



[Current Issue](#) | [From the Editors](#) | [Weblog](#) | [Editorial Board](#) | [Editorial Policy](#) | [Submissions](#)  
[Archives](#) | [Accessibility](#) | [Search](#)

*Composition Forum* 37, Fall 2017

## New Jersey City University's College of Education Writing Assessment Program: Profile of a Local Response to a Systemic Problem

SHARE ...

Audrey Fisch

**Abstract:** This profile presents New Jersey City University's Writing Assessment Program from its creation in 2002 to its elimination in 2017. The program arose as an attempt to raise the writing skills of the diverse, first-generation teacher certification candidates in the College of Education. Despite political missteps, the program gained greater administrative support in 2009, and in this second stage, the program capitalized on greater institutional support to use data-driven analysis to inform policy. In 2014, however, New Jersey moved to require the Praxis CORE, and the Writing Assessment Program became obsolete. This profile discusses the many ways in which a locally developed, student-centered, and instruction-driven assessment program can raise student skills and the losses involved in a shift from local to national assessment.

In 2002, the Dean of our College of Education received a letter from a local public school superintendent referencing correspondence from a job applicant seeking employment as a teacher. The superintendent wrote: "I was appalled by the number of errors in the correspondence, especially since the author is seeking employment as a teacher in my district. I am confident that your institution produces bright, energetic people fully capable of entering the academic profession; however, letters such as the one attached negate the educational and professional criteria taught by your university."

The superintendent was no academic snob concerned with the odd proofreading errors of a less than careful job applicant. He enclosed a copy of the job letter, and it demonstrated widespread writing issues in a variety of areas. I quote only one portion to demonstrate the range of writing issues:

Having the utmost interest in your school feeling I would be an asses to your school. I am dedicated and , reliable, and will use every possible asses to teach the students at your school. As my enclosed resume indicates I have previous experience working in different class room settings, process a NJ Teacher of the Handicapped Certification. Being a dedicated teacher, enjoy flexible work and helping others find success.

Certainly, the student may have been a dedicated and reliable teacher, with much to offer potential students, but the job letter submitted demonstrated some combination of a lack of writing proficiency and a lack of understanding of the kind of polish necessitated by this particular writing task.

While the letter was an anomaly in the fact that it made its way back to a university administrator, many of us working in the College of Education had concerns about the writing skills of our teacher candidates. We had already begun work to revamp a flawed instrument designed to measure and ensure the writing competence of students seeking certification. This letter confirmed the need for a new exam and a wholesale writing-support program.

The following program profile describes the stages of the development of our Writing Assessment Program, including our creation of an exam and various permutations of support for students. I describe our political challenges and missteps, and the ways in which institutional support and data-driven decision making allowed us to create a successful program. Finally, I discuss how the current reality of the Praxis CORE, an ETS exam now required of all certification students in New Jersey and in many states across the country, rendered our program obsolete. At the

same time, I reflect on the costs, to our students and to the cause of improving the quality of teacher education and of public education broadly, of moving from local, home-grown assessments like ours at NJCU to national ones like the Praxis CORE.

## Institutional Background

New Jersey City University (NJCU) is a comprehensive, public university with approximately 6,300 undergraduate students and 2,000 graduate students. The University offers two doctoral, 27 graduate, and 43 undergraduate programs. We have approximately 130 full-time faculty, with 45% of all courses taught by part-time, adjunct instructors.

NJCU's mission is "to provide a diverse population with an excellent university education" ("Mission"). Designated as both an Hispanic-serving and a minority-serving institution, our undergraduate student population is broadly diverse: 25% White, 21% Black, 34% Hispanic, and 9% Asian. The average undergraduate is a 26-year old woman from a working-class family. 77% of our students receive financial aid, with approximately 64% receiving Pell Grants. Many are the first in their families to attend college ("NJCU Profile and Accreditations"). Most of our undergraduates commute from the surrounding urban areas where they were raised in communities where Academic English is not the dominant discourse. For many, English is their second or non-primary language. They work many hours off campus (30% work 21 or more hours per week; 70% work 11 or more hours per week), and many manage extensive family responsibilities on top of their school work.

The diversity of our student body is its strength. The wide variety of cultures, backgrounds, languages, and experiences our students bring into the classroom makes for a rich and vibrant learning environment. But our students present significant challenges in terms of their skills and their preparation for college. As of Fall 2013, 70% of incoming students enrolled in one or more developmental course. In Fall 2015, average SATs for full-time students were 446 Reading ("NJCU Progress Card") and 438 Writing (average SATS for part-time students as of Fall 2014 were 365 Reading, 361 Writing) ("Institutional Profile").

