spokespersons for the group, to prevent new teachers from being attacked individually.³³*

Soon after, CPS removed CASE from every high school in CPS. Katie Hogan, one of the "Curie 12," reflected that,

We were not twelve angry teachers who wanted to tell CPS to shove it. We were just twelve regular teachers on different ends of the activism spectrum who wanted to end a really bad test. We never expected to win; we only expected to fight. Isn't that what we should teach our students every day?³⁴

Other community-driven efforts have openly challenged the effectiveness of high-stakes testing. For example, Parents United for Responsible Education (PURE) has developed a set of public advocacy action steps to push for assessment that is accountable to the issues and concerns of parents and students, including:

- *Learn about some of the problems with testing. The FairTest website (www.fairtest.org) is a great place to start.*
- *Think about how testing has affected the lives of you, your child and/or someone else you know.*
- *Create a message about testing to share in various contexts, such as letters to the editor.*
- *Inform others.*
- *Tell your local, state, and federal legislators what you think about the current national obsession with testing.*
- *Join with others to learn more about the issues, share what you know, and take joint action.³⁵*

They also suggest that when people are contacting legislators or media outlets to use "clear, one-sentence statements about what you oppose in assessment."³⁶ These small collective steps are critical in getting the word out to families who may sense that something is wrong with test-prep curriculum, but may not have the language or talking points to express their concerns. Key to this process is the recognition of collaborative efforts throughout the city to address over-dependence on high-stakes tests. Numerous activist groups like the Kenwood-Oakland Community Organization, the Grand Boulevard Federation, and Blocks Together have taken a collective stance on the dangers of test-prep curriculum.

By developing documents in this vein, CReATE seeks to support the movement of everyday students, teachers, families, and community members to stand against high-stakes testing. While research has the potential to benefit the lives of the aforementioned groups, it can only serve effectively if paired with concerted, on-the-ground, community-based action.

Notes

- 1 McCracken, M., & McCracken, H. T. (2001). Teaching in the time of testing: What have you lost? *The English Journal*, 91(1), p. 30.
- 2 Ronell, A. (2005). *The test drive*. Urbana: University of Illinois Press, p. 17.
- 3 Taubman, P. M. (2009). *Teaching by numbers: Deconstructing the discourse of standards and accountability in education*. New York, NY: Routledge.
- 4 Wainer, H. (2011). *Uneducated guesses: Using evidence to uncover misguided education policies*. Princeton, NJ: Princeton University Press.
- 5 Ibid.
- 6 Kliebard, H. (2004). *The struggle for the American curriculum: 1983-1958*. New York, NY: Routledge/Falmer; Spring, J. (2012). *Education networks: Power, wealth, cyberspace and the digital mind*. New York, NY: Taylor & Francis; Pinar, W., Reynolds, P., Slattery, P., & Taubman, P. (2008). *Understanding curriculum*. New York, NY: Peter Lang.
- 7 Thorndike, R. M. (2011). *Measurement and evaluation in psychology and education* (8th ed.). Upper Saddle River, NJ: Prentice Hall.
- 8 Taubman, 2009.
- 9 Kohn, A. (2000). *The case against standardized testing: Raising the scores, ruining the schools*. Boston, MA: Heinemann; Sacks, P. (1999). *Standardized minds: The high price of America's testing culture and what we can do to change it*. Cambridge, MA: Perseus.
- 10 Wainer, 2011.
- 11 Kohn, A. (2004). *What does it mean to be well educated?* Boston, MA: Beacon Press.
- 12 Popham, W. J. (2011). *Classroom assessment: What teachers need to know*. Boston, MA: Pearson, p. 61.
- 13 From Thorndike, 2011.
- 14 Popham, 2011, p. 83.
- 15 Thorndike, 2011.
- 16 National Research Council. (1999). *High stakes: Testing for tracking, promotion, and graduation*. Washington, D.C.: National Academy Press; American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (1999). *Standards for psychological and educational testing*. Washington, D.C.: AERA; Joint Committee on Testing Practices. (2004). *Code of fair testing practices in education*. Washington, D.C.: American Psychological Association.
- 17 Lipman, P. (2004). *High stakes education: Inequality, globalization, and urban school reform*. New York, NY: RoutledgeFalmer.
- 18 Personal communication, 2006.
- 19 Personal communication, 2012.
- 20 Personal communication, 2011.
- 21 Luppescu, et al. (2011). *Three eras of Chicago school reform*. Chicago, IL: Consortium on Chicago School Research.
- 22 Illinois Interactive Report Card data, 2011.
- 23 Personal communication, 2012.
- 24 Sahlberg, P. (2011). *Finnish lessons: What can the world learn from educational change in Finland?* New York, NY: Teachers College Press; Darling-Hammond, L. (2010). *The flat world and education: How America's commitment to equity will determine our future*. New York, NY: Teachers College Press; Laukkanen, R. (2008). *Finnish strategy for high-level education for all*. In N. C. Soguel & P. Jaccard (Eds.), *Governance and performance of education systems* (pp. 305-324). Dordrecht, the Netherlands: Springer.
- 25 See also Davidson, C. N. (2011). *Now you see it: How the brain science of attention will transform the way we live, work and learn*. New York, NY: Viking.
- 26 Chicago Teachers Union. (2010). *CTU position paper: Debunking the myths of standardized testing*.
- 27 www.projectappleseed.org/assessment.html
- 28 Bryk, A. S. (2003). *No Child Left Behind, Chicago style*. In P. E. Peterson & M. R. West (Eds.), *No Child Left Behind?* Washington, D.C.: Brookings Institution Press.
- 29 Ibid.
- 30 Ibid.
- 31 Hogan, K. (2008). *The Curie 12: A case for teacher activism*. In W. Ayers, G. Ladson-Billings, G. Michie, & P. Noguera (Eds.), *City kids city schools: More reports from the front row*. New York, NY: The New Press.
- 32 Ibid.
- 33 Ibid.
- 34 Ibid.
- 35 Parents United for Responsible Education. (2010). *Action steps for better assessment: What you can do to help win the testing battle*.
- 36 Ibid.

FOR MORE INFORMATION ON CReATE activities, or membership, or to read this full report and view other reports, please visit <http://createchicago.blogspot.com> or email createchicago@gmail.com

This briefing paper was prepared by Isabel Nuñez, Peter Hilton, Eileen Quinn Knight, Kevin Kumashiro, Jeff Kuzmic, Greg Michie, Brian Schultz, Noah Sobe, and David Stovall.

CReATE

Public education in a democratic society is based on the principle that every child is of equal and incalculable value. This guiding principle requires the fullest development of every member of our nation. Effective public schools are necessary to enable every member of our nation to reach his or her fullest potential. Schools in a democracy aim to prepare the next generation to be knowledgeable and informed citizens and residents; to be critical thinkers and creative problem solvers; to be prepared to contribute positively to communities, workplaces, and societies that are characterized by diversity and inequities; and to be healthy, happy, and prepared to support the well-being of others with compassion and courage. The children and youth of Chicago deserve no less.

CReATE, a volunteer group of Chicago-area education researchers, conducts, reviews, and distributes studies to address the needs of our students, parents and schools, and promotes citywide learning and dialogue about educational issues through free public events. CReATE partners with educators, public interest groups and community-based organizations. Our university-based scholars and educational workers are available for comment and discussion on topics raised here and in our policy statement, available online at: <http://createchicago.blogspot.com>