Growth through Language Arts and the Conundrum of Teacher Evaluation

Peter Smagorinsky

I have adapted the title of this essay from John Dixon, whose Growth through English (1967) was among the books written in the wake of the Anglo-American Conference at Dartmouth College in 1966. This event led to major changes in the teaching of English and language arts in the United States. The ideas generated at Dartmouth helped to found the National Writing Project in 1974, brought Rosenblatt’s transactional theory of literary response greater attention, helped to shift teachers’ attention from learning product to learning process, and introduced to North America other changes away from formalism—an emphasis on the final form of texts—and toward the processes and experiences of students as they engaged with texts and each other. According to this “growth model,” the purpose of engagement with an English or language arts curriculum is to promote the personal growth of individual learners. Dixon and other British participants argued that emphasizing texts at the expense of the learner prevented students from growing as people through their reading, writing, talking, performing, and other activities available through language arts.

The monolith that Dixon and colleagues worked against sounds quite a bit like the formalist curriculum that our policy leaders are imposing on teachers today. He said, for instance, that when culture undergoes rapid change, “there is a tendency to panic, to define an external curriculum—a system into which teacher and pupil must fit—instead of helping teachers, in departments and larger groups, to define for themselves the order and sequence that underlies their best work” (p. 84). This description could fit many present-day school systems that are adopting some version of the Common Core State Standards (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) or other national curriculum in order to force uniformity upon a population that is increasingly diverse (United States Census Bureau, 2012). Dixon continues, “[I]t seems an elementary mistake to demand a list of skills, proficiencies, and knowledge as the basis for an English curriculum. Demands of this kind produce two wrong kinds of answer: answers so detailed that we determine, let’s say, the books every child should read by a particular stage; or answers so general that the skills, etc., described are not amenable to being put in order one after the other” (p. 5).

Plus ça change, plus c’est la même chose.

But some things do change. One is the definition of “growth model.” I live in Georgia, where the state Department of Education (DOE) has now announced a new way of measuring student and school progress and unveiled an interactive Internet tool that provides easy access (gastudentgrowth.gadoe.org). According to the DOE, it produces “[t]he metric that will help educators, parents, and other stakeholders better understand and analyze the progress students make year to year.” Sounds fantastic, doesn’t it?

The “Student Growth Model,” as it is officially known, relies on two measurements. One is based on the percentage of students who meet or exceed state standards on, for elementary and middle school students, Criterion-Referenced Competency...
Tests; high school students’ growth is measured by scores on End-of-Course tests. The second measurement is designed to assess year-to-year progress of each student, in comparison both to students in other Georgia schools and students at the national level in “academically similar” schools in terms of demographic and socioeconomic statistics. And here’s where things get predictably dicey: these measurements contribute in a big way to the state’s new teacher assessment system because the model assumes that there is a one-to-one causal relationship between individual teachers and individual students in terms of their test scores, scores that serve as a proxy for learning, for growth, and for teacher effectiveness.

My primary source of information on the Student Growth Model has been the Atlanta Journal-Constitution, which in general does a good job covering education and features one of the nation’s best blogs, Maureen Downey’s Get Schooled. Their coverage (Bloom, 2014; Downey, 2014) and the releases from the DOE are replete with the language of advancement: assertions about “progress,” “learning,” “achievement,” and “growth.” Damian Betebenner, the consulting statistician who designed the model (Betebenner, 2011) that Georgia’s system is based on, is reported to have said, “You may have a teacher that’s in a classroom and the kids aren’t growing. We’re not saying that you’re necessarily a bad teacher, but it’s just not working here.” By factoring “growth” into these measurements, however, the tests indeed designate teachers whose students do not improve their test scores relative to local and national peers as bad teachers.

Betebenner is an associate at The National Center for the Improvement of Educational Assessment in Dover, New Hampshire, who often writes critiques of bad policy statements for the National Education Policy Center, a group whose efforts I generally admire. This Student Growth Model, then, was developed by a reputable statistician and has generally gotten praise from people whose opinions I would tend to respect. Why, then, do I consider it to be such a wrong-headed way of thinking about human growth, and in the context of education, a horrible way of assessing teacher effectiveness?