Most of our students were educated in the surrounding urban schools, and our certification students intend to return to those schools to serve their local communities. In so doing, they represent significant and important social change. They bring back to their home or neighboring districts their personal experiences, their newly acquired passion for their subject matter, and their commitment to their local communities. They are neither surprised nor intimidated by the challenges of urban public education. For most, the public school salaries, even at lower-paying charter schools, are a gateway into the middle class. Moreover, our graduates of the College of Education are ultimately responsible for preparing the pool of future NJCU undergraduates. So, most of the faculty share in the understanding that it is incumbent on us at NJCU and in the College of Education to produce teacher candidates who are highly competent in literacy and in their subject matter.

I should note that the recognition of writing deficiencies in our students was not an indictment of our Composition Program (housed in the English department). The University Composition Program, in which students complete 6-8 credits in first-year composition, is only one piece of our students' writing preparation. Approximately half the students in our certification programs are transfer students, who enter from a variety of feeder institutions, including several different community colleges with their own composition programs and requirements. In addition, a portion of our students move through the English as a Second Language Program's composition program with its own different requirements and standards. Moreover, as the average undergraduate age mentioned above indicates, our students have spent time away from school before and during college, and this only magnifies the variety in skills. Their progress towards their degrees is slow and uneven, with substantial time off for work, family, etc. Our four-year graduation rate, a very low 5% as of Fall 2015, for example, indicates how non-traditional our students' pathway through college is and how frequently their responsibilities outside of college have caused them to stop out or veer away from the straight pathway other more traditional students take ("NJCU Progress Card"). Whether or not the students had acquired reading and writing competency early on in their college education, we were not confident that our students were still in full possession of the writing skills we wanted as they embarked on their careers in education. Our task, then, was clear.

## Stage One: Our Initial Initiative

When a small group of us came together to think about the issue of writing proficiency, in an initiative driven by faculty but with some administrative support, we were cognizant of the dangers and social consequences of writing assessments, particularly on the underrepresented populations that make up the majority of our student population.

We were well aware that “the history of writing assessment is about the existence of power” (Elliot 349). We were cautious about developing an assessment that would operate as a proxy for some over-generalized notion of writing ability: the assumption that “an ability to answer selected-response items about grammatical conventions or to produce an impromptu belletristic essay were sufficient proxies for writing ability.... [an] assumption [that] still operates in the most powerful US large-scale writing assessments” (Dryer and Peckham 36). We wanted our locally-based assessment to be “situated in social practice” and “context” (Anson 116). We were inspired by narratives about local faculty claiming and building their own assessment tools (Haswell and Wyche-Smith 221). Finally, we were optimistic that an assessment could help us build broader investment in writing (Huot et al 511): including resources for development courses, for our new writing center, and for broader faculty engagement with writing in courses. From early on, we articulated a vision of a writing assessment program that centered around helping students acquire stronger skills rather than weeding students out.

Our first step was to convene an initial group that included key faculty from the important constituencies who would support the Writing Assessment Program: Literacy Education, English, Special Education, Multicultural Education, Elementary and Secondary Education, and Early Childhood Education. Our group then expanded to include the newly hired Director of Composition. Our preliminary discussions centered around the rejection of a holistic assessment (Huot “Reliability”) in favor of a focus on what Haswell and Wyche-Smith call a “diagnosis” (229). We wanted to be able to use our assessment to identify and define “future paths of instruction” (229) in the areas we felt best reflected the skills we wanted for our future teachers.

We designed our exam around a fairly standard approach. We gave students a medium-length reading prompt (600-900 words) and a meaty two-hour time frame in which to read the prompt and write their responses. Students were asked to offer analysis, guided by a choice of three different sets of questions. Students were invited but not required to include their own ideas and personal experiences in their responses. ([Appendix 1 - Sample Exam](#))

In many ways, our assessment fit well with the CCCC guidelines (1995 and revised and affirmed 2006, 2009, 2014). We gave students ample time to plan and write, although only in one sitting; we publicized the purpose of the assessment; we gave students substantial feedback about their results; and we provided a clear avenue for appeal (CCCC 434-435). Our greatest difficulty was providing a writing task that was “developed from the curriculum and grounded in ‘real-world’ practice” (CCCC 434) because our assessment could not be embedded in any particular class. To ameliorate this difficulty, we worked continuously to find readings and questions that were “appropriate to and appealing to the particular students being tested” (CCCC 431). We strove to use authentic topics (Condon 149) with relevance to the students’ experiences in education (the summer slide, caring for undocumented or homeless students, the role of substitute teachers, etc.). We also worked to choose readings that would be maximally engaging and accessible, with just “enough grist for the writer’s mill” but not too much depth to challenge unduly the reader and hence “undermine her performance in writing” (Condon 143).

The rubric we developed assessed students in two general areas: grammar/vocabulary and organization/development ([Appendix 2 - Rubric](#)). We hoped that our identification of these areas, as Stock and Robinson put it, would make “explicit the values inherent in the set of expectations that assessors bring to the act of evaluation” (100) and further drive instruction in these areas throughout the College of Education and university.

