

The Impact of Early Childhood Testing

Kevin Russell

The “one-size-fits-all,” “drill and kill” approach being pushed is creeping its way into early childhood education. Chicago Public Schools has expanded the required and optional testing for young children, placing importance on continuous assessment through standardized testing and benchmarks. These tests include REACH, Teaching Standards GOLD and Quarterly Benchmark Performance Tasks. Matching the standardization trend in high school education, testing is being used as the foremost indicator of student performance.

Assessment-driven education plays a large role in the new CPS preschool funding system. Known as *Ready to Learn!*, the new system forces district schools, charters, parochial schools and community centers to compete for the same pool of preschool funding. Mirroring the standardization approach adopted by CPS in other areas, rubrics and lengthy application packets have been created to grade proposed programs. For CPS, testing is a key part of “program quality.” When asked to describe the accomplishments of the early childhood program, CPS specifically asks for results on standardized tests, such as Teaching Standards GOLD and the Kindergarten Readiness Test. While the entire *Ready to Learn!* program is troubling, the added emphasis on standardized testing for young children deserves special analysis.

Questionable Validity

With so much riding on testing scores, one would expect clear evidence that the tests are, in fact, evaluating what they intend. It is important to remember that young children are constantly developing. This fact would question the predictive validity of tests in early childhood, since a 5-year-old may be completely different after completing the school year than he was when he took an introductory test. Samuel Meisels of the Erikson Institute writes:

“Given that young children are undergoing significant changes in their first eight years of life in terms of brain growth, physiology, and emotional regulation, and recognizing that children come into this world with varied inheritance, experience, and opportunities for nurturance, it is not difficult to imagine that a brief snapshot of a child’s skills and abilities taken on a single occasion will be unable to capture the shifts and changes in that development. To draw long-term conclusions from such assessments seems baseless.” (Meisels, 9).

Meisels uses two studies to support his claims. LaParo and Pianta (2000) found that only a quarter of variance on academic/cognitive skills on first and second grades tests were accurately predicted by preschool or kindergarten tests (LaParo & Pianta, as quoted by Meisels, 9). Social/behavioral variations were even more difficult to predict, with only ten percent also being shown in preschool or kindergarten testing. The authors conclude that due to the rapidly developing nature of children, standardized testing has little consistent predictive validity. For young children, the authors argue that “instability or change [in cognitive and behavioral ability] may be the rule rather than the exception during this period” (LaParo & Pianta, as quoted by Meisels, 9). Kim and Suen (2003) performed a similar study and found that “the predictive power of any early assessment from any single study is not generalizable, regardless of design and quality of research” (Kim & Suen, as quoted by Meisels, 9). While some results may indeed

turn out as predicted, both of these studies show that it is not enough data to form any generalizable correlation for predictability given the constant developmental changes of a child's brain.

The Effect on Curriculum

In a paper published in the *Journal of Education*, Louisiana State Professor Renee Casbergue argues that standardized tests incorrectly assume that the “complex processes like those required for reading and writing can be simplified into component parts, each of which can be tested separately” (Casbergue, 18). The development skills needed in early childhood cannot simply be considered by small, individual aspects, but must instead consider the whole picture.

Standardized tests only assess “constrained skills,” which are skills that are limited to small sets of knowledge that are mastered in relatively brief periods of time” (Paris, as quoted in Casbergue, 13). “Unconstrained skills”—skills that continuously learned and developed throughout life, such as vocabulary, symbolism and story themes—are rarely evaluated in these tests.

Meisels uses the Head Start National Reporting System (NRS), a standardized test to be administered every two years to *all* four-year olds in Head Start), as a case study. The study would cost \$25 million annually. While usage of NRS was quickly suspended after widespread criticism, it had many of the same problems as current standardized tests. Despite not yet (although it was eventually planned to be) being a high-stakes test that would greatly affect the allocating of Head Start funding, the U.S. General Accountability Office (GAO) warned in 2005 that 18 percent of programs studied had altered their curriculum in order to accommodate the skills tested by NRS. GAO also found that the necessary analysis to establish validity and reliability of the study was not done (GAO, as quoted by Meisels, 13). The test used complex questions involving causality, subtraction, metric unions, and subjunctive case—all of which were viewed by child development experts as beyond the understanding of a four-year old (Meisels, 12). Overall, the test represented a misunderstanding of the way young children learn:

“[NRS] is a model of passive reception, of pouring into a vessel knowledge and skills that are needed for competence, rather than recognizing learning as active and teaching as a joint process of interaction between child and adult. An active view of learning, fundamentally based on enhancing relationships between teachers, children, and challenging materials, is nowhere to be seen in this test. ... Yet, when you know that the results of a test will be used to make decisions that may affect your program's continuation and other things you value, you are sorely tempted to begin teaching to the test. ... It was a failure because it ignored the complexity of early development that teaches us that no single indicator can assess a child's skills, achievements, or personality.” (Meisels, 13, 16).