Readers of Language Arts should not be surprised to hear that my primary objection concerns the use of standardized tests as the sole measure of human growth and, consequently, as an indicator of teacher quality. I have spent much of the last twenty-five years exploring the notion of human development, primarily through the lens provided by Lev S. Vygotsky (see Smagorinsky, 2013). Vygotsky is difficult to distill into a few key points, but essentially, he explored the ways in which people grow into members of their societies by engaging with others in “cultural practices.” For a Southern Baptist, for instance, this growth might involve learning, through faith-based texts and adult guidance, Biblical precepts so as to walk a righteous path according to the church’s teachings. This path is, above all, going somewhere. As James Wertsch has noted, any developmental psychologist needs to address the question, development toward what? And thus by implication, development by what means?

The pathway and destination of a Southern Baptist adherent would be quite different from those of an Israeli Orthodox Jew or a Mexican Rarámuri Shaman. I use these religious and spiritual examples because they provide such clear distinctions, but within any group, idiosyncratic directions are available and undoubtedly represent the rule rather than the exception. In pluralistic schools like those in most of the United States, a developmental growth model measuring only one type of learning cannot possibly assess the developmental progress of students whose families are raising them toward a variety of destinations through their culture’s means of engagement. What the Student Growth Model implemented in Georgia—and no doubt other

---

1Some might find it a conflict of interest that I’m a frequent contributor to this blog. But read on.

2Wertsch said this at a conference, but I can’t remember which one.
states given the national trend toward statistical accountability— attempts to achieve, then, is so narrow as to belie the concept of human development suggested by the model’s name. Growth, progress, achievement, learning: we all want these attributes in our children and expect our teachers to promote them. But the new Student Growth Model measures do not measure up to what most people hope for in their child’s developmental course—their development into good human beings according to some cultural definition of a quality life.

I will now return to ideas from Dixon (1967). Of course, the sixties are now viewed by many as The Problem in today’s society, especially from those who believe that the decentering of authority during that period has produced an enduring failure of societal institutions (e.g., Bloom, 1987). Contrary to the idea that the sixties undermined the foundation of American life, the decade (and its preceding and following years) produced the civil rights movement, the women’s liberation movement, and a heck of a lot of great music. Let’s see what the growth model of the sixties offered in terms of language arts pedagogy. Dixon (1967) provides a catalogue of practices that promote human growth under the aegis of the language arts curriculum that I’ll modify in light of current policies:

• Classrooms should be characterized by discussions that involve the students speaking to one another about things that matter to them. Topics should be open-ended so that students push their thinking and sharpen their opinions into well-formulated ideas, and the process should allow for expressive or exploratory talk: talk that involves thinking through an unresolved problem rather than expressing a fully-formed idea.

• Just as discussion should be open-ended and exploratory, so should writing, so that writing potentially enables the generation of new insights. Further, writing should, especially in early drafts, be more concerned with ideas than form. Finally, writing should represent the range of writing done outside school, including personal, creative, and other forms often dismissed as frivolous by those who believe that school should be very serious and exclusively oriented to preparing students for the workforce—even beginning in kindergarten, where recess has been jettisoned in favor of yet more preparation for testing.

• Learners’ lives should be central to their education, such that writing might explore personal issues, and discussions might go outside what Common Core State Standards architect David Coleman calls “the four corners of the text.” This writing should draw on students’ experiences, and reading should also be informed by what readers bring to the experience, with this transaction situated in a classroom that values the integration of personal and academic knowledge. The curriculum should thus be meaning-centered rather than form-centered, and should concern students’ emotions as opposed to the cold analytic textual experiences inscribed in the CCSS.

• In a very sixties turn, teachers should be less authoritarian in order to allow “the liberation of pupils from the limits of the teacher’s vision” (p. 48) and, in general, from established rules, which may be more limiting than enabling to young learners. As a corollary, students should help set the direction and content of the classroom, thus moving away from standardization and toward personal pathways. I have argued that Dixon and colleagues were a bit too obsessed with individuals at the expense of social responsibilities, and so would add the dimension that children’s personal preferences should always take the social context of activity, including ground rules, into account (see Smagorinsky, 2002).