In terms of organization/development, we looked for the basics of essay writing in relation to a reading prompt: organization (introduction and conclusion; paragraphing and transitions); the ability to understand and respond to an argument using textual evidence; and some development with examples. In relation to grammar/vocabulary, we looked for overall competence, without patterns of error or excessive errors that inhibit meaning. One singularity of our rubric was the requirement that a student exhibit “adequate performance” in the area of grammar/vocabulary, regardless of his/her overall score. In other words, a student who could construct a competent, organized, and well-developed essay but had extensive problems (patterns) in specific areas of grammar (sentence construction, verb endings, pronoun use) or vocabulary (spelling, homonym usage, word choice) would fail. The key here was not the number of scattered errors but the presence of a pattern of errors in a particular area, indicating, for example, that the student had never learned to use an apostrophe. Always, our focus was on identifying areas for instruction, so that we could better prepare the students we were sending out.

Our rubric changed over time and much discussion, and the narrow three-point scale evolved to meet our needs without making our task of grading needlessly difficult. Because our assessment was trying to identify basic competence rather than differentiate and rank students, “excellent,” “adequate,” and “weak” performance were sufficient performance indicators. The differentiation between excellent and adequate only became important when a student fell down in one area but demonstrated superior performance in another, which allowed an overall passing score. This scenario only came into effect when a student struggled in an area on the bottom of the rubric, such as failing to include an introduction or conclusion, but otherwise demonstrated strong enough facility in another area (responsiveness to the question and comprehension accuracy, for example) sufficient to boost his or her overall

score.

Our confidence with our writing task and our scoring instrument was compounded by our inter-reader reliability and collective approach to scoring. We were particularly attentive to what Dryer and Peckham call the “ecology of scoring,” including issues related to “imbalances of authority, expertise, and assertiveness” among graders (33). For this reason, we encouraged a good amount of off-topic table-chat (about families and other work issues) that built a level of intimacy in our grading community. This sense of community served as an important background to any questions about scores that arose during readings and created a model environment that “emphasize[d] the communal nature of reading” (Dryer and Peckham 34).

As a result, our practice of reading - the standard two readers per paper, unless there was disagreement - was far less likely to result in the need for third readings. More commonly, two readers who initially disagreed about or perhaps were simply unsure about how to score a paper would immediately confer, sometimes with a third person, in order to come to a consensus. In this way, a lively and loud working environment promoted continuous conversations about what a student’s essay reflected in terms of skills and whether a student might benefit from some development work (more on this below). We also had the opportunity to consistently remind each other that our goal was to assess writing competence: to identify areas where students did not have adequate command of an aspect of writing that we felt confident they could be taught and could master. The community of scorers was quick to bring back on track anyone who reacted to what appeared to be a lack of overall intelligence (again, not within the purview of our assessment) in a response or an unwelcome expression of values or ideas in relation to education. We also had the benefit of our local experts within our scoring group to help us navigate around potentially distracting second-language interference and markers reflective of special-education issues.

While we developed our assessment tool, we also experimented with different plans for remediation in relation to student results. We distinguished between what we identified as the F1 and the F3 paper. This was based on our initial creation of 1- and 3- credit developmental classes. We distinguished papers that demonstrated a general failure to construct a coherent essay from those that exhibited isolated problems in the areas of grammar or vocabulary (patterns of error in sentence construction, verb endings, pronoun use or spelling, homonym usage, word choice). This determination of what constituted an F1 or F3 paper was made holistically, in relation to both the score and the content of the paper, and like the score itself required agreement among the graders. Always, our emphasis was on how we thought the student could best acquire the necessary competency. According to our latest data from 2015, 28% of first-time test takers scored an F3, and most of these also demonstrated substantial patterns of errors in the areas of grammar and vocabulary. 22% of first-time test takers scored an F1.

Initially, we created two different classes, but it quickly became apparent from those of us who taught that class that 1-credit was overkill for most students who wrote F1 papers. We modified our structure and instead sent students with an F1 to our Writing Center for a minimum of four sessions of tutoring. For our 3-credit class, we repurposed an existing, credit-bearing class in Literacy Education. We were able to secure institutional support to run this class with low enrollments, capped at 12 students. Instructors for this class used the rubric and mock exams to offer intensive instructional work to address the writing issues identified by the test and the rubric. (See [Appendix 3: Course Guidelines and Sample Syllabi](#))

Many of our classes are staffed by adjunct faculty who are naturally more difficult to fully embed in the school culture and who may have received late assignment to and last-minute training about the class. Because we wanted to give these faculty and the students in the 3-credit class additional support, we designed the class to include a midterm mock exam, scored by the Writing Assessment Committee. To score these midterms, regular Writing Assessment Committee graders joined with all the instructors of the 3-credit classes to grade and discuss papers. No student’s paper was graded by her own instructor, and we paired instructors with regular graders. This allowed the Committee and the instructors to share in mid-semester norming of instruction and expectations. The faculty always appreciated the mid-semester reassurance that they and their students were “on track.” The midterm mock exam also allowed students to get comments from the “real” committee and in so doing reinforced the feedback being presented to them by their instructor.