Given the increased stakes for teachers on most standardized tests, educators have had to alter their curriculum to concentrate mostly on constrained skills while putting valuable unconstrained skills—those that will be especially important as children age into high school and college—on the back burner and “can homogenize what might otherwise be a very heterogeneous curriculum (Meisels, 3, 8). This consequence, Casbergue argues, “neglects the consideration of the whole

child and the full range of environmental influences that shape learning” and these test-influenced programs “compromise time spent on authentic, constructive activities that are more in keeping with longstanding traditions in early childhood education and with current understanding of how young children learn” (Casbergue, 13, 18). The closed-ended questions of a standardized test fail to engage or access a child’s natural curiosity and creativity, an essential part of developing future academic and social skills (Meisels, 2).

Preparing for these tests can lead to teachers restricting or even eliminating creative lessons from their curriculum. A pre-kindergarten teacher in the Lake Calumet Elementary Network said that because of the importance placed on REACH and other testing, she has been forced to spend less time developing her students’ creativity. Activities that stimulate children’s creativity, such as allowing them to formulate and articulate their own stories, are sometimes neglected to focus on the more academic tasks assess in REACH. The teacher explained that it “seems like fun has been taken out of pre-K” because of the overwhelming amount of testing.

These standardized tests are “unable to describe individual patterns of learning” in the way that teachers can be simply observing their students (Meisels, 2). Observational assessment, which analyzes children’s behaviors in connection with their academic and developmental growth, is a far more effective teaching method. Casbergue argues that for early literacy, observational assessment by trained teachers better encompasses the full skill set needed to develop properly:

“Observing this experimentation, whether it occurs as children are playing or engaging in shared reading and writing activities with caregivers or teachers, affords an opportunity to document their understandings about literacy. Documenting what children do when they are actually reading or writing enables educators to understand what children know and are able to do. The vocabulary children use as they describe pictures they are drawing or reenact a story they have heard provides insight into how they understand and use language. The forms of writing that appear in their drawings, journals, signs, and notes to each other illustrate what they understand about the forms, functions, and meanings of print. Their attempts at reading aloud, whether during pretend reading or as they follow print during shared, guided, or independent reading, reveal their understanding of the connections between sounds and symbols, as well as their knowledge of how print connects to meaning.” (Casbergue, 18)

Observational learning is also a better method to develop spelling abilities and understand the typical narrative structure (Casbergue, 19).

Test-taking procedure

When assessing young children’s reading ability, standardized literacy tests often fail to replicate the skills needed to succeed at reading. The style of reading and writing required in standardized testing “often bears little resemblance” to that which children are generally exposed and is more likely to “capture a snapshot of a child’s literacy—one that may not actually resemble the child as a literate being” when faced with works of literature (Casbergue, 16). True literary skill is not adequately measured in standardized testing because “what dominates the whole enterprise when children actually take the tests is test-taking behavior—filling in bubbles, moving the marker,

making sure everyone is in the right place. These activities may be related to test taking, but they have nothing to do with reading” (Stallman & Pearson, from Casbergue, 17).

Complex instructions and devices can also confuse children and lead them to answer incorrectly despite having the knowledge being assessed. In a report published by the Erikson Institute, Meisels argues:

“Young children are developmentally unreliable test takers. They have a restricted ability to comprehend such assessment cues as verbal instructions, aural stimuli, situational cues, or written instructions. Further, questions that require complex information-processing skills—giving differential weights to alternative choices, distinguishing recency from primacy, or responding correctly to multistep directions—may cause a child to give the wrong answer.” (Meisels, 8)

The integration of technology into testing adds yet other variable to challenge the validity of scores. Many children, especially those living in poor communities, have most likely never used a computer, tablet or other devices being used to administer tests. This unfamiliarity and the aura of a new, advanced object can distract children from the material of the assessment.

Along with new and bizarre test-taking procedures, there are other extraneous factors that affect a young child’s performance on standardized tests. Short attention span, been asked to sit longer than usual and boredom, fatigue, hunger, illness and anxiety can skew test results (Salinger, as quoted in Casbergue, 17; Meisels, 8).

Challenges to Self-Esteem and the Issue of Labeling

Given the emphasis on identifying learning disabilities in standardized testing for young children, there is also the danger of incorrectly “labeling” students. These tests can miss students who may perform well on phonics and the alphabet but “have relatively little understanding of how those skills relate to actual reading and writing” (Casbergue, 18). Since teachers are forced to use much of their time on the skills measuring on the test, children with problems in other areas may not be identified or helped until far later in school, making it more difficult to fully solve the problem.

The opposite is also true. These tests can result in “false positives” which children incorrectly identified as having a particular disability. As previously discussed, a poor result on a test may be more indicative of a misunderstanding of test-taking procedure than actually struggling with the skill being assessed. Being mislabeled in early childhood can disrupt students’ entire developmental process, with children performing less challenging tasks that lead to falling behind their peers. Labeling children as ineffective at one isolated skill may also lead to excess focus on the “deficit area, at the expense of activities that address the needs of the whole child” (Casbergue, 18).