• School should include more drama—not so much formal theater, but interpretive performance. Barnes (1992) writes that this notion of drama “differs from other talk in three
ways: movement and gesture play a larger part in the expression of meaning; a group working together upon an improvisation needs more deliberately and consciously to collaborate . . . the narrative framework allows for repetition and provides a unity that enables the action more easily to take on symbolic status” (p. 37).

Well, maybe the sixties weren’t so nutty after all. These beliefs undoubtedly sound familiar to anyone who has regularly read *Language Arts* in the last thirty to forty years. Unfortunately, just as the Dartmouth participants fought the monolith of formalism and authoritarianism nearly a half-century ago, today’s teachers face a revival of authoritarianism from the US and state departments of education, the corporate interests behind Big Testing and Big Curriculum, and the various bullying techniques of those in power who would force teachers away from what they believe to be effective means of teaching dedicated to the authentic learning needs of students.

One unfortunate consequence of the conception of growth in the Georgia Growth Model is its competitive nature. The one whose test scores rise the fastest is the winner, and others and their teachers are the losers. I’ll quote a Georgia educator’s appraisal of this phenomenon (Bacallao, 2014):

This new “growth” model is a zero sum game. If all the students in the state of Georgia advanced their reading level by 3 years in just one year’s time, it would not register as growth because everyone improved. In addition, since the tests measure rigid grade level skills, they cannot measure academic gains outside of the narrowly defined government standards. With this system, for every “winner” there has to be a “loser.” For every student that shows “growth” using this model, another student needs to show “decline” relative to the other students. If every student in the state failed the test, the “growth” model would not register the poor performance because the “growth” measure is based on the performance of the students compared to other students. The measure is not objective and cannot be correlated with actual academic growth. It creates a “survival of the fittest” numbers game where it is impossible for every student in the state of Georgia to show growth, even if students do make significant academic gains. Only 50% of the students will be able to show “growth.” It makes it impossible for half of our students and half of our teachers to show “growth.” (n. p.)

Although I disagree with the premise that any measure, or any human judgment at all, can be considered “objective,” I find this commentary to be otherwise on target. The notion of objectivity, however, is presumed in the statistical model developed by Betebenner (2011) and is viewed in the policy world as more “rigorous”—the *de rigueur* term for all “reform”—than subjective teacher evaluations of students and student evaluations of teachers.

Indeed, the whole idea that standardized tests measure growth relies on the subjective opinion that these tests are valid and reliable. They might indeed be valid and reliable from a statistician’s point of view, but I’m more concerned with ecological validity—the everyday validity of whether or not something works in practice. Very subjective, of course—just like everything in the human world, including conclusions drawn from statistics.

So, what is human growth, in my subjective opinion? I don’t think it’s the same for everyone, because people are headed in different directions. Even those headed in the same direction often take different pathways, follow different paces, integrate that pathway with different goals, and otherwise follow Thoreau’s (1845) wisdom: “If a man does not keep pace with his companions, perhaps it is because he hears a different drummer. Let him step to the music which he hears, however measured or far away.” In today’s world of pseudo-objectivity, if students were given this item on a test, their response would not be measured by how they follow the advice, but on which definition of “pace” is being employed:

1. the speed at which someone or something moves
2. the speed at which something happens
3. a single step or the length of a single step
4. a manner of walking

*If you agree that this model does not measure either student growth or teacher effectiveness, and if you agree that making students accountable for growth is a good idea, what might be a better alternative in terms of developing teacher effectiveness measures?*
If you agree that this model does not measure either student growth or teacher effectiveness, and if you agree that making students accountable for growth is a good idea, what might be a better alternative in terms of developing teacher effectiveness measures? Statisticians’ solutions are admirable in their ability to reduce assessments to single numbers, and thus are prized in the policy world. Teachers’ solutions tend to be much knottier, because they work with kids of delightful variety and hope to help each one realize his or her potential in an appropriate way. In prior writing on teacher evaluation (Smagorinsky, 2014), I’ve argued for more local models of evaluation, conducted by teachers who understand kids, schools, life, and teaching. Evaluation would be both formative on an annual basis, and more formal and consequential at greater intervals of 3–5 years. Teachers would have a much greater say in the criteria by which their colleagues are evaluated—much more so than the detached policymakers who determine evaluation criteria from distant offices removed from kids, parents, and bees flying in the classroom window. Although my emphasis in that plan is on secondary school teachers, I see it as appropriate for primary school teachers as well.