At the end of the semester, students took the writing assessment exam as their final for the course, and their score on the exam counted as a substantial part of their final grade. If a student did not succeed in passing the exam, the instructor could give him/her a grade no higher than B-, depending on the overall quality of his/her work in the class. In this way, the class did not represent a large potential threat to the student’s GPA, as it might if a failing score on the final meant an F in the class.

Thus our Writing Assessment Program took shape. Students seeking certification in the College of Education were required to pass the exam before they embarked on their first (of two) field experiences (usually in their junior year). We allowed all students to retest without any remediation, with the idea that the first failure might simply indicate that

the student had had a bad testing day, in which he/she had, for whatever reason, failed to display his/her abilities. We required remediation (tutoring for the F1 paper or the 3-credit class for the F3 paper) after the second failure. Students were allowed to retest as many times as they wanted.

## Stage Two: Setbacks and Growth

The exam, as I stated at the outset, grew out of a recognition that our students had substantial writing issues that we wanted to address, particularly in terms of patterns of error or fundamental difficulties in constructing an argument or working with text. As the numbers above indicate (22% score F1, 28% score F3), approximately 50% of initial test-takers failed the exam. These results correlate with our students' results on other standard assessments. Our students with a score below 400 on the Reading SAT, for example, had an initial pass rate of 28%; students with a score below 400 on the Writing SAT had an initial pass rate of 21%. Students with scores of 500 or above, passed at more than 80%. These numbers are generally unsurprising; they point to the challenges of our university's urban mission and the needs of our underprepared student body. Indeed, 37% of test takers in our data set have Reading SAT scores below 400.

To be clear, however, many of our students who enter with low scores on standardized assessments grow and thrive at NJCU, and we know that standardized assessments are not the best measure of either their skills or potential. But were these students acquiring *all* the writing skills we wanted? One piece of data on this issue was troubling: students with a GPA of 3.7-4 had a passing rate of only 67%. These students were thriving academically, yet our exam indicated that many still had some writing issues. Not surprisingly, faculty and administrators were defensive and uncomfortable with this data. More broadly, because our exam was institutionalized as a barrier to the educational field experience, students who were not able to pass (or to pass in a timely manner) found their progress towards certification and their degree stymied. This, in turn, began to affect enrollments and to do so unevenly, as certain programs discovered "their" students to be disproportionately affected by the standards put into place. We didn't, I think, do well enough in sharing information about the exam and our expectations. Nor did we do well in addressing the departmental insecurities and very real enrollment concerns unleashed by the exam. Our exam was decried as a flawed instrument, and much institutional drama ensued.

In 2009, however, a member of the Writing Assessment Committee was appointed Dean. He championed our initiative with much-needed additional institutional support. For example, running our intensive 3-credit developmental classes had always been a tricky endeavor. Students often waited until the last minute to take the exam. Then they delayed registering for remediation, so our Dean needed to fight to keep these classes open and to allow them to run them with low numbers. (Conversely, late enrollment also meant that our instructors had to be generous in allowing extra students in above the cap, which was also often necessary.) Only an administrator with a view of the big picture could defend these courses.

The committee also had a second chance to build a broader understanding of and commitment to the Writing Assessment Program across campus. We worked harder to develop a broader group of committed full-time and part-time faculty and staff from across institutional units who support the program as instructors or members of the Writing Assessment Committee responsible for grading. We gathered greater representation from departments in the College of Education, more participation from the English department and people working in the Composition Program, and invited involvement from key staff from some of our specialized student-support programs (e.g., our Opportunity Scholarship Program for Economically Disadvantaged Students).

We also worked with our Writing Center, sharing samples of retired exams to be used to help prepare students. The Writing Center developed and offered writing assessment bootcamp classes, designed to familiarize students with the exam and offer quick tips and tricks and often conducted by peer tutors right before the exam was offered. We created a website where we posted scored sample exams, complete with scored rubrics and student-friendly explanations for why the samples received the scores they did (see [Appendix 4: Scored Exams and Rubrics](#)). We also posted tips (it's okay to use "I," mark up the reading passage, and leave time for revision and proofreading) (see [Appendix 5: Tips](#)). Overall, we worked hard to demystify the process, so our expectations were transparent to students, faculty, and staff.

We also focused attention on advisement in order to help students take the exam both more seriously and earlier in their academic careers. This allowed more time for retesting and remediation and in turn meant fewer students delayed their academic progress; students were able to get the support they needed earlier and more easily. As our Writing Assessment Program evolved, some students who self-identified as struggling writers even elected to take the 3-credit developmental class in preparation for taking the exam. We met with instructors from our feeder community colleges to share information, materials, and expectations. We also shared sample exams, scored

essays, and data with the broader university community during sessions at our Center for Teaching and Learning.