Being labeled as behind one’s peers can have devastating effects for a child’s self-worth and self-perception. Young children can be stigmatized by being “tracked into low achieving groups that will further confirm their sense of powerlessness and limited potential. Their estimates of their own abilities—their self-perceptions and their motivation and ultimately their achievement—are likely to suffer as a result” (Meisels, 8).

The frequency of testing, even if it was valid, can also have negative effects on children. With testing becoming a major part of schooling, and even early schooling, students may form long-term negative attitudes toward school. This hinders both academic progress and a student's sense of self (Meisels, 3). Students can also become less interested in schooling, lessening the impact of good teaching. A CPS pre-kindergarten teacher stated that she notices a marked decline in engagement and enthusiasm among her students when preparing for testing. Luckily, the teacher said that during breaks in testing when she can maximum her curriculum's creativity, students become reengaged with the material and enjoy learning. It is this opportunity cost of testing—sacrificing a good teacher's creative lesson plan that sparks students' interest in lifelong learning—that is perhaps most troubling for a young child's development.

Limited Teaching Time and the Control of the Classroom

The emphasis on standardized testing has given teachers across the district more work and less time with their students. Teaching Standards GOLD, a computer program that allows teachers to track the progress of their students on 36 individual goals, requires teachers to input results for each student. While teachers have reported that GOLD's goals do help guide academic and socio-emotional observation of their students, this data entry takes approximately 10 hours and must be completed up to three times per year.

REACH seems to be even more time consuming for teachers. A Lake Calumet Elementary Network pre-K teacher reported that it takes at least 30 minutes to administer the REACH assessment, which is given twice a year, to each child. Teachers must give their complete attention to the child being tested, leaving the other children mostly on their own. While some classrooms are lucky enough to have PSRPs or parent volunteers to watch over the other children, many teachers have to leave their students alone while administering testing.

Teachers without PSRPs or volunteers must remain in their classrooms when giving the REACH to a student to ensure that the other children are not left completely unattended. Because of this, students being tested may be distracted, not focusing on the questions of the assessment. Teachers have also reported that they cannot give assignments that are too engaging during REACH. If children become too absorbed in the assignment, they will not want to leave to take the assessment. It is these types of unintended consequences of testing that the district does not seem to adequately prepare for or anticipate, but can result in a teacher's loss of management of his or her classroom.

Professional development, an integral part of improving teaching, has also been polluted by excessive testing. Rather than using this essential collaboration time to share teaching strategies and creative lesson plans, this time must be spent preparing for testing and discussing ways to increase test scores.

Conclusion

This test-driven approach can lead to a narrowing of students' "opportunities for richer learning and development as the professionals strive to support those aspects of learning they surmise to be most important to school success" (Casbergue, 13-14). This "narrowing of opportunity" has a disproportionate effect on low-income children who have few other opportunities to develop

these unconstrained skills (Hart and Risley, as quoted in Casbergue, 14). With more and more evidence showing that elementary and high schools with a high percentage of low-income students have difficulty receiving the resources needed to educate the whole child, we cannot afford an expansion of the socioeconomic “achievement gap” that puts poor children at a significant disadvantage, even in pre-kindergarten. It has been shown that two years of high-quality preschool—a program that focuses on the “whole child”—can reduce the achievement gap by close to half (Mathis, 2).

Underlying the push for increased testing is the fallacy of a “one-size-fits-all” solution. Policymakers and academics see problems in American education and are looking to use their influence to implement a top-down solution. Rather than empowering highly trained teachers to evaluate their own children through observational assessment, education “reformers,” many of whom have little experience in education, have developed theories for a formula that will uniformly produce a well-educated child. Standardized tests have been created to assess if children are at the proper stage of this education “machine.” These theories are, however, simply theories. There is very little evidence to support that the skills early childhood standardized tests assess correlate with ultimate academic or developmental success. “Children differ so greatly from one another in their early experiences, opportunities to learn, genetic inheritance and family structure,” that it’s impossible to achieve uniform solutions (Meisels, 6-7). Children of different experiences and income levels cannot be put together using a nuts-and-bolts approach. Standardized testing simply takes time away from teachers who have been trained to observe and correct behavior in their students in a way that a multiple choice tests cannot.

Kevin Russell is a Research Consultant at Chicago Teachers Union.

Works Cited

- Casbergue, Renee. "Assessment and Instruction in Early Childhood Education: Early Literacy as a Microcosm of Shifting Perspectives." *Journal of Education*, 190, no. 2 (2010): 13-20.
- Mathis, William. "Preschool Education: Research-Based Options for Education Policymaking." National Education Policy Center. Nov. 2012: 1-6.
- Meisels, Samuel, The Erikson Institute. "Accountability in Early Childhood: No Easy Answers." Occasional Paper, no. 6 (2006): 1-28.