This essay has not solved the problem of developing an ecologically valid assessment system, which I describe in much greater detail in the aforementioned 2014 article in *English Education*. But I hope that it at least provides some direction and substance in coming up with better options than our policy lords are currently imposing on us.

### References


Peter Smagorinsky is Distinguished Research Professor of English education at The University of Georgia and can be reached at smago@uga.edu.

*Language Arts*, Volume 93, Number 2, November 2015
Starting with Why in Teacher Preparation: Community

Sean Kottke, Paula McPhee, Arlene Lents, and Meredith Dodson

What makes a good teacher? How should teachers’ performance be assessed—or should it be assessed at all? What does it mean to call a teacher “effective” or “proficient,” and can first-year teachers ever hope to be “proficient” in the first place? These questions frame much of educational discourse in 2015, as evidenced by the collection of essays you’re holding in your hands (or perusing on a screen), and their answers form the core of myriad interventions for teacher preparation, induction, and professional development.

Indeed, American education policy conversations have perennially—but most heatedly in the NCLB/Race to the Top era—been shaped by how proponents of teacher accountability on the one hand and proponents of teacher autonomy on the other come down on these issues. However, as Green (2014) argues, both perspectives tend to give some degree of credence to the well-worn “Myth of the Natural-Born Teacher” (p. 6), with the former group trusting high-stakes assessments to sort out good teachers from bad and the latter group trusting that professional educators, left to their own devices, will do the right thing. The problem is, neither perspective attends in much detail to how great teachers do what they do or how struggling teachers might grow into great ones. These questions merely focus on the what and how of ensuring educator effectiveness. Beyond pious invocations of what “all children” deserve, neither reformers nor traditionalists consistently approach these considerations by, as Simon Sinek (2009) would advise, starting with why.

Why make good teachers? Answering this question in a way that actually supports the making of good teachers—that generates the what and the how of teacher preparation—requires more than a carefully composed mission statement or program claim. It requires the thoughtful consideration and engagement of actual teachers in actual contexts of teaching and learning. In short, it requires the involvement of an authentic community. The authors of this commentary have had the privilege of participating in just such a professional community over the past decade. Each of us has taken multiple stakeholder positions in the work of teacher preparation at The Robert B. Miller College’s Elizabeth H. Binda School of Education, from college professor and administrator to elementary school teacher and principal to state Department of Education representative. Our experiences in facilitating the entrance of seventy-one new teachers into the profession over this period give us warrant for pragmatic responses to the above questions, that is, useful whats and hows for educator preparation that have yielded positive outcomes. However, Miller College started with why; all those whats and hows developed organically from our specific answer to the question of why make good teachers: community.

At the turn of the millennium, the community in which Miller College was founded, Battle Creek, Michigan, had no locally based avenues for attaining credentials beyond an associate’s degree. As a consequence, Battle Creek had one of the lowest per capita rates of adults holding bachelor’s degrees or higher in Michigan, and at the dawn of the Great Recession, it was the Michigan metropolitan area with the lowest percentage of adults between ages 25 and 34 holding bachelor’s degrees (Erickcek & Watts, 2008, slide 40). The lack of a local teacher preparation institution in Battle Creek resulted in shortages of teachers in critical areas like elementary and special education, a shallow pool of certified educators to respond to the increasing need for highly qualified teachers created by retirements and local demographic changes, and no local avenues for supporting the continuing education of practitioners.

American education policy conversations have perennially . . . been shaped by how proponents of teacher accountability . . . and proponents of teacher autonomy . . . come down on these issues.
and professional development of teachers and paraprofessionals.