Institutional support also meant careful analysis of our exam data, and we were able to initiate important policy changes driven by this analysis. Initially, we felt that any student, regardless of the score received on his/her first attempt, should be allowed to retest. However, our 2012 data made clear that we were doing a disservice to students and making more grading work for ourselves with this policy. Data revealed that students scoring F1 on the first attempt, those with limited problems, typically in the areas of grammar and/or spelling/word use, passed at a rate of 53% on the second attempt. For those scoring F3 on their first attempt, those with more widespread writing issues, the second attempt pass rate was substantially lower, 36%. Taking the developmental course after the first attempt, moreover, was beneficial to students scoring F3 on the first attempt. We were able to identify a 25% difference in pass rate between those who scored F3 and who did not take the class after the first attempt (22%) and those who did take the class (47%).

As a result, we made a policy decision: students who scored F3 on the first attempt were no longer allowed the bad-day second chance. In order to retake the exam, they were forced to take the 3-credit class. Our subsequent data made clear that this was a beneficial decision. Prior to our policy change, first-time test takers who scored an F3 and retook the test had an overall pass rate of 40% on their second attempt. After our policy change and as of 2014, the pass rate for the second attempt (which now meant an attempt made after taking the 3-credit class) improved to 58%. In other words, by requiring that our F3 scorers take the 3-credit class, we helped them succeed.

We also officially limited students to three test attempts (although this limit was regularly waived by the Dean). The limit helped underscore the stakes of the exam and ensured that students were more prepared to take seriously each attempt. As a result, students who were under some kind of duress (stress or illness, for example), were more likely to excuse themselves from the exam (with no penalty) rather than continue under less than ideal circumstances and earn a confidence-damaging F score. The testing limit also caused more students to elect to take the 3-credit class, even if they only scored an F1, if they felt that tutoring was not helping them make adequate progress on their writing. Rather than allowing our students to test, retest, and flounder without adequate preparation, our policy changes provided the students with the guidance they needed to improve their chances for success. Again, given the importance of students making adequate and timely progress in their completion of program requirements, it was critical to give students this kind of guidance.

Overall, based on our 2015 data analysis, our initial pass rates continued to be low (49% pass on the first attempt—and this number generally held steady through more than a decade-long work at this enterprise); cumulatively, 93% of students passed before or on their third attempt and 99% of those who continued to take the test on a fourth or fifth attempt passed (these numbers have been in the single digits, as students were required to obtain administrative permission to retake the 3-credit class and exam after the third attempt). In other words, most students who persisted did eventually pass the exam. This was an important glass half-full observation, although it is indisputable that the exam was perceived by some students to be an insurmountable barrier: some did not make additional attempts either because they left NJCU or changed educational pathways.

All of which is to say that with greater institutional support from the Dean's office, the English Department, the Writing Center, the Advisement Center, various Student Service Offices, and the Office of Institutional Effectiveness, and stronger political navigation, which in turn engendered more support from the departments in the College of Education, we were able to create an effective program. Some students still exhibited skepticism about the pressure to achieve the level of proficiency required by the exam ("I'm going to be a math teacher, why do I need to know this stuff?"), but they felt supported by the instruction to become more confident and more competent writers ("why wasn't I taught some of this stuff earlier?").

### **Stage Three: Looking Ahead to Changes in Teacher Education**

While our local intervention persevered, the national conversation about raising standards in teacher education escalated. Across the country, while pay and treatment of public school teachers remains sub-par, calls have grown louder for higher standards for teachers and in teacher education. On June 14, 2014, New Jersey, like many states, adopted a version of these higher standards, requiring for teacher-certification candidates a higher college GPA (3.0, up from 2.75) and a minimum SAT, ACT, or GRE score ("approximately equal to the top-third percentile score for all test takers in the year the respective test was taken"—560 verbal and 540 math on the SAT) or passage of a new Praxis CORE, an ETS exam fashioned along the lines of the SAT, with sub-tests in reading, writing, and math ("Teacher Candidate").

In many ways, these new standards align with the efforts of our group. They attempt to ensure that new teachers enter the classroom with stronger academic skills so that they can impart those skills to the next generation of

students. The approach is laudable, but it also poses many of the same and some new challenges to an urban, public institution like my own where the study body is academically under-prepared. The differences between our local, home-grown and home-administered assessments and the CORE, for example, illuminate the broader challenges that remain in raising standards in teacher education.

One difference between our local assessment and the Praxis CORE is cost. Our Writing Assessment exam was free. While we floated the idea of a minimal charge for the exam to help cover the cost of exam administration (proctors and copying costs, for example), we never moved in that direction. Beyond developmental classes run with very low enrollment numbers, the broader institutional costs of our program included some release-time credits for our entire committee and its leaders. But because one of the founding members of our Writing Assessment Committee became Dean of the College of Education, we had an institutional champion, and we were not forced to pass along any costs to our students.

The Praxis CORE, in contrast, presents a substantial financial cost borne only by the student. Students can take all three parts (reading, writing, and math) together for a cost of \$150; each section costs \$90. This cost, like many now tacked on to the overall cost of certification (the subject area Praxis, substitute teacher licensing and fingerprinting fees required for student teaching, and internship assessments like EdTPA), make teaching certification a substantial investment, particularly for working class students who are struggling to manage tuition payments, commuting costs, books, housing, and food in a costly higher education environment.

Another difference between our exam, which was admittedly high-stakes in terms of its consequence to students, and national standardized exams like the Praxis CORE and the SAT is in terms of fairness. While our students may indeed lack the writing skills we want them to have, the Praxis CORE and the SAT are not accurate measures of those skills. A substantial body of research indicates that the SAT is most accurate in assessing student demographic characteristics rather than skills. As Lani Guinier writes in *The Tyranny of the Meritocracy: Democratizing Higher Education in America*, SATs serve as “accurate reflectors of wealth and little else” (11). These standardized tests “reliably measure a student’s household income, ethnicity, and level of parental education” (22). Guinier’s remarks here reflect a well-known reality, which is among many reasons the college admissions universe has moved away from what Guinier calls the tyranny of the “testocracy” (18).

Students who score above the cut-off threshold on the SAT (or ACT or GRE) and thus do not need to pass the Praxis CORE possess greater family wealth and educational background but not necessarily stronger reading and writing skills than the typical NJCU student. The problematic nature of standardized exams, in other words, is heightened by the demographic of our student body. Using them, as we now must, for admission into our teacher preparation programs is heart-wrenching. Remember that we too felt that we wanted to raise the skills of our teacher candidates, and we also used an assessment to do that work. But our local writing assessment, unlike the SAT and the CORE, was designed solely to identify areas of deficiency so that we could then address them.

Moreover, we strove for accessibility and transparency. Because we know our study body, their interests and their knowledge base, we were well prepared to choose topics that would allow the students to best display their skills. When we chose a poor essay topic (which did happen, of course), it was always obvious from the responses. For this reason, we typically piloted new topics as midterms in the 3-credit class where they served only as formative assessments and did not “count.” For example, one reading centered around the idea of low pay for teachers as an impediment to teacher quality in the field of education; our students didn’t tackle this topic well. They lacked background knowledge about the range of salaries for different professions, including teaching, and the mindset of thinking about education as a poorly paid profession. Indeed, the middle-class bias of the article was confusing for the students and detracted from their ability to demonstrate their writing skills. An article about sexual education in schools was right in their wheel house; they wrote with confidence and enthusiasm about the necessity for schools to address young adult sexuality!

In aiming at transparency, we provided, on the university website, in the Writing Center, and in the developmental classes, multiple opportunities for students (and faculty) to see what we were looking for. Students could place requests online for copies of their failing exams and scored rubrics; score appeals, which were initiated with the Associate Dean and then forwarded to committee members, were also routinized.

Again, because our goal as an educational institution was for our students to gain proficiency rather than for us to engage in some kind of sorting exercise, we were invested in our students’ success. To be sure, there were still issues: our second language students continued to struggle as did students in certain majors, and the work of helping our students attain the skills they needed was not easy. But we felt quite proud of the fact that nearly all the students who persisted grew as writers and passed our exam.

Faced with the new reality of the Praxis CORE, our Committee began studying it carefully. There is much that could

be said about the specifics of the content of the exams. The Praxis CORE, like the new SAT, is more strongly aligned with the Common Core State Standards in an emphasis on what the standards call informational text (non-fiction) and a diverse range of readings from various disciplines (notably science and social science). The Writing portion of the CORE includes a variety of multiple-choice questions in the areas of usage, sentence correction, revision in context, and research skills. Most of these questions are relatively familiar in the standardized-testing world; questions in the research skills section are more novel, including those focused on documentation style and about validity and reliability of sources. There are two 30-minute essays: an argumentative essay written in response to a brief prompt and a “source-based” (“Core Academic” 35) essay, written in response to two brief readings (about three paragraphs). The latter, like our writing assessment exam, requires students to quote (“incorporate information”) from the sources and to draw connections (“effective links”) between the sources and other examples and details (“Core Academic” 35). The goals of the CORE, then, if not the structure of the assessment (multiple-choice grammar questions, for example) seem generally in concert with our local attempt to raise standards, with the Common Core standards, and with generally shared values about what we want teacher candidates to know and be able to do.

But does passage of the Praxis CORE really reflect “higher” academic standards? To answer this question, we tried to determine the passing scores (cut scores). We conducted a small and marginally-scientific experiment. Several members of the Writing Assessment Committee (faculty) took the exam (we also employed a math faculty member to take the math section and help us with this data analysis). For each section, we had one (presumably well-equipped) test taker complete 50% of the questions, one test taker complete 75% of the questions, and one test taker complete all the questions. Then we used the data released in our score reports (both the raw scores and the scaled scores) to extrapolate about the passing scores. This extrapolation is difficult, because each exam has a number of field test questions that do not count towards the final score. But based on our marginally-scientific research and our colleague’s admittedly simplistic and flawed analysis, we concluded that a passing score on the exam equates to about a 60% raw score. This relatively low bar presumably correlates to the top third percentile score required for exemption from the CORE.