The Binda School of Education was developed by community members to address this need and has been a collaboration of education professionals and various stakeholders throughout its first decade of operation. Miller’s teacher preparation philosophy, curriculum, and pedagogy emerged from the work of a coalition of K–12 stakeholders—principals, master teachers, and central office administrators with decades-long records of effective instructional leadership in Battle Creek’s public schools—under the guidance of professional teacher educators from Aquinas College. It’s important to note that this mentorship involved many critical areas of program development and allowed Miller to develop its own framework for teacher education, with assistance from Aquinas in identifying local priorities for the preparation of Battle Creek’s next generation of teachers. This collaborative spirit has been so pervasive and palpable in Miller’s operations that during a 2013 site team visit associated with reaffirmation of Miller College’s accreditation with the Higher Learning Commission of the North Central Association of Colleges and Schools, one of the peer reviewers remarked that Miller was, in essence, an “upper division community college.” This was not meant in a pejorative sense by any means, but rather was intended to be complimentary of Miller’s engagement with the needs of its community.

The first priority for teacher education programming at Miller has been a strong practitioner focus, privileging the wisdom of practice—and more specifically, the wisdom of practice derived from extensive experience in Battle Creek area schools—over an accumulation of traditional scholarly achievements when assembling an advisory team and faculty. Recognizing that the absence of a local pipeline for induction of new teachers obliged Battle Creek schools to seek new teachers from outside of the community, the program’s founders were explicit about the value of a “grow your own” partnership between K–12 schools and teacher education. By attracting to its faculty locally well-respected master teachers with extensive track records for successful instruction and school improvement, Miller ensured that its candidates would learn effective pedagogical strategies from instructors who modeled such practices in the college classroom. It further nurtured a strong, trusting relationship with local schools, whose leaders could be assured that the teaching practices they wished to foster among their faculties would be infused into the instructional toolboxes of the next generation of teachers.

This trust and buy-in from the community was critical in order for Miller to realize its second priority for teacher education: a deep commitment to field-based instruction. The linchpin of Miller’s practitioner-based approach, field experiences are integrated into nearly all of the courses in the teacher preparation program, beginning with the very first courses taken by candidates prior to formal admission to the Binda School of Education. Mentor teachers for these placements are carefully vetted and selected by the College in consultation with principals to ensure that the most effective mentors are selected.

It’s also important to note that whereas some programs stop at merely defining the number of hours that candidates spend in the field (see Tyson, 2014), the field experiences built into Miller’s teacher education program exist on a continuum of legitimate peripheral participation (Lave & Wenger, 1991). In early foundational courses, candidates primarily observe practicing teachers and reflect upon their observations in prescribed journal writing activities specifically linked to course content, live and online class discussions, formal observation papers, and oral presentations. However, they also have narrowly focused opportunities to teach single lessons in field classrooms, which are recorded on video and critiqued by peers, instructors, and mentors.

In the core content area methods courses that form the middle phase of the program, candidates begin taking increasingly active roles in assessment and instruction of students under close supervision.
of course instructors and mentor teachers. This is particularly well developed within the language arts methods sequence, in which candidates begin studying and observing best practices in literacy assessment and instruction, followed by working with a single child, administering reading assessments, and developing instructional interventions informed by the data and progress. Later, they practice assessment and instruction across a full semester with groups of students of varying abilities in a methods course taught in partnership with a practitioner onsite in a local school. With multiple opportunities to enact the teaching and learning cycle across all core content areas in authentic settings, candidates arrive at their culminating internship experience ready to accept the responsibility for the learning and achievement of individuals and whole classes of students.

The mutual trust and sense of a cohesive, K–16 educational community fostered by the practitioner-oriented and field-based paradigms Miller College emphasizes allow for a community-based approach to candidate performance assessment, the third pillar on which the program is based, and the one that provides the strongest warrant for confidence in the quality of teachers prepared by Miller. From the first foundation course in the program to the final internship, candidates’ performance in field settings is assessed independently by multiple parties using common rubrics. In the internship experience alone, as many as six separate assessors provide feedback on multiple observations of a single candidate’s teaching across all core content areas. Unlike the framework for assessing intern teaching in some larger institutions, Miller College’s assessors are not contingent employees with a sole responsibility for supervising student teachers. Rather, they are the same practitioner-oriented faculty and administrators (yes, even the Dean of the Binda School of Education conducts observations) who oversee and instruct the foundations and methods courses.