The bar, then, doesn’t seem unreasonably high. Indeed, one of our test takers submitted only one essay (because her time ran out and she was unable to press submit in time), and she still passed the writing portion of the exam. Presumably, she did well enough on the multiple-choice writing questions to offset this writing deficiency. But her results point us back to our initial questions: how good a measure is this exam and what exactly is being measured? Will the new Praxis CORE requirement result in higher standards in teacher education and more skilled teacher candidates?

At NJCU, we faced an intellectual and pedagogical question: how does the Praxis CORE compare with our Writing Assessment in terms of raising standards? We also faced a short-term policy question: given the reality of the Praxis CORE requirement, what would we do with our writing assessment exam moving forward?

We began trying to analyze our limited data. Our numbers were small, but our results indicated that no student had failed our Writing Assessment and passed the CORE. Some of our students passed our exam but failed the CORE. Some failed both; some passed both. These results are no surprise. Remember our CORE experiment and my colleague who completed only one essay but still passed? The mix of multiple choice questions and essays on the CORE makes it a different assessment from ours: the ability to spot and correct writing errors is a different skill from the ability to write good sentences. Moreover, as discussed above, some of what the CORE is measuring is our students’ demographically determined sub-par test-taking ability and test preparation (based on their family’s income and education). Our local assessment presented a different, probably easier, and to our minds’ more valid challenge for our students than the CORE. In the end, given the state mandate of the CORE, we all agreed that continuing our local exam would be a perfect textbook example of overtesting. As a result, the committee recommended that we phase out our Assessment Program, and we did so as of December 2016.

I am far from sanguine that the new CORE will improve the skills of our certification students. Many who pass the CORE, I conjecture, will have writing issues that remain unaddressed (and may undermine them in the job market or in their work as teachers). The exam is not designed to identify the student who is generally a proficient writer but never learned to use an apostrophe. Like most national assessments, it is not designed as an instructional tool.

Most worrisome, however, is the plight of the student who fails the CORE. After all, when our students fail our local assessment, we have their essays. We can review their errors in our developmental classes, talk to our students about their strengths and weaknesses during office hours, and offer our students individual feedback and support at our Writing Center. This work is difficult, but it is the work of instruction an educational institution like ours is set up to provide. The students who fail the CORE now face a far more daunting task in getting help, and so do we as an institution in trying to support them. To start with, ETS score reports, as is typical with these sorts of standardized exams, are nearly meaningless. On the writing section of the CORE, for example, test takers receive two separate scores: one for the multiple choice section and another for the essays section. These two scores are broken down



further to indicate the raw points earned, the raw points available, and the average performance range. This reported information, however, gives failing students and their educational institutions little to go on as they attempt to improve their performance.

The test preparation industry has already begun to supply classes to support the CORE. For NJCU students, however, these classes, like the CORE exam itself, represents one more substantial financial obstacle that deepens the inequality on the educational playing field. As teacher Karen Lewis writes in *More Than A Score* about her experience with Kaplan test-prep classes, “kids who already had advantages would now [with these classes] have even more” (80). NJCU has already begun to put in place some free test-preparation classes, but we are not a test-prep service provider, so, especially given the opacity of the tests and the score reports, I am not wholly optimistic about our ability to help every student who persists to succeed.

## What I Wish I'd Known

I am a firm believer in public education, and I strongly support the idea that our nation's teachers should be highly-skilled, particularly in the area of literacy. Colleges of Education, in particular, have a responsibility to produce these highly-skilled teachers, which is no small task given the fact that many of our strongest students steer clear of the politically-demonized and relatively low-paying world of public education.

At NJCU, when we began our Writing Assessment Program, we naively hoped to be able to make a difference: to raise standards while providing a surfeit of support to help our students to reach these higher standards. The task was never easy. It was expensive and labor-intensive, and we did not begin with nor did we do enough to cultivate the broad institutional understanding of our project necessary for its success. In particular, we underestimated the degree to which some members of our community saw any and all testing as a form of gatekeeping.

Moreover, even in the successful second phase of our program, we never created a community of shared values around literacy. For example, we hoped faculty in the College of Education would use our rubric with some of their class assignments. They did not. For some skeptical faculty and students, our program's expectations of writing proficiency remain unreasonable and unnecessary.

Strong administrative leadership and a committed group of faculty and staff certainly made a difference: our local assessment helped our students improve their literacy skills. But as our local initiative is subsumed by the national push towards higher standards and the Praxis CORE, I fear that we will find it significantly more difficult, both at NJCU and at schools like ours, to train first-generation, under-represented college students to be the highly-skilled teachers our public school children need and deserve. As a result, I fear it will become more difficult for us to ameliorate the broader inequities in K-16 public education and in our society today.