Mock interviews with panels of local administrators invited to campus at the end of the internship experience support the preservice to inservice transition by preparing candidates for their imminent job search. They also provide a regular feedback loop between the College and key K–12 stakeholders, as does the return of alumni to campus each semester to advise interns on what to expect in their first years of teaching. These practices maintain strong ties between the College and K–12 faculties, facilitate effective triangulation of assessment information on candidates, and, more important, build a community of practitioners around each candidate to support their induction into the profession.

These three pillars—practitioner-oriented faculty, field-based instruction, and shared assessment of candidate performance—work in concert to maintain a strong, interdependent educational community. That idea of community engagement should be central to any teacher preparation project, no matter what specific curricula or tools for crafting or assessing the skills of early career teachers a program deploys. It is telling that in their recent comments on hearings related to the implementation of edTPA in New York, Sharon Robinson and Mark LaCelle-Peterson’s first main observation is not about edTPA itself, but about the context of effective performance assessment: “Clinical development of new teachers and rigorous performance assessment require real partnerships with schools and school districts” (2014, para. 7).

Miller College’s experiences in developing an approach to educator preparation in close concert with its community bear out the truth of this observation. In a period of recession and teacher cutbacks, Miller College boasts a job placement rate of over 80%, with most of its graduates employed in the Battle Creek area in the same schools in which they completed field experiences and internships. Their immersion in the local educational community during their initial preparation empowers familiarity with district curricula, pedagogy, procedures, and policies, minimizing the shock new teachers typically experience in moving from the academic environment to the teaching field.

As more of its graduates are inducted into the local teaching force, concentrated pockets of Miller-prepared teachers in individual buildings promise to transform those schools into professional development schools, perhaps even realizing the dreams of many Miller faculty for a true lab school. Chief among the beneficiaries of this cycle are the children.
in the schools assigned to new teachers who are more familiar with their teaching expectations and can therefore focus their efforts on the needs of their young students. Our hope is that a portion of those well-educated young students aspire to become educators, take advantage of the education academy experience available in high school, attend college in this community, and remain to continue the cycle by contributing their skills to the local school districts.

Our intention in this commentary has been to present not a blueprint, but rather a mindset for teacher preparation. We have been careful to minimize detailed discussion of the whats of Miller College’s programming in favor of emphasizing the why—community—and key ways in which the why affects the how. To be sure, a casual perusal of the Miller College catalog may not strike the reader as anything special, with straightforward, even square, course names like “Classroom Management,” “Introduction to Education,” and “Teaching Reading and Language Arts I.” Let’s face it, those course names wouldn’t seem out of place in any average American school of education of virtually any decade.

Similarly, in contrast to trends championed by many contemporary education reformers to shorten pathways to teacher certification, Miller’s program seems positively retrograde, with most candidates earning bachelor’s degrees in excess of 130 credit hours. And, as currently structured, the program is more boutique than franchise, with an upper limit to the number of candidates that can be effectively served at any given time. However, what makes Miller College unique among its peers in Michigan’s teacher preparation landscape is not its what, but its why: community, and the value it seeks both to add and to derive from its engagement with it. Its current president frequently invokes Miller as “Battle Creek’s College.” We can think of no better moniker to evoke the commitment to community upon which all of its teacher preparation efforts are based.

References

Sean Kottke is an education consultant with the Michigan Department of Education’s Office of Professional Preparation Services and can be reached at kottkes@michigan.gov. Arlene Lents is the Field Placement Coordinator for the Binda School of Education and can be reached at lentsa@millercollege.edu. Paula McPhee is part of the education faculty in the Binda School of Education and can be reached at mcpheep@millercollege.edu. Meredith Dodson is Assistant Dean for the Binda School of Education at Miller College and can be reached at dodsonm@millercollege.edu.

Chicago Public Schools and Performance Assessments for Teacher Evaluation

Peter Leonard, Claudinette (Didi) Swartz, and John Barker

In January 2010, the State of Illinois passed the Performance Evaluation Reform Act (PERA), which required Chicago Public Schools (CPS) to begin implementing an enhanced teacher and principal
evaluation system in the 2012–13 school year. Following passage of PERA, CPS conducted focus groups of CPS educators and other stakeholders to gather their thoughts and ideas on teacher evaluation and actively engaged CPS teachers and their Chicago Teachers Union representatives in the development of a new teacher evaluation program. The product of those efforts is CPS’s system of educator evaluation and support called “REACH Students” (Recognizing Educators Advancing Chicago Students). Launched in 2012–13, REACH has been phased in incrementally. The 2014–15 school year marked the first time that nearly all educators received a REACH Students Rating. Teachers will receive their ratings based on administrator observations of teacher practice and the academic growth of their students.

Student growth is measured in two ways for CPS teachers. The first is a value-added measure that attempts to isolate the impact of the teacher on student growth along the MAP (Measures of Academic Progress, an elementary level assessment) and EPAS (Educational Planning and Assessment, for high school) assessments. This metric meets the requirement for a statewide or nationally administered reliable assessment instrument. According to PERA, the second student growth measure must be rigorous and aligned to course curriculum. CPS and the Chicago Teachers Union (CTU) collaboratively designed the REACH Performance Task program to satisfy this condition for all district teachers. The REACH Performance Task program is an assessment creation and administration program managed by the district’s Department of Student Assessment.

REACH Performance Tasks are CPS-teacher-created performance assessments that measure student growth over the course of the academic year. Two test administrations, one at the beginning of the year (BOY) and one at the end of the year (EOY), capture student progress within one to two core standards in a disciplinary grade level. The BOY assessment is written at an end-of-year expectation level so that an accurate “apples to apples” comparison can be made between the two assessment administrations. Because of this structure, students are not expected to perform exceptionally well at the beginning of the year. They are, however, expected to improve in the knowledge and skills necessary to succeed in that standard by the EOY task administration. Teachers administer a Performance Task to one of their classrooms at the BOY and EOY. The percentage of a teacher’s students achieving growth along a standardized 0–3 scale is factored into a teacher’s evaluation score.

Performance Tasks are currently designed for 11 content areas across PreK through 12th grades—arts, career and technical education, education technology, library science, literacy, mathematics, physical education, preschool, science, social science, and world languages. Teachers input 449,000 student task scores into CPS’s student information system during the 2014–2015 BOY testing window. The literacy assessments were the most commonly administered Performance Tasks, with more than 184,000 students completing a literacy BOY task during the 2014–15 school year, approximately 12,000 of which were administered in Spanish.

In the process of creating standardized Performance Tasks, CPS has learned many lessons about how to effectively use student performance assessments as teacher evaluation tools. Two of these lessons are that the teachers must own the process and the assessments must serve multiple purposes.

CPS has learned many lessons about how to effectively use student performance assessments as teacher evaluation tools. Two of these lessons . . . teachers must own the process and the assessments must serve multiple purposes.

Teachers Must Own the Process

REACH Performance Tasks are “by teachers, for teachers.” CPS teachers write the assessments that will influence their colleagues’ evaluations based on student performance. The Department of Student Assessment implements a rigorous application process to select participants for the task-writing teams. These exemplary teachers devote
the necessary time and effort to create high-quality tasks; they own their responsibility to their fellow CPS teachers.

The task/creation process begins with assessment/design training for all teacher task writers, co-facilitated by the assessment team and CPS content specialists. The training focuses on how to design relevant and rigorous performance assessments that are accessible for CPS students. Teams are provided tools that help them apply general design principles to their specific content areas. The Department of Student Assessment organizes a full day of release time for all Performance Task writers and pays for schools’ substitute teachers on that day. In the design of literacy tasks, teams work in four grade bands: K–2, 3–5, 6–8, and 9–12. Each band is supported by a literacy specialist, with the grade-band configuration ensuring consistency and alignment between tasks.