**Acknowledgments:** I am appreciative to Greg Giberson, Thomas Sura, and the readers at *Composition Forum* for their wonderful support and feedback on this work. I am also grateful to Elise Lemire and Caroline Wilkinson for their suggestions. The Writing Assessment Program would never have thrived without the support of Allan De Fina, Dean of the College of Education, Lourdes Sutton, Associate Dean, and my many colleagues in the enterprise, especially Elba Herrero, Alex Kim, Tracy Amerman, Sai Jambunathan, Irma Maini, Matt Sutton, Tamara Cunningham, Michael Basile, and Ann Wallace.

## Appendices

1. [Appendix 1 - Sample Exam Reading Prompt: Sample Exam Questions](#)
2. [Appendix 2 - Rubric](#)
3. [Appendix 3 - Course Guidelines and Sample Syllabi](#)
4. Appendix 4 - Sample Exams and Rubrics
  - a. [Sample Exam A and Rubric A](#)
  - b. [Sample Exam B and Rubric B](#)
  - c. [Sample Exam C and Rubric C](#)
5. [Appendix 5 - Tips](#)

## Works Cited

Anson, Chris M. “Closed Systems and Standardized Writing Tests.” *College Composition and Communication*, vol.

60, no. 1, Sep. 2008, pp. 113-128.

CCCC Committee on Assessment. "Writing Assessment: A Position Statement." *College Composition and Communication*, vol. 46, no. 3, Oct. 1995, pp. 430-437.

CCCC Committee on Assessment. "Writing Assessment: A Position Statement." *Conference on College Composition and Communication*, Nov. 2006, Revised March 2009, Reaffirmed November 2014, Oct. 1995, <http://www.ncte.org/cccc/resources/positions/writingassessment>.

Condon, William. "Looking beyond judging and ranking: Writing assessment as a generative practice." *Assessing Writing*, vol. 12, 2009, pp. 141-156.

"Core Academic Skills for Educators: Writing." Educational Testing Service. <https://www.ets.org/s/praxis/pdf/5722.pdf>.

Dryer, Dylan B. and Irvin Peckham. "Social Contexts of Writing Assessment: Toward an Ecological Construct of the Rater." *WPA: Writing Program Administration*, vol. 38, no. 1, Fall 2014, pp. 12-39.

Elliot, Norbert. *On a Scale: A Social History of Writing Assessment in America*. Peter Lang, 2005.

Fisch, Audrey A. "The Difficulty of Raising Standards in Teacher Training and Education." *Pedagogy*, vol. 9, issue 1, 2009, pp. 142-152.

Guinier, Lani. *The Tyranny of the Meritocracy: Democratizing Higher Education in America*. Beacon Press, 2015.

Haswell, Richard and Susan Wyche-Smith. "Adventuring into Writing Assessment." *College Composition and Communication*, vol. 45, no. 2, May 1994, pp. 220-236.

Huot, Brian. "Reliability, Validity, and Holistic Scoring: What We Know and What We Need to Know." *College Composition and Communication*, vol. 41, no. 2, May 1990, pp. 201-213.

Huot, Brian, Peggy O'Neill and Cindy Moore. "A Usable Past for Writing Assessment." *College English*, vol. 72, no. 5, May 2010, pp. 495-517.

Lewis, Karen. "Testing Nightmares." *More Than A Score: The New Uprising Against High-Stakes Testing*, edited by Jesse Hagopian, Haymarket Books, 2014, pp. 77-84.

"Institutional Profile, 2014-2015." *New Jersey City University*. [http://njcu.efkgroup.com/sites/default/files/njcu\\_institutionalprofile\\_2014-2015.pdf](http://njcu.efkgroup.com/sites/default/files/njcu_institutionalprofile_2014-2015.pdf).

"Mission Statement." *New Jersey City University*. <http://njcu.edu/about/mission-statement>.

"NJCU Profile and Accreditations." *New Jersey City University*. <http://www.njcu.edu/about/njcu-profile-and-accreditations>.

"NJCU Progress Card." *New Jersey City University*. April 2016. <http://www.njcu.edu/sites/default/files/njcuprogresscardapril2016.pdf>.

Stock, Patricia and Jay L. Robinson. "Taking on Testing: Teachers as Tester-Researchers." *English Education*, vol. 19, 1987, pp. 93-121.

"Teacher Candidate Basic Skills Requirement." *New Jersey Department of Education*, Nov. 2015, <http://www.state.nj.us/education/educators/rpr/preparation/BasicSkillsExemptionCutScores.pdf>.

"Profile of a Local Response to a Systemic Problem" from *Composition Forum 37* (Fall 2017)

© Copyright 2017 Audrey Fisch.

Licensed under a [Creative Commons Attribution-Share Alike License](https://creativecommons.org/licenses/by-sa/4.0/).



Return to [Composition Forum 37 table of contents](#).