Over the course of six months, these teacher teams write BOY assessments, pilot them with their students, and revise them based on student and teacher feedback. They passionately debate question wording, standards language, and rubric construction, with a constant eye toward task quality and validity. Satisfied with the BOY product, teams then create the parallel EOY assessments. These EOY tasks measure the same comprehension and analysis skills as the BOY but with new texts. In school year 2013–14, fourth-grade students demonstrated their ability to identify the main idea of a passage by reading about George Washington Carver at the BOY and Benjamin Franklin at the EOY.

Approximately 275 teachers contributed to the design of school year 2015–2016 Performance Tasks through task writing and piloting. Teacher ownership of Performance Tasks results in many task writers returning to create tasks in subsequent years. Through repeated participation, teachers build their assessment design skills and collaborate with accomplished peers.

Although it is widely publicized that Performance Tasks are created by district educators, CPS is reinforcing this point in school year 2015–2016 by printing “DESIGNED BY CPS TEACHERS” on every task distributed across the district.

**The Assessments Must Serve Multiple Purposes**

REACH Performance Tasks are used for teacher evaluation, but their utility extends beyond their main accountability function. These assessments communicate expectations to teachers of what students should know and be able to do by the end of the school year in a particular content area and grade level. Through the BOY administration, tasks provide teachers with baseline information about their students’ proficiencies in foundational standards.

In language arts, the assessments reflect the three major shifts present in the Common Core State Standards (CCSS; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) for English Language Arts (ELA): regular practice with complex texts and their academic language; reading, writing, and speaking grounded in evidence from texts, both literary and informational; and building knowledge through content-rich nonfiction. The quality of the literacy assessments begins with the selection of grade-appropriate, complex texts. In the eleventh-grade BOY task, students read a speech delivered by Elie Wiesel at the White House Millenium Lecture Series in 1999. In the address, Wiesel uses powerful and philosophical language to warn against the “perils of indifference.” Invoking both personal and world history, he argues for the demonstration of empathy in human affairs. Wiesel’s words force students to grapple with sophisticated ideas expressed through academic vocabulary.

In the selection of texts such as Wiesel’s speech, the literacy task writers focus on the evaluative criteria delineated in Appendix A of the CCSS for ELA. The task-writing teams review both the qualitative and quantitative dimensions of texts and consider how students will engage with the passages. Levels of meaning, language conventions, Lexile level, and student prior knowledge are some of the factors considered in this comprehensive view of grade-level-text appropriateness. Students interact with these passages and ground their responses in textual evidence. Students in second grade listened to their teacher read The Tortoise and...
the Hare; they then identified the story’s moral and cited a detail from the book to support their position. Grounding answers in evidence is an expectation across grade levels, from early childhood through high school. The literacy Performance Tasks make visible how that expectation builds, from second graders responding to a familiar fable to ninth graders analyzing John Brown’s courtroom speech after his Harper’s Ferry raid.

In the Brown and Wiesel speeches, students encounter content-rich nonfiction texts that address social, political, and historical themes. In other grade levels, scientific topics are addressed. In the eighth-grade BOY task for school year 2014–15, for instance, students read about vivisection, the practice of dissecting a live animal. REACH Performance Tasks model what good assessment looks like and provide teachers with a blueprint to design their own high-quality tasks. Performance Tasks are meaningful tools for CPS teachers because they deliver important data on student performance and reflect districtwide instructional shifts.

The existence of teacher evaluation systems provides an opportunity to create and distribute tools like REACH Performance Tasks that help teachers see their students’ abilities and plan for what comes next in the classroom. Through careful construction, clear purpose, and ongoing iteration, CPS is building a sustainable program that aims to support teacher evaluation, develop teacher practice, and advance student learning.

Reference

Peter Leonard is Program Manager for REACH performance tasks and can be reached at pyleonard1@cps.edu. Claudinette (Didi) Swartz is the Director of Assessment at Chicago Public Schools and can be reached at cmswartz@cps.edu. John Barker is the Chief Accountability Officer for Chicago Public Schools and can be reached at jrbarker@cps.edu